STUTTERING TREATMENTS: WHAT IS HAPPENING TO THE ACOUSTIC SIGNAL?

Ann Packman, Janis van Doorn, and Mark Onslow
School of Communication Disorders
The University of Sydney

ABSTRACT: Identifying the critical variables in stuttering treatment will increase treatment effectiveness and contribute to understanding the nature of the disorder. A research program is described which investigates the notion that changes in the variability of acoustic segment durations, in particular the variability of vowel length, may contribute to the success of various stuttering treatments.

INTRODUCTION

Stuttering is a disorder in which the flow of speech is disrupted by repetitions of syllables, words, and phrases, prolongation of sounds, and silent blocks (see Bloodstein, 1987, for a review of stuttering phenomenology). Stuttering usually starts before the age of five, although controversy surrounds the identification of early stuttering because of its similarity to "normal" dysfluencies which occur at this stage of speech development. It is estimated that only about 30 to 50% of young children "grow out" of stuttering (Ingham, 1983). Stuttering may interfere with a child's social development, and can impede academic and vocational progress. Individuals who stutter may report feeling frustrated and embarassed by the disorder.

The distal cause of stuttering remains a mystery, although the literature abounds with explanatory models and theories. During this century, anatomical, physiological, neuropsychological, linguistic, emotional, and cognitive explanations for the disorder have been put forward, but there has been little empirical support for this theoretical activity. Currently, there is considerable interest in explaining stuttering as a disorder of speech motor control (see Peters, Hulstijn, and Starkweather, 1991). One reason for this is the growing availability of speech analysers, and the development of sophisticated technology for the measurement of speech breathing, phonation, and articulation. An important conceptual shift to emerge from the speech motor control research is that stuttering seems to be a continuous phenomenon (Smith, 1990) rather than a series of speech "events".

One way of finding out more about the nature of stuttering involves exploring the variables which influence the occurrence and severity of the disorder. Any model which might develop from this type of research should be able to explain both treatment efficacy, and the fluctuations in severity which characterise the disorder. Indeed, the variability of stuttering is one of its most intriguing aspects. At onset, periods of stuttering may be interspersed with weeks or months of stutter-free speech. In more advanced stuttering, severity is more likely to vary from one speaking situation to another. Stuttering severity is also known to be sensitive to particular speaking conditions such as delayed auditory feedback (DAF), auditory masking, and speaking in time to a rhythm. Wingate (1969) suggested that changing the way the voice is produced is a common factor in these ameliorative speaking conditions.

TREATMENT FOR STUTTERING

It is only in the last few years that treatment for early stuttering has been advocated. Until recently, treatment was considered unneccessary because it was believed that most children "grow out of" stuttering. There was also a long standing belief that drawing attention to stuttering in young children exacerbated it. It is now clear that neither of these positions is necessarily true. Early stuttering has been shown to be amenable to simple behavioural management based on giving feedback about stuttered speech, particularly when parents are trained to do this in the child's natural environment (Onslow, Costa and Rue, 1990). Recent research suggests that this type of treatment may also be effective for children up to the age of about 12 years.

During the last 40 years, treatment for stuttering in adolescents and adults has exploited the ameliorative speaking conditions described above. Portable masking devices were developed during the 1960's to extend the benefits of auditory masking outside the clinic. Syllable-timed speech, which involves saying syllables in time to a regular beat, also produced virtually stutter free speech outside the clinic. Understandably, these treatments were not very popular with the consumers! Currently, there are a number of treatment procedures which have evolved from prolonged speech, which is the stretched-out speech pattern produced when speaking under DAF. This speech pattern reduces speech rate and alters typical respiratory, phonatory, and articulatory activity. During "prolonged speech" treatment programmes, intensive and lengthy programmed instruction shapes the bizarre speech pattern towards more natural sounding speech. This novel speech pattern is then used to control stuttering outside the clinic. Despite the considerable popularity of prolonged speech as a treatment for advanced stuttering, it has long been recognised that the resulting stutter-free speech is distinguishable from that of normally fluent speakers (Ingham and Packman, 1978) and is likely to be perceived as unnatural (see Onslow and Ingham,1987, for a discussion of speech naturalness and stuttering).

Treatments for stuttering can be described as either simple or complex. Complex treatments like prolonged speech set out to alter an individual's typical speech pattern while simple treatments, based on giving feedback on the individual's speech performance, do not. There has been no research into why simple treatments are effective in reducing stuttering. There has been more research, however, into the effects of complex treatments on the acoustic speech signal. This research has been descriptive and most studies have used a similar methodology: Speech which is *perceptually stutter-free* is analysed before and after treatment to find changes in the speech pattern which might indicate how stuttering is being controlled. Earlier research indicated that speech rate reduces, and the duration of acoustic segments such as vowels and voice onset time increases, after complex treatment. These factors are thought to contribute to unnatural sounding speech after treatment (Metz, Schiavetti, and Sacco, 1990).

THE PRESENT RESEARCH

The original goals of the research program described in this paper were (i) to investigate further the critical acoustic variables in prolonged speech, and (ii) to explore how these findings relate to any acoustic changes which might occur after simple treatments. One study has been completed and a number of other studies are underway, for which preliminary findings are reported.

1. Speech samples

Apart from a study by Adams (1987) previous acoustic treatment research has analysed oral reading or very short utterances. To increase external validity in the present studies, the duration of acoustic segments is measured in spontaneous speech. The conditions under which speech samples are recorded are matched as closely as possible. In a 5min sample of spontaneous speech, every measurable segment is located. Large numbers of segments are gathered because variables such as phonetic context and speech rate are not controlled.

2. Instrumentation

High quality audiotape recordings are played on a Nakamichi CR5 cassette audiotape recorder and analysed using a Kay DSP 5500 Sonagraph. The wide band spectrogram of 1s of speech is displayed on the lower screen, and the waveform is displayed on the upper screen with a time axis of 3.12 ms/cm. Segments can be quickly and easily located by scrolling through the buffer and using the playback facility. When a segment is located, cursors are positioned on the waveform at the start and end of the segment according to guidelines described below. The cursor positions can be stored in a spreadsheet on an IBM-compatible PC linked to the Sonagraph. The time between the cursors is also calculated automatically by the Sonagraph. In one laboratory study, an electroglottograph (Kay) was used to measure intervals of voicing. In this procedure, two electrodes were held in place at the level of the vocal folds by a neckband, and voicing produced changes in impedance in the high frequency electrical signal as it passed across the glottis. The EGG signal and the audio signal were recorded simultaneously during five minute speaking sessions onto two tracks of a Sony DAT recorder (TCD10PRO).

3. Measurement of acoustic segments and articulation rate

The acoustic segments selected for measurement are vowel duration (VD), intervocalic interval (IVI), and voiced and voiceless voice onset time (VOT). Guidelines are being developed for measuring acoustic segments and articulation rate, as follows:

a. The onset of the vowel (or diphthong) is the peak (up or down*) of the first discernible waveform shape of the segment. The offset of the vowel (or diphthong) is the peak (up or down*) of the last

discernible waveform shape of the segment. (* The waveform shape for the entire vowel or diphthong is examined to ascertain whether the peaks are upward or downward going and the decision to use upward or downward peaks will apply to both onset and offset measures). Only vowels and diphthongs surrounded by consonants are measured.
b. The intervocalic interval (IVI) is the interval between the offset of one vowel or diphthong, and the onset of another vowel or diphthong, where this interval contains only voiceless articulatory activity.
c. Voice onset time is the interval from the onset of the plosive burst on the waveform, to the peak of the first subsequent discernible period of vocal fold activity.
d. Articulation rate (AR) is measured from the wideband spectrogram of stutter-free speech samples lasting 1s or more. Samples may contain pauses which are shorter than the longest IVI for the particular speaker. To calculate AR, the number of syllables *perceived to have been spoken* is divided by elapsed time.

Not all segments in spontaneous speech are measureable. The most common reasons for this are absence of an identifiable plosive burst, syllable elision, and short vowels with negligible high frequency energy. There are two sources of error : (i) the accuracy of the cursor placement and (ii) the decision about which period best represents the onset and offset of a vowel. The following procedure was used to estimate measurement error. Audiotape recordings were made of six normal speakers using the Sony DAT recorder and Sony microphone (ECM-979). Recordings were made during conversation with another person. Three age groups were represented; below 5 years, 8 to 12 years, and over 18years. There was a male and a female speaker in each of these age groups. From these recordings, a number of utterances were selected so that there would be two examples of VD, IVI, voiceless VOT (-VOT) and voiced VOT (+VOT) for each speaker. The utterances containing the 48 segments were recorded onto the Sonagraph and stored in an IBM PC using Kay software. The three investigators independently measured the duration of the segments and AR for the six speakers.

For each of the 48 segments, the measures of the three investigators for 44 (92%) segments were within 10ms. For another 3 segments (98%) the measures were within 20ms, and two out of three of these measures were always within 10ms. Percentage error is likely to be greater for shorter segments than for longer ones, so this was investigated for measures under 30ms. The only measures under 30ms were those of the 12 +VOT segments . In the most extreme case, a difference of 4.7ms on a short +VOT resulted in a percentage error of 27.3%. On the other hand, for three of the +VOT segments, the measures of the three investigators were identical. The error rate for AR was also calculated. The differences between all three investigators on their measures of AR were below 0.4 sylls/s. Overall, the findings suggest that the error in these acoustic measures is generally one period or less, but the error for +VOT may at times be large enough to confound experimental effects.

4. Studies of complex treatments

The aim of Study 1 (Onslow, van Doorn, and Newman, 1992) was to replicate previous findings that acoustic segments are longer after successful treatment with prolonged speech. In addition, this study controlled for variables not considered in previous studies by (i) using children who had not had previous complex treatment and (ii) analysing spontanous speech rather than reading or short utterances. Subjects were a group of 10 children who stuttered, ranging in age from 10 to 14 years. Programmed instruction was used to shape prolonged speech towards natural-sounding speech during the three week intensive treatment program. The acoustic measures were made from spontaneous speech samples recorded immediately before and after treatment. The predicted increase in mean segment duration and speech rate after treatment was not found. The only difference in acoustic segment duration after treatment was a decrease in the variability of VD. There was no increase or decrease in AR associated with this finding. The percentile distribution of the VD measures for the 10 subjects before and after treatment is shown in a scattergram in Figure 1. The figure shows that for the group, there were fewer very short and very long vowels after treatment. This finding was unexpected and its significance unclear, because reduced variability had not been a treatment target. Reduced variability may have been important in reducing stuttering, or it may simply have been a bi-product of the treatment, particularly the programmed instruction.

The second prolonged speech study, Study 2, is in progress. The aim of this study is to investigate the effects on the acoustic signal when prolonged speech is learned without programmed instruction. In single subject experiments, prolonged speech is used to reduce stuttering in the laboratory under controlled conditions. Acoustic and EGG recordings are being analysed in a small number of subjects. The study is exploring the relationship between changes in the duration of acoustic measures and the duration of short intervals of phonation (PI) as detected by the EGG. The frequency of very short PIs has been shown in laboratory experiments to influence the frequency of stuttering (see Gow and Ingham, 1992).

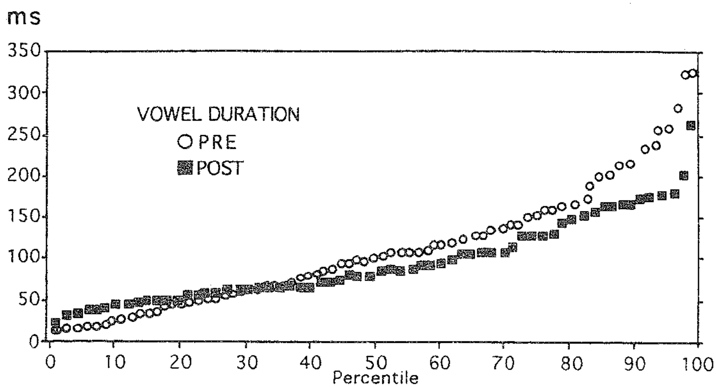

Figure 1. Percentile scattergram of pretreatment and posttreatment vowel durations for 10 subjects.

5. Studies of simple treatments

Two young children aged 4 and 5 years, whose stuttering was diagnosed by a speech pathologist, received the behavioural treatment described above (see Onslow et al, 1990). Parents learned the principles of the treatment in the clinic with the speech pathologist. Short speaking sessions were conducted by the parent each day in the home, during which the child was engaged in conversational tasks. Stutter-free speech was followed by praise and other rewards, and occasionally the child was asked to repeat a stuttered word or phrase without stuttering. The parent also gave immediate verbal feedback to the child about the child's speech at various times during the day. Stuttering frequency in both children reduced to about 1% as a consequence of the treatment, which took place over 2 to 3 months. The measurement of acoustic segment durations in these two children before and after treatment comprise Study 3. There were some changes in the duration of all segments but the most consistent were the changes to VD.

The VD measures for the two children are shown in Figure 2. The percentile scattergrams show that for one child (Subject 1) both duration and variability reduced after treatment. This reduction in variability was apparent at the high end of the range. For the other child (Subject 2), there was no reduction in overall duration, but there was a reduction in variability. This was apparent across the distribution.

These preliminary data suggest that changes to the acoustic signal may occur after behavioural treatment, as well as after treatment with prolonged speech. In particular, reduction in the variability of VD was a common finding. However, the interpretation of the results of Study 3 is problematic because of the relationship between age, speech rate, segment duration, and variability (Kent and Forner, 1980). This means that any reduction in segment duration or variability after treatment in young children may be confounded by maturation, especially if the treatment extends over a number of months. To find out more about these maturational changes and normal speech development, a study of 4 normally fluent children is underway (Study 4). Children aged 2.5, 3.5, 4.5, and 5.5 years will be studied over a six month period to get more information on the relationship between maturation and the length and variability of acoustic segments. To minimise the effects of maturation, Study 5, which is in progress, is investigating whether a group of older children, matched in age to the children in Study 1, show reduced VD variability after simple treatment for their stuttering.

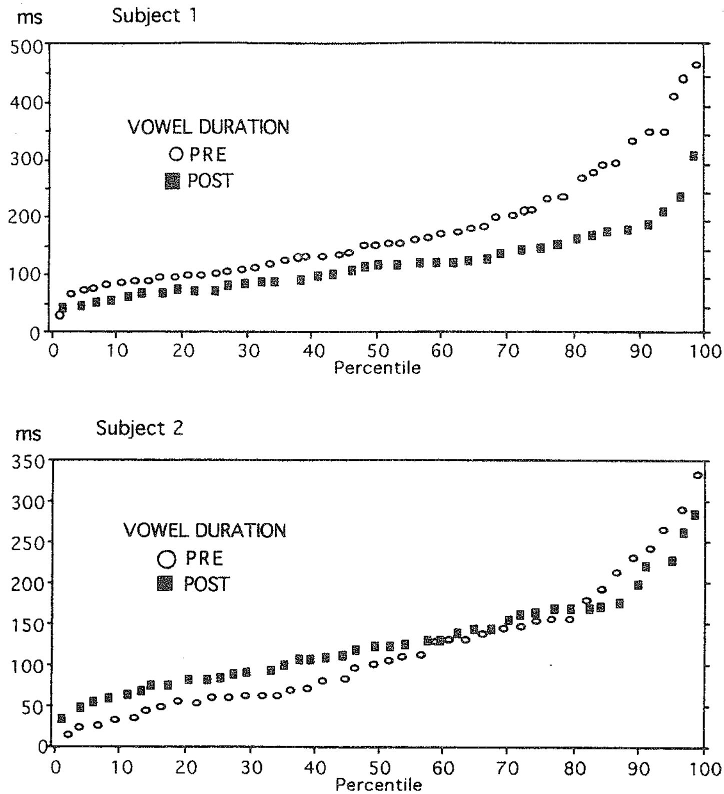

Figure 2. Percentile scattergrams of pretreatment and posttreatment vowel duration for 2 subjects.

IMPLICATIONS OF THE PRESENT RESEARCH

The impetus for the present research program came from the findings of Study 1. This study gave the first indication that reduction in the variability of acoustic segments may be associated with successful "prolonged speech" treatment for stuttering. From the preliminary findings of subsequent studies it seems that this redistribution of acoustic segments, particularly vowel length, may also be associated with simple stuttering treatment. Thus, both simple and complex treatments appear to produce similar changes in the acoustic signal, despite the fact that they are quite dissimilar in their execution and in their theoretical underpinnings.

A possible explanation for these findings is that the variability of acoustic segments co-varies with stuttering. If stuttering is a continuous phenomenon, the ongoing disruption caused by stuttering may increase the variability of acoustic segments in perceptually stutter-free speech. In this case, any treatment which brings about a reduction in stuttering severity is also likely to reduce the variability of acoustic segments. Another explanation is that there is a treatment variable (or variables) which reduces both stuttering and variability. There are many factors such as stress, speech rate, and prosody which may influence vowel length and consequently variability (Crystal and House, 1988). A third possibility is that reducing the variability of acoustic segments has a critical role in reducing stuttering. If this is the case, simple and complex treatments must somehow provide individuals with the means to reduce the variability in their speech in order to reduce stuttering. How this might occur is unknown. Discovering more about the nature of the relationship between stuttering and acoustic segment variability is the goal of future research, which will investigate variability in other ameliorative speaking conditions and in normally fluent speakers. It is hoped that further research into how individuals control their stuttering will make treatments more effective, and at the same time increase understanding of this puzzling disorder.

REFERENCES

Adams, M.R. (1987) *Voice Onsets and Segment Durations of Normal Speakers and Beginning Stutterers,* Journal of Fluency Disorders, 12, 133-139.

Bloodstein, O. (1987) *A Handbook on Stuttering (4th ed.),* (National Easter Seals Society: Chicago)

Crystal, T.H. & House, A.S. (1988) *A Note on the Variability of Timing Control,* Journal of Speech and Hearing Research, 31, 497-502.

Gow, M. L. & and Ingham, R.J. (1992) *Modifying Electroglottograph-Identified Intervals of Phonation: The Effect on Stuttering,* Journal of Speech and Hearing Research, 35, 495-511.

Ingham, R.J. (1983) *Spontaneous Remission of Stuttering: When Will the Emperor Realize He Has No Clothes On?* In Prins, D. & Ingham, R.J. (eds.) Treatment of Stuttering in Early Childhood, (College Hill Press: San Diego).

Ingham R.J. & Pckman, A. C. (1978) *Perceptual Assessment of Normalcy of Speech Following Stuttering Therapy,* Journal of Speech and Hearing Research, 21, 63-73.

Kent, R.D. & Forner L.L. (1980) *Speech Segment Durations in Sentence Recitations by Children and Adults,* Journal of Phonetics, 8, 157-168.

Metz, D.E., Schiavetti, N., & Sacco, P.R. (1990) *Acoustic and Psychophysical Dimensions of the Perceived Speech Naturalness of Nonstutterers and Postreatment Stutterers,* Journal of Speech and Hearing Disorders, 55, 516-525.

Onslow, M., Costa, L., & Rue, S. (1990) *Direct Early Intervention with Stuttering: Some Preliminary Data* , Journal of Speech and Hearing Disorders, 55, 405-416.

Onslow, M, van Doorn, J., & Newman, D. (1992) *Variability of Acoustic Segment Durations After Prolonged-Speech Treatment for Stuttering,* Journal of Speech and Hearing Research, 35, 529-536.

Onslow, M. & Ingham, R.J. (1987) *Speech Quality Measurement and the Management of Stuttering,* Journal of Speech and Hearing Disorders, 52, 2-17.

Peters, H.F.M., Hulstijn, W., & Starkweather, C.W.(eds.) (1991) Speech Motor Control and Stuttering, (Elsevier: Amsterdam)

Smith, A. (1990) *Towards a Comprehensive Theory of Stuttering: A Commentary,* Journal of Speech and Hearing Disorders, 55, 398-401.

Wingate, M. (1969) *Sound and Pattern Change in "Artificial" Fluency,* Journal of Speech and Hearing Research, 12, 677-686.