The relationship between acoustic indices of speech motor control variability and other measures of speech performance in dysarthria

Frits van Brenk and Anja Lowit

Speech and Language Therapy

School of Psychological Sciences and Health

Strathclyde University

76 Southbrae Drive

Glasgow G13 1PP

Abstract

Previous studies suggested that variability indices based on information extracted from the acoustic signal are potentially useful in assessing dysarthric speech. Because of the ease of data collection, this method is especially applicable in the clinical setting. This study assessed the relationship between variability indices of sentence repetitions obtained by Functional Data Analysis with intelligibility ratings and maximum performance tasks in groups of speakers with hypokinetic dysarthria and ataxic dysarthria. The results showed significant correlations between selected parameters, which varied with dysarthria type. For the speakers with ataxic dysarthria, the variability measure mainly reflected differences in intelligibility, while for the group with hypokinetic dysarthria, there was a stronger relationship between variability indices and DDK performance. Lack of stronger correlations between variability measures and intelligibility ratings and maximum performance tasks are possibly due to heterogeneity of severity across and within speaker groups. This study provides further evidence that variability measures such as the FDA might be sensitive to speech performance of speakers with dysarthria, and can potentially differentiate between dysarthria types.

Introduction

Researchers and clinicians are always looking for objective and reliable ways to measure speech impairment. This is important in order to characterize the speech disorder and compare with others, as well as to measure change in speech performance due to degenerative processes or improvement in therapy. There are many ways of assessing the nature and severity of speech disorders. The most commonly used clinical tools are intelligibility scores and maximum performance tasks (Duffy, 1995; Hustad, 2008). Current research is aimed at finding more refined measures of articulation performance, for example long-term average spectra (Utianski, Liss, Lotto, & Lansford, 2012), voice onset time (Fischer & Goberman, 2010), formant Centralization Ratio (Sapir, Ramig, Spielman, & Fox, 2010) and vowel space (Kim, Hasegawa-Johnson, & Perlman, 2011).

Another measure of articulation performance is the variability index. This index is a representation of the variability of a particular articulatory movement, e.g. lip opening and closing movements, over a series of repetitions of the same movement sequence. Smith, Goffman, Zelaznik, Ying, & McGillem (1995) developed the Spatiotemporal Index (STI) to measure variability in lower lip movements in younger adults. Lucero, Munhall, Gracco, & Ramsay (1997) developed a nonlinear version of the STI, the Functional Data Analysis (FDA) of speech properties. FDA enables the separation of variability in the temporal dimension from variability in the spatial dimension. The temporal dimension represents the timing of actions, while the spatial dimension represents the degree of articulator excursion. While articulatory variability was initially obtained by tracking lip movements (Smith & Goffman, 1998), Howell, Anderson, Bartrip, & Bailey (2009) succeeded in extracting similar information from the acoustic signal, such as root sum squared amplitude and formant variation. Compared to kinematic data collection and analysis, information obtained from

acoustic data has the advantage of portability, making it more suitable for the clinical setting. Preliminary studies suggest that variability indices have the potential to differentiate between disorder types and severities in dysarthric speech (Anderson, Lowit, & Howell, 2008; McHenry, 2003, 2004). However, more research is necessary to establish their usefulness as a clinical or research outcome measure.

In order to investigate this further, this study aimed to correlate FDA measures taken from audio recordings with common severity measures, in this case intelligibility ratings and maximum performance tasks, in two types of dysarthric speakers.

Method

Participants

Two groups of speakers participated. The first group consisted of 23 speakers with Parkinson's Disease (PD) and mild to moderate hypokinetic dysarthria (18 male, 5 female; age 40-81, M = 66.8, SD = 10.6 years). The second group consisted of 8 speakers with mild to severe ataxic dysarthria (AD) due to a variety of neurological disorders, predominantly Multiple Sclerosis, cerebellar and spino-cerebellar ataxia (5 male, 3 female; age 37-70, M = 49.5, SD = 12.5 years). The presence and severity of dysarthria in speakers of each group had been assessed by the referring speech and language therapist (SLT) or by the second author where participants were not recruited through the SLT services.

Tasks and Instrumentation

All participants engaged in a series of speaking tasks. DDK tasks of /pə/, /tə/, /kə/ and /pətəkə/ were performed to assess variability in maximum movement rates. Intelligibility was

assessed by reading a set of unpredictable sentences (McHenry & Parle, 2006), reading the Grandfather passage (Darley, Aronson, & Brown, 1975) and a monologue about a holiday. Approximately 20 repetitions of the phrase "Tony knew you were lying in bed" were recorded for FDA variability analysis. As all three intelligibility tasks correlated well with each other and showed no differences in their relationship to the other measures, only the results of the monologue will be reported here. Audio recordings were taken using a portable wave recorder (Edirol R-09HR) and a head-mounted condenser microphone (AKG C-420) spaced about 4 cm from the speaker's mouth.

Analysis

For the DDK tasks the rate of repetition as well as the coefficient of variation (CoV) of syllable durations was calculated to capture movement speed and variability.

15 undergraduate SLT students participated in the intelligibility judgments. They orthographically transcribed the unpredictable sentences on which basis the average number of correctly transcribed words per sentence was calculated. The Grandfather Passage and a fragment of the monologue were scored using a nine-point scale incorporating intelligibility and listener effort (Dobinson, 2007). These scores were subsequently converted into percentage values.

For the variability analysis, the intensity envelope (SPL), fundamental frequency envelope (F0) and first (F1) and second formant envelope (F2) were extracted with Speech Filing System (University College London, http://www.phon.ucl.ac.uk/resource/sfs/), and subsequently the spatial and temporal variability were calculated using FDA registration software (McGill University, http://www.psych.mcgill.ca/misc/fda/software.html), illustrated by figure 1. Further procedural details can also be found in Anderson et al. (2008).
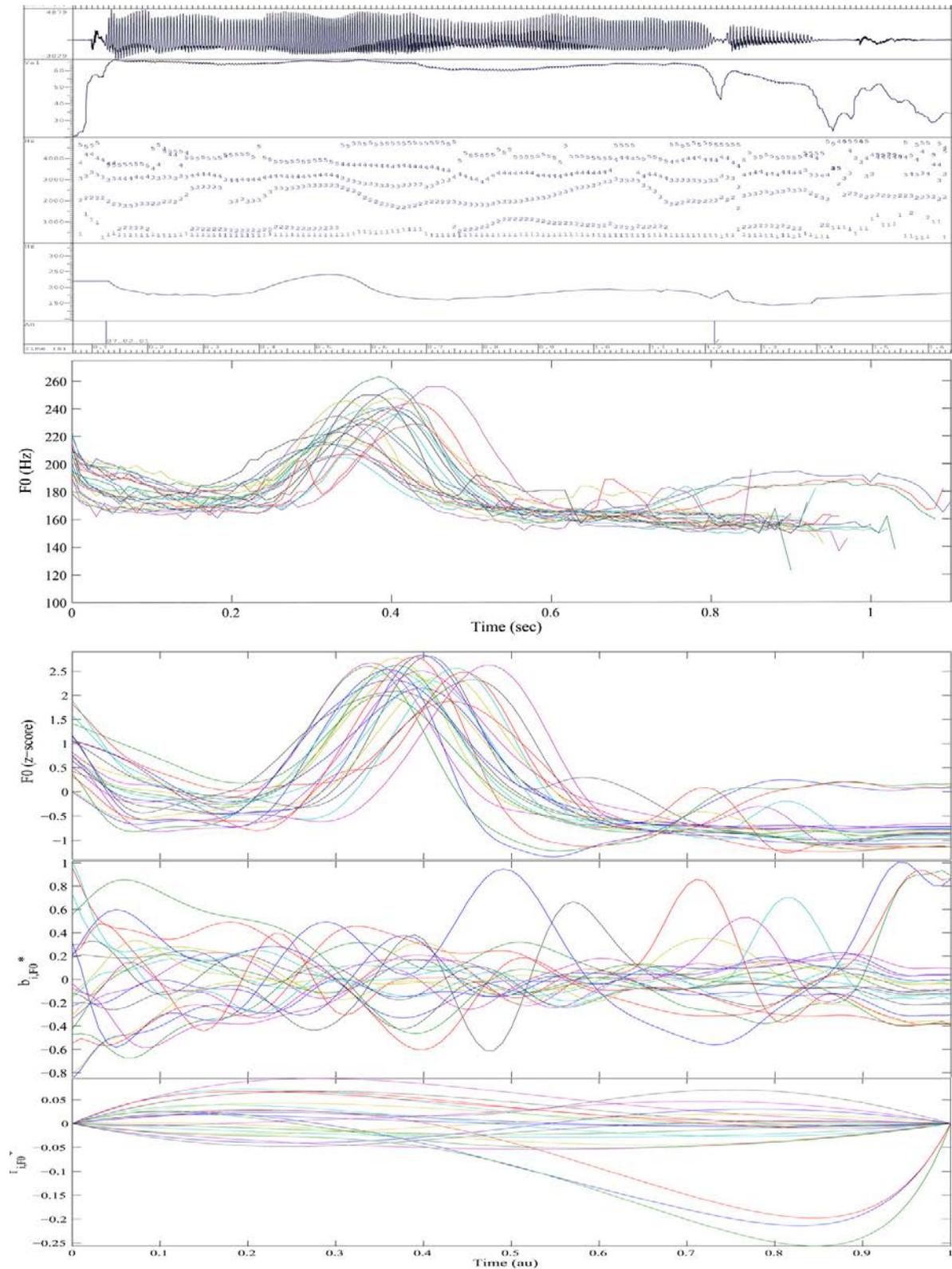
Figure 1: Example results of FDA variability analysis. The first panel displays the SFS annotation interface showing respectively the waveform, amplitude envelope, F1-F4 envelopes, F0 envelope and annotation grid of the sentence "Tony knew you were lying in bed". The second panel displays the F0 envelopes of approximately 20 repetitions. The third panel shows the F0 envelopes after normalizing and FDA registration. The fourth panel displays the pattern error, i.e. the distortion in amplitude to align all records in the spatial direction. The bottom panel displays the phase error, i.e. the distortion in time to align all records in the temporal direction.

6

The variability results for the two speaker groups were compared by a repeated measures analysis of variance with Group as between-group factor, and speech parameter (SPL, F0, F1 and F2), separately for Temporal Variability and Spatial Variability. A one-way analysis of variance was carried out to compare the two speaker groups in the intelligibility task and the four DDK tasks. A linear regression analysis was performed to assess the degree of correlation between the FDA analysis and the DDK and intelligibility assessments.

Results

Group Comparison

First, the AD group was compared with the PD speakers across the various measures. There was no significant difference in intelligibility scores, indicating that the groups were comparable based on the most global measure of severity. Despite this, some differences emerged for the other analyses. There was a marginal difference for DDK performance, where the PD group showed greater variability of syllable durations (CoV) than the AD speakers ($F (1, 29) = 4.94$, $p = .034$), but only for /pə/. No significant differences could be found for the other syllable types. While there was some suggestion that the PD group might be more variable in the DDK task, the variability indices for sentence repetition showed the opposite. Both for spatial variability, and in particular, temporal variability, the AD group had significantly higher FDA values than the PD speakers (Group effect for Spatial Variability: $F (1, 29) = 4.17$, $p = .050$, Temporal Variability: $F (1, 29) = 8.32$, $p = .007$).

Correlation Analyses

There was no significant correlation between the intelligibility ratings and the CoV of syllable durations in the DDK task.

The comparison of the FDA analysis with the DDK performance yielded few significant results for the AD group, which showed a positive correlation only between Temporal Variability of F2 and /kə/: r (6) = .768, p = .026. More significant results were found for the PD group. A negative correlation was present between Spatial Variability of F0 and /pə/: r (21) = -.464, p = .026. On the other hand, a positive correlation was present between Temporal Variability of F1 and /kə/: r (21) = .607, p = .002 as well as /pətəkə/: r (21) = .542, p = .008, and of F2 and /pə/: r (21) = .422, p = .045. The positive correlations suggest that higher variability in sentence repetitions was reflected in greater variance across syllable repetitions.

The comparison of the FDA analysis with the intelligibility ratings yielded significant results only for the AD group. The results showed a negative correlation between the monologue intelligibility rating and Temporal Variability of F0: r (6) = -.825, p = .012, as well as Temporal variability of F2: r (6) = -.814, p = .014, suggesting that higher variability was associated with lower intelligibility.

Discussion

The FDA analysis showed significantly higher temporal and spatial variability for the AD group than the PD group despite similar levels of severity, as suggested by the lack of difference in intelligibility ratings. Further group differences emerged in the relationship between the results of the FDA analysis and the other speech tasks. The FDA results were a better prediction for intelligibility ratings in the AD group, compared to the PD group. On the other hand, the PD group showed a closer relationship between the FDA results and

variability in the DDK tasks. This result was reflected to some degree in the greater variability of /pə/ syllable durations found in the PD compared to the AD group.

Closer investigation of the measurement parameters that correlated best revealed that for the AD group, an increase in Temporal Variability of F0 and F2 was associated with a decrease in intelligibility ratings, showing that variability of pitch and tongue movements across repetitions of a sentence might be associated with intelligibility in this client group. In relation to the DDK tasks, an increase in syllable duration variability was associated with an increase in Temporal Variability of F1 for the PD group and of F2 for the PD and AD groups. This indicates that measurements of temporal articulatory control of tongue body height and frontness can reflect syllable duration variability in DDK tasks. These results are in line with results from a study by McHenry (2003), who found a relationship between the Spatiotemporal Index of lower lip movements and severity of dysarthria. No significant relationship was found between DDK performance and intelligibility ratings, confirming earlier reports of limited relationship between a single dimension of articulatory accuracy in DDK and overall intelligibility (Kent, Kent, Duffy, Thomas, Weismer, & Stuntebeck, 2000; Ozawa, Shiromoto, Ishizaki, & Watamori, 2001).

Although results were not as clear as expected, the comparison of FDA analysis with DDK performance and intelligibility ratings indicated that selected parameters correlated significantly with each other in at least one of the speaker groups. It is difficult to determine exactly why the type of dysarthria seems to have an influence on correlations, but this was most likely due to the fact that there was a larger spread of severities across the AD group. The higher degree of variability found in the DDK performance of the PD group was also surprising given that the literature generally associates poor movement control more with ataxic rather than hypokinetic dysarthria. However, previous studies have not investigated the

CoV in DDK tasks extensively, and this measure might thus be interesting to investigate further in future.

In conclusion, the current study provides further evidence that variability measures such as the FDA might be sensitive to overall speech performance of speakers with dysarthria, and have the potential to show differences between different types of dysarthria (see also Anderson et al., 2008). In addition, our previous reports on these speakers have shown that the FDA analysis was more sensitive than other speech measures in differentiating the speakers with dysarthria from healthy controls (Brenk & Lowit, 2011). Although these results will have to be confirmed with further studies on different dysarthria types and with more clearly delineated severity groups, the current results indicate that FDA analyses are a promising tool to add to the current understanding of speech motor control in disordered populations. Whilst this premise extends to clinical use as a tool for more refined differential diagnosis and treatment outcome measure, one has to caution that currently the analysis cannot be sufficiently automated to allow a quick and easy, but reliable analysis by clinicians for everyday use. More developments in signal processing are thus required to turn this method into a fully functioning clinical tool.

References

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*(4), 173-180.

Darley, F. L., Aronson, A. E., & Brown, J. R. (1975). *Motor speech disorders*. Philadelphia: Saunders.

Dobinson, C. (2007). *An investigation into the use of computers for conducting home practice of speech exercises for people with dysarthria.* Unpublished PhD Thesis, Newcastle University, Newcastle upon Tyne.

Duffy, J. R. (1995). *Motor speech disorders: Substrates, differential diagnosis, and management*: St. Louis: Mosby.

Fischer, E., & Goberman, A. M. (2010). Voice onset time in Parkinson disease. *Journal of Communication Disorders, 43*(1), 21-34.

Howell, P., Anderson, A., Bartrip, J., & Bailey, E. (2009). Comparison of Acoustic and Kinematic Approaches to Measuring Utterance-Level Speech Variability. *Journal of Speech, Language and Hearing Research, 52*(4), 1088-1096.

Hustad, K. C. (2008). The Relationship Between Listener Comprehension and Intelligibility Scores for Speakers With Dysarthria. *Journal of Speech, Language and Hearing Research, 51*(3), 562-573.

Kent, R. D., Kent, J. F., Duffy, J. R., Thomas, J. E., Weismer, G., & Stuntebeck, S. (2000). Ataxic Dysarthria. *Journal of Speech, Language and Hearing Research, 43*(5), 1275-1289.

Kim, H., Hasegawa-Johnson, M., & Perlman, A. (2011). Vowel Contrast and Speech Intelligibility in Dysarthria. *Folia Phoniatrica et Logopaedica, 63*(4), 187-194.

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*(5), 1111-1117.

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*(3), 702-710.

McHenry, M. A. (2004). Variability Within and Across Physiological Systems in Dysarthria: A Comparison of STI and VOT. *Journal of Medical Speech - Language Pathology, 12*(4), 179.182.

McHenry, M. A., & Parle, A. M. (2006). Construction of a Set of Unpredictable Sentences for Intelligibility Testing. *Journal of Medical Speech - Language Pathology, 14*(4), 269.

Ozawa, Y., Shiromoto, O., Ishizaki, F., & Watamori, T. (2001). Symptomatic Differences in Decreased Alternating Motion Rates between Individuals with Spastic and with Ataxic Dysarthria: An Acoustic Analysis. *Folia Phoniatrica et Logopaedica, 53*(2), 67-72.

Sapir, S., Ramig, L. O., Spielman, J. L., & Fox, C. (2010). Formant Centralization Ratio: A Proposal for a New Acoustic Measure of Dysarthric Speech. *Journal of Speech, Language and Hearing Research, 53*(1), 114-125.

Smith, A., & Goffman, L. (1998). Stability and patterning of speech movement sequences in children and adults. *Journal of Speech, Language and Hearing Research, 41*(1), 18-30.

Smith, A., Goffman, L., Zelaznik, H. N., Ying, G., & McGillem, C. (1995). Spatiotemporal stability and patterning of speech movement sequences. *Experimental Brain Research, 104*(3), 493-501.

Utianski, R. L., Liss, J. M., Lotto, A. J., & Lansford, K. L. (2012). *The use of long-term average spectra (LTAS) in discriminating dysarthria types.* Paper presented at the Conference on Motor Speech, Santa Rosa, CA, USA.

Van Brenk, F, & Lowit, A. (2011). *Assessing dysarthria using variability measures from audio recordings.* Paper presented at the 6th International Conference on Speech Motor Control, Groningen, The Netherlands.