Can we detect altered disease state in early stage Huntington’s Disease using acoustic markers of speech?

Vogel, A.P.1*, Shirbin, C.2, Churchyard, A.J.2, Stout, J.C.2

1 Speech Neuroscience Unit, Melbourne Medical School, University of Melbourne, Australia
vogela@unimelb.edu.au
2 School of Psychology and Psychiatry, Monash University, Melbourne, Australia

Abstract

Altered prosody has been described in symptomatic Huntington’s Disease (HD) individuals, however, the extent to which acoustic analysis of speech is sensitive to gene positive pre-manifest (PreHD) individuals is unknown. Speech samples were acquired from 30 individuals carrying the mutant HTT gene (13 PreHD, 17 early stage HD) and 15 matched controls. Participants read a passage, produced a monologue, counted from 1-20 and said the days of the week. Data were analysed acoustically for measures of timing, frequency and intensity. Tasks were compared to determine their relative sensitivity to disease state. Tasks with greater cognitive demand appeared more suitable for detecting Huntington’s disease compared to tasks with high levels of automaticity.

Index Terms: acoustic analysis, clinical marker, Huntington’s disease

1. Introduction

Huntington’s disease (HD) is an autosomal-dominant neurodegenerative disorder caused by an expansion mutation exceeding 36 CAG repeats in the IT15 gene [1, 2], leading to whole brain atrophy and widespread neuronal degeneration of both white and grey matter, initially arising within the caudate in premanifest individuals [3]. Clinically, the disorder is characterized by progressive motor dysfunction including chorea, loss of balance and bulbar dysfunction [4-6], psychiatric disturbances [7, 8] and cognitive deficits characteristic of a subcortical or executive dementia [9, 10]. The presence of oral motor deficits [11] and language difficulties [12] observed in diagnosed patients with HD, suggests that speech may be a suitable candidate for detecting disease onset or marker of disease progression.

2. Methods

Speech samples were acquired from 30 participants carrying the mutant HTT gene and 15 controls matched with the HD group for age and sex (see Table 1). The PreHD group was defined as ‘presymptomatic’ individuals carrying the mutant HTT gene. PreHD participants required a diagnostic confidence score of equal to or less than one in the Total Motor Score [13] of the Unified Huntington’s Disease Rating Scale (UHDRS), indicating no substantial motor signs. Diagnostic confidence levels are based on clinician judgment following administration of the UHDRS and are aligned with the confidence a clinician has in diagnosing HD in a particular participant (grade 0 = no abnormalities; grade 1 = non-specific motor abnormalities) [13]. The early stage HD group included symptomatic individuals with a recent (within the past three years) diagnosis of HD made by a specialist neurologist based on a positive neurological examination in a participant with confirmed HD genetic status. Participants were included in the HD groups (PreHD or HD) if they had CAG lengths in one of the HD alleles greater than 39. Similarly, healthy controls were required to have CAG repeats less than 30. Potential participants were excluded if they presented with any other major neurological or endocrine disease other than HD; they had a history of serious alcohol or drug abuse within the previous year; they had a history of learning disability and/or intellectual impairment; or were pregnant. The burden of pathology scores were calculated using a formula developed by Penney et al [14], (age × [CAG–35–5]), that uses age and length of CAG repeat to provide an estimate of pathological burden irrespective of disease progression or symptomatology.

2.1. Stimuli

Participants provided four different speech samples: i) they read a phonetically balanced paragraph, the Grandfather passage (129 words, 169 syllables) [GRAN]; ii) produced a monologue with positive content (i.e., happy memory, amusing story, topic of interest to participant) for approximately one minute (the broad semantic category was stipulated to reduce the potential influence of emotive content on the participants’ performance) [FREE]; iii) produced an ‘automated’ sample (the days of the week) that allowed analysis of a sample that was identical for all groups, yet did not rely on reading ability [DAYS]; and iv) counting from 1 to 20 [C120]. Participants were asked to repeat the reading and automated speaking tasks to reduce unfamiliarity effects often observed in speech research [15]. Samples were recorded using a standard laptop computer and a head-mounted USB connected microphone. Head mounted microphones eliminate the affect of head movement on mouth to microphone distance. This configuration provides adequate fidelity for the acoustic measures under investigation [16].

Table 1: Demographic, clinical and genetic information of participants

<table>
<thead>
<tr>
<th></th>
<th>Controls (n=15)</th>
<th>PreHD (n=13)</th>
<th>HD (n=17)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>57.9 (12.2)</td>
<td>38.4 (8.1)</td>
<td>57 (12.2)</td>
</tr>
<tr>
<td>Sex (male)</td>
<td>8 (53%)</td>
<td>5 (39%)</td>
<td>9 (53%)</td>
</tr>
<tr>
<td>Burden of Pathology scores [14]</td>
<td>278.9 (98.2)</td>
<td>390.9 (61.6)</td>
<td></td>
</tr>
<tr>
<td>Burden of Pathology [14]</td>
<td>77.5 – 389.5</td>
<td>302.5 – 506</td>
<td></td>
</tr>
<tr>
<td>UHDRS Total Motor Score [13]</td>
<td>1.8 (1.32)</td>
<td>3.39 (4.82)</td>
<td>31.47 (14.69)</td>
</tr>
</tbody>
</table>

Data are mean (standard deviation or percentage), range of values
2.2. Acoustic Analysis

Quantitative data were extracted acoustically for measures of timing (e.g., speech rate), fundamental frequency variation (e.g., $f_0$ coefficient of variation, which is the ratio of the standard deviation about the mean fundamental frequency), and spectral tilt using the alpha ratio [comparing intensity at the higher end of the spectrum (above 1kHz) with intensity at the lower end of the spectrum (below 1kHz)] using freely available acoustic analysis software, PRAAT [17]. Given the small group size, only measures independent of speaker sex were utilised, for example direct measures of fundamental frequency were not employed because these metrics are based on laryngeal physiology. The focus on measures of timing were restricted to total silence time, total speech time (seconds) and speech rate, the percentage of silence in the sample, and speech rate (number of syllables/total signal time).

3. Results

There was a clear effect of group across many of the acoustic measures, so that speech performance often differed in-line with disease state. Non-parametric comparisons between groups revealed significant differences between the control and the early stage HD group on measures of timing including speech rate, total silence time, total speech time, and the proportion of silences produced within phrases. Trends toward significance were observed between the control and PreHD groups. Broken down by task, significant differences were observed on the monologue and reading tasks, but not the automated stimuli. Table 2 details the standardised mean difference between groups using Cohen’s d (effect size) and highlights the differences between controls and early stage HD, as well as PreHD and HD on speech rate, total speech and silence times on the reading task. A modest, non-significant difference was observed between PreHD and controls.

A similar pattern was observed during the monologue task. Early stage HD presented with significantly shorter speech times and longer pauses than controls and PreHD participants. Data also showed that differences between HD and controls were greater than differences in performance between PreHD and HD on measures of timing during extemoporous speech. The automated tasks yielded small to large effect sizes between control, PreHD and HD groups respectively. Participants in the control group spoke with a faster rate than both the HD and PreHD groups. Similarly, silences made up a smaller proportion of the sample compared to the other groups.

Table 2: Effect sizes and raw scores for speech outcomes broken down by task and group

<table>
<thead>
<tr>
<th>Measure</th>
<th>Task</th>
<th>Mean (SD)</th>
<th>Effect size (Cohen’s $d$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Control</td>
<td>PreHD</td>
</tr>
<tr>
<td>Total speech time (seconds)</td>
<td>Reading</td>
<td>35.58 (3.5)</td>
<td>36.91 (6.55)</td>
</tr>
<tr>
<td></td>
<td>Monologue</td>
<td>31.47 (2.68)</td>
<td>30.94 (3.71)</td>
</tr>
<tr>
<td></td>
<td>Days</td>
<td>3.81 (0.71)</td>
<td>4.05 (0.88)</td>
</tr>
<tr>
<td></td>
<td>Counting</td>
<td>9.72 (1.51)</td>
<td>10.5 (2.36)</td>
</tr>
<tr>
<td>Total silence time (seconds)</td>
<td>Reading</td>
<td>9.69 (3.14)</td>
<td>9.3 (2.38)</td>
</tr>
<tr>
<td></td>
<td>Monologue</td>
<td>8.23 (2.41)</td>
<td>8.52 (3.26)</td>
</tr>
<tr>
<td></td>
<td>Days</td>
<td>0.4 (0.49)</td>
<td>0.63 (0.79)</td>
</tr>
<tr>
<td></td>
<td>Counting</td>
<td>2.95 (1.96)</td>
<td>4.9 (5.43)</td>
</tr>
<tr>
<td>% silence</td>
<td>Reading</td>
<td>21.2 (5.85)</td>
<td>20.16 (4.54)</td>
</tr>
<tr>
<td></td>
<td>Monologue</td>
<td>20.75 (6.08)</td>
<td>21.65 (8.21)</td>
</tr>
<tr>
<td></td>
<td>Days</td>
<td>8.7 (8.01)</td>
<td>11.53 (9.67)</td>
</tr>
<tr>
<td></td>
<td>Counting</td>
<td>21.7 (10.7)</td>
<td>25.2 (17.25)</td>
</tr>
<tr>
<td>Speech rate (syllables/total signal time)</td>
<td>Reading</td>
<td>3.9 (0.39)</td>
<td>3.88 (0.67)</td>
</tr>
<tr>
<td></td>
<td>Days</td>
<td>3.74 (0.82)</td>
<td>3.44 (0.9)</td>
</tr>
<tr>
<td></td>
<td>Counting</td>
<td>2.46 (0.85)</td>
<td>2.22 (0.9)</td>
</tr>
<tr>
<td>$f_0$ CoV (mean/SD)</td>
<td>Reading</td>
<td>0.15 (0.04)</td>
<td>0.13 (0.03)</td>
</tr>
<tr>
<td></td>
<td>Monologue</td>
<td>0.14 (0.04)</td>
<td>0.13 (0.05)</td>
</tr>
<tr>
<td></td>
<td>Days</td>
<td>0.11 (0.05)</td>
<td>0.11 (0.04)</td>
</tr>
<tr>
<td></td>
<td>Counting</td>
<td>15.77 (8.55)</td>
<td>15.77 (10.32)</td>
</tr>
<tr>
<td>Alpha Ratio</td>
<td>Reading</td>
<td>0.21 (0.19)</td>
<td>0.23 (0.21)</td>
</tr>
<tr>
<td></td>
<td>Monologue</td>
<td>0.3 (0.19)</td>
<td>0.35 (0.18)</td>
</tr>
</tbody>
</table>
moderate effect sizes between HD and controls on both the reading and monologue tasks. Observed differences on measures of spectral tilt during the reading and monologue tasks highlighted a clear linear relationship across groups, where scores for the control group were demonstrably different from the PreHD and HD groups’ data. Broken down by task, Figures 1-6 highlight the differences between PreHD and HD groups. On most cases the performance of HD participants is differentiated from the control and PreHD groups. The degree to which groups varied in their production varied as a function of task.

Figure 1: Speech Rate by task and group

![Figure 1](image1)

Figure 2: Total silence time by task and group

![Figure 2](image2)

Figure 3: Total speech time by task and group

![Figure 3](image3)

Figure 4: Percentage of silence by task and group

![Figure 4](image4)

Figure 5: Coefficient of variation of f0 by task and group

![Figure 5](image5)

Figure 6: Alpha Ratio by task and group

![Figure 6](image6)

4. Discussion

The speech of early stage HD differed significantly from controls on some measures of timing. The sensitivity of acoustic measures for discriminating between groups was largely determined by task type. Tasks that required greater cognitive load, i.e., the reading and contemporaneous speech tasks, appeared to separate the pathological and control participants. Whereas, tasks that were highly automatic in their content and nature, i.e., counting and saying the days of the week, did not differentiate the groups. The fact that more complex tasks demonstrated higher sensitivity to disease state is not surprising given the previously documented cognitive deficits presenting in early stage HD [23]. However, the sensitivity of more complex tasks over highly automatic stimuli suggests that the speech deficits presenting in early stage and prodromal HD are the result of cognitive decline rather than a purely motor manifestation of symptoms. If the effect on speech production in HD was the result of purely motor dysfunction, we would anticipate differed performance on all speech tasks.

From these data we can see that variation within groups played an important role in demonstrating the value of each stimulus. The HD group presented with predominately greater variation (SD) on almost all measures compared to the control and PreHD groups. Variability has been previously identified as a key feature of disease state [24], often making comparisons difficult due to the heterogeneity of comparative cohorts. The large amount of variation within groups on the counting task, as well as the non-linear direction of the differences between groups (i.e., the PreHD productions were distinctly different from the HD and Controls) rendered this particular stimuli insensitive to disease state.

For the reading and contemporaneous speech, and to a lesser degree, the days of the week task, differences between groups tended to reflect anticipated performance based on disease.
state. PreHD individuals, although not reaching significance, tended to lie between the performance of controls and early stage HD. This is a noteworthy result given the small sample size and the broad range of burden of pathology scores in the PreHD group. Participants carrying the mutant HTT gene presented with slower rates of speech, took longer to say words and produced greater silences between and within words compared to healthy controls. PreHD and HD participants also produced speech with lower levels of energy in frequencies above 1kHz compared to controls during connected speech tasks. In this regard, the speech of gene positive individuals was slower, less precise and contained different energy levels in the components of speech that help to define the ‘personality’ of our voice. Overall, these data suggest that changes in speech production are developing prior to diagnosis but that effect sizes are modest and thus need a larger sample size to be replicated and reliably measured.

5. Conclusions
Speech has the potential to contribute to existing assessment batteries given its utility and objectivity in neurodegenerative disease clinical trial research. The nature of this study necessitates its expansion to include greater participant numbers, tracking individuals’ over time and direct comparison with existing sensitive markers of disease symptomatology (e.g., rates of change in chorea, finger tapping grip and tongue force, cognition and emotion recognition). The heterogeneity of the burden of pathology scores in the PreHD group could also be controlled to only include individuals with disease burden scores greater than 250 to optimise the recruitment of individuals with the best chance of detecting differences between groups [26].

6. Acknowledgements
TBC.

7. References
Author/s: Vogel, AP; Shirbin, C; Churchyard, AJ; Stout, JC

Title: Can we detect altered disease state in early stage Huntington’s Disease using acoustic markers of speech?

Date: 2012

Citation: Vogel, A. P., Shirbin, C., Churchyard, A. J. & Stout, J. C. (2012). Can we detect altered disease state in early stage Huntington’s Disease using acoustic markers of speech?. Australasian Speech Science and Technology Association (ASSTA)

Persistent Link: http://hdl.handle.net/11343/221218

File Description: Published version