Chapter 12

Speech motor timing and fluency

Peter Howell, Andrew Anderson, and Jorge Lucero

Abstract

Smith's spatio-temporal index (STI) is widely used to assess variability in motor timing performance in various speaking conditions. STI has been shown to be a sensitive index of developmental changes and as a way of assessing performance differences between speakers with fluency disorders and controls. STI typically takes records obtained from an articulator (e.g., the lower lip) for repeated attempts at the same utterance. STI aligns the set of records linearly, normalizes the amplitude axis, and obtains the standard deviation at 50 points along the aligned time axis which are then averaged to give the index. This chapter uses the functional data analysis (FDA) method that aligns features on the time axis (it is a non-linear method). FDA allows separate estimates of timing and amplitude deformations. When two or more signals are obtained concurrently on utterances (here the lower lip and speech energy are used), the timing deformations can be compared to estimate their degree of inter coordination. The method is described and applied to see whether a group of speakers who stutter have poorer inter coordination than a group of fluent speakers.

12.1 Introduction

This chapter describes work on speech motor timing and fluency. Its main aims are: (1) to describe a newly developed method for measuring interarticulator coordination (inter coordination); and (2) to apply these methods to see whether a group of speakers who stutter have poorer inter coordination than a group of fluent speakers. The methods are a development of functional data analysis (FDA), which can be used to measure the variability across a set of records when a speaker attempts to repeat the same utterance a number of times (Lucero et al., 1997; Ramsay and Silverman 1997).

12.1.1 Estimating timing variability using FDA

To estimate timing variability in speech, usually ten or more records that represent repetitions of the same articulatory movement are obtained. Due to natural variation in speech, similar events are likely to be misaligned in their relative timing. FDA is used to manipulate the time lines of records nonlinearly to bring their features into alignment across the set. Variation in the degree of adjustment necessary to bring the records into alignment provides an estimate of temporal variability over the course of the records. After the set of records has been aligned, differences on the amplitude axis provide an estimate of amplitude variability over time. This allows separate statistics to be formulated indexing temporal and amplitude variability in an analogous way to that in Smith's spatio-temporal index (STI) (Smith et al., 1995). Phase variability can be indexed by estimating the standard deviation (*sd*) in phase adjustments over the normalized time axis and averaging across the extract. The *sd* of the amplitudes can be processed in a similar way to index amplitude variability (see also Lucero 2005). The procedure for obtaining the STI score is basically the same as that described for FDA except that the adjustments to the time base are linear (proportional stretching or squeezing) for STI rather than aligned according to features in the signal. Consequently, the amplitude variability cannot be separated from the timing variability. STI is a statistic that reflects joint temporal and amplitude variation.

Fluent speech requires a set of quasi-autonomous articulatory systems to work in coordination. For example, when a voiceless plosive is produced, the speaker has to control the pulmonary and laryngeal systems to ensure there is a transition from aspirated to voiced excitation and this has to be coordinated with the vocal tract opening gesture. If a speaker repeated a phrase involving voiceless plosives, as described above, and records associated with the laryngeal, pulmonary, and vocal tract systems were obtained, each record could be processed independently by FDA. If the three components were exactly coordinated, the time deformations would be the same. When coordination is poor (as would occur when a speaker is dysfluent or when materials are selected which are hard to produce), there will be some departure from this situation. The analysis procedure described here uses the time deformations to obtain a statistic that indicates the inter coordination between different articulators.

As has been seen, multiple records associated with a single activity (speech or otherwise) are required to apply the procedures necessary to obtain inter coordination. Here we used a lower lip (L) kinematic record obtained directly by using a movement transducer and a record derived from the audio waveform, the amplitude of the energy envelope (E). The lower lip record does not require further introduction as it is widely used in contemporary studies of speech variability (Smith and Goffman 2004). The amplitude envelope is computed from an audio recording obtained concurrently. It is a continuous signal that mainly represents the pulmonary system, but also includes the activity of the laryngeal and vocal tract systems. Because of its integrated nature, it incorporates aspects of all the major structures that must be controlled in order to produce speech, so if problems in control of any articulator occur across ages or clinical groups, it should be present in this record.

12.1.2 Inter coordination ability in speakers with fluency problems

In this section, we outline how articulatory coordination has featured in some perspectives about stuttering. Several authors have hypothesized that speakers with fluency problems, such as stuttering, have difficulty with inter coordination compared to fluent speakers. For instance, Alfonso and van Lieshout (1997) reported that speech gestures are not coordinated appropriately for production by speakers who stutter. Max and Caruso (1997) also used syllable- and phrase-level estimates for assessing temporal performance in detail. However, in a study of the relative timing of acoustic measures of stop gap and voice onset time, stuttering and fluent individuals were not found to differ significantly (Max and Gracco 2005).

Other groups have also hypothesized that speakers who stutter have inter coordination problems that lead to difficulties (Packman et al., 1996). Recently, some of these authors in their syllable initiation hypothesis of stuttering (Packman et al., 2007) proposed that fluency-enhancing techniques such as syllable-timed speech reduce levels of linguistic stress, which is thought to trigger stuttering in speakers who stutter. Support for these hypotheses was drawn from studies that showed that decreasing speech motor variability reduced stuttering in adults (Packman et al., 1994) and preschool children (Packman et al., 1992).

All the above hypotheses focus on variability in the motor system. Howell's (2002, 2004) EXPLAN model includes linguistic and motor factors and predicts that there are specific things the motor system does that cannot be performed by the linguistic system (such as inter coordination). Some general details about EXPLAN are required before the hypothesis that inter coordination would affect fluency in this model is discussed.

EXPLAN maintains that linguistic processing (PLAN) and motor programming (EX) are involved during speech production. These operate together and have to be synchronized in time for speech to proceed fluently. An utterance starts with the speaker planning the first element. When the PLAN is complete, it is input to EX where a fluent motor programme is generated and the speech is output. While the first element is being programmed for output, PLAN of the next element occurs. As we have said, speech is fluent when the successive plans are sent for motor programming on time, so that the sequence can proceed continuously. This process is most likely to fail and speech to be dysfluent when there is a sequence where the element currently being executed is programmed for output rapidly, and/or the next element takes a long time to plan. Sequences like this only allow a short planning-time and can arise because PLAN is slow and/or EX is rapid.

A single junction where planning is slow and prior execution is rapid occurs in a prosodic word (PW) in which the obligatory content word is preceded by at least one function word (Selkirk 1984), as in a sequence such as 'the swing'. The properties of the elements at the quick-to-execute/slow-to-plan junctures determine whether speech is fluent or not. When problems arise, they can be dealt with in one of two ways: (1) Stalling, where speech before the juncture is interrupted by pausing or repeating one or more of the preceding, already planned, words; (2) Advancing speech and using the completed part of the linguistic plan on the difficult word coming up to generate motor output. This leads to dysfluencies on parts of words, such as part-word repetitions, prolongations, and word breaks.

According to EXPLAN, the major questions to address are what affects: (1) language planning timing; and (2) motor programming timing. To date, most work has been conducted on factors affecting language planning. Here we look at inter coordination of lip movement (L) with the speech amplitude envelope (E) in speakers who stutter or who are fluent. Next we describe a method we have developed to assess inter coordination. We then apply this procedure to assess inter coordination between lip and energy records and to see whether this differs between people who stutter and fluent controls. We consider a number of specific questions about the temporal variability measures for L and E signals. First, do the speakers with high temporal variability for L also tend to have high temporal variability on E? This would be the case if variability affects all speech systems, not any particular one. We addressed this question by examining whether temporal variability indices in L and E (temporal quantities like these indices use the term 'phase') correlate for each of the participant groups. Second, we asked whether temporal variability in L or E differed between the participant groups. Third, we explored whether inter coordination between L and E differs between the participant groups.

12.2 **Method**

12.2.1 **Participants**

There were 24 participants who were secondary referrals to a specialist stuttering clinic in London. They were confirmed as stuttering by a speech-language pathologist. There were 12 participants who stuttered, with a mean age of 14 years 9 months (*sd* was 3 years 3 months). Two were female and the rest were male. Stuttering Severity Instrument version three, SSI-3, (Riley 1994), scores ranged from very mild to severe. There were 12 fluent participants (seven females and five males), with a mean age of 16 years 11 months (*sd* was 6 years 7 months).

12.2.2 **Procedure**

Participants repeated the phrase 'buy Bobby a puppy' 20 times at their normal speech rate in a quiet room. At least ten fluent repetitions were made. Participants repeated the phrase as exactly as possible, paying attention to timing and amplitude control. Participants adopted their most comfortable rate and level. Lip movement and speech records were obtained concurrently. The microphone-to-mouth distance was monitored for each participant by the experimenter to ensure it was kept at a constant distance of approximately 20cm. Any amplitude fluctuations due to slight differences in microphone-to-mouth distance or because of different voice levels of participants were removed when *z*-transformations were made. Phrases that contained word repetition, phrase repetition, pause, prolongation, part-word repetition, and word break were omitted from the analysis.

12.2.3 **Lower lip movement and energy envelope data processing**

The lower lip movement signal (L) was obtained by a cantilever/strain gauge transducer arrangement. The transducer was suspended from a headcage that consisted of an adjustable low-mass tubular aluminum assembly (Barlow et al., 1983). The apparatus matched that described by Abbs and Gilbert (1973) and the headcage superstructure and transducers were positioned and stability checks were made in the same way as reported by Barlow et al. (1983).

The strain gauge output was connected to an integrated circuit socket whose output was passed through an amplifier and low-pass filtered (four-pole Butterworth with a cut-off of 10 Hz) and then captured by a PC for signal processing. An extra channel on the converter recorded speech (transduced by a Sennheiser K6 microphone). The output from the microphone was low-pass filtered at 3.5kHz through a four-pole Butterworth filter. The strain gauge and speech signals were each sampled at 8kHz.

The original audio oscillogram and L track were uploaded into Speech Filing System (SFS) files (http://www.phon.ucl.ac.uk/resource/sfs/). The audio record was processed after capture to obtain the energy over time envelope (E). E was obtained by rectifying and low-pass filtering the signal at 15 Hz. The track of E over time was calculated at every millisecond along the waveform as the average of the sum of the amplitude values within that millisecond frame.

The oscillogram, L and E records can be displayed in alignment, one beneath the other in SFS. The onset of the first /b/ in the phrase was located on the oscillogram and the point where voicing ceased in 'puppy' was marked, again using the oscillogram. These start- and end-points were used as pointers to the L and E tracks in order to select the appropriate section of the respective record for analysis. Subsequently each L and E records were amplitude normalized by transforming to z-scores (as is standard for the STI).

12.2.4 **FDA Registration**

Non-linear time normalization (also termed registration) using FDA is described before the extension, which allows inter articulator coordination to be estimated. Readers who are already familiar with FDA procedures can go directly to the section entitled 'Phase and amplitude deformations and inter coordination'.

It is assumed that variability in speech motor control can be decomposed into independent components of amplitude and time, and that the difference between observed and expected articulation patterns can be determined in terms of amplitude-time deviations across the length of the observed records. Thus, an observed pattern $O_i(t)$, $i = 1,2,...,N$, can be modelled as:

$$O_i(t) = E[t + \phi_i(t)] + \beta_i(t) \tag{12.1}$$

Where t is the closed interval [0,1] (records were linearly normalized in time to satisfy this requirement, as is done with the STI), $\phi_i(t)$ is the expected (estimated) target pattern, and $\beta_i(t)$ and $\phi_i(t)$ are the set of temporal and amplitude deformations between the observed and the expected pattern.

$E(t)$, $\phi_i(t)$, and $\beta_i(t)$ can be estimated by optimizing a set of non-linear time transformations (or warps) $h_i(t)$ that best align the set of observed records, whilst satisfying the constraints that transformed time: (1) is smooth (parts of records are not unreasonably stretched or squashed); and (2) increases monotonically (i.e., parts of records are not reordered).

The degree of alignment can be measured in terms of the variance in the amplitude displacement pattern across the set of aligned records. An aligned record is obtained from $O_i[h_i(t)]$) and the form of $h_i(t)$ can be constrained by expressing it in terms of the second order differential equation:

$$\frac{d^2h_i(t)}{dt^2} = w_i(t)\frac{dh_i(t)}{dt} \tag{12.2}$$

The terms in the equation can be rearranged and it can then be seen that $w_i(t)$ is the ratio of the curvature of the transform to its incline. Therefore, manipulation of $w_i(t)$ allows control over smoothness (and by penalizing large values in the optimization, unwanted undulations in the time transformation can be restricted).

The standard approach to solving equation 12.2 is to integrate it twice giving:

$$h_i(t) = C_0 + C_1\int_0^t [\exp\int_0^u w_i(v)dv]du \tag{12.3}$$

C_0 and C_1 are set so they keep $h_i(t)$ in the range[0,1].

With this solution, records can be aligned by optimizing $h_i(t)$ to minimize the cost function given in equation 12.4. This function is formed from the variance in amplitude of the aligned records and a roughness penalty $\lambda\int_0^1 w_i(t)^2$. λ stipulates the penalty (large λ inhibits undulation in the time transform). λ was set at 0.001 as done by Lucero et al. (1997).

$$Cost = \sum_{i=1}^{N}\int_0^1 [O_i[h_i(t)] - E(t)]^2 dt + \lambda\int_0^1 w_i(t)^2 dt \tag{12.4}$$

The expected pattern is the mean of the aligned records as follows:

$$E(t) = N^{-1}\cdot\sum_{i=1}^{N} O_i[h_i(t)] \tag{12.5}$$

As cost is defined by $E(t)$, which is estimated using $h_i(t)$ minimization must follow an iterative process with the following two steps repeated until the differences between successive costs is negligible: (1) recalculate $E(t)$; (2) re-optimize $h_i(t)$. To make optimization tractable, the standard approach is to represent $\int_0^u w_i(v)dv$ in equation 12.3 in terms of a linear summation of B-splines. For further information, see Ramsay and Silverman (1997) and Lucero et al. (1997).

MATLAB functions that perform all processing discussed are available at ftp://ego.psych.mcgill.ca/pub/ramsay/FDAfuns/.

12.2.5 Phase and amplitude deformations and inter coordination

After the records have been non-linearly aligned, the phase and amplitude deformations of each record from the target can be calculated ($\phi_i(h_i(t))$ and $\beta_i(h_i(t))$ respectively). The phase deformation is given by the difference in actual time and transformed time as in equation 12.6 (see Lucero 2005).

$$\phi_i(h_i(t)) = t - h_i(t) \tag{12.6}$$

The amplitude deformation for two kinematic signals is obtained by differencing the amplitude of each aligned record and the expected record as shown in equation 12.7:

$$\beta_i(h_i(t)) = O_i[h(t)] - E(t) \tag{12.7}$$

The amplitude deformation for these signals is not reported in this study as our focus is on timing.

If records are available for two signals p and q, then if each is aligned independently, the temporal inter coordination can be estimated by comparing their phase deformations. Temporal inter coordination (also termed asynchrony) may be estimated as the difference between phase deformations (equation 12.8):

$$\alpha_i(h_i(t)) = \phi_{i,p}(h_i(t)) - \phi_{i,q}(h_i(t)) \tag{12.8}$$

where a positive value indicates that an event occurs relatively later than planned in p than q. $(\alpha_i(h_i(t))$ is the temporal offset necessary to bring into synchrony with p.

12.2.6 Indices of variability

The statistic that summarizes variability in timing or asynchrony is denoted here by the function *IV*. This corresponds to the average variability in the respective measure over the time line. Variability in phase is indicated in equation 12.9, and variability in amplitude would be given by replacing ϕ with α. $IV(\phi)$ is equivalent to the Index of Phase Variability (IPV) in Lucero (2005) and is similar to the STI (Smith et al., 1995)

$$IV(\phi) = (N-1)^{-1} \cdot \sum_{i=1}^{N} \int_{t=0}^{1} [\phi_i(h_i(t)) - \overline{\phi}(h_i(t))]^2 \tag{12.9}$$

Figure 12.1 shows an example of FDA registration for the E (left) and lip displacement (right) signals. The panels are the same for the two signals and are described for the E signal on the left. The superimposed tracks are given in the first row and these same tracks after non-linear registration are given in the second row. Phase variability is shown in the third row. Here, the x-axis represents a linear scaling of time so that all records fit on the same (arbitrary) time frame. The y-axis is the non-linear deformation of the x-axis resulting from FDA registration. If the records were identical, this would be a single line with a slope of one. The width of the stripe gives a visual impression of phase variability across records. Phase variability is quantified as the average stripe width as indicated in equation 12.9. The last row gives the mean of the aligned records over transformed time with +/− 1 *se* indicated.

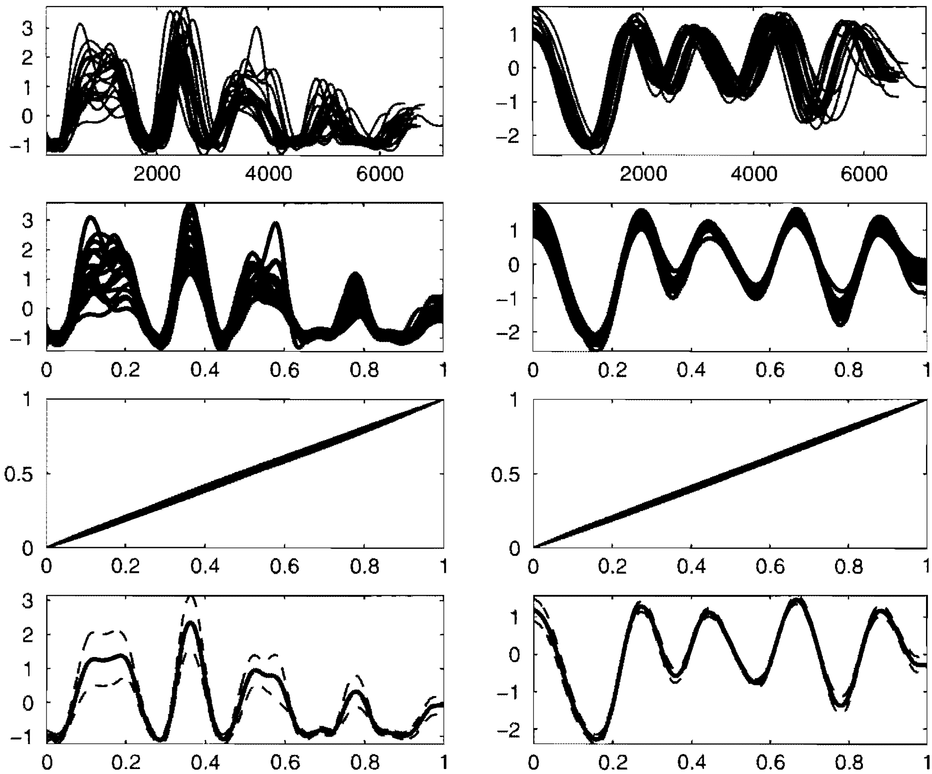

Fig.12. 1 Example results for FDA registration. The left column corresponds to speech energy and the right column to lower lip displacement. The first row are the observed records overlaid (low pass filtered at 15Hz and amplitude transformed to z-scores). Units on the *x*-axis correspond to sample number (frequency 8 kHz). The second row shows the same records post registration [$O_i h(t)$]. Units on the *x*-axis correspond to the artificial time line [0,1] as for the bottom two rows. The third row plots the new time lines [$h_i(t)$] on the *y*-axis against *t*. If there were no warping the transform would be a straight diagonal line gradient 1. The bottom row shows the mean of the aligned records [$E(t)$] +/− standard deviation in amplitude. Units on the *y*-axis are *z*-scores.

12.3 **Results**

The results reported are: (1) correlation of phase variability in L against E; (2) the differences in L and E phase variability indices for the participant groups; and (3) inter coordination between L and E for the participant groups.

12.3.1 **Correlation of phase variability in lips and energy across participants**

Figure 12.2 gives plots of phase variability for E (abscissa) against phase variability for L (ordinate). Each point represents one participant, and the participant group can be identified from the caption in the inset. The correlations were significant for fluent participants ($r = 0.79$, df = 10, $p = 0.002$, 2-tail) and participants who stutter ($r = 0.64$, df = 10, $p = 0.03$, 2-tail). Participants who have high temporal variability on one measure also tend to have higher temporal variability on the other.

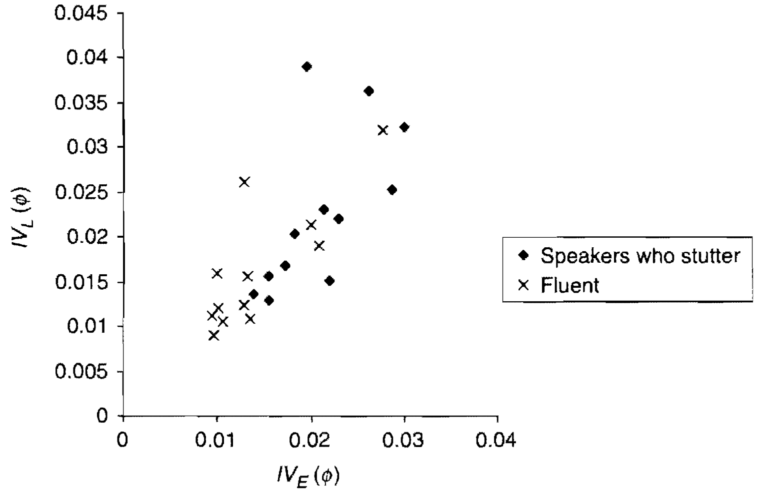

Fig. 12.2 Correlation between phase variability in energy and lips.

12.3.2 **Differences between fluency groups in phase variability between L and E**

The L phase variability of the participants who stuttered differed by t-test from the participants who were fluent ($t = 1.938$, df = 22, $p = 0.03$, 1-tail). Similarly, the mean E phase variability of the participants who stuttered differed by t-test from the participants who were fluent ($t = 2.92$, df = 22, $p < 0.005$, 1-tail). These results are shown in Fig. 12.3 where group mean and +/− 1 *se* are plotted for E (abscissa) and L (ordinate).

12.3.3 **Inter coordination differences between fluency groups**

The mean phase inter coordination scores between L and E for each participant group and +/−1 *se* are plotted on the abscissa of Fig. 12.4. A *t*- test showed that the participant groups differed

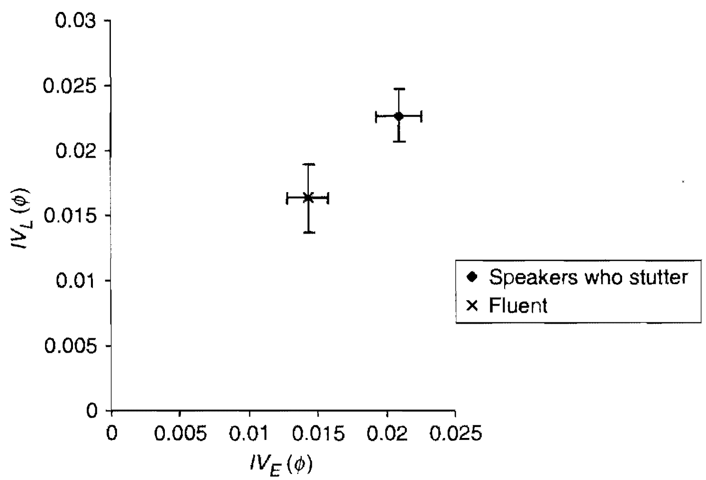

Fig. 12.3 Comparison of mean +/− *se* (top) phase variability in energy vs. lips.

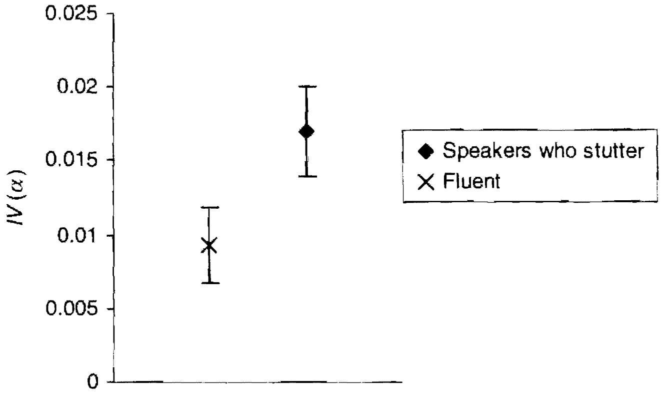

Fig.12.4 Comparison of mean +/− *se* phase asynchrony.

significantly in their ability to inter coordinate the temporal component of FDA ($t = 2.278$, $df = 22$, $p = 0.02$, 1-tail).

12.4 **Discussion**

Three results were reported: (1) Phase variability in L and E correlated (both for the groups of participants who stuttered and controls); (2) Phase variability was significantly different across participant groups; and (3) Inter coordination between L and E differed between the participant groups.

The first finding may be interpreted as showing that when a participant's timing is variable, it affects diverse structures (L and E) not specific ones (L or E). Second, the higher variability in people who stutter than in controls indicates that the control of the lower lip and the mechanisms that generate the amplitude envelope are more problematic for participants who stutter than for fluent controls. Third, the novel method for assessing inter coordination (here between L and E) shows that participants who stutter are more variable in the way they coordinate lip and pulmonary control. The latter finding, in particular, offers support for theories which propose that a motor deficit alone (Alfonso and van Lieshout 1997; Max and Caruso 1997; Packman et al., 1992, 1994, 1996), or in conjunction with a language system deficit (Howell 2002, 2004), occurs in people who stutter. The latter view allows language factors (which are well documented) to affect stuttering as well as motor ones like inter coordination. This comprehensiveness comes at the price of: (1) specifying how the independent language and motor systems link together; and (2) how variability in motor performance could affect fluency. The EXPLAN account of these matters is considered briefly.

Howell (in press) addressed the first of these issues. He argued that gestures serve as output from the PLAN system and that they drive the programming in the EX system. According to this view, gestures are symbolic representations in the language planning system and specifications of the goals of motor programming. In EXPLAN, motor-programming time has been treated as proportional to language planning time for simplicity. For example, in Howell's (2007) spreading activation version of EXPLAN, motor programming is directly related to planning time. In this account, time for activation to build up represents planning, and this activation decays at the same rate over the time that the word was executed. Function and content words had different decay and activation rates. The goal of this work was to show how serially organized inputs representing the words in a PW could lead to elements being triggered out of order and how this could lead to stalling and advancing dysfluency patterns. A threshold activation level was stipulated which automatically specified when a motor programme had to be initiated and what programme from the elements in the current PW should be selected. The selection procedure allows an entire past word, or part of the next word, to be selected for motor programming which gives rise to stalling and advancing behaviour depending on parameter settings such as number of function words prior to the content word and relative activation rates of the function and content words. The model also allowed motor factors to influence fluency by varying speech rate (e.g., when there is a long sequence of function words, rate will accelerate). The current results suggest that motor timing may fluctuate more in speakers who stutter. This could be incorporated into the model by treating execution as a stochastic process (the variability of which must be taken into account in planning) and examining how word sequencing is affected by modifying motor variability. This would also offer insight into whether motor timing impairments alone serve as a possible explanation for the symptoms of stuttering.

Turning to some general issues about the procedure, comparison across kinematic and audio records is possible providing the temporal signal can be obtained separately. In turn, this commends use of FDA registration (decomposition of temporal and amplitude information is not possible with STI). Further justification for use of FDA registration is that it gives more accurate estimates of amplitude variability for synthetic kinematic sequences where the outcome is known (Lucero 2005). Average articulator trajectories calculated from non-linearly aligned records also appear closer to the pattern seen in original records than those aligned linearly (Lucero et al., 1997). FDA and possibly the inter coordination statistic are useful for looking at other fluency problems. For example, Anderson et al.(2008) have used FDA coordination to examine McHenry's (2003) hypothesis that Parkinson Disease and ataxic patients are more variable on different dimensions. Anderson et al. (2008) found Parkinson Disease patients had more variability on the amplitude domain, and ataxic patients more on the temporal domain, relative to controls.

The method needs to be generalized to estimate inter coordination for other articulatory and acoustic measures of speech control. Strictly speaking, the current results favour a motor contribution to problems of speech control in participants who stutter and they rule out theories which exclude such a possibility. As no examination has been made of linguistic processing, they are neutral about whether there is a linguistic problem that co-occurs with the motor problem (Howell 2002, 2004). However, Smith has investigated language-speech motor interactions in her work using a similar paradigm as was employed here (Smith and Goffman 2004). Language factors and motor factors can be manipulated factorially. For instance, the 'buy Bobby a puppy' phrase can be inserted in phrases which differ in syntactic complexity, and participants required to speak them at different rates. Such studies would help inform interaction theories like EXPLAN, which admits a contribution to fluency control for both of these levels.

Acknowledgements

This work was supported by grant 072639 from the Wellcome Trust to Peter Howell.

References

Abbs, J.H., & Gilbert, B. (1973). A strain gauge transduction system for lip and jaw motion in two dimensions: Design criteria and calibration data. *Journal of Speech and Hearing Research*, 16, 248–256.

Alfonso, P.J., & van Lieshout, P.H.H.M. (1997). Spatial and temporal variability in obstruent gestural specification by stutterers and controls: Comparisons across sessions. In W. Hulstijn, H.F.M.

Peters & P.H.H.M. van Lieshout (Eds), *Speech Production: Motor Control, Brain Research and Fluency Disorders* (pp.151–160). Amsterdam: Elsevier.

Anderson, A., Lowit, A., & Howell, P. (2008). Temporal and spatial variability in speakers with Parkinson's Disease and Friedreich's ataxia. *Journal of Medical Speech Language Pathology*, 16, 173–180.

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

Howell, P. (2002). The EXPLAN theory of fluency control applied to the treatment of stuttering by altered feedback and operant procedures. In E Fava (Ed.), *Pathology and Therapy of Speech Disorders.* (pp. 95–118). Amsterdam: John Benjamins.

Howell, P. (2004). Assessment of some contemporary theories of stuttering that apply to spontaneous speech. *Contemporary Issues in Communicative Sciences and Disorders*, 39, 122–139.

Howell, P. (2007). A model of serial order problems in fluent, stuttered and agrammatic speech. *Human Movement Science*, 26, 728–741.

Howell, P. (in press). Language processing in fluency disorders. In J. Guendouzi, F. Loncke and M. Williams (Eds). *The Handbook on Psycholinguistics and Cognitive Processes: Perspectives on Communication Disorders.* London: Taylor & Francis.

Lucero, J.C. 2005. Comparison of measures of variability of speech movement trajectories using synthetic records. *Journal of Speech, Language and Hearing Research*, 48, 336–344.

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

Max, L., & Caruso, A. J. (1997). Acoustic measures of temporal intervals across speaking rates: Variability of syllable- and phrase-level relative timing. *Journal of Speech, Language and Hearing Research*, 40, 1097–1110.

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

McHenry, M. A. (2003). The effect of pacing strategies on the variability of speech movement sequences in dysarthria. *Journal of Speech, Language, and Hearing Research*, 46, 702–710.

Packman, A., Code, C., & Onslow, M. (2007). On the cause of stuttering: Integrating theory with brain and behavioral research. *Journal of Neurolinguistics*, 20, 253–262.

Packman, A., Onslow, M., & van Doorn, J. (1994). Prolonged speech and the modification of stuttering: Perceptual, acoustic, and electroglottographic data. *Journal of Speech, Language and Hearing Research*, 37, 724–737.

Packman, A., Onslow, M., Richard, F., & van Doorn, J. (1996). Syllabic stress and variability: A model of stuttering. *Clinical Linguistics and Phonetics*, 10, 235–263.

Packman, A., van Doorn, J., & Onslow, M. (1992). Stutteirng treatments: What is happening to the acoustic signal? *Proceedings of the Fourth Australian International Conference on Speech Science and Technology*, Brisbane, Australia.

Ramsay, J.O., & Silverman, B.W. (1997). *Functional Data Analysis.* New York: Springer-Verlag.

Riley, G.D. (1994). *Stuttering Severity Instrument for Children and Adults, Third Edition.* Austin, TX: Pro-Ed.

Selkirk, E. (1984). *Phonology and Syntax: The Relation Between Sound and Structure.* Cambridge, MA: MIT Press.

Smith, A., & Goffman, L. (2004). Interaction of motor and language factors in the development of speech production. In B. Maasen, R. Kent, H. Peters, P. van Lieshout, & W. Hulstijn (Eds), *Speech Motor Control in Normal and Disordered Speech* (pp. 227–252). Oxford: Oxford University Press.

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