Analyses of Sustained Vowels in Down Syndrome (DS): A Case Study Using Spectrograms and Perturbation Data to Investigate Voice Quality in Four Adults With DS

***Tracy Jeffery, †Stuart Cunningham, and †Sandra P. Whiteside,** **Lincoln and †Sheffield, UK*

Summary: Objectives. Automatic acoustic measures of voice quality in people with Down syndrome (DS) do not reliably reflect perceived voice qualities. This study used acoustic data and visual spectral data to investigate the relationship between perceived voice qualities and acoustic measures.

Study design. Participants were four young adults (two males, two females; mean age 23.8 years) with DS and severe learning disabilities, at least one of whom had a hearing impairment.

Methods. Participants imitated sustained /i/, /u/, and /a/ vowels at predetermined target pitches within their vocal range. Medial portions of vowels were analyzed, using *Praat*, for fundamental frequency, harmonics-to-noise ratio, jitter, and shimmer. Spectrograms were used to identify the presence and the duration of subharmonics at onset and offset, and mid-vowel. The presence of diplophonia was assessed by auditory evaluation.

Results. Perturbation data were highest for /a/ vowels and lowest for /u/ vowels. Intermittent productions of subharmonics were evident in spectrograms, some of which coincided with perceived diplophonia. The incidence, location, duration, and intensity of subharmonics differed between the four participants.

Conclusions. Although the acoustic data do not clearly indicate atypical phonation, diplophonia and subharmonics reflect nonmodal phonation. The findings suggest that these may contribute to different perceived voice qualities in the study group and that these qualities may result from intermittent involvement of supraglottal structures. Further research is required to confirm the findings in the wider DS population, and to assess the relationships between voice quality, vowel type, and physiological measures.

Key Words: Down syndrome–Diplophonia–Subharmonics–Voice–Phonation.

INTRODUCTION

Individuals with Down syndrome (DS) are described as having voices that are characteristic of the syndrome,[1,2] with descriptors commonly identifying "harsh," "guttural," and "raucous" qualities.[3–6] Additionally, DS voice is often perceived as being atypically breathy and rough[2,7] and low-pitched.[2,3,7] Several studies have investigated voice qualities in adults and children with DS, using automatic measurements of fundamental frequency (f_0), jitter (frequency perturbation), shimmer (amplitude variation), and harmonic-to-noise ratio (HNR; additional noise in the harmonic spectrum).[2,5,8–13] Historically, these measures have been used as indicators of atypical phonation, although there is increasing evidence that they are unreliable for pathological voices.[14] Research using such measures has failed to determine the role of phonation in perceived voice qualities in DS populations. Inconclusive and contradictory findings from acoustic measures have led some researchers to suggest that supralaryngeal factors, such as the properties of the vocal tract, may contribute to the characteristic voice qualities more than laryngeal factors.[10,15]

Several studies have compared acoustic data to the perceived vocal qualities of people with DS, with conflicting results.[9–11] Lee et al[10] analyzed the vowels from continuous speech of nine British adults with DS, aged between 17 and 24 years, and those of typically developing (TD) controls, matched for age and sex. They reported no clear difference between groups in jitter and shimmer values. Using data derived from spoken words, Albertini et al[11] reported higher mean f_0 and lower spectral energy in DS adults, compared to controls, and reduced shimmer in adult males with DS. In contrast to both studies, Moura et al[9] reported elevated measures of jitter and shimmer in the sustained vowels (/a/, /e/, /i/, /o/, /u/) of 66 Portuguese children with DS, aged 4–8 years. In comparison to data collected from TD children, the DS children produced sustained vowels at a lower f_0, with greater deviation, higher perturbations in shimmer and jitter, and with increased noise in the signal in comparison to the strength of the phonation. The DS children performed statistically differently on all voice measures except for the f_0 of the vowel /u/. In comparison to TD controls, the children with DS were found to have lower measures in spectral tilt,[9] a measure of the energy across the frequency ranges. Spectral tilt can indicate creakiness (strong positive slope) or breathiness (strong negative slope).[16] The authors suggested the finding indicated higher than typical levels of breathiness and more "forceful" phonation.

Few studies have examined sustained vowel production in adult DS populations. Of these, there is agreement that mean f_0 is high, compared to controls,[2,8,13] but findings are mixed and are difficult to reconcile with perceived qualities. Moran and Gilbert[2] compared acoustic data to auditory-perceptual judgments in 16 adults with DS. They reported elevated jitter (<6%) in three of their participants and additional noise in the harmonic

Accepted for publication August 1, 2017.

From the *Bishop Grosseteste University, Lincoln, UK; and the †Department of Human Communication Sciences, University of Sheffield, 362 Mushroom Lane, Sheffield S10 2TA, UK.

Address correspondence and reprint requests to Tracy Jeffery, Bishop Grosseteste University, Longdales Rd, Lincoln LN1 3DY, UK. E-mail: Tracy.jeffery@bishopg.ac.uk

Journal of Voice, Vol. ■■, No. ■■, pp. ■■-■■

0892-1997

http://dx.doi.org/10.1016/j.jvoice.2017.08.004

spectrum of nine participants. Despite an atypically high mean f_0 in the DS group, low pitch was perceived by 70% of the judges in five of the participants. For females, mean f_0 correlated with perceived pitch, but there was low correlation between perceived pitch and f_0 for DS males. The authors suggest that the perceived voice quality in DS males is affected by the interaction of several factors, including HNR, breathiness, and laryngeal tension. A mismatch between perceived voice quality and acoustic data was also reported in a more recent study.[13] Seifpanahi et al[13] analyzed the sustained /a/ vowels of 22 adults with DS. Jitter was lower than controls, but shimmer was comparable. Their data did not indicate perturbation, despite agreement from three speech and language therapists (SaLTs) that all participants were "moderately hoarse."

An early study by Beckman et al used spectrographs to examine the sustained /a/ that were produced by a woman (aged 22.67 years) and a man (27.5 years) with DS.[8] Both participants had normal hearing levels and voices that were described as breathy, with an imbalance in oral/nasal resonance. They identified that in six of the nine vowels produced by the female, subsequent cycles of voicing were more variable in duration (jitter) and in amplitude (shimmer). They reported that regular alternations in periods of the waveform resulted in an effective halving of f_0 and a perceived octave drop in pitch. The female was subsequently identified as having diplophonia, which is the generation of two audible pitches.[17–19] Beckman et al[8] suggested that subglottal variations in pressure or laryngeal pathology might have contributed to the phenomenon in their subject.

Beckman et al[8] suggested that a high incidence of diplophonia in the DS population might explain reports of a lower perceived pitch. A decade earlier, Novak[4] had proposed that ventricular voice, which is caused by the continued oscillation of the ventricular folds, is the cause of the perceptually harsh voice quality in DS subjects. Beckman et al[8] argued that hypotonia ruled out the probable engagement of the ventricular folds in their subject. However, it is now known that people with DS apply more energy than controls with healthy voices to trigger contraction in the surface of the laryngeal musculature.[15] The habitual use of excessive effort in producing voice can result in hyperfunctional voice disorders[20,21] in which supraglottal structures, such as the ventricular folds, are employed during phonation.[17,22–24] One recent study with children and young people with DS used auditory-perceptual evaluation to identify laryngeal tension and diplophonia.[12] Diplophonia was not perceived, although elevated levels of laryngeal tension were perceived. However, diplophonia is not reliably perceived by auditory evaluation alone.[18]

Recent studies of DS voice rely on automatic acoustic measures, and have not included data derived from visual inspection of individual samples. Perturbation analysis is unreliable for aperiodic voice.[17] Therefore, visual inspection may be necessary to confirm whether samples are valid for automatic analysis, and to provide additional information regarding the possible presence of diplophonia.[18] The current study examines the acoustic characteristics of phonation in sustained vowels produced by four adults with DS, severe learning disabilities (SLDs), and hearing impairment (HI). The study uses voice perturbation data (HNR, jitter, and shimmer) and mean f_0 alongside evidence from visual data to explore the nature of phonation in sustained vowels, and to consider how the data link to the participants' habitual voice qualities. Although sustained vowels are not a reliable indicator for voice quality in speech[25] the use of sustained vowels allows for examination of phonation when articulation and processing demands are low. As such, it can be expected that any difficulties that are revealed in these conditions will be exacerbated in connected speech.[17,25,26]

METHODS

Participants

The study was a multiple case-study design that involved four participants. Data were collected from four young adults (mean age: 23.8 years; SD = 0.37) with DS and SLD, who were participating in a larger, explanatory study that examined voice production in speech and song. Explanatory case studies seek to explain causal links between phenomena that cannot be understood through experimental studies.[27] They are appropriate for investigating underresearched or poorly understood aspects of behavior, such as voice production in DS, and heterogeneous groups, such as those with DS. Data were generated for each participant separately. Participants acted as their own controls.

Demographic data

As part of the larger study, the group had participated in a range of standard and non-standard cognitive tasks. This established that the individuals in the group were of similar abilities in verbal mental age, as measured by the British Picture Vocabulary Scale, version 1 (Tables 1–4). The British Picture Vocabulary Scale test[28] measures receptive vocabulary, and correlates to verbal intelligence. Participants are required to listen to a word and identify the corresponding picture, from a set of four. It has been standardized for use for children and adults aged between 3 and 18.11 years, and has been successfully used with people with DS to estimate mental age across a range of ability levels. Information about voice quality and hearing ability was provided by the participants' SaLT. Descriptors of voice quality for each participant were also given by the SaLT, based on her existing knowledge and historical evaluations of each participant. To establish vocal range, participants were asked to imitate an ascending and descending vocal glide on /a/ vowel. To assess speaking range and mean f_0, participants were asked to describe a picture[29] and to provide positive comments on a peer's performance of a song.

Recruitment

Ethics approval was granted by the Human Communication Sciences' Ethics Committee, University of Sheffield. Participants gave informed consent to take part in the study and for their data to be used.

All participants were resident in long-term care, and informed consent was sought and obtained from the residential care home in the first instance. The principal of the care home identified possible participants using the following inclusion criteria: