Abstract

Background: Patients with intelligibility difficulties associated with quiet voice are often prescribed a voice amplifier. This study examined whether artificial voice amplification improved intelligibility in people with Parkinson's (pwP) and whether there was an optimum increase that brought about best improvement.

Methods: Twelve pwP (four mild, eight moderate intelligibility difficulties) and five controls read low predictability sentences in their habitual voice. Audio-recordings were digitally manipulated to create samples at 2.3, 5 and 10 decibels amplification. Listeners transcribed the recorded sentences. Percentage words correctly identified was compared across levels of amplification and groups.

Findings: Moderately affected pwP were significantly less intelligible than controls in all conditions. Moderately, but not mildly affected pwP showed higher intelligibility in the amplified conditions, though statistically significantly only at +2.3dB. No other significant effects of intensity or interactions with groups were found. At an individual level some participants showed clear advantages of amplification.

Conclusion: Based on current participants, potential benefits of amplification cannot be promised to all pwP. Nevertheless several provisos regarding methods employed suggest the question can gainfully be pursued using broader measures to assess effects of amplification with more varied groups of pwP and with other aetiologies where voice production can be an issue.

Keywords: voice; loudness; Parkinson's; amplification; intelligibility

Introduction

How loud one's voice is depends on the degree of subglottal air pressure generated on expiration and the level of resistance that airflow encounters when the vocal cords close together. The greater the subglottal pressure and the tighter the vocal cord approximation, the greater the intensity of someone's voice will be and the louder they will be perceived to be speaking. People with acquired neurological speech disorders and a variety of other conditions that affect lung capacity, expiratory control, airway integrity or laryngeal function may face difficulties producing or sustaining voice loud enough for listeners to reliably understand. Apart from the obvious immediate barriers this poses for successful communication it can also exert considerable indirect influence on communication through psychosocial impact on speakers and added burden on carers (Miller *et al.*, 2006; Miller *et al.*, 2011; Mackenzie *et al.*, 2013). Intervention typically entails training the speaker to use a louder voice. Indeed increasing the volume of one's voice represents an (un)consciously adopted strategy by any speaker if they are not being heard/understood. Such training has proved successful for certain speakers, situations and conditions (de Angelis *et al.*, 1997; Tjaden and Wilding, 2004; Yorkston *et al.*, 2007; Fox *et al.*, 2012; Tjaden *et al.*, 2014; Wight and Miller, 2015). In particular, for people with Parkinson's (pwP) a quieter voice is an early and restricting handicap (Miller *et al.*, 2006; Sapir, 2014). A key active intervention for pwP is the Lee Silverman Voice Treatment programme (Fox *et al.*, 2012), with its slogan 'Think Loud'. This focuses exclusively on speaker initiated resetting and monitoring of voice intensity associated with the impaired sense of effort in voice production in Parkinson's.

For a variety of reasons such programmes may be contraindicated or unsuccessful and artificial amplification may be contemplated as an alternative or adjunctive strategy to bypass the cognitive and/or physical hurdles the individual faces (Cariski and Rosenbek, 1999; Spencer *et al.*, 2003; Hargrove *et al.*, 2009; Adams *et al.*, 2010; Kim and Jo, 2013). However, speaker-initiated increase in volume and artificial amplification have different modes of operation. In particular self-initiated intensification can, alongside sound pressure level increases, alter speech breathing patterns, stress and intonation parameters, rate of speech, acoustic variables and articulatory accuracy, all of which separately may heighten intelligibility independently of intensity change (Tjaden and Wilding, 2004; Bunton, 2006; Huber and Chandrasekaran, 2006; Watson and Hughes, 2006). These multiple spontaneous adjustments doubtless contribute to the success of speaker-controlled intensity change.

The same changes, however, are believed not to occur naturally in artificially amplified speech, which alters only listener perceived loudness, not temporal or acoustic aspects of the speech signal. This underlines why the two methods cannot be considered equivalent (Neel, 2009; Kim and Kuo, 2012). Use of artificial amplification to overcome vocal asthenia is supported (Green and Watson, 1968; Adams, 1997; Bain *et al.*, 2005), though with some reservations, at least as far as single word intelligibility is concerned (Turner *et al.*, 2008). Several issues concerning for whom and how much this might apply remain open. This study addresses two related questions in relation to artificial amplification: whether simply amplifying the voice does produce greater intelligibility; and, if it does, what magnitudes of increase are required to make a significant difference to intelligibility? Through this it is hoped to contribute to understanding when artificial amplification may be helpful and what increases should be targeted to bring about improved communication.

Variables affecting intelligibility

Clearly the suitability of different voice intensity levels is dependent on a whole range of speaker, listener and environmental influences. These include how severely a person's voice intensity is impaired, whether the problem is restricted to voice production or involves articulatory problems too. It can depend on the amount of ambient background noise, reverberation within a room, the distance from speaker to listener, the listener's hearing acuity and auditory processing abilities (D'Innocenzo *et al.*, 2006). It differs according to the topic being discussed; the semantic (meaning), syntactic (grammar) and phonological (sounds) redundancy of the message; single word vs connected speech tasks; as well as familiarity of the listener with disordered speech, with the specific speaker, whether audio vs audiovisual clues to meaning are present, and so forth (Miller, 2013).

To gain a degree of control over these variables in the present study, certain choices were made. To achieve some uniformity of participants and comparability of outcomes across speakers and groups we involved unimpaired speakers and people with Parkinson's, a condition noted for voice intensity impairment (Sapir, 2014). Participants were divided into those with no or minimal articulatory changes and those with more pronounced decline. Rather than asking listeners to rate adequacy or level of perceived loudness (measures that have been linked with reliability issues (Kreiman *et al.*, 2007)) we compared transcription intelligibility scores (Hustad, 2006) across different amplification conditions. To remain closer to natural speech but still maintain some control over speaker output (Miller, 2013) sentence intelligibility tasks were favoured over single words. To minimise signal independent factors in intelligibility we utilised sets of unpredictable sentences in an audio only mode (McHenry and Parle, 2006).

Questions

The following questions were posed: does speech that is synthetically amplified in intensity differ in intelligibility from habitual production? Is the relationship between intensity and intelligibility change linear (i.e. the greater the amplification, the greater the intelligibility gain)? Do effects vary depending on presence or not of neurological speech impairment and its severity?

METHODS

This study was conducted in accordance with University (withheld for blind review) ethics committee approved procedures which, amongst other stipulations, ensured voluntary, informed, anonymous participation, with right to leave without reason at any stage in the research.

Participants

Participants who provided the sentence recordings were twelve pwP) and five neurologically healthy speakers matched in overall age to the pwP (Table 1). Inclusion criteria were: reading ability self-reported as normal; vision and hearing normal or corrected normal; no history of neurological conditions (except solely Parkinson's for the pwP) and, for pwP no self or clinician reported prominent changes to articulation. PwP were recruited from a Parkinson's support group. PwP were in Hoehn and Yahr stages I-III (Goetz, Poewe, Rascol et al, 2004), caring for themselves and able to make their way to the support group. Control participants were friends/ family members of pwP.

Thirty-three listeners (ages 16-71 years, mean 41; 15 male) listened to the recordings to provide the intelligibility scores. They were native English speakers, with self-reported normal (or corrected to normal) vision and hearing and no professional or personal experience of dysarthria or rating speech.

Speech sample recording and stimulus preparation

We employed low predictability sentences (McHenry and Parle, 2006), all seven words long (mean 5.5 letters per word) and suitable for 6.6 grade reading level (e.g. Animals often wander across wooded grassy paths. Nice men usually grill better fresh vegetables). Sentences were thereby long enough to tax speaking ability of individuals with dysarthria, short enough to transcribe without requiring repetition, and minimalized clues to words heard from semantic and syntactic context. In their original validation study McHenry et al presented readers with twenty-five sets of 70 sentences with one word missing in varying positions. They asked 25-30 readers to supply the missing words for each set. Only one sentence had zero predictable words in any position. Fifty-six sentences met the operational definition of unpredictable as any word that was not guessed by more than 20% of respondents. These 56 sentences were transcribed from free-field loudspeakers. Fifty sentences were transcribed with greater than 80% accuracy by everyday listeners and these are the items employed in the current study.

From these we assigned the first 10 to the first speaker, the second 10 to the second speaker and so forth. When the bottom of the list was reached it was (re)randomised and the process repeated. Stimuli were printed one per sheet, Calibri font size 18. Participants familiarised themselves with the words and sentences before recording commenced. Audio-recording employed an Edirol R-09HR digital voice recorder set at a sampling frequency of 44.1 kHz and an input volume level of 60, with an AKG C520 head-mounted microphone to minimise head movement artefact and control mouth to microphone distance. Recordings took place at the participant's home with environmental sounds reduced to minimum.

The audio-recordings were digitised in .wav format at a 16 bit resolution. Each sentence by each speaker was saved as a separate file and imported into Audacity (1.3 beta) software program. Four of the 170 files (17 participants, 10 sentences each) were discarded due to unacceptable levels of distortion. The remainder were 'cleaned' to remove device feedback and background noise, by applying the Noise Removal and High Pass Filter (150.0Hz) effects. The resultant tracks were saved as the habitual/baseline condition.

Each of these baseline items then had the Amplify effect applied to produce tracks amplified by +2.3dB, +5.0dB and +10dB. These increases in sound pressure level (SPL) represent the approximate minimum and maximum post voice loudness therapy increases reported in the literature (Wight and Miller, 2015) and an approximate midpoint between them. They also represent a span of perceptible change where +1dB is imperceptible, +3dB about perceptible, +5dB clearly noticeable, and +10 perceived as twice as loud (Chasin and Russo, 2004). The baseline recordings and amplified sentences were stored as individual files (n 664, four SPL levels per sentence per speaker). These were randomised and the first 66 were assigned to a first listener, the second 66 to a second listener and so on. When the bottom of the list was reached the list was re-randomised and the process repeated to minimise possible order effects.

Intelligibility scoring

Listeners carried out the task alone with the researcher in a sound deadened room. They were blind to the object of the investigation, the participants and the content of sentences. Before commencing the investigatory task they transcribed a set of practice sentences spoken by a dysarthric speaker, separate from the main study stimuli and participants, to minimise effects of unfamiliarity with the exercise and possible practice effects over the initial investigatory transcriptions. Each investigatory item was transcribed verbatim by three listeners, with no repetitions permitted. Each sentence therefore received a total intelligibility score out of 21 (7 words per sentence x 3 listeners).

Recordings were played through a Toshiba Satellite Pro laptop with a Realtek HD Audio output soundcard with an external loudspeaker positioned 1m from the listener to reflect the approximate distance between speakers in a 1:1 conversation. An identical playback level was retained for all listeners. Progress to the next item happened when the transcriber indicated they were ready. One point was awarded for each correctly transcribed word (accepting homophone and spelling errors clearly related to the target, e.g. *threw* for *through*, *instructers* for *instructors*).

Data processing and analysis

Raw scores were total words identified per speaker per condition. The pwP were divided into mild vs moderate groups using a cut off score of 80% words recognised based on the intelligibility score from listeners transcribing the habitual (non-amplified) speech samples. Data was checked for normality of distribution and equality of variance. Analyses were conducted in SPSS v17.00. A one-way ANOVA determined whether there were any significant differences in intelligibility between loudness settings (+0dB, +2.3dB, +5db, +10dB). A second two way ANOVA (pwP mild, pwP moderate, control speakers x dB levels) examined whether there was an interaction between loudness level and groups. Alpha was set at p 0.05, with Bonferroni corrections for multiple comparisons.

RESULTS

Participants' age and gender and mean intelligibility in the habitual condition appear in Table 1. Mean age of pwP was 75.83 years (SD 8.12, range 64-88) and speakers without Parkinson's mean 65 years (SD 20.77, range 33-91), with no significant difference in age (Mann Whitney, p 0.29).

Table 1 about here

Four of the baseline sentences had to be discarded due to distortion, leaving speaker 5 with maximum score 168 (7 words per sentence x 3 listeners x 8 sentences) and speakers 9 and 13 maximum 189. To maintain comparability of scores across speakers total words recognised was converted to a percentage of the maximum score possible. Summary mean intelligibility scores per group appear in table 2. Figure 1 presents the individual score profiles.

Table 2 about here

Figure 1 about here

A one-way ANOVA was conducted on the combined speaker groups (n 17) and the 4 x SPL levels to determine whether synthetically amplified speech differed in intelligibility from habitual speech. There was no effect of loudness ($F_{(3,64)} = 0.283$, $p > 0.05$), indicating no significant difference in intelligibility between the four loudness conditions. Within groups the only significant differences between levels concerned a significant difference for the moderately affected pwP between habitual and +2.3dB conditions (p 0.025). For the moderately affected group mean intelligibility was higher than habitual in the +5dB and +10dB conditions, but not statistically significantly (+5dB, p 0.07; +10dB, p 0.21). There were no other within group significant differences and no interactions between groups and levels.

Table 3 summarises the differences in intelligibility between groups at their habitual sound pressure level and other levels of amplification. It indicates no significant differences between the control speaker group and mildly affected pwP at any intensity levels; a significant difference between control speakers and moderately affected pwP at all SPL levels; significant differences between mildly vs moderately affected pwP in habitual and +10dB conditions, but not in the others.

Table 3 about here

Discussion

Participants without Parkinson’s had significantly higher intelligibility than the moderately affected group in all conditions, as expected. The mildly affected pwP performed within ranges comparable with the control speakers. Mildly affected pwP were perceived as more intelligible than moderately affected pwP at habitual and +10dB levels. The non-significant differences between Parkinson’s groups at +2.3 and +5dB presumably reflect the improved intelligibility of the moderate group at these levels.

In answer to the question of whether synthetically amplified speech differs in intelligibility from habitual productions, the data indicate not. Neither was there any evidence that there is a linear improvement according to intensity gain. Only speakers 3, 9 and 11 (figure 1) displayed anything approaching a linear increase across the intensity conditions, but this was not apparent in any other participants. Outcomes thereby support others’ findings that simply artificially increasing speech intensity alone may produce at most limited gains in intelligibility (Turner *et al.*, 2008; Neel, 2009; Kim and Kuo, 2012).

Although only one comparison produced a statistically significant difference across intensity levels, the possibility exists, given the multiple comparisons, that this was a chance finding. However, data from the two other amplification conditions, where mean scores were higher for the moderately affected group, though not significantly so, suggest a more definitive decision on the issue might be pursued with larger groups, more items and maybe with addition of listener perceived impressions of adequacy of communication in addition to the transcription intelligibility scores.

As regards possible interactions between severity of neurological speech impairment and increasing speech intensity the present study found no such effects in terms of intelligibility. However, a salient factor here may have been the distribution of intelligibility levels. The control speakers and mildly affected pwP had little room for improvement on scores in the amplified conditions and this may have created a ceiling effect. In audio recording based intelligibility tasks even control speakers do not attain 100% (Miller *et al.*, 2007), as was clear in the current data. There was no conclusive relationship between age and intelligibility in the control speakers for those attaining less than 90% intelligibility (older age may be associated with voice-speech changes). Reasons for lower than 100% intelligibility may therefore lie with the task (unpredictable sentences are more difficult than day to day speech) and the number of items employed (mishearing 8 or more words from total possible 70 would bring a score below 90%). To add more sensitivity at the milder end of the intelligibility spectrum further research in this field may add tasks such as listening in noise or speakers producing speech within a dual task paradigm (Bunton and Keintz, 2008; Dromey *et al.*, 2010). A greater number of words to be transcribed may also ameliorate the effects small changes in words recognised having a disproportionate effect on percentage scores.

On the other hand, the samples may under-reflect the severity of the participants' unintelligibility. PwP can temporarily raise their speech performance when concentrating in short bursts on an important activity (Miller *et al.*, 2006), and healthy speakers can show increased intensity simply by being seated before a microphone (Goberman *et al.*, 2010). Either way, future research into this area should aim to include speech samples from a wider severity spread than the present study. They should also include participants with different aetiologies for their quieter voices, not just in terms of other neurological conditions, but also from other conditions impacting speech-voice output. Added data that may have assisted interpretation would have been details of speaker habitual sound pressure levels (SPL). This could have given a clearer indication than sentence intelligibility alone of how affected people were by their Parkinson's. We did not have access to the technological facility to measure SPL across varied home settings, and in particular not for measuring variability in connected speech. A future laboratory based study would enable this data to be placed alongside intelligibility scores.

Gains in intelligibility with voice amplification have generally relied on or included speaker generated amplification. As noted previously, for a gain in intensity to translate to a gain in intelligibility it appears that changes need to encompass all the levels of speech production, covering rate, articulatory excursions and preciseness, intensity and pitch variation and subglottal air pressure. Thus, clinically self-initiated volume increase may be gainfully combined with amplification, and future work should examine effects of artificial amplification alone and in combination with listener-initiated gain.

The pwP were selected to have no or minimal perceived articulatory imprecision. If some imprecise sounds were present the main problem in communication had to be voice intensity. This screening was based on reports from speech and language clinicians and speaker and carer perceptions. Effects of articulatory changes on intelligibility should thus have been minimal. This does not preclude that there were nevertheless some changes (e.g. to rate, fluency, and some imprecise vowels and consonants) and these contributed to listener scores. This would reflect the findings of e.g. (Turner *et al.*, 2008) who commented that as hypophonia in pwP rarely occurs in isolation from other alterations to speech output, artificially increasing speech intensity will only amplify any speech sound distortions without repairing them, leading to little difference in speech intelligibility. On the other hand, if the conclusions of Turner et al were fully the case the control group at least should have shown the expected pattern – though again, attention is drawn to possible ceiling effects here. This issue could be addressed if a repeat study gathered objective data on speech sound and sound contrast production.

An additional possibility concerns the quality of artificially amplified speech. Research into the correction of hearing loss has noted that linear amplification often causes soft input sounds not to be amplified enough whilst causing loud input sounds to be amplified too much, often exceeding an uncomfortable listening level (Sockalingam *et al.*, 2011). Furthermore, linear amplification is limited by the occurrence of peak clipping when the amplifier is pushed beyond its maximum output. This type of limiting causes various forms of distortion that have been found to hinder intelligibility and impair subjective quality of speech (Bray and Nilsson, 2009). Even though tracks were screened for clipping, it remains a possibility that these inadequacies of linear amplification contributed to the results of the present study.

To address the issue of concomittant articulatory distortion work has begun to combne amplification with signal clarification (Cariski and Rosenbek, 1999; Bain *et al.*, 2005; Kain *et al.*, 2007). However, more widely results remain mixed. Enhancement can be effective to some degree in increasing intelligibility, but degree of improvement varies greatly between individuals. Furthermore, results must be interpreted cautiously due to methodological limitations in these studies such as use of imprecise measurement procedures, lack of inferential statistics, insufficient sample size and listener familiarity with the dysarthric participants.

People with reduced voice amplitude report that amplification enables them to address larger audiences or speak in situations where the listener is not immediately next to them. We assessed only in a simulated 1:1 condition. Future work may add simulation of greater speaker-listener distances or presence of different levels and types of competing background noise.

An additional factor may concern the intelligibility task. Understanding what a speaker says is a different perceptual task to accurately transcribing each word that is said (Tjaden and Wilding, 2011). Although the significant interaction found between the percentage intelligibility scores and severity groups suggests that the task was accurately representing speech severity, the transcription task may not have been representative of the processes used to understand natural spontaneous speech (Hustad, 2008). Further, the short, relatively simple utterances did not require listeners to process linguistic or emotional parameters and lacked broader context. Future research could examine a range of speech and voice abilities by eliciting a variety of task types and complexities as well as including a measure of listener comprehension to better measure any functional change in intelligibility (Hustad, 2008; McLeod *et al.*, 2012).

Finally, whilst randomisation of conditions, speakers and sentences was carried out to minimise learning and order effects for listeners, it is inevitable that some listeners will have heard a given sentence more than once in their selected items. Though this may have given some clues to target words, it is unlikely to be a major distorting factor in scores here. Nevertheless, a replication of the work could employ more listeners to minimise or remove the possibility of auditing a same sentence more than once.

Conclusions

Whilst the present study provides little support for the use of artificial amplification in isolation to circumvent intelligibility issues, at least in the participants here, there are sufficient variables to incorporate into future work that mean the question remains a live issue. These have been mentioned in the discussion above. Thus, the results do not preclude considering artificial amplification. They do suggest, however, that it will not work, or work equally, for everyone in all circumstances. Consideration of whether to introduce artificial amplification must weigh up the whole speech-voice and cognitive-language profile of the speaker and the situations in which they might need to apply amplified voice. Results also underline that when employing amplification methods, performance needs to be closely monitored through a variety of objective methods that quantify intelligibility and functional communication in a range of tasks and situations.

Key phrases

Artificially increasing speech intensity does not necessarily lead to increased intelligibility in people with Parkinson's disease and clinician advice to individuals should clarify this fact.

There is not a linear relationship between increasing speech intensity and intelligibility.

It is essential for clinicians to examine a person's whole communication profile when considering artificial implication as an intervention.

Further research is needed to explore the effects of amplification on intelligibility, for both people with Parkinson's disease and a range of other conditions that affect speech intensity.

Conflict of interest

The authors declare no conflicts of interest, financial or otherwise.

.

References

Adams, S. (1997) 'Hypokinetic dysarthria', in McNeil, M. (ed.) *Clinical management of sensorimotor speech disorders*. New York: Thieme, pp. 261-285.

Adams, S.G., Winnell, J. and Jog, M. (2010) 'Effects of Interlocutor distance, multi-talker background noise, and a concurrent manual task on speech intensity in Parkinson's Disease', *Journal of Medical Speech-Language Pathology*, 18(4), pp. 1-8.

Bain, C., Ferguson, A. and Mathisen, B. (2005) 'Effectiveness of the speech enhancer on intelligibility: A case study', *Journal of Medical Speech-Language Pathology*, 13(2), pp. 85-95.

Bunton, K. (2006) 'Fundamental frequency as a perceptual cue for vowel identification in speakers with Parkinson's disease', *Folia Phoniatrica*, 58(5), pp. 323-339.

Bunton, K. and Keintz, C. (2008) 'Use of dual task paradigm for assessing speech intelligibility in clients with Parkinson's disease', *J Medical Speech-Language Pathology*, 16(3), pp. 141-155.

Cariski, D. and Rosenbek, J. (1999) 'The effectiveness of the Speech Enhancer (TM)', *Journal of Medical Speech-Language Pathology*, 7(4), pp. 315-322.

Chasin, M. and Russo, F.A. (2004) 'Hearing AIDS and music', *Trends in Amplification*, 8(2), pp. 35-47.

D'Innocenzo, J., Tjaden, K. and Greenman, G. (2006) 'Intelligibility in dysarthria: Effects of listener familiarity and speaking condition', *Clinical Linguistics & Phonetics*, 20(9), pp. 659-675.

de Angelis, E.C., Mourao, L.F., Ferraz, H.B., Behlau, M.S., Pontes, P.A. and Andrade, L.A. (1997) 'Effect of voice rehabilitation on oral communication of Parkinson's disease patients', *Acta Neurologica Scandinavica*, 96(4), pp. 199-205.

Dromey, C., Jarvis, E., Sondrup, S., Nissen, S., Foreman, K.B. and Dibble, L.E. (2010) 'Bidirectional interference between speech and postural stability in individuals with Parkinson's disease', *International Journal of Speech-Language Pathology*, 12(5), pp. 446-454.

Fox, C., Ebersbach, G., Ramig, L. and Sapir, S. (2012) 'LSVT LOUD and LSVT BIG: Behavioral treatment programs for speech and body movement in Parkinson Disease', *Parkinsons Disease*. Article ID 391946, http://dx.doi.org/10.1155/2012/391946

Goberman, A.M., Recker, B. and Parveen, S. (2010) 'Performance effect: does the presence of a microphone influence Parkinsonian speech?', *Journal of Medical Speech-Language Pathology*, 18(4), pp. 40-45.

Goetz, C.G., Poewe, W., Rascol, O., Sampaio, C., Stebbins, G.T., Counsell, C., Giladi, N., Holloway, R.G., Moore, C.G., Wenning, G.K., Yahr, M.D. and Seidl, L. (2004) 'Movement disorder society task force report on the Hoehn and Yahr staging scale: Status and recommendations', *Movement Disorders*, 19(9), pp. 1020-1028.

Green, M. and Watson, R. (1968) 'The value of speech amplification in Parkinson's disease patients', *Folia Phoniatrica*, 20, pp. 250-257.

Hargrove, P., Anderson, A. and Jones, J. (2009) 'A critical review of interventions targeting prosody', *International Journal of Speech-Language Pathology*, 11(4), pp. 298 - 304.

Huber, J.E. and Chandrasekaran, B. (2006) 'Effects of Increasing sound pressure level on lip and jaw movement parameters and consistency in young adults', *J Speech Lang Hear Res*, 49(6), pp. 1368-1379.

Hustad, K.C. (2006) 'Estimating the intelligibility of speakers with dysarthria', *Folia Phoniatrica*, 58(3), pp. 217-228.

Hustad, K.C. (2008) 'The relationship between listener comprehension and intelligibility scores for speakers with dysarthria', *J Speech Lang Hear Res*, 51(3), pp. 562-573.

Kain, A.B., Hosom, J.-P., Niu, X., van Santen, J.P.H., Fried-Oken, M. and Staehely, J. (2007) 'Improving the intelligibility of dysarthric speech', *Speech Communication*, 49(9), pp. 743-759.

Kim, S.J. and Jo, U. (2013) 'Study of accent-based music speech protocol development for improving voice problems in stroke patients with mixed dysarthria', *NeuroRehabilitation*, 32(1), pp. 185-190.

Kim, Y. and Kuo, C. (2012) 'Effect of level of presentation to listeners on scaled speech intelligibility of speakers with dysarthria', *Folia Phoniatrica*, 64(1), pp. 26-33.

Kreiman, J., Gerratt, B.R. and Ito, M. (2007) 'When and why listeners disagree in voice quality assessment tasks', *Journal of the Acoustical Society of America*, 122(4), pp. 2354-2364.

Mackenzie, C., Kelly, S., Paton, G., Brady, M. and Muir, M. (2013) 'The Living with Dysarthria group for post-stroke dysarthria: the participant voice', *International Journal of Language & Communication Disorders*, 48(4), pp. 402-420.

McHenry, M.A. and Parle, A.M. (2006) 'Construction of a set of unpredictable sentences for intelligibility testing', *Journal of Medical Speech-Language Pathology*, 14(4), pp. 269-271.

McLeod, S., Harrison, L.J. and McCormack, J. (2012) 'The Intelligibility in Context Scale: validity and reliability of a subjective rating measure', *J Speech Lang Hear Res*, 55(2), pp. 648-656.

Miller, N. (2013) 'Measuring up to speech intelligibility', *International Journal of Language and Communication Disorders* 46(6), pp. 613-624.

Miller, N., Allcock, L., Jones, D., Noble, E., Hildreth, A.J. and Burn, D.J. (2007) 'Prevalence and pattern of perceived intelligibility changes in Parkinson's disease', *J Neurol Neurosurg Psychiatry*, 78(11), pp. 1188-1190.

Miller, N., Andrew, S., Noble, E. and Walshe, M. (2011) 'Changing perceptions of self as a communicator in Parkinson's disease: a longitudinal follow-up study', *Disability & Rehabilitation*, 33(3), pp. 204-10.

Miller, N., Noble, E., Jones, D. and Burn, D. (2006) 'Life with communication changes in Parkinson's disease, *Age Ageing*, 35(3), pp. 235-239.

Neel, A.T. (2009) 'Effects of loud and amplified speech on sentence and word intelligibility in Parkinson Disease', *Journal of Speech Language and Hearing Research*, 52(4), pp. 1021-1033.

Sapir, S. (2014) 'Multiple factors are involved in the dysarthria associated with Parkinson's Disease: a review with implications for clinical practice and research', *Journal of Speech Language and Hearing Research*, 57(4), pp. 1330-1343.

Sockalingam, R., Lundh, P. and Schum, D.J. (2011) 'Severe to profound hearing loss: What do we know and how do we manage it', *Hearing Review*, 18(1), pp. 30-33.

Spencer, K.A., Yorkston, K.M. and Duffy, J.R. (2003) 'Behavioral management of respiratory/phonatory dysfunction from dysarthria: A flowchart for guidance in clinical decision making', *Journal of Medical Speech-Language Pathology*, 11(2), pp. XXXIX-LXI.

Tjaden, K., Sussman, J.E. and Wilding, G.E. (2014) 'Impact of clear, loud, and slow speech on scaled intelligibility and speech severity in Parkinson's Disease and Multiple Sclerosis', *Journal of Speech, Language, and Hearing Research*, 57(3), pp. 779-792.

Tjaden, K. and Wilding, G. (2011) 'Effects of speaking task on intelligibility in Parkinson's disease', *Clinical Linguistics & Phonetics*, 25(2), pp. 155-168.

Tjaden, K. and Wilding, G.E. (2004) 'Rate and loudness manipulations in dysarthria: acoustic and perceptual findings', *J Speech Lang Hear Res*, 47(4), pp. 766-783.

Turner, G., Martin, H.J. and de Jonge, R. (2008) 'Influence of amplification on word intelligibility and phonetic contrast errors in dysarthria', *Journal of Medical Speech-language Pathology*, 16, pp. 267-273.

Watson, P.J. and Hughes, D. (2006) 'The relationship of vocal loudness manipulation to prosodic F0 and durational variables in healthy adults', *Journal of Speech Language and Hearing Research*, 49(3), pp. 636-644.

Wight, S. and Miller, N. (2015) 'Lee Silverman Voice Treatment for people with Parkinson's: audit of outcomes in a routine clinic', *International journal of language & communication disorders / Royal College of Speech & Language Therapists*, 50(2), pp. 215-25.

Yorkston, K.M., Hakel, M., Beukelman, D.R. and Fager, S. (2007) 'Evidence for effectiveness of treatment of loudness, rate, or prosody in dysarthria: A systematic review', *Journal of Medical Speech-Language Pathology*, 15(2), pp. XI-XXXVI.

Table 1: Age (years), gender and mean intelligibility in habitual condition for participants with Parkinson's and without (participants 13-17).

Partic-ipant	Gender	Habitual	Age
1	F	77.14	64
2	M	75.71	69
3	M	56.67	71
4	M	73.33	85
5	M	82.14	78
6	M	93.81	84
7	F	82.38	77
8	M	71.43	70
9	F	74.07	67
10	F	88.10	72
11	M	52.38	88
12	M	74.29	85
13	F	89.42	91
14	F	97.14	70
15	M	91.43	33
16	F	85.71	65
17	F	85.24	66

Table 2: Summary mean percentage intelligibility scores per group (1 pwP mild; 2 pwP moderate; 3 control) and spl (sound pressure level). Hab = habitual level

Group and spl conditions		**N**	**Mean**	**SD**	**95% Confidence Interval** Lower Bound	Upper Bound	**Minimum mean**	**Maximum mean**
Hab	1.00	4	86.61	5.53	77.80	95.41	82.14	93.81
	2.00	8	69.38	9.39	61.53	77.23	52.38	77.14
	3.00	5	89.79	4.85	83.76	95.81	85.24	97.14
	Total	17	79.44	12.13	73.20	85.67	52.38	97.14
+2.3dB	1.00	4	85.98	4.59	78.68	93.28	81.55	91.90
	2.00	8	76.52	8.99	69.01	84.02	63.33	85.71
	3.00	5	89.14	6.76	80.75	97.54	83.33	99.52
	Total	17	82.46	9.24	77.71	87.21	63.33	99.52
+5dB	1.00	4	81.28	9.23	66.59	95.97	68.45	90.48
	2.00	8	75.26	6.52	69.81	80.71	65.71	82.54
	3.00	5	92.06	3.86	87.27	96.86	85.71	95.71
	Total	17	81.62	9.63	76.67	86.57	65.71	95.71
+10dB	1.00	4	86.52	1.51	84.11	88.92	85.12	88.57
	2.00	8	73.13	6.57	67.63	78.62	63.81	83.60
	3.00	5	88.12	4.63	82.36	93.87	82.86	95.24
	Total	17	80.69	8.89	76.12	85.26	63.81	95.24

Table 3: Comparison of intelligibility levels between groups at different levels of amplification

spl = sound pressure level

SPL	**Group**	**Comparator group**	**p**
Habitual	PwP mild	PwP moderate	.007
		Control	1.000
	PwP moderate	Control	.001
+2.3dB	PwP mild	PwP moderate	.185
		Control	1.000
	PwP moderate	Control	.034
+5dB	PwP mild	PwP moderate	.478
		Control	.088
	PwP moderate	Control	.002
+10dB	PwP mild	PwP moderate	.003
		Control	1.000
	PwP moderate	Control	.001

Figure 1 Percentage total words recognised by listeners per condition for all speakers