Newcastle University ePrints

Miller N, Pennington L, Robson S, Roelant E, Steen N, Lombardo E. [Changes in voice quality after speech-language therapy intervention in older children with Cerebral Palsy](#). *Folia Phoniatrica et Logopaedica* 2013, 65(4), 200-207.

Copyright: The definitive version of this article, published by S. Karger AG, 2013, is available at: <http://dx.doi.org/10.1159/000355864>

Always use the definitive version when citing.

Further information on publisher website: <http://www.karger.com>

Date deposited: 13th June 2014

Version of file: Author final

This work is licensed under a [Creative Commons Attribution-NonCommercial 3.0 Unported License](#)

ePrints – Newcastle University ePrints

<http://eprint.ncl.ac.uk>

Authors' copy of final draft. Paper published as

Changes in Voice Quality after Speech-Language Therapy Intervention in Older Children with Cerebral Palsy

Miller N, Pennington L, Robson S, Roelant E, Steen N. & Lombardo E.

Folia Phoniatrica et Logopaedia 2013;65:200-207 (DOI: 10.1159/000355864)

<http://www.karger.com/Article/FullText/355864>

Copyright: Karger Medical and Scientific Publishers

Changes in voice quality after speech-language therapy intervention in older children with cerebral palsy

Nick Miller (corresponding author), Lindsay Pennington, Sheila Robson, Ella Roelant
Nick Steen, Eftychia Lombardo
Institute of Health & Society, Newcastle University

Address for correspondence:

Nick Miller

Institute of Health and Society

Speech and Language Sciences

George VI Building

Newcastle University

Newcastle-Tyne NE1 7RU, Great Britain

Telephone: +44 191 222 56 03, Email: Nicholas.Miller@ncl.ac.uk

Running title: Changes in voice quality

Abstract:

Objective: We examined whether perceived voice quality is altered in a group of children with cerebral palsy (CP) following an intervention focusing on respiration and phonation, and whether possible improvements might be associated with increased intelligibility levels.

Methods: Sixteen individuals with CP and dysarthria (nine girls, mean age 14 years, SD = 2; nine with spastic type cerebral palsy, two dyskinetic, four mixed, one Worster Drought) completed intelligibility assessments on separate days twice before intervention, at termination of treatment and at six week follow-up using 50 words from the Children's Speech Intelligibility Measure lists, and describing cartoon strips. Experienced speech-language pathologists rated voice quality employing GRBAS scales.

Results: There was no clear evidence that change in voice quality pre-post intervention was large compared with change in the pre-intervention or post-intervention periods. Asthenia demonstrated largest improvement (effect size of 0.4). Intelligibility correlated weakly with Grade, Breathiness and Asthenia, but not to Roughness or Strain. A deterioration of 1 unit on the Grade and Asthenia scales was associated with an approximately 11% decrease in intelligibility.

Conclusion: Perceived changes in voice quality were small compared to changes in intelligibility. Further investigations must examine other variables potentially associated with intelligibility gain to better understand the links between the respiratory-phonatory intervention and improved intelligibility.

Keywords: cerebral palsy; treatment; voice; intelligibility

Introduction

Cerebral palsy (CP) represents the most common cause of motor disorder in childhood [1,2]. Motor speech disorders are associated with all types of CP, though most prevalent in dyskinetic forms. Approximately 50% of children with CP present with a speech-voice disorder [3,4].

All vocal tract subsystems can be implicated in CP, giving poor control over any or all of respiration, phonation, resonance, articulation and prosody [5]. Controlled studies of intervention for dysarthria in CP have not been extensive, though a recent systematic review [6] indicated several phase II intervention studies supporting the potential effectiveness of speech therapy for children with CP. In particular, findings suggested a focus on control of respiration, phonation and speech rate is associated with changes at both the impairment and activity levels.

Such an approach is supported from a theoretical position, as well as from work in germane areas where work on voice, prosodic and rate variables has been associated with gains in voice quality, loudness and articulation (even when the latter is not specifically targeted) [7-10]. Improving breath control can increase respiratory support and bring better coordination of on- and offset of breathing with phonation. This brings the potential for less air wastage, improved subglottal driving pressure and longer utterances. Better breath control aids phonation, and achievement of greater voice stability, amplitude/perceived loudness and pitch range. These gains pave the way for work on prosody. This in itself can benefit intelligibility through added suprasegmental cues to meaning [11,12]. Speaker initiated increases in loudness/amplitude also bring changes to rate and improved articulatory precision, both of which can be linked to better intelligibility [13-15]. Taking these altogether one would expect a focus on respiration and voice to be linked to improvements in intelligibility.

Our previous work [16-18] showed that therapy on respiration, phonation and slowing rate, with no focus on articulation (which has been a more traditional target of speech treatment for children with CP), can increase intelligibility. We found gains in intelligibility in single words and connected speech in the region of up to 15 percentage points for some young people aged 11-18 years and in younger (5-11 years) children with different CP types and severities of speech disorder. Some doubled their intelligibility levels. Fox and Boliek [19] employed Lee Silverman Voice Treatment (LSVT), a programme developed for people with Parkinson's, focusing on phonation and respiratory effort and monitoring, to provide preliminary information on voice and speech outcomes with children with spastic CP. When listeners compared pre-intervention with post-treatment audiorecordings they preferred predominantly post-treatment ones in terms of loudness, articulatory precision (even though this had not been explicitly targeted in therapy).

Despite predictions from other studies on why one might achieve these gains, studies leave open the question of the basis for improvements in intelligibility if speech is not a target of intervention. One supposition from an intervention emphasizing better breath control and stability of phonation is that voice quality is improved and this contributes to raised intelligibility.

In this paper we extend our data from [16] to examine whether the gains in intelligibility were associated with improvements in overall or subspects of perceived voice quality as measured by the GRBAS (Grade, Roughness, Breathiness, Asthenia, Strain) scales. These scales have been used in other studies with a focus on perceived changes to voice, though not widely in children with CP. Their advantages and disadvantages have been widely described [20]. Of interest to us was the possibility to examine differing dimensions in voice quality, as opposed to one global voice measure.

Methods

The study was conducted in accordance with principles approved by Sunderland UK NHS Research Ethics Committee. Participants' guardians provided written informed consent for participation and participants also gave written or supervised verbal consent. The participants and data are the same as reported in [16].

Participants

We recruited 16 individuals with CP and dysarthria (12-18 years, 9 girls, mean 14 years, SD = 2) from schools in the North East of England. Details appear in table 1.

Table 1 Children's characteristics about here

Inclusion criteria comprised: diagnosis of CP by a paediatric neurologist; referral to speech-language pathologist (SLP); dysarthria diagnosed by SLP and rated as moderate to severe by the child's local SLP using subjective, consensus rating scales; aged 11-19 years; attending special or mainstream schools. Exclusion criteria: audiological assessment showing bilateral hearing impairments greater than 50 dB; diagnosis of visual impairments not correctable with spectacles (stated in medical notes); profound cognitive impairments as classified by an educational or clinical psychologist; SLP diagnosed profoundly delayed language comprehension, even for grammatically simple instructions. Gross Motor Function Classification Scale (GMFCS) [21], ratings for the group ranged from 1-5, (Median 4), indicating most of the children needed adaptive seating to maintain sitting posture and hand control and may at best walk short distances with support for turning and maintaining balance.

Intervention

All speakers received six weeks of speech therapy at school, comprising three 35-40 minute individual sessions per week, on different days, delivered by a research SLP. Therapy focused on achieving and maintaining a suitable posture for breathing and phonation; stabilising students' respiratory and phonatory effort and control, including coordination of respiration-phonation-articulation onset and offset; speech rate and phrase length/syllables per breath; all with a view to improving intelligibility (see [16] for full details). Articulation was not directly targeted.

Measures

Figure 1 about here

Each speaker was recorded on two days at each of four time points (times 1 to 4): week one six weeks before therapy (Recording A = day 1 and Recording B = day 2); week five one week before therapy commenced (Recordings C and D); one week after therapy termination (week 12, recordings E and F); and six weeks after termination (week 17, recordings G and H). The pathway through the project is illustrated in figure 1. Between weeks 1-5 participants received SLP 'input as normal' (which may have included no intervention); between weeks 12-17 participants received no SLP intervention nor any other therapy that directly or indirectly targeted respiration or voice. Recordings took place in a quiet room at school using an EDIROL R1 digital recorder and an AKG C420 headmounted microphone.

At each assessment session participants repeated 50 single words from the Children's Speech Intelligibility Measure (CSIM) lists [22], and described cartoon strips (three pictures), to elicit connected speech performance. Separate lists from the CSIM were randomly assigned to each participant across each of the eight recording times; four separate cartoon strips were randomly allocated. We used the GRBAS rating scales [20] (0 normal to 3 severely impaired) to evaluate possible changes to voice quality on scales of overall Grade (degree of voice abnormality), Roughness (extent of irregular vocal cord vibration affecting steadiness of pitch and loudness), Breathiness (extent of excess air leakage between the vocal cords), Asthenia (lack of power) and Strain (excessively tense and high pitched).

Data processing

Separate sound files for each participant for each task (single words; connected speech) and assessment point were saved as wave files. Sixteen SLPs rated voice quality. All were specialists in voice disorder and experienced in using the GRBAS scales. As there were no significant differences between intelligibility scores for the two recordings at each time point (A vs B, C vs D, etc), GRBAS ratings were based on evaluation of recordings from one of the two days at each of the four time points. Which of the two recordings was used was determined at random for each child.

Recordings were allocated to SLPs using constrained randomisation: i.e., each SLP heard one recording from each participant and each voice recording was rated by four different SLPs. Each therapist therefore rated 18 recordings (each participant plus two repeats for intra-rater purposes). Therapists were blind to all speaker and time point information. SLPs rated tracks using high fidelity playback equipment. They rated each recording independently but could listen to tracks as many times as they wished to make a decision. They were not permitted to alter volume control. Mean ratings across the four therapists per recording formed the basis for group comparisons between time points.

Analysis

To investigate change over time, mean voice quality ratings across all raters on GRBAS subscales for each child were determined for the four time points. These mean ratings were then analysed with repeated-measures analysis of variance assuming a normal error structure. Within this framework we fitted three contrasts:

- 1) Intervention effect: difference between pre-post intervention (recordings A, B, C, D vs E, F, G, H)
- 2) A change within each of the pre and post periods (A, B, E, F versus C, D, G, H) – in conjunction with (1) this allows us to assess whether there is a trend over the period of the study rather than a step change in voice quality following intervention.
- 3) A change within time points (A, C, E, G versus B, D, F, H) – to examine for any systematic difference between first and second recordings at each time point.

Interval estimates of differences in voice quality corresponding to each of these effects are reported. Agreement between raters was assessed using a variance components model considering each recording as a sample and incorporating a therapist effect. Confidence intervals were estimated using 200 bootstrap samples.

To examine the relationship of voice quality to intelligibility correlations between GRBAS subscale ratings and speech intelligibility were calculated. Intelligibility scores were based on the mean percentage intelligibility score per participant from the single word and connected speech (cartoon strip) results obtained from the multiple unfamiliar listeners in our previous study [16]. The mean scores were derived by taking the mean of the single word and connected speech intelligibility scores for each listener, then calculating the mean intelligibility score across the three intelligibility raters for each recording. These mean percentage intelligibility scores were correlated with the mean GRBAS scores from the four SLPs who rated each recording. For correlational purposes we employed the intelligibility assessment that matched the day for which GRBAS ratings were obtained for any particular child. A linear mixed effects model with recordings nested within children with intelligibility as dependent variable and GRBAS subscale ratings as predictors was fitted to investigate if voice quality predicted the child's intelligibility.

Results

Inter and intrarater reliability

For interrater reliability calculations each recording was considered as a sample rated by four therapists. Confidence limits were calculated using 200 bootstrap samples. Levels of agreement ascertained were: Grade 0.35 (95% confidence interval 0.27-0.60); Roughness 0.36 (0.30-0.59); Breathiness 0.26 (0.21-0.51); Asthenia 0.36 (0.30-0.58); Strain 0.26 (0.21-0.54). Intra-rater agreement for scores across the two repeated samples gave coefficients and significance levels, adjusting for chance agreement (Cohen's kappa), of G: 0.27, p 0.026; R: 0.21, p 0.086; B: 0.02, p 0.85; A: 0.29, p 0.01; S: 0.11, p 0.28.

Voice quality

Averaged ratings across the four listeners per track for the first two time points indicated that before therapy all participants had a voice quality perceived as different to normal on all subscales: Grade mean 1.78 (standard deviation 0.49), Roughness 1.02 (0.54), Breathiness 1.26 (0.50), Asthenia 1.35 (0.65), Strain 1.20 (0.55).

We considered a repeated-measures ANOVA with a contrast for the difference between before and after therapy, a contrast for the difference within the pre and post periods (times 2 and 4 versus times 1 and 3) and a contrast corresponding to the systematic difference between the two recordings at each of the four time points. The mean differences (and 95% confidence interval) averaged over the four raters for the different GRBAS subscales for before to after therapy were Grade 0.03 (CI -0.21, 0.28), Roughness -0.09 (-0.34, 0.16), Breathiness -0.01 (-0.30, 0.27), Asthenia -0.26 (-0.53, 0.00), Strain 0.17 (-0.09, 0.42). Apart from a borderline significant difference in the asthenia rating (F 1, 43, 3.94; p 0.053) no other differences approached significance.

In general there was no evidence that the change pre-post intervention was large in comparison either with the change within the pre and post periods or with the mean difference between the first and second recordings across the four time points. The largest observed change, in Asthenia, corresponded to an effect size of 0.4 (mean change approximately 0.4 times standard deviation of baseline scores).

Relationship voice quality- speech intelligibility

The mean intelligibility score for each child across single word and connected speech based on unfamiliar listener ratings was correlated (Pearson's) with the GRBAS ratings per child averaged across the four SLP raters. Results are summarised in Table 2.

Table 2 about here

Intelligibility scores were weakly to moderately correlated with grade, breathiness and asthenia, but not to roughness and strain. A deterioration of 1 unit on the grade and asthenia scales was associated with approximately an 11% decrease in speech intelligibility. Using a mixed effects model with intelligibility as outcome and Asthenia, Grade and Breathiness as predictors, the asthenia score was the most important predictor of intelligibility (p =0.017). Once the relationship between asthenia and intelligibility was allowed for the additional effects of grade (p =0.945) and breathiness (p =0.226) were very small.

Discussion

We previously noted [16] changes in speech intelligibility corresponding to effect sizes in the range 0.5 to 0.75 (effect size defined as mean change divided by standard deviation of baseline scores). We aimed to establish whether these changes might be associated with improvements in perceived voice quality. Results here suggest little evidence of corresponding improvements in voice quality. The largest observed change (in asthenia) corresponds to an effect size of 0.4. Furthermore, whilst asthenia change was the variable most strongly associated with intelligibility variation, it is unlikely that this was sufficient to account for the extent of intelligibility change observed in the group as a whole.

Other studies employing interventions similar to the present study (principally those using LSVT) have reported significant positive outcomes for loudness/ sound pressure level and overall voice quality [8,19,23]. However, studies have not always related these to intelligibility changes, and, where they have, have used widely divergent ways of measuring this, still leaving open the question of the relationship between voice quality – intelligibility. Fox et al [19], for instance, asked listeners to choose which one of paired pre-post recordings had greater loudness, better overall voice quality and better articulatory precision. However, listeners were given no firm definitions of what constituted elements of voice quality or articulatory precision and they did not examine associations between these variables.

Other studies have employed visual analogue scales or direct magnitude estimation (DME) to assess intelligibility before-after voice-respiration therapies or restricted themselves to acoustic analyses of articulation change (where assumptions of intelligibility gains have been implicit but not formally demonstrated) and global measures of voice quality/ loudness. This renders them poorly comparable to the present work in terms of gaining insight into therapeutic variables. Where diagnostic intelligibility testing or transcription accuracy has been used as in our work, findings regarding the relationship of voice and articulation changes remain unsettled. Wenke and colleagues [23,24] (using LSVT with speakers after brain injury or stroke) found significant gains in loudness and articulatory precision (based on DME) and word and sentence transcription improvement. They did not, however, examine for strength of association between loudness and intelligibility. Immediately post intervention significant intelligibility gains (transcription) fell back to nonsignificance at follow up, whilst overall intelligibility (DME) and articulatory precision were observed to have even greater gains. Cannito et al [8] examined the relationship of loudness gains and sentence intelligibility after LSVT in people with Parkinson's. Whilst there was general improvement in voice not all speakers improved in intelligibility.

Given the current study focused on improved breath and voice control, the fact that asthenia emerged as a more prominent factor than other voice quality ratings is not unexpected. However, the small effect size suggests other variables were at play in effecting intelligibility change, including elements of the study methods and design.

Although intuitively one might expect raised loudness/amplitude to be associated with increased intelligibility, their relationship is not entirely clear. Studies report strong or no links [15,25,26]. Partly this relates to artificially manipulated versus speaker controlled manipulation. In the latter, added effort to increase loudness brings about alterations to other speech and voice parameters [27,28] rather than increased SPL alone. Additionally whilst voice quality per se may not alter significantly (and therefore changes on GRBAS scales would not be expected), increased respiratory and phonatory control has been linked to better vowel differentiation and prosodic variables [9,29-31]. Improved prosody may be associated with an improved match between phrase length and breath control – e.g. individuals are not trying to speak on residual air. Further, more reliable prosodic cues may offer listeners better information on word and phrase boundaries. To understand whether

these speculations apply to the current data further analyses are required to examine for possible relationships between phrase length, pause placement, speech rate and intelligibility before and after intervention.

Whilst the mean baseline GRBAS ratings were in the mild-moderate range, the group as a whole were nevertheless relatively severely affected by their CP (GMFCS median level IV) and pretherapy intelligibility levels were predominantly <60% (table 1). The possibility exists that the present intervention failed to lift the voice quality ratings significantly because of the broader motor context, there may have been a floor effect, and/or the GRBAS scales were insufficiently sensitive to gauge perceived changes. Fox et al [19] raised a similar possibility in relation to their participants. Despite statistically significant perceptual gains, following intervention children still remained around 2 standard deviations below their typically developing peers on maximum performance tasks. These authors suggested that the available envelope for improvement on certain parameters is narrow for this clinical population, even despite intervention. Another factor at play in our study may concern the intelligibility measures utilised for correlations with GRBAS ratings. We combined the single word and connected speech measures. At the severe (around <30%) and mild (around >70%) ends of the intelligibility spectrum changes in single word vs connected speech intelligibility may display a linear correlation; but within the mid-range of intelligibility that relationship is more curvilinear. By conflating the two measures a distortion in change metrics may have been introduced. When we examined this in preliminary analyses there was no significant difference in correlations when single word and connected speech were examined separately.

A strong proviso when interpreting outcomes in this study arises from the modest levels of inter-rater agreement, although they are not dissimilar to values obtained in comparable studies [19,32-34]. Intra-rater agreement was more favourable if agreement within one scale point on related evaluations was measured, but fell to low to no agreement once exact and chance agreement were considered. Typically intra-rater scores are better than interrater judgements. However, intrarater agreement here fell below that found for experienced GRBAS raters in other studies and may relate to a number of factors.

Firstly, although the GRBAS scales have been employed to judge voice quality in people with neurological disorders, their prime use rests in evaluating isolated vocal cord/ voice quality dysfunction, where acceptable levels of inter and intra-rater agreement have been achieved [20,35], especially when based on prolonged vowel stimuli. In this study judgements were based on whole words/phrases where one may expect several voice dimensions to be simultaneously impaired, and individual speaker variability across items to have been raised. This would render judgements more inconsistent and add to the multiplicity of factors influencing therapists' ratings and distorted judgements [36,37]. Task effects on judgements of voice quality and loudness are present [38,39]. This raises a question of comparability across tasks and how vowel vs connected speech based ratings might relate differently to intelligibility change. Isolated prolonged vowels deliver a more stable signal for voice quality judgements/analyses, but are more distant from live speaking tasks, where voice production interacts with a range of other variables. Judgements on connected speech sacrifice this stability, but may offer a more ecologically valid comparison of day to day performance. A future study would benefit from employing both prolonged vowel and connected speech assessments, or include having judges evaluate pre-post pairs of utterances (similar to [19]) for preference on GRBAS subscales separately to still retain some insights into specific voice quality changes.

Even though we employed experienced listeners, training on the tasks with materials from children with CP may have improved agreement [37]. A method similar to Kreiman et al [40] might also lead to more stable and objective ratings of voice quality and loudness. They used synthesized signals. Listeners adjusted these until they matched their perceptual evaluation of real samples. This enabled acquisition of objective measures on different acoustic parameters linked more directly to listener perceived values.

A further possible explanation for the lack of association between GRABS and intelligibility changes could relate to methodological procedures. The intelligibility improvement with which voice measures were being compared may have been overinflated if effects of listener familiarity with materials or speaker familiarity with tasks and items were at work. We went to some lengths to assure randomisation of word lists and pictures to speakers for testing purposes and to listeners when scoring. For each track we employed groups of listeners, rather than one individual rater or single group of raters per child, together with blinding of samples. This should have neutralised or minimised order, task and learning effects in evaluation.

There remains an outside possibility that familiarity effects with materials and tasks by the children being tested exercised some influence on gains. The fact that analyses showed no stepwise increment within day 1 vs 2 at each assessment point or across the four assessment times would suggest such a learning or familiarity effect was not at work, nor a maturation effect. The fact that differences pre- vs post intervention in intelligibility were observed also argues against possible floor or ceiling effects in the intelligibility measures. To settle issues here a future study could gainfully include a control group who undergo assessments at the same time points as the intervention group but only receive intervention later.

A last factor that must be considered concerns inter-speaker variability. Inspection of individual profiles gives a strong indication of why no clear group effects of voice quality change were found. Performance within individuals across time is highly variable. There is no clear group trend in profiles for any of the GRBAS scales. A future fruitful line of enquiry would be to examine profiles of individuals where there is a strong effect of treatment versus weaker or no effect, in order to gain insights into for whom the particular therapy regime might be best suited.

Conclusions

We showed that perceived changes in voice quality following intervention focusing on respiration and phonation control are small compared to the observed changes in speech intelligibility. Based on current data, the association between intelligibility gains and voice quality change on individual voice quality parameters is at best weak-moderate. Further investigations must examine other variables potentially associated with intelligibility gain – e.g. prosody, acoustic variables, speech rate, sound pressure level – to better understand the links between the respiratory-phonatory intervention and improved intelligibility. This was a preliminary study employing relatively small group numbers with some individual variability in response to the treatment. Future work may also look at speech, language and cerebral palsy group factors that may link to differential effects.

Acknowledgements

We thank the children, parents and listeners who helped in this study. Work was funded by The Children's Foundation, UK.

Conflicts of interest

The authors report no conflicts of interest

References

- 1 Andersen E, Fledelius HC, Fons M, Haugsted R: An epidemiological study of disability in 4-year-old children from a birth cohort in frederiksborg county, denmark. *Danish Medical Bulletin* 1990;37:182-185.
- 2 Stanley F, Blair E, Alberman E: *Cerebral palsy: Epidemiology and causal pathways*. London, MacKeith Press, 2000.
- 3 Bax M, Tydeman C, Flodmark O: Clinical and mri correlates of cerebral palsy: The european cerebral palsy study. *JAMA* 2006;296:1602-1608.
- 4 Kennes J, Rosenbaum P, Hanna SE, Walter S, Russell D, Raina P, Bartlett D, Galuppi B: Health status of school-aged children with cerebral palsy: Information from a population-based sample. *Dev Med Child Neurol* 2002;44:240-247.
- 5 Levin K: Babbling in infants with cerebral palsy. *Clin Linguist Phon* 1999;13:249-267.
- 6 Pennington L, Miller N, Robson S: Speech therapy for children with dysarthria acquired before three years of age. *Cochrane database of systematic reviews* 2009
- 7 De Bodt MS, Huici M, Van De Heyning PH: Intelligibility as a linear combination of dimensions in dysarthric speech. *Journal of Communication Disorders* 2002;35:283-292.
- 8 Cannito MP, Suiter DM, Beverly D, Chorna L, Wolf T, Pfeiffer RM: Sentence intelligibility before and after voice treatment in speakers with idiopathic parkinson's disease. *Journal of voice : official journal of the Voice Foundation* 2012
- 9 Lee T, McCann C: A phonation therapy approach for mandarin-english bilingual clients with dysarthria. *Clin Linguist Phon* 2009;23:762 - 779.
- 10 Fox CM, Ramig LO, Ciucci MR, Sapir S, McFarland DH, Farley BG: The science and practice of lsvt/loud: Neural plasticity-principled approach to treating individuals with parkinson disease and other neurological disorders. *Seminars in Speech and Language* 2006;27:283-299.
- 11 Patel R, Campellone P: Acoustic and perceptual cues to contrastive stress in dysarthria. *J Speech Lang Hear Res* 2009;52:206-222.
- 12 Ballard KJ, Djaja D, Arciuli J, James DGH, van Doorn J: Developmental trajectory for production of prosody: Lexical stress contrastivity in children ages 3 to 7 years and in adults. *J Speech Lang Hear Res* 2012;55:1822-1835.
- 13 Tjaden K, Wilding GE: Rate and loudness manipulations in dysarthria: Acoustic and perceptual findings. *J Speech Lang Hear Res* 2004;47:766-783.
- 14 Darling M, Huber JE: Changes to articulatory kinematics in response to loudness cues in individuals with parkinson's disease. *J Speech Lang Hear Res* 2011;54:1247-1259.
- 15 Neel AT: Effects of loud and amplified speech on sentence and word intelligibility in parkinson disease. *J Speech Lang Hear Res* 2009;52:1021-1033.
- 16 Pennington L, Miller N, Robson S, Steen N: Intensive speech and language therapy for older children with cerebral palsy: A systems approach. *Developmental Medicine & Child Neurology* 2010;52:337-344.
- 17 Pennington L, Smallman C, Farrier F: Intensive dysarthria therapy for older children with cerebral palsy: Findings from six cases. *Child Language, Teaching and Therapy* 2006; 22:255-273.

- 18 Pennington L, Roelant E, Thompson V, Robson S, Steen N, Miller N: Intensive dysarthria therapy for younger children with cerebral palsy. *Developmental Medicine & Child Neurology* 2013;55:464-471.
- 19 Fox CM, Boliek CA: Intensive voice treatment (lsvt loud) for children with spastic cerebral palsy and dysarthria. *J Speech Lang Hear Res* 2012;55:930-945.
- 20 Karnell MP, Melton SD, Childes JM, Coleman TC, Dailey SA, Hoffman HT: Reliability of clinician-based (grbas and cape-v) and patient-based (v-rqol and ipvi) documentation of voice disorders. *Journal of Voice* 2007;21:576-590.
- 21 Palisano R, Rosenbaum P, Bartlett D, Livingston M: Gross motor function classification system expanded and revised. Toronto, CanChild Centre for Childhood Disability Research, McMaster University, 2007.
- 22 Wilcox K, Morris S: Children's speech intelligibility measure. San Antonio, Psychological Corporation, 1999.
- 23 Wenke RJ, Cornwell P, Theodoros DG: Changes to articulation following lsvt® and traditional dysarthria therapy in non-progressive dysarthria. *International Journal of Speech-Language Pathology* 2010;12:203-220.
- 24 Wenke R, Theodoros D, Cornwell P: The short- and long-term effectiveness of the lsvt for dysarthria following tbi and stroke. *Brain Inj* 2008;22:339-352.
- 25 D'Innocenzo J, Tjaden K, Greenman G: Intelligibility in dysarthria: Effects of listener familiarity and speaking condition. *Clin Linguist Phon* 2006;20:659-675.
- 26 Kim Y, Kuo C: Effect of level of presentation to listeners on scaled speech intelligibility of speakers with dysarthria. *Folia Phoniatri Logop* 2012;64:26-33.
- 27 Hustad KC, Sassano K: Effects of rate reduction on severe spastic dysarthria in cerebral palsy. *J Med Speech-Lang Pathol* 2002;10:287-292.
- 28 Yorkston KM, Hammen VL, Beukelman DR, Traynor CD: The effect of rate control on the intelligibility and naturalness of dysarthric speech. *Journal of Speech and Hearing Disorders* 1990;55:550-560.
- 29 Patel R: Acoustic characteristics of the question-statement contrast in severe dysarthria due to cerebral palsy. *J Speech Lang Hear Res* 2003;46:1401-1415.
- 30 Kim H, Hasegawa-Johnson M, Perlman A: Vowel contrast and speech intelligibility in dysarthria. *Folia Phoniatri Logop* 2011;63:187-194.
- 31 Sapir S, Spielman JL, Ramig LO, Story BH, Fox C: Effects of intensive voice treatment (the lee silverman voice treatment [lsvt]) on vowel articulation in dysarthric individuals with idiopathic parkinson disease: Acoustic and perceptual findings. *J Speech Lang Hear Res* 2007;50:899-912.
- 32 Bunton K, Kent RD, Duffy JR, Rosenbek JC, Kent JF: Listener agreement for auditory-perceptual ratings of dysarthria. *J Speech Lang Hear Res* 2007;50:1481-1495.
- 33 Bhuta T, Patrick L, Garnett JD: Perceptual evaluation of voice quality and its correlation with acoustic measurements. *Journal of Voice* 2004;18:299-304.
- 34 Kelchner LN, Brehm SB, Weinrich B, Middendorf J, deAlarcon A, Levin L, Elluru R: Perceptual evaluation of severe pediatric voice disorders: Rater reliability using the consensus auditory perceptual evaluation of voice. *Journal of Voice*;24:441-449.
- 35 Webb AL, Carding PN, Deary IJ, MacKenzie K, Steen N, Wilson JA: The reliability of three perceptual evaluation scales for dysphonia. *Eur Arch Oto-Rhino-Laryn* 2004;261:429-434.
- 36 Stevens SS, Galanter EH: Ratio scales and category scales for a dozen perceptual continua. *Journal of experimental psychology* 1957;54:377-411.
- 37 Kreiman J, Gerratt BR, Ito M: When and why listeners disagree in voice quality assessment tasks. *Journal of the Acoustical Society of America* 2007;122:2354-2364.
- 38 Zraick RI, Wendel K, Smith-Olinde L: The effect of speaking task on perceptual judgment of the severity of dysphonic voice. *Journal of Voice* 2005;19:574-581.
- 39 Zraick RI, Marshall W, Smith-Olinde L, Montague JC: The effect of task on determination of habitual loudness. *Journal of Voice* 2004;18:176-182.

- 40 Kreiman J, Gerratt BR: Comparing two methods for reducing variability in voice quality measurements. *J Speech Lang Hear Res* 2011;54:803-812.

Table 1: Participant data with composite mean intelligibility scores from unfamiliar listeners pre vs post intervention.

SD spastic and dyskinetic, S spastic, A athetosis, WD Worster Drought

GM = Gross Motor Function Classification System

Cog = Cognitive score (Surveillance of Cerebral Palsy in Europe database)

SC = Sessions completed

Child	Age	Sex	CP Type	GM	Cog	SC	Mean intelligibility	
							Pre	Post
1	15	F	SD	4	2	16	11.75	34.65
2	15	M	A	2	1	18	48.09	72.51
3	14	F	SD	5	2	17	44.77	61.28
4	15	M	S	4	1	16	67.37	90.46
5	12	M	S	4	2	17	29.43	41.73
6	12	M	S	2	2	18	10.84	19.27
7	13	F	SD	3	2	18	33.73	54.48
8	17	M	S	3	2	11	18.49	26.21
9	18	F	A	2	1	17	44.64	52.25
10	11	M	S	4	1	16	19.90	41.86
11	17	F	S	5	2	14	28.96	40.36
12	13	F	S	4	2	14	8.92	6.58
13	18	F	SD	5	2	14	8.67	16.67
14	14	M	S	4	2	15	27.05	57.10
15	13	F	S	2	1	15	39.43	71.04
16	16	F	WD	1	1	14	33.22	44.83

Table 2: Association between speech intelligibility and GRBAS scores. *based on mixed effects model with recordings nested within children; dependent variable = intelligibility; explanatory variable = GRBAS scale.

GRBAS scale	Bivariate correlation between Intelligibility - GRBAS			Change in intelligibility associated with unit change in GRBAS score*		
	R	95% CI		mean	95% CI	
Grade	-0.431	-0.612	-0.199	-10.0	-17.3	-2.6
Roughness	-0.051	-0.297	0.202	-3.6	-11.0	3.8
Breathiness	-0.371	-0.566	-0.130	-8.7	-15.1	-2.3
Asthenia	-0.392	-0.582	-0.154	-10.7	-16.6	-4.7
Strain	-0.111	-0.350	0.143	-1.3	-8.6	6.0

Figure 1: Assessment and intervention schedule

week 1	weeks 2-4	week 5	weeks 6-11	week 12	week 17
<ul style="list-style-type: none">•Day 1•CSIM list a•Cartoon strip a <ul style="list-style-type: none">•Day 2•CSIM list b•Cartoon strip b•Intervention as usual	<ul style="list-style-type: none">•Intervention as normal	<ul style="list-style-type: none">•Day 1•CSIM list c•Cartoon strip c <ul style="list-style-type: none">•Day 2•CSIM list d•Cartoon strip d•Intervention as usual	<ul style="list-style-type: none">•new intervention 3x 35-40 mins per week	<ul style="list-style-type: none">•Day 1•CSIM list e•Cartoon strip a <ul style="list-style-type: none">•Day 2•CSIM list f•Cartoon strip b•No intervention as usual	<ul style="list-style-type: none">•Day 1•CSIM list g•Cartoon strip c <ul style="list-style-type: none">•Day 2•CSIM list h•Cartoon strip d•afterwards return to intervention as usual