

Lowit, Anja and Anderson, Andrew and Howell, Peter (2008) Temporal and spatial variability in speakers with Parkinson's Disease and Friedreich's Ataxia. *Journal of Medical Speech-Language Pathology*, 16 (4). pp. 173-180. ISSN 1065-1438

<http://strathprints.strath.ac.uk/26453/>

Strathprints is designed to allow users to access the research output of the University of Strathclyde. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. You may not engage in further distribution of the material for any profitmaking activities or any commercial gain. You may freely distribute both the url (<http://strathprints.strath.ac.uk>) and the content of this paper for research or study, educational, or not-for-profit purposes without prior permission or charge. You may freely distribute the url (<http://strathprints.strath.ac.uk>) of the Strathprints website.

Any correspondence concerning this service should be sent to The Strathprints Administrator: eprints@cis.strath.ac.uk

Published in final edited form as:

J Med Speech Lang Pathol. 2008 December ; 16(4): 173–180.

Temporal and spatial variability in speakers with Parkinson's Disease and Friedreich's Ataxia

Andrew Anderson¹, Anja Lowit², and Peter Howell¹

¹*University College London Gower St., London WC1E 6BT*

²*University of Strathclyde*

Abstract

Speech variability in groups of speakers with Parkinson's disease (PD) and with Friedreich's ataxia was compared with healthy controls. Speakers repeated the same phrase 20 times at one of two rates (fast or habitual). A non-linear analysis of variability was performed which used some of the principles behind the spatio-temporal index (STI). The STI usually employs variation in lip displacement over repetitions of the same utterance and a linear analysis of such signals is conducted to represent the combined variation in spatial and temporal control. When working with patients, audio measures (here we used speech energy) are preferred over kinematics ones as they are minimally disruptive to speech. Non-linear methods allow spatial variability to be estimated separately from temporal variability. The results are tentatively interpreted as showing that PD speakers were distinguished from healthy control speakers in spatial variability and ataxic speakers were distinguished from controls in temporal variability. These findings are consistent with the speech symptoms reported for these disorders. We conclude that the non-linear analysis using the speech energy measure is worth investigating further as it is potentially revealing of the differences underlying these two pathologies.

Introduction

Parkinson's disease (PD) and ataxia may lead to more variable speech performance in such patients compared to healthy controls. Previous work using variability measures has employed the spatio-temporal index (STI) introduced by Anne Smith (Smith, Goffman, Zelaznik, Ying, & McGillem, 1995). In her original work, Smith applied the STI to a lower lip kinematic signal obtained on a phrase rich in bilabials ("Buy Bobby a puppy"). Participants repeated the phrase as precisely as they could 10-20 times, and lower lip movement was measured. During analysis, the individual lip-movement records were normalized in amplitude by transforming each record to z scores and normalized in time by stretching or compressing them to a common length using a linear scale factor. The standard deviation (sd) was then obtained at 2% intervals on the normalized time axis, and the computed quantities were summed to give the STI score.

A small number of studies have used STI in speakers with dysarthria. Kleinow, Smith and Ramig (2001) investigated the motor performance of speakers with PD in different rate (habitual, slow and fast) and loudness conditions (loud and soft voice). All participant groups showed elevated STI levels in the slow rate condition. The PD speakers did not differ significantly from the age-matched healthy control group, but the PD participants were of very mild severity. The STI also showed that the loud voice condition resulted in more consistent

Address Correspondence to: Peter Howell, Department of Psychology, University College London, Gower Street, London WC1E 6BT, England. Telephone Number: +44 (0)20 7679 7566 Fax number +44 (0)20 7436 4276 Email address: p.howell@ucl.ac.uk.

articulation in all groups, lending some support to the benefits of LSVT® (Ramig, Pawlas & Countryman, 1995) treatment. In another study McHenry (2003) compared participants with different dysarthria severities across rate conditions. STI values were related to the observed differences in severity. Despite encouraging results for the use of the STI in clinical populations, insufficient data are currently available to make predictions on how STI measures are affected in speech disorders, mainly because the current methodology is too invasive to gain data from a large number of speakers. There is also the question of whether the current STI measure is sufficiently discriminative to reflect differences in the underlying pathology. This exploratory study therefore employed a different methodology that is more suited to disordered speech investigations: first, we employed audio rather than kinematic measures to estimate variability. Howell, Anderson, Bartrip and Bailey (in press) have shown that speech energy variation over time obtained directly from audio recordings (signal rectified and low-pass filtered) can be used to provide a suitable signal for estimating STI. Second, non-linear functional data analysis (FDA) procedures were used to estimate spatial and temporal variability separately (Lucero, 2005; Ramsay & Silverman 1997). The FDA procedure manipulates the time lines of signals so as to bring their features into alignment with each other. The degree of adjustment necessary to bring one signal into alignment with the remainder of the set provides an estimate of temporal variability. Once signals have been aligned, differences on the amplitude axis provide an estimate of spatial error. It is possible to provide separate estimates of temporal and spatial variability by taking the standard deviation of both temporal error and spatial error across the time line, and averaging each to provide indices analogous to those involved in the joint measure of both components in the STI.

This study compared speakers with PD and Friedreich's Ataxia (FA) as well as healthy control participants for phrase repetitions at two speech rates. Temporal and spatial variability indices were extracted separately from the audio signal to identify whether the disordered speakers showed more spatial and/or temporal variability than their respective control participants.

Method

Participants

Four speakers with hypokinetic dysarthria resulting from PD and three speakers with ataxic dysarthria resulting from FA participated. The two groups were comparable in relation to dysarthria severity, and speakers within each group presented with similar speech symptoms. Each participant was matched with a healthy control participant of the same gender and similar age. It was not possible to match across disordered groups for age and gender. However, previous research evidence on gender and age differences suggests that this would not influence the results. Background information on all participants is presented in Table 1.

Procedure

Participants repeated the nonsense phrase “well we'll will them” twenty times at their habitual rate, and 20 times twice as fast as normal at a comfortable loudness level. This provided a minimum of ten fluent repetitions for each participant in each rate condition. Participants phonated continuously over the phrase, and repeated it as exactly as possible in terms of rate and loudness. The microphone-to-mouth distance was kept constant at approximately 20cm.

Audio signal data processing

Each repetition of the utterance was examined using Speech Filing System (SFS) software (<http://www.phon.ucl.ac.uk/resource/sfs/>). Samples with discontinuous phonation, fluency breakdown or extraneous noises were excluded. For the repetitions that remained, the onset of the first /w/ and the point where voicing started on “them” were marked on oscillograms. The energy envelope (E) was calculated from the rectified and low-pass filtered (10 Hz) audio

signal as the average of the sum of the squared pressure values (at every millisecond along the waveform). FDA registration used software available from <ftp://ego.psych.mcgill.ca/pub/ramsay/FDAfuns/>.

Results

Average speech rate across conditions for PD, ataxic and healthy control participants

1) Mean duration over repetitions of the phrase, and 2) the standard deviation of durations for each speaker and rate condition, were obtained. Separate analyses were performed for each measure and for each participant/control group. The same mixed-model ANOVA was employed in all cases and had the factors speech rate (habitual/fast condition as a repeated measure within participants), and speaker category (patient group/corresponding healthy controls as an independent group factor).

For mean durations of the PD participants, the only significant difference was between fast and habitual speech rate conditions ($F(1,6) = 19.24, p < 0.01$). For the standard deviations, there were no significant differences on any of the factors or the interaction. These analyses showed that PD and control participants were comparable in terms of mean duration and its standard deviation at each speech rate and that all participants spoke significantly slower at the habitual, than at the fast, rate (to about the same extent).

For mean durations of the ataxic participants, the ANOVA revealed the only significant difference was between fast and habitual rate conditions $F(1,4) = 10.25, p < 0.04$. The lack of differences across speaker types may be because the sample size was small. The analysis of the *sds* of speech durations, showed highly significant effects of speaker group $F(1,4) = 46.92, p < 0.01$, but no other effects were significant.

Though there was no significant difference in mean duration between groups (probably because of the small N), all ataxic participants were slower than controls in the fast rate condition. Indeed, there was a significant differences between the average duration of ataxic fast rate speech and that of healthy controls' ($t = -2.862, df = 4, p = 0.046, 2$ -tail) as is apparent from Figure 1 (mean extract duration ± 1 *sd* for each of the ataxic and healthy control participants at both speech rates). A further feature that is apparent from Figure 1 is that the ataxic participants' fast rate speech was close to the control speakers' habitual rate speech. This was supported statistically: There was no difference between fast rate ataxic speech and that of healthy control speech at habitual rate ($t = -0.86, df = 4, p = 0.33, 2$ -tail). This has implications about what conditions should be compared in the FDA analysis reported later.

Registration graphs for PD, ataxic and healthy control participants

The four panels of Figure 2 show registration results for a PD patient (top left) and his control (bottom left) and an ataxic patient (top right) and her control (bottom right). Within each speaker-panel, the section in the top half is at fast rate and that in the bottom half at habitual rate. For each rate within each speaker-panel, the superimposed energy tracks are given in the first row and these same tracks after non-linear registration are given in the second row. The temporal variability can be seen in the last row of the plots. Here, the x axis represents a linear scaling of time so 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 temporal variability across records. The average stripe width quantifies the temporal variability.

The second rows of the plots at left in Figure 2 show that spatial variability in the PD participant was greater than that of the control participant at both speech rates. Spatial variability was

greater in the PD participant at fast than habitual rate whereas there was little difference in spatial variability between speech rates for the control participant. The final row of the control participant's plots indicates that he had marginally higher temporal variability at habitual rate than at fast rate. This trend was reversed in the PD participant, who showed higher temporal variability at fast, than habitual, rate. Furthermore, the PD participant showed higher temporal variability than the control at both rates. Notably there was a bigger difference between habitual and fast rate speech for the control participant.

An illustrative set of registration records for an ataxic (top) and her healthy control (bottom) at both speech rates is shown in Figure 2. Comparison of ataxic and control performance is deferred until the next section because of the average duration differences between notionally the same speech rate conditions (discussed above).

Comparison of variability indices for PD, ataxic and healthy control participants

Plots of the temporal (x axis) versus spatial (y axis) estimates are given in Figure 3 with PD/control and ataxic/control participants on the left and right respectively. 1) Spatial and 2) temporal variability were obtained for all participants and analyzed separately for each measure and for each participant/control group by ANOVA. All ANOVAs had habitual/fast speech rate condition as a repeated measure factor within participants, and patient group/control as an independent group factor. For the PD participants in Figure 3, spatial variability of speech energy appears higher in the fast rate condition than for the controls. The ANOVA showed there were significant main effects of speech rates $F(1,6) = 10.92, p < 0.02$, and of PD versus healthy control participants $F(1,6) = 13.15, p < 0.02$. Further examination using independent t tests showed that at fast rates the PD participants exhibited significantly different spatial variability to controls ($t = -5.36, df = 6, p < 0.01, 2$ tail) but not at habitual speech rate. Paired t tests of variability across speech rate conditions where rate conditions from the same participants were compared showed that PD participants had significantly different spatial variability in energy at fast, than at habitual rate ($t = 14.86, df = 3, p = 0.001, 2$ tail)

The individual profile with the highest temporal variability was shown by a PD participant. However, this was not a standard feature of the disorder as several PD participants showed similar levels of temporal variability to the control participants. The ANOVA on the temporal variability data did not show significant main effects or interactions on either factor.

It is apparent from the right-hand plot of Figure 3 that there was little to separate ataxic and control participants in terms of spatial error variability, as the results of the two groups overlap to a large extent along the y axis. The exception is the one outlier from a control participant at fast rate (this same participant also showed comparatively high temporal variability in this condition). The ANOVA confirmed this as neither of the main effects was significant in the ANOVA using spatial variability of energy.

Figure 3 suggests the ataxic group may be differentiated from the control speakers using temporal variability. However, this was not supported by the ANOVA (the only significant difference in the ANOVA for temporal error variability was between speech rates $F(1,4) = 14.69, p = .019$). The small group size and the fact that the control group included an outlier might explain the non-significance of speaker group. As the earlier analyses showed there were no differences between fast rate ataxic and habitual rate control conditions, the temporal error variability was compared between these conditions. There was a significant effect of speaker group ($t = -4.82, df = 4, p = .009, 2$ tail). Thus, there is some evidence for a temporal processing difference between ataxic and control speakers when equivalent duration conditions are compared.

Discussion

These are the first results with PD and ataxic patients that use non-linear registration techniques. The techniques distinguished the two types of participants. The PD speakers showed higher spatial variability. For the ataxic speakers and their controls, conditions which had equivalent durations differed in temporal variability.

The results for the PD and ataxic participants fit with clinical views about these disorders (note that the literature quoted for ataxic dysarthria below is based on a number of different pathologies, not just FA). Speech rate in hypokinetic dysarthria can be slower or faster than healthy control speech, with many speakers falling within the normal range (Duffy, 2005; Kent & Kent, 2000). In ataxic dysarthria, on the other hand, rate tends to be reduced (Duffy, 2005; Kent & Kent, 2000; Schalling, Hammarberg & Hartelius, 2007) and these speakers are often unable to increase rate (Kent et al., 1997; Kent & Kent, 2000). Our results reflect these differences, with the PD speakers falling within the control range. Whilst the ataxic speakers also fell within the normal range when notionally the same rate conditions were compared, there was suggestive evidence of a temporal problem (their fast rate speech was significantly slower than the controls fast rate speech but, statistically, their fast rate speech was not different from controls habitual rate speech). Moreover, when the conditions which did not differ in duration between the ataxic patients and their controls were compared, there was a significant difference in temporal variability. We caution that these findings need examining with larger group sizes.

Ataxic dysarthria is also often characterised by disruptions in speech rhythm (Schalling & Hartelius, 2004; Henrich, Lowit & Mennen, 2006; Kent, Kent, Duffy, Thomas, Weismer & Stuntebek, 2000). Speakers have also been found to vary more across repeated productions of the same utterance than control participants (Ackermann & Hertrich, 1994). These findings may reflect the suggestive temporal variability effect found in our study. The fact that the PD group did not also show greater temporal variability than the healthy speakers could be due to the fact that this group of speakers did not have any significant timing problems (their speech rate fell within the normal range, there was a significant difference between speech rates but no hint of differences in rate across PD and healthy control participants).

Although the current data are consistent with previous research evidence on speech impairment in PD and ataxia, there are a number of qualifications about the study. First, the number of participants was small. Further investigations are necessary to see whether the observed differences are consistently associated with the current groups. Second, there is the question whether the ataxic speakers should be compared across material matched for actual or notionally equivalent speech rate. There are arguments for both sides. On the one hand, as the speech rate of the ataxic speakers' fast rate condition was virtually half that of their healthy controls in some instances, the increased temporal variability may have been a by-product of the slower rate rather than a sign of an inherent disturbance of timing due to cerebellar pathology. When this problem was controlled for by comparing the ataxic participants' fast speech with the control participants' habitual rate, other influences might be involved, e.g. speech at maximum rate might involve different motor control processes than that at habitual tempo, in which case the two samples would not be comparable (Adams, Weismer & Kent, 1993).

It should also be noted that speech rate in PD participants did not differ relative to healthy control participants. The fact that PD participants specifically showed higher spatial variability than their controls commends the value of the FDA procedures used in this investigation. Future work needs to examine more patients and age- and gender-match speaker types so comparison can be made across disorders.

Acknowledgements

This work was supported by grant 072639 from the Wellcome Trust and a grant from the Parkinson's Disease Society, UK.

References

- Ackermann H, Hertrich I. Speech rate and rhythm in cerebellar dysarthria: An acoustic Analysis of Syllabic Timing. *Folia Phoniatica et Logopedica* 1994;46:70–78. [PubMed: 8173615]
- Adams SG, Weismer G, Kent RD. Speaking rate and speech movement velocity profiles. *Journal of Speech, Language, and Hearing Research* 1993;36:41–54. [PubMed: 8450664]
- Duffy, JR. *Motor Speech Disorders. Substrates, Differential Diagnosis and Management*. Vol. 2nd ed.. St Louis, MO: Mosby; 2005.
- Henrich J, Lowit A, Schalling E, Mennen I. Rhythmic disturbance in ataxic dysarthria: A comparison of different measures and speech tasks. *Journal of Medical Speech Language Pathology* 2006;14:291–296.
- Howell P, Anderson AJ, Bartrip J, Bailey E. An acoustic method for measuring variability in speech tokens. *Journal of Speech, Language and Hearing Research*. in press
- Kent RD, Kent JF. Task-based profiles of the dysarthrias. *Folia Phoniatica et Logopedica* 2000;52:48–53. [PubMed: 10474004]
- Kent RD, Kent JF, Duffy JR, Thomas JE, Weismer G, Stuntebek S. Ataxic dysarthria. *Journal of Speech, Language and Hearing Research* 2000;43:1275–1289. [PubMed: 11063247]
- Kent RD, Kent JF, Rosenbek JC, Vorperian HK, Weismer G. A speaking task analysis of the dysarthria in cerebellar disease. *Folia Phoniatica et Logopedica* 1997;49:63–82. [PubMed: 9197089]
- Kleinow J, Smith A, Ramig LO. Speech motor stability in IPD: effects of rate and loudness. *Journal of Speech, Language, and Hearing Research* 2001;44:1041–1051.
- Lucero JC. Comparison of measures of variability of speech movement trajectories using synthetic records. *Journal of Speech, Language and Hearing Research* 2005;48:336–344. [PubMed: 15989396]
- McHenry M. The effect of pacing strategies on the variability of speech movement sequences in dysarthria. *Journal of Speech, Language, and Hearing Research* 2003;46:702–710.
- Ramig, LO.; Pawlas, AA.; Countryman, S. *The Lee Silverman Voice Treatment (LSVT®): A practical guide to treating the voice and speech disorders in Parkinson disease*. Iowa City: National Center for Voice and Speech; 1995.
- Ramsay, JO.; Silverman, BW. *Functional Data Analysis*. New York: Springer-Verlag; 1997.
- Schalling E, Hartelius L. Acoustic analysis of speech tasks performed by three individuals with spinocerebellar ataxia. *Folia Phoniatica et Logopaedica* 2004;6:367–380. [PubMed: 15557775]
- Schalling E, Hammarberg B, Hartelius L. Perceptual and acoustic analysis of speech in individuals with spinocerebellar ataxia (SCA). *Logopedics Phoniatics Vocology* 2007;32(1):31–46. [PubMed: 17454658]
- Smith A, Goffman L, Zelaznik HN, Ying G, McGillem C. Spatiotemporal stability and the patterning of speech movement sequences. *Experimental Brain Research* 1995;104:493–501. [PubMed: 7589300]

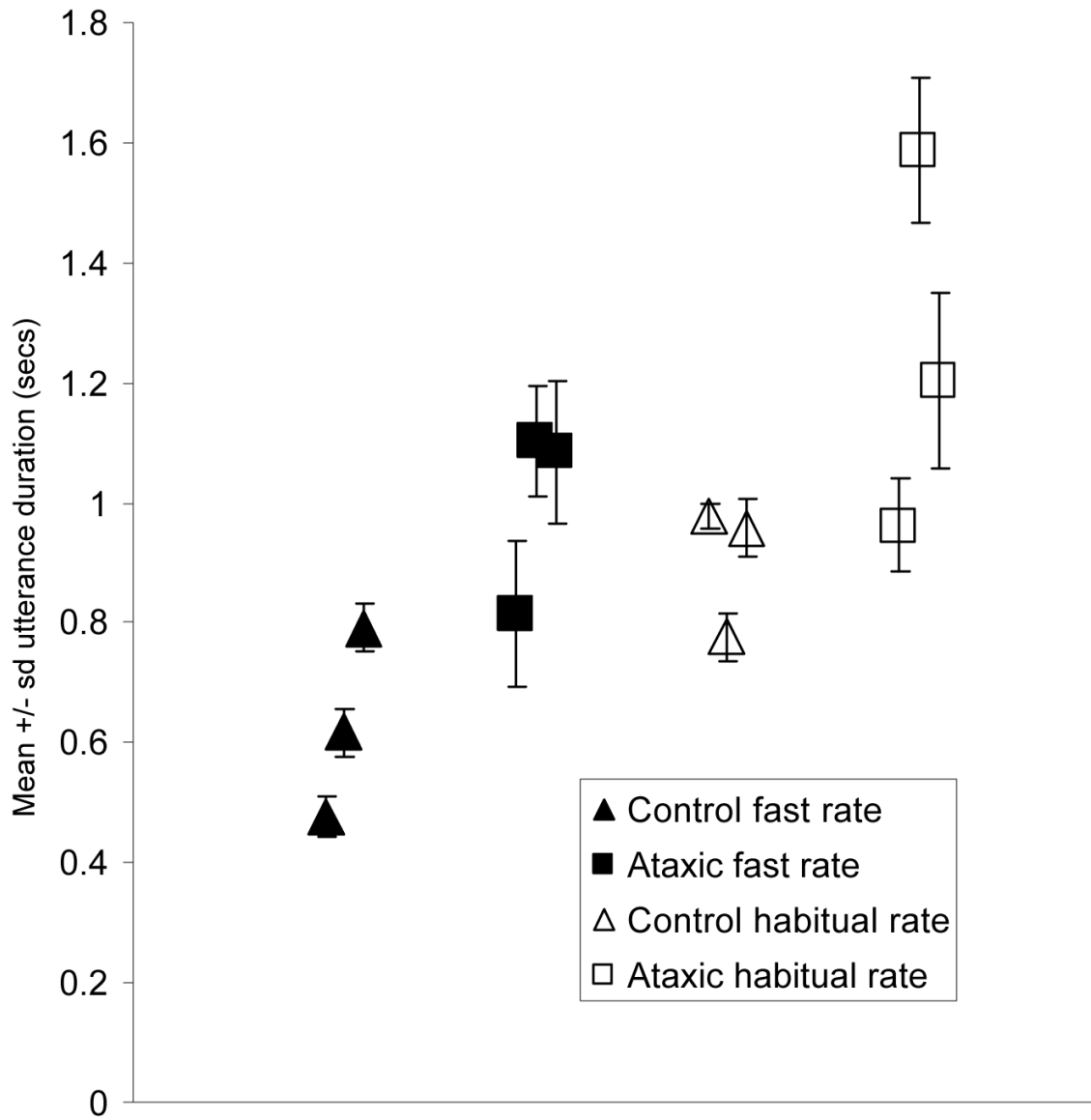


Figure 1. Mean phrase duration ± 1 *sd* for each of the ataxic speakers and their healthy controls at both speech rates. Participant type and rate condition can be identified by the key in the inset.

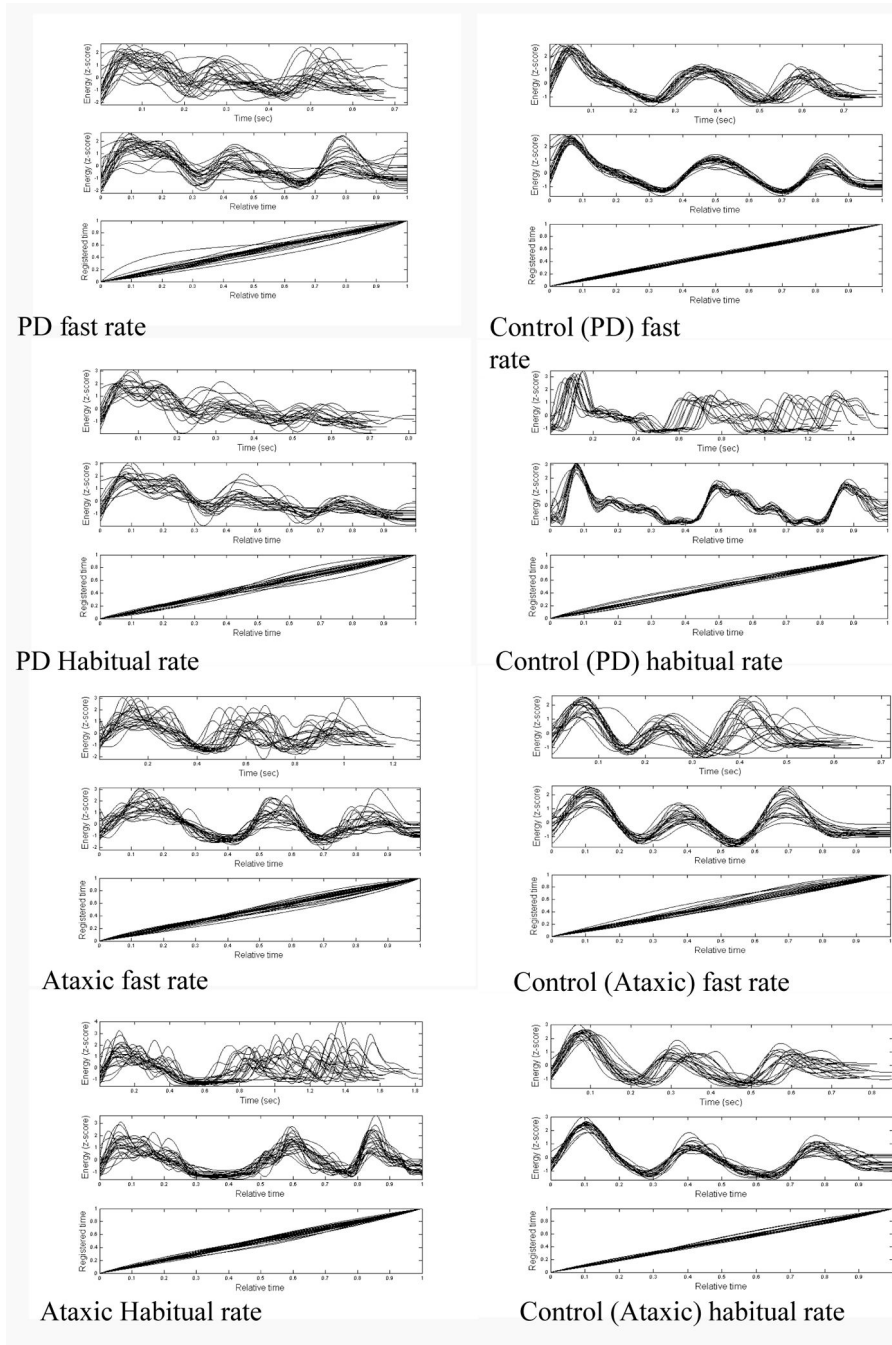


Figure 2. FDA registration graphs for a PD participant (top left) and his healthy control (top right) and for an ataxic participant (bottom left) and her healthy control (bottom right). For each of the four participants shown, fast rate is at the top and habitual rate at the bottom. For each speaker by rate panel, the top row shows the superimposed energy tracks, the second row shows these same tracks after non-linear registration. In the last row, the x axis represents a linear scaling of time so all records fit on the same (arbitrary) time frame and the y axis is the non-linear deformation of the x axis resulting from FDA registration. Full details of what is plotted are given in the text.

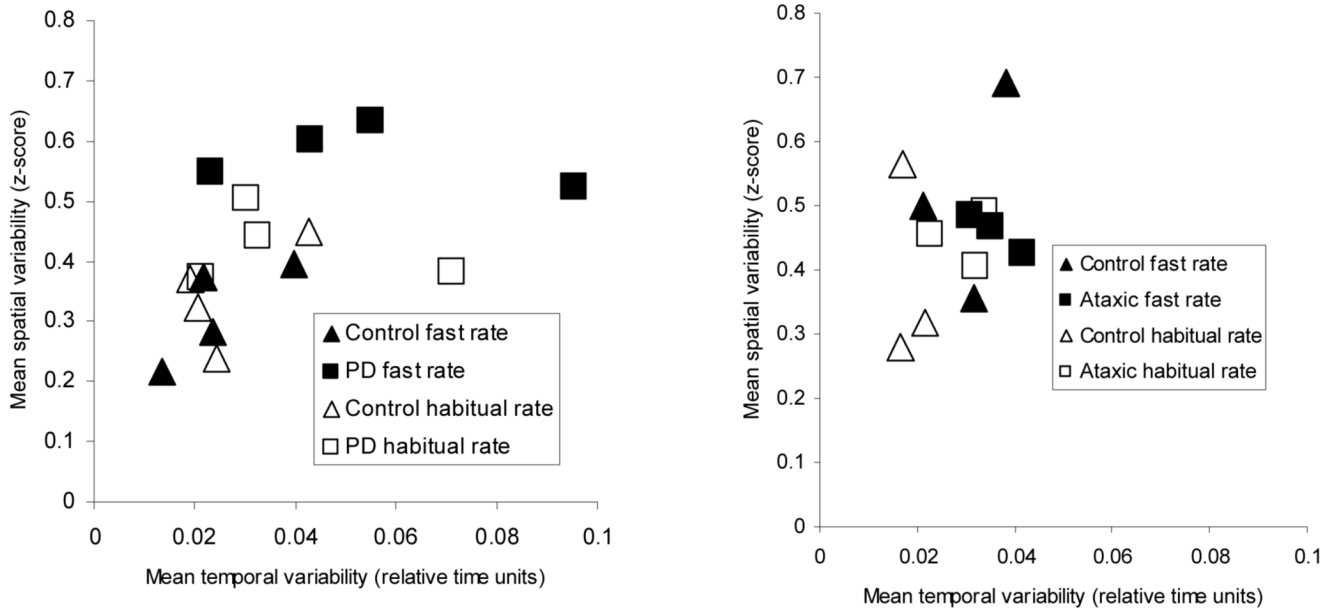


Figure 3. Temporal (abscissa) versus spatial variability (ordinate) for each participant and rate condition for the PD participants and their controls (left) and ataxic participants and their controls (right). Participant type and rate condition can be identified by the key in the inset.

Table 1

Participant information including disease designation (where appropriate), age, gender, and speech characteristics.

Speaker	Code	Age	Gender	Speech Characteristics
Parkinson's Disease (PD)	PD1	52	Male	Mild to moderate intelligibility problem, slightly accelerated tempo, normal volume
	PD2	46	Male	Mild intelligibility problem, normal tempo, normal volume
	PD3	66	Male	Mild to moderate intelligibility problem, slightly accelerated tempo, normal volume
	PD4	64	Male	Mild intelligibility problem, normal tempo, normal volume
PD control	CON1	58	Male	NA
	CON2	46	Male	NA
	CON3	66	Male	NA
	CON4	60	Male	NA
Friedreich's Ataxia (FA)	FA1	39	Female	Mild to moderate intelligibility problem, slightly reduced tempo, mild rhythmic abnormality
	FA2	51	Female	Mild intelligibility problem, slow tempo
	FA3	55	Female	Mild intelligibility problem, slow tempo, mild rhythmic abnormality
FA control	CON5	38	Female	NA
	CON6	51	Female	NA
	CON7	54	Female	NA