Research Article

Loudness Perception of Pure Tones in Parkinson's Disease

Defne Abur,[a] Ashling A. Lupiani,[a] Ann E. Hickox,[b]
Barbara G. Shinn-Cunningham,[c] and Cara E. Stepp[a,c,d]

Purpose: Reduced intensity is a hallmark of speech production in Parkinson's disease (PD). Previous work has examined the perception of intensity in PD to explain these speech deficits. This study reports loudness ratings of pure tones by individuals with PD and controls, all with normal thesholds for older adults.
Method: Twenty individuals with PD and 23 age- and sex-matched controls rated the loudness of pure tones from 1 (*very soft*) to 7 (*uncomfortably loud*). Tones at 500, 750, 1000, 2000, and 4000 Hz were presented from 35 to 80 dB HL (or until a rating of 7 was given). A mixed-model analysis of variance was performed on ratings to assess

the effects of group, frequency, sound intensity, and ear. Loudness growth slopes were determined for each participant and analyzed by group.
Results: The mean loudness growth slopes of the control and PD groups did not significantly differ.
Conclusions: No difference was found in loudness growth slopes in response to externally generated tones in PD. This is in contrast with the findings of previous studies of self-generated speech and externally presented speech. The underlying causes for impaired perception and production of loudness in PD require further investigation.

P arkinson's disease (PD) is the second most common neurogenerative disorder (Mayeux, 2003; Tanner & Aston, 2000), and it presents with both motor and nonmotor complications (Jankovic, 2008). The lifetime risk of PD is estimated to be between 1% and 2%, with a higher risk in men than in women based on patient data from the Rochester Epidemiology Project (Elbaz et al., 2002). The incidence of PD worldwide has been reported to increase with age (Pringsheim, Jette, Frolkis, & Steeves, 2014), and previous work has projected that the number of people affected by PD will reach 8.67 million worldwide by 2030. The disease is commonly characterized by the first occurring motor symptoms, tremor or rigidity (Hoehn & Yahr, 1998), but nonmotor symptoms that increase with age and PD disease severity (Chaudhuri, Healy, & Schapira, 2006) also negatively affect quality of life (Global Parkinson's

Disease Survey Steering Committee, 2002; Martinez-Martin, Rodriguez-Blazquez, Kurtis, & Chaudhuri, 2011; Witjas et al., 2002). Among nonmotor symptoms, hypokinetic dysarthria impacts daily communication (McNamara & Durso, 2003; Miller, Noble, Jones, & Burn, 2006). Hypokinetic dysarthria is a motor speech disorder affecting vocal characteristics, such as prosody, articulation, and sound pressure level (Aronson & Brown, 1975). Hypokinetic dysarthria is widespread in PD, with 89% affected (Logemann, Fisher, Boshes, & Blonsky, 1978). Most commonly, this presents as hypophonia (Duffy, 2013), or reduced sound pressure level of speech, and an impairment in self-perception of voice loudness (Ho, Bradshaw, & Iansek, 2000). Although successful speech therapy protocols have been developed to address hypophonia, maintaining the treatment effects in the long term can be challenging (Fox, Morrison, Ramig, & Sapir, 2002), and the physiological causes of changes in vocal loudness production and perception in PD are still not well understood.

Previous work has examined the perception of speech sound pressure level in PD in both active speech production and passive listening tasks, although these studies have varied substantially in terms of experimental methods and the degree of experimental control over hearing status. Several studies have found evidence for atypical perception of active, self-generated, speech sound pressure level in PD (Fox & Ramig, 1997; Ho et al., 2000; Ho, Iansek, Marigliani,

[a]Department of Speech, Language, and Hearing Sciences, Boston University, MA
[b]Decibel Therapeutics, Boston, MA
[c]Department of Biomedical Engineering, Boston University, MA
[d]Department of Otolaryngology–Head and Neck Surgery, Boston University School of Medicine, MA

Correspondence to Defne Abur: defneabur@gmail.com

Bradshaw, & Gates, 1999; Marigliani, Gates, & Jacks, 1997; Marsden, 1982). Notably, Ho et al. (2000) showed an overestimation of voice intensity in typical-hearing individuals with PD compared to typical-hearing controls both during active self-productions and during a passive task in which participants adjusted a knob to match loudness levels while listening to a playback of their previously self-generated speech through headphones. Further investigation into the perception of loudness of externally presented stimuli has followed with inconclusive findings. Clark, Adams, Dykstra, Moodie, and Jog (2014) investigated loudness perception using a passive task with speech stimuli that were not self-generated in individuals with PD with normal hearing sensitivity who presented with hypophonia. Participants used a direct magnitude estimation (DME) task in which prerecorded sentences were played from a loudspeaker. A difference in the loudness growth slopes between groups was found: The PD group had a significantly shallower loudness slope compared to the controls. De Keyser et al. (2016) investigated loudness perception using a similar rating task in PD; however, the methods differed slightly from Clark et al. (2014). Participants were not screened for hearing status, hypophonia was not used as an inclusion factor, the participant's own recorded speech was used for playback perception, and a visual analog scale (VAS) was used to acquire loudness ratings. Results showed a similar trend to the results of Clark et al. (2014), in which individuals with PD tended to have higher loudness ratings for lower intensities (60 and 65 dB SPL) and lower loudness ratings for higher intensities (75 and 80 dB SPL) compared to controls. However, no significant difference was found between the mean loudness growth slopes of the PD group and the control group. Finally, Dromey and Adams (2000) examined loudness perception of nonspeech stimuli during a passive task in PD. Participants with PD as well as controls used a DME task to quantify the loudness of warble tones at 500 Hz presented from a loudspeaker at varying sound pressure levels. Participants' hearing thresholds were then correlated with the loudness perception data. No statistically significant differences were found between the loudness ratings by group, and no significant interactions were found between hearing sensitivity and loudness ratings.

The differing modes of playback (headphones or loudspeaker) and measurement scale (DME or VAS), as well as different criteria for participant hearing status used in previous studies, may have affected findings, making it difficult to determine whether or not voice perception deficits in PD extend to sounds that are not self-generated speech, such as externally generated speech or tones. It is important to better characterize whether there is a general perceptual deficit of sound intensity in PD in response to external sounds, in addition to self-generated speech, to provide more insight into whether auditory neural pathways are affected in PD and whether any dysfunction in auditory perception may underlie alterations in intensity of speech production.

The current study reports a systematic investigation of the loudness perception of pure tones over a large range of frequencies (500–4000 Hz) and a large range of sound intensities (35–80 dB HL) using a bounded categorical scale (Gu, Halpin, Nam, Levine, & Melcher, 2010). Careful audiometric screenings for typical older-adult hearing (Schow, 1991) were conducted prior to the experimental task. Tones were then presented unilaterally, beginning with the left ear, through insert earphones. A bounded categorical scale was chosen in order to keep the range of loudness ratings the same across participants. Given the trends reported in Clark et al. (2014) and De Keyser et al. (2016), we hypothesized that individuals with PD would have significantly shallower loudness growth slopes to pure tones compared to healthy controls across the tested frequencies. We additionally hypothesized that the loudness growth patterns in PD would correlate with disease progression as measured by the Movement Disorder Society-sponsored revision of Unified Parkinson's Disease Rating Scale (MDS-UPDRS; Goetz et al., 2008), years since PD diagnosis, and voice difficulties as estimated with the Voice-Related Quality of Life (V-RQOL); Hogikyan & Sethuraman, 1999) questionnaire. Specifically, we hypothesized that individuals with more advanced disease progression and more vocal difficulties would have shallower loudness growth.

Method

Participants

Twenty individuals with PD and 23 healthy controls were recruited to participate in the study. Two individuals with PD and four healthy controls were excluded because their hearing thresholds did not meet criteria for inclusion (see Hearing Thresholds section).

Thus, 18 individuals with PD with a mean age of 64.5 years (9 female, 9 male) and 19 healthy controls with a mean age of 64.6 years (11 female, 8 male) completed the experimental tasks (see Table 1). The participant data were included in the study if they met the criteria for normal older-adult hearing thresholds by frequency. All individuals with PD were diagnosed with idiopathic PD by a neurologist prior to participation in the study and were receiving daily levodopa (L-dopa)/carbidopa therapy. In addition, many were regularly taking the following medications: dopamine agonists ($n = 4$), monoamine oxidase B (MAO-B) inhibitors ($n = 3$), catechol-O-methyl transferase (COMT) inhibitors ($n = 2$), amantadine ($n = 2$), anticholinergics ($n = 1$), anti-epileptics ($n = 1$), antidepressants ($n = 4$), and citrulline ($n = 1$). Controls reported no history of neurological, speech, hearing, cognitive, or language disorders. All participants provided written consent in compliance with the Boston University Institutional Review Board.

Movement Disorder Society (MDS)-Sponsored Revision of the Unified Parkinson's Disease Rating Scale (UPDRS)

Part III (Motor Examination) of the Movement Disorder Society–sponsored revision of the UPDRS (MDS-UPDRS;

Table 1. Participant characteristics.

Characteristic	Parkinson's disease	Controls
Age (years)	*M* = 64.5, *SD* = 6.1, range: 52.0–72.0	*M* = 64.6, *SD* = 6.1, range: 50.0–77.0
Sex	9 female, 9 male	11 female, 8 male
Voice-Related Quality of Life standardized score	*M* = 82%, *SD* = 14%, range: 45%–100%	*M* = 97%, *SD* = 5%, range: 80%–100%
Unified Parkinson's Disease Rating Scale Part III score (total motor score)	*M* = 41.5, *SD* = 11.5, range: 16.0–49.0	
Time since diagnosis (years)	*M* = 6.6, *SD* = 4.9, range: 1.5–21.0	

Goetz et al., 2008) was administered to each participant with PD to determine the extent of motor difficulties resulting from PD (e.g., abnormalities in walking, upper and lower extremity movements, and speech). Each examination was scored per protocol by the first author, who is certified to administer MDS-UPDRS. The mean UPDRS motor score and the participant's report of their time since PD diagnosis (see Table 1) were used as correlates of disease progression for each participant with PD in analyses.

V-RQOL

All participants completed the V-RQOL questionnaire to determine self-reported impacts of their voice on their daily life (Hogikyan & Sethuraman, 1999). The questionnaire consists of 10 questions about voice-related problems that can affect quality of life. The experimenter read the standard instructions and the 10 questions aloud. Participants were asked to rate their responses from 1 (*none, not a problem*) to 5 (*problem is as "bad as it can be"*) based on their voice over the previous 2 weeks. The V-RQOL raw scores were calculated for each participant by tallying the total points assigned for each of the 10 responses. The raw scores were then standardized per the scoring guidelines detailed by Hogikyan and Sethuraman (1999) to determine the resulting percentile score between 0% and 100% (see Table 1).

Hearing Thresholds

Each participant underwent pure-tone hearing threshold testing (using the threshold measurement procedure listed in American Speech-Language-Hearing Association, 2005) at 500, 750, 1000, 1500, 2000, and 4000 Hz using 3M E-A-RTONE Gold 3A insert earphones and the Grason-Stadler GSI 18 Screening Audiometer. Data for each participant were included in the study if that participant had normal hearing thresholds for older adults at the tested frequency (using a cutoff of 25 dB HL for frequencies 1000 Hz and below and 35 dB HL above 1000 Hz; Schow, 1991). If a participant did not meet criteria for normal hearing thresholds at any of the tested frequencies, they were excluded from the study. Hearing thresholds by frequency and ear, with corresponding V-RQOL scores and age, are listed for each participant with PD (see Table 2) and control

participant (see Table 3); all values in bold indicate frequencies that were not tested for a particular participant. None of the participants wore hearing aids.

Loudness Discomfort Level Task

A loudness discomfort level protocol adapted from Cox, Alexander, Taylor, and Gray (1997) and Gu et al. (2010) was administered using the same audiometer and insert earphones as the hearing screening. Participants were instructed to judge the loudness of different sounds with increasing sound pressure level. Each participant was given a printed copy of the loudness categories (ranging from 1 = *very soft* to 7 = *uncomfortably loud*; see Table 4).

Participants were then instructed to rate each tone using the loudness categories and were informed that the tones would be presented unilaterally, always beginning with the left ear. Participants were also instructed that testing for each condition would stop if a rating of 7 (*uncomfortably loud*) was given. The maximum sound intensity presented was 80 dB HL. Experimenters presented pure tones that were 1 s in duration. Tones increased in level beginning at 35 dB HL, in 5-dB steps, up to 80 dB HL (or lower if the participant rated the sound as "uncomfortably loud"). Tones were tested at 500, 750, 1000, 1500, 2000, and 4000 Hz, in the same frequency order. The exact set of frequencies tested varied across participants, based on the hearing screening criteria (see Tables 2 and 3). For each participant, testing progressed through each intensity level at a particular tone frequency, and participants responded with the corresponding loudness category, which was recorded by the experimenter. If a rating of 7 was reached before the 80 dB HL tone was presented, all subsequent sound intensities were coded as 7 for that tone frequency. When either this maximum rating or a presentation level of 80 dB HL was reached, the next tone frequency was tested.

Statistical Analysis

A mixed-model analysis of variance was performed on the loudness ratings to assess the effects of group (between participants; individuals with PD vs. healthy controls), frequency (500, 750, 1000, 2000, and 4000 Hz), sound intensity of the presented tone (35–80 dB HL), ear (left or right), and all interactions. Factor effect sizes were quantified using η_p^2 (Witte & Witte, 2010). An alpha of .05 or less was

Table 2. V-RQOL scores, age, and hearing threshold information for participants with Parkinson's disease.

Subject	V-RQOL	Age (years)	500 Hz Left	500 Hz Right	750 Hz Left	750 Hz Right	1000 Hz Left	1000 Hz Right	1500 Hz Left	1500 Hz Right	2000 Hz Left	2000 Hz Right	4000 Hz Left	4000 Hz Right
PDT1	97.5%	70	10	15	15	**30**	5	15	5	10	5	10	30	30
PDT2	55.0%	68	10	20	10	15	10	5	10	20	20	15	35	35
PDT3	92.5%	72	20	20	20	25	20	25	20	25	30	30	35	**40**
PDT4	100.0%	70	20	20	20	20	20	20	20	20	20	20	20	20
PDT6	85.0%	65	25	25	25	25	15	25	15	20	20	20	15	**40**
PDT7	97.5%	59	10	20	15	25	15	15	0	0	5	0	10	10
PDT8	85.0%	69	20	20	15	15	10	10	5	10	5	20	10	15
PDT9	90.0%	60	5	10	10	5	5	10	15	15	5	10	25	15
PDT10	92.5%	65	25	10	**30**	15	25	20	30	20	25	15	35	25
PDT11	87.5%	69	15	10	10	10	10	10	0	15	15	0	10	10
PDT12	72.5%	70	15	20	15	20	10	10	10	5	25	15	**50**	30
PDT13	45.0%	59	0	10	5	20	5	25	15	25	30	**45**	80	65
PDT14	80.0%	71	10	10	10	20	10	15	20	10	25	25	35	20
PDT15	77.5%	56	10	5	10	10	10	10	10	10	10	5	15	10
PDT16	82.5%	58	20	25	25	25	15	15	10	10	25	15	30	15
PDT17	75.0%	59	20	20	15	20	15	20	10	20	20	20	10	15
PDT18	72.5%	52	10	10	5	5	10	10	10	10	5	10	5	10
PDT21	87.5%	65	25	25	15	20	20	20	30	15	30	25	25	15

Note. Bold values indicate frequencies that were not tested for a particular participant. V-RQOL = Voice-Related Quality of Life.

determined to be statistically significant. Loudness ratings were plotted as a function of dB HL for all tested frequencies, by ear, for each participant. Loudness growth slopes were then determined using a linear fit. Two two-sample *t* tests were then performed on mean loudness growth slopes for each ear, pooled across frequency, comparing the slopes between the two groups. In the PD group only, the mean slopes of the left and right ears were averaged and compared to the UPDRS Part III Motor scores, the years

since PD diagnosis, and the V-RQOL scores using Pearson product–moment correlation coefficients.

Results

All four factors showed significant main effects on the loudness ratings (group, frequency, tone intensity, and ear; see Table 5). Group had a small effect size ($\eta_p^2 = .01$). Frequency had a medium effect size ($\eta_p^2 = .10$) on loudness

Table 3. V-RQOL scores, age, and hearing threshold information for control participants.

Subject	V-RQOL	Age (years)	500 Hz Left	500 Hz Right	750 Hz Left	750 Hz Right	1000 Hz Left	1000 Hz Right	1500 Hz Left	1500 Hz Right	2000 Hz Left	2000 Hz Right	4000 Hz Left	4000 Hz Right
CT1	100.0%	77	5	14	5	15	10	10	10	0	15	15	25	35
CT3	100.0%	70	10	10	10	15	10	10	0	0	0	0	15	5
CT4	92.5%	69	**30**	10	**35**	15	25	15	20	15	15	20	20	15
CT5	100.0%	65	5	5	10	5	5	5	5	5	0	5	15	5
CT6	100.0%	72	15	20	10	20	10	10	5	5	0	15	25	20
CT7	80.0%	65	15	20	15	15	10	25	10	15	20	20	30	10
CT8	100.0%	59	10	**30**	10	**30**	15	20	10	15	20	15	25	15
CT9	100.0%	68	0	0	0	5	10	10	20	20	10	15	15	20
CT11	100.0%	60	15	25	15	25	15	15	15	15	15	20	**45**	**45**
CT12	100.0%	65	20	10	20	10	15	15	15	10	20	20	25	20
CT13	96.5%	62	10	20	15	15	15	20	20	10	15	15	20	20
CT14	97.5%	62	0	5	0	5	10	10	−10	0	5	0	0	10
CT15	92.5%	64	**30**	25	25	25	20	20	5	20	15	15	5	15
CT16	90.0%	59	10	5	5	10	5	10	10	10	10	10	25	**60**
CT17	100.0%	64	5	0	5	15	15	10	0	5	5	5	35	25
CT18	100.0%	71	0	0	10	10	0	0	0	0	−5	0	15	25
CT19	95.0%	61	5	10	10	10	10	15	15	10	30	25	25	35
CT20	100.0%	50	15	15	15	20	15	15	5	10	15	20	25	25
CT22	90.0%	61	5	20	5	10	10	15	0	15	10	20	15	20

Note. Bold values indicate frequencies that were not tested for a particular participant. V-RQOL = Voice-Related Quality of Life.

Table 4. Categorical loudness levels.

7. *Uncomfortably loud*
6. *Loud, but OK*
5. *Comfortable, but slightly loud*
4. *Comfortable*
3. *Comfortable, but slightly soft*
2. *Soft*
1. *Very soft*

ratings: For frequencies up to 1000 Hz, as frequency increased, loudness ratings decreased. To visualize responses by frequency, mean loudness ratings by group were plotted for each frequency as a function of dB HL (see Figure 1). Loudness ratings appear to remain consistent across frequencies, with the exception of 500 Hz where higher ratings are seen in the control group for the left ear compared to other frequencies. Sound intensity had a large effect size ($\eta_p^2 = .79$), with loudness ratings increasing as sound intensity increased. Ear had a very small effect size ($\eta_p^2 < .001$), with lower loudness ratings in the right ear relative to the left ear. A significant interaction was found between sound intensity and group, between ear and frequency, and the combined interaction of ear, frequency, and group, but all interactions had very small effect sizes ($\eta_p^2 < .01$). The effect of ear and the significant interactions with ear were likely due to the results in the right ear for the control group at 500 Hz (see Figure 1). There was no significant interaction between the sound intensity and ear or between the sound intensity and frequency. No significant interactions were found between ear and group or between frequency and group. There were also no significant interactions of sound intensity, ear, and frequency or sound intensity, ear, and group.

The two-sample t tests showed no significant difference in mean loudness growth slope between the control group and the PD group for both the left ear ($p = .371$)

Table 5. Results of mixed model analysis of variance on loudness ratings.

Effect	df	η_p^2	F	p
Group	**1**	**.01**	**0.35**	**< .001**
Participant	**35**	**.58**	**155.20**	**< .001**
db HL	**9**	**.79**	**1666.79**	**< .001**
Ear	**1**	**.00**	**14.97**	**< .001**
Freq (Hz)	**5**	**.10**	**83.73**	**< .001**
dB HL × Ear	9	.00	1.32	.220
dB HL × Freq (Hz)	45	.01	1.16	.213
dB HL × Group	**9**	**.01**	**5.27**	**< .001**
Ear × Freq (Hz)	**5**	**.01**	**4.65**	**< .001**
Ear × Group	1	.00	8.63	.003
Freq (Hz) × Group	5	.00	1.36	.237
dB HL × Ear × Freq (Hz)	45	.01	0.47	.999
dB HL × Ear × Group	9	.00	0.31	.972
dB HL × Freq (Hz) × Group	45	.01	0.49	.999
Ear × Freq (Hz) × Group	**5**	**.01**	**4.37**	**.001**
dB HL × Ear × Freq (Hz) × Group	45	.00	0.31	1.000

Note. Bold values indicate factors that had significant effects at the p less than or equal to 0.001 level. Freq = frequency.
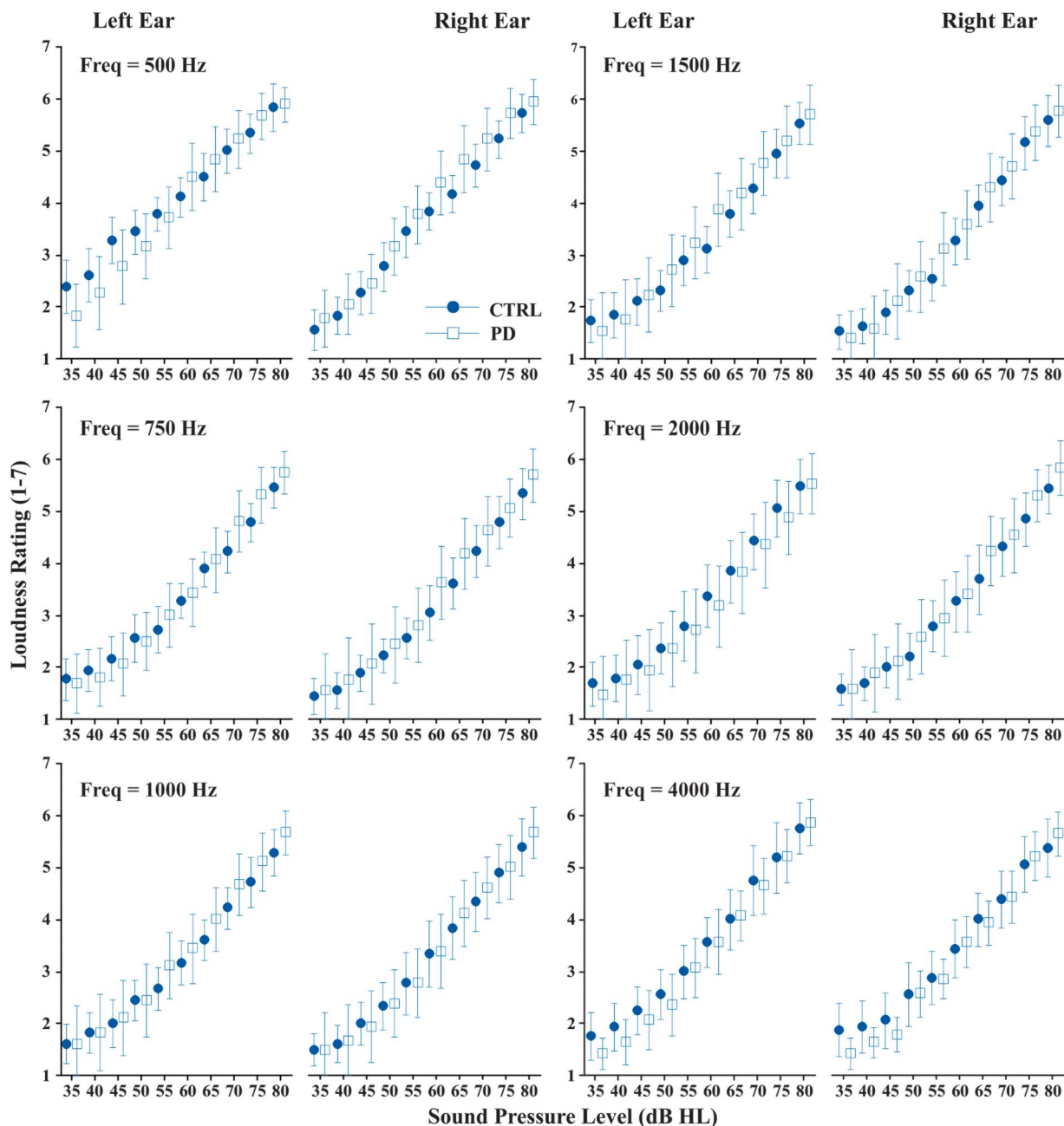
and the right ear ($p = .556$). To visualize individual participant responses for each group, loudness ratings were averaged across frequencies and plotted as a function of dB HL. The mean loudness ratings and 95% confidence intervals are shown for both the left and right ears for the two groups (see Figure 2). Although no significant differences were found in mean loudness growth slopes, the PD group's loudness rating slopes showed higher variability in both the left ear ($M = 0.094$, $SD = 0.027$) and the right ear ($M = 0.097$, $SD = 0.027$) compared to the control group's left ear ($M = 0.087$, $SD = 0.024$) and right ear ($M = 0.092$, $SD = 0.027$). Using the methods in Dienes (2014), Bayes factors were calculated using effect sizes from previous work (.976 based on Clark et al., 2014; .304 based on Dromey & Adams, 2000). All calculated Bayes factors were below 1/3 for both the left ear (.0130 based on Clark et al., 2014; .0416 based on Dromey & Adams, 2000) and the right ear (.0121 based on Clark et al., 2014; .0386 based on Dromey & Adams, 2000), supporting the null hypothesis that there is no difference between the PD group and the control group loudness rating slopes (Dienes, 2014).

The individual mean loudness ratings for each participant, across frequencies, are shown for each group by ear in Figure 3 to highlight the increased variability in the PD group. Specifically, there was one participant in the PD group who gave higher loudness ratings and a shallower slope than all other participants (shown in red triangles; Figure 3). This participant's data were within 3 SDs from the mean and thus were not classified as an outlier. This response was not explained by any of the PD participant measures (intelligibility score, V-RQOL score, hearing thresholds, and disease progression). Two participants with PD (gold diamonds, gray squares; Figure 3) had shallower loudness growth slopes than other participants; however, none of the participant measures explained the shallow slopes for these two participants. In the PD group, the UPDRS Part III Motor scores were mild to moderate, ranging from 16 to 49 ($M = 41.5$, $SD = 11.5$). The years since PD diagnosis ranged from 1.5 to 21 years ($M = 6.61$, $SD = 4.95$; see Table 1), and the V-RQOL normalized scores ranged from 45% to 100% ($M = 82\%$, $SD = 14\%$). No significant correlations were found between loudness rating slopes and UPDRS scores ($p = .548$), years since PD diagnosis ($p = .388$), or V-RQOL scores ($p = .669$).

Discussion

No significant differences were found in the loudness growth slopes between the PD and control groups. Bayes factor analysis yielded values below 1/3, providing evidence for the null hypothesis that there is no difference in loudness growth slopes between the PD and control groups. These results are consistent with a previous study that found no significant difference in loudness perception of warble tones using DME in healthy adults and adults with PD (Dromey & Adams, 2000). Although Clark et al. (2014) found that a PD group had significantly shallower mean loudness growth slopes than a control group in response to a DME task using

Figure 1. Mean and 95% confidence interval for loudness ratings for each frequency and ear as a function of sound intensity for healthy controls (CTRL; circles) and individuals with Parkinson's disease (PD; squares). Freq = frequency.
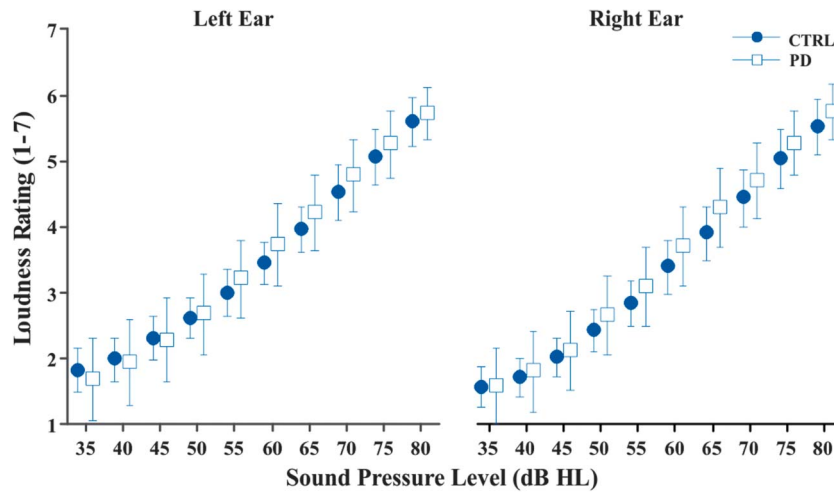

external prerecorded speech stimuli, De Keyser et al. (2016) conducted a similar study using speech stimuli recorded from the participants and found no statistically significant difference between the slopes of loudness estimations of the PD group and the control group during a VAS task.

Stimulus Type

The stimuli in this study differed from those used in previous studies, which included prerecorded speech, playback

of self-productions, and warble tones at 500 Hz. In healthy listeners, varying stimuli types yield similar loudness growth patterns, as long as the stimuli are matched in terms of intensity (Holte & Margolis, 1987; Ricketts & Bentler, 1996). However, this may not be true for PD listeners; thus, it is important to consider this difference when comparing previous studies of sound intensity perception in PD. Both Dromey and Adams (2000), who used warble tones at 500 Hz, and the current study, which used audiometric pure tones at varying frequencies, did not find significant differences
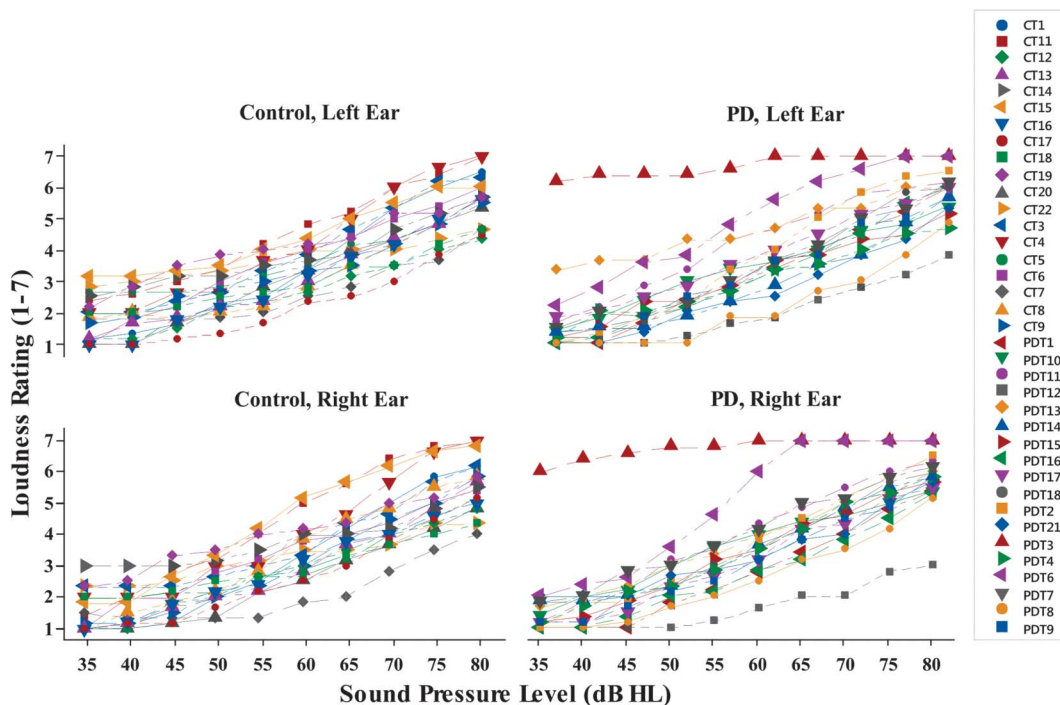
Figure 2. Mean loudness rating as a function of stimulus sound intensity in dB HL for healthy controls (CTRL; circles) and individuals with Parkinson's disease (PD; squares) for the left and right ears (collapsed across frequency).

in loudness growth slopes to external tones in a PD group and a control group. This suggests that reduced loudness in PD speech is not due to atypical loudness perception of externally generated sounds. Using speech stimuli, studies of loudness perception in PD have had variable results. Significantly shallower loudness slopes for prerecorded

speech (Clark et al., 2014) and overestimations of intensity for playbacks of self-generated speech intensity (Ho et al., 2000) have been seen in PD groups compared to control groups, but one study also reported no significant difference in loudness slopes when rating a playback of self-generated speech (De Keyser et al., 2016). However, higher loudness

Figure 3. Individual participant loudness ratings, averaged across frequencies, by sound intensity for the control group (left column) and Parkinson's disease group (PD; right column) for the left ear (top) and right ear (bottom).

ratings for PD groups relative to controls at lower intensities (60 and 65 dB SPL), but not higher intensities (75 and 80 dB SPL), were found for both prerecorded speech and playback of self-generated speech (Clark et al., 2014; De Keyser et al., 2016). Taken together, these findings suggest that variable results cannot fully be explained by stimulus differences, as similar stimuli have resulted in contradicting results even when the stimuli are matched for sound pressure level.

Inclusion Criteria: Voice and Hearing Status

Previous work examining loudness perception in PD has differed in terms of participant inclusion criteria related to voice and hearing status. Clark et al. (2014) and Ho et al. (2000) used hypophonia as an inclusion criteria, whereas the current work, Dromey and Adams (2000), and De Keyser et al. (2016) did not. Thus, the latter studies investigated loudness perception in a sample with more variability in voice symptoms, which may have affected the results. However, no significant correlation was seen between V-RQOL scores and loudness growth slopes in the current work. In addition, this study and most of the previous studies controlled for hearing status (Clark et al., 2014; Dromey & Adams, 2000; Ho et al., 2000), whereas De Keyser et al. (2016) did not administer hearing screenings to their participants. Their participants (aged 52–83 years old) may have had hearing impairment, which can affect loudness ratings because hearing loss often results in loudness recruitment or steeper loudness growth slopes (Al-Salim et al., 2010; Brand & Hohmann, 2001; Elberling, 1999; Launer, 1995; Rasetshwane et al., 2015). Thus, any within-group or between-groups differences in loudness growth slopes due to differences in hearing status cannot be ruled out in their study.

Other Methodological Differences

In comparing the current work to previous studies, there were other important methodological differences, such as stimulus presentation, the type of measurement scale used for perception, and range of sound pressure levels employed. In previous work, either headphones (Ho et al., 2000) or a loudspeaker (Clark et al., 2014; De Keyser et al., 2016; Dromey & Adams, 2000) was used to present the stimuli to both ears simultaneously. In the current study, stimuli were presented to each ear independently using insert earphones. Presenting stimuli to both ears bilaterally, as opposed to sequentially unilaterally, may lead to differences in loudness ratings due to binaural loudness summation (when loudness is perceived twice as loud when presented to both ears compared to one; Marks, 1978). Therefore, measuring loudness perception in both ears simultaneously could potentially result in higher loudness ratings for lower sound pressure levels when compared to loudness ratings for stimuli presented to each ear separately; this could result in shallower loudness growth slopes (Epstein & Florentine, 2009). In the current work, tones were presented as constant stimuli

(increasing from 35 to 80 dB HL). Previous work that compared loudness functions using adaptive, randomly presented stimuli and constant stimuli have reported lower loudness function variability measured using adaptive methods (Brand & Hohmann, 2002). However, in subjects with normal hearing, loudness function variability was very similar across stimulus presentation methods for intensity values below 80 dB HL (Brand & Hohmann, 2002), which was the intensity limit for the current work. Thus, it is unlikely that an adaptive protocol would have changed current results.

The measurement scale used to determine loudness perception also varied in previous work. Using a DME (Clark et al., 2014; Dromey & Adams, 2000) task, in which participants assign an unbounded value in relation to a certain modulus stimulus (Stevens, 1956), could cause high intersubject variability (as listeners have been shown to associate numbers and loudness as absolute rather than relative values; Hellman & Zwislocki, 1961), especially in a population with cognitive difficulties such as PD (Dubois & Pillon, 1996; Verbaan et al., 2007). Measures such as VAS (in which participants mark a line where each location corresponds to a numerical value; De Keyser et al., 2016) or the set categories used in the current study (see Table 4) are bounded ratings. This makes the range of possible responses the same for all participants; however, it also assumes linearity in the rating of loudness. When adjusting a volume knob with a sound pressure level limit (Ho et al., 2000), the rating is bounded and linearity is not assumed; however, variability may be higher because there are no defined categories. Given these considerations, loudness ratings reported in previous work were likely somewhat impacted by the measurement scale used to quantify loudness perception. Measuring with unbounded scales or without defined categories could have created response differences due to the higher response range and variability in the measurement system itself instead of pure differences in loudness perception.

Last, when comparing reports of loudness perception, the range of sound intensity used for testing should be considered. In the studies using speech stimuli, a 20-dB range was tested (Clark et al., 2014; De Keyser et al., 2016), or the range was determined by the participant's own productions played back to them (Ho et al., 2000). In the current study, a range of 45 dB was tested at each frequency, whereas Dromey and Adams (2000) tested a range of 60 dB at 500 Hz. The range of sound pressure levels tested affects the steepness of the loudness perception slope, so this would contribute to differences between studies measuring loudness growth over a smaller range (Clark et al., 2014; De Keyser et al., 2016) compared to a wider range (as in the current work and Dromey & Adams, 2000). Furthermore, both Clark and De Keyser reported trends for higher loudness ratings by the PD group relative to controls at lower intensities, but not higher intensities. Differences in loudness functions for low compared to high intensities have also been observed for categorical loudness scaling measurements (Brand & Hohmann, 2002; Oetting, Brand, & Ewert,

2014). When looking at the current results over the range of low intensities ranging from 35 to 50 dB HL, for example, no clear trend is seen, but the slopes appear more variable in both participant groups. Thus, only testing in this region of low intensities (and over a more restricted level range of 15 dB) would have resulted in different overall loudness slopes. The higher variability at lower intensities in the current data may also be explained by the lower intensities being tested first and increasing familiarization with the categorical scale with increasing stimulus level.

In summary, differences in methodology could help explain differences between previous results and the current study of perception of loudness in PD. Regardless of these differences, none of the reported results, in either previous work or the current work, provide strong evidence that the loudness perception deficit in PD drives hypophonia in PD. Although hypophonia has been reported extensively, with large effect sizes, a similarly large difference in loudness perception of external sound in PD has not been found (even when hypophonia is an inclusion criteria; Clark et al., 2014).

Limitations and Future Work

In the current work, the ear order was not counterbalanced during stimulus presentation, and the PD group had high variability in terms of disease severity and symptoms, both of which may have affected the data reported in the study. There was no counterbalancing of ear order during stimuli presentation; for each participant, the left ear was always tested first. As tones were always presented to the left ear first and 500 Hz was the first frequency presented, it is possible that ratings were lower in the right ear due to a familiarization process with the rating scale. There may have been effects of the scale used as well. A bounded categorical scale was chosen to provide the same range of ratings for all participants; however, this method also assumes linearity of loudness ratings, which could have impacted the results seen here. Another limitation of this study is the heterogeneity of the participants in the PD group in terms of disease severity and effects of medication. The PD group included a large range of MDS-UPDRS Part III Motor scores (16–49), and all participants were receiving daily L-dopa therapy. Previous work also included participants with PD who were taking anti-Parkinson medication (Clark et al., 2014; De Keyser et al., 2016; Dromey & Adams, 2000; Ho et al., 2000). The effects of L-dopa can vary based on several factors including the PD symptoms and disease severity of the patient when the L-dopa therapy began (Goetz, Stebbins, & Blasucci, 2000).

Conclusion

The current results show no significant difference in loudness growth slopes of audiometric pure tones in individuals with PD compared to controls performing a bounded categorical rating task. These findings agree with the findings of Dromey and Adams (2000), who also found no significant differences in loudness perception during an unbounded rating task using 500 Hz warble tones in a PD group compared to controls. Taken together, these studies suggest that loudness perception of external tones in PD is not atypical. However, this conclusion is at odds with previous work that reports that individuals with PD had significantly shallower loudness growth slopes in response to speech stimuli (Clark et al., 2014). Considering the conflicting results, it is possible that the stimulus type (external tone vs. external speech) may have an effect on the loudness perception, but this must be investigated further. Given that hypophonia is widespread in PD, but that existing studies come to inconsistent conclusions about whether or not loudness perception of external sounds is atypical in PD, it seems unlikely that loudness perception differences are a primary cause of hypophonia in PD.

References

Al-Salim, S. C., Kopun, J. G., Neely, S. T., Jesteadt, W., Stiegemann, B., & Gorga, M. P. (2010). Reliability of categorical loudness scaling and its relation to threshold. *Ear and Hearing, 31*(4), 567–578.

American Speech-Language-Hearing Association. (2005). *Guidelines for manual pure-tone threshold audiometry*. Rockville, MD: Author.

Aronson, A. E., & Brown, J. R. (1975). *Motor speech disorders*. Philadelphia, PA: WB Saunders Company.

Brand, T., & Hohmann, V. (2001). Effect of hearing loss, centre frequency, and bandwidth on the shape of loudness functions in categorical loudness scaling. *Audiology, 40*(2), 92–103.

Brand, T., & Hohmann, V. (2002). An adaptive procedure for categorical loudness scaling. *The Journal of the Acoustical Society of America, 112*(4), 1597–1604.

Chaudhuri, K. R., Healy, D. G., & Schapira, A. H. V. (2006). Non-motor symptoms of Parkinson's disease: Diagnosis and management. *The Lancet Neurology, 5*(3), 235–245.

Clark, J. P., Adams, S. G., Dykstra, A. D., Moodie, S., & Jog, M. (2014). Loudness perception and speech intensity control in Parkinson's disease. *Journal of Communication Disorders, 51*, 1–12.

Cox, R. M., Alexander, G. C., Taylor, I. M., & Gray, G. A. (1997). The contour test of loudness perception. *Ear and Hearing, 18*(5), 388–400.

De Keyser, K., Santens, P., Bockstael, A., Botteldooren, D., Talsma, D., De Vos, S., . . . De Letter, M. (2016). The relationship between speech production and speech perception deficits in Parkinson's disease. *Journal of Speech, Language, and Hearing Research, 59*(5), 915–931.

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology, 5,* 781.

Dromey, C., & Adams, S. (2000). Loudness perception and hypophonia in Parkinson disease. *Journal of Medical Speech-Language Pathology, 8*(4), 255–259.

Dubois, B., & Pillon, B. (1996). Cognitive deficits in Parkinson's disease. *Journal of Neurology, 244*(1), 2–8.

Duffy, J. R. (2013). *Motor speech disorders: Substrates, differential diagnosis, and management* (3rd ed.). St. Louis, MO: Elsevier Health Sciences.

Elbaz, A., Bower, J. H., Maraganore, D. M., McDonnell, S. K., Peterson, B. J., Ahlskog, J. E., . . . Rocca, W. A. (2002). Risk tables for parkinsonism and Parkinson's disease. *Journal of Clinical Epidemiology, 55*(1), 25–31.

Elberling, C. (1999). Loudness scaling revisited. *Journal of the American Academy of Audiology, 10,* 248–260.

Epstein, M., & Florentine, M. (2009). Binaural loudness summation for speech and tones presented via earphones and loudspeakers. *Ear and Hearing, 30*(2), 234–237.

Fox, C. M., Morrison, C. E., Ramig, L. O., & Sapir, S. (2002). Current perspectives on the Lee Silverman Voice Treatment (LSVT) for individuals with idiopathic Parkinson disease. *American Journal of Speech-Language Pathology, 11*(2), 111–123.

Fox, C. M., & Ramig, L. O. (1997). Vocal sound pressure level and self-perception of speech and voice in men and women with idiopathic Parkinson disease. *American Journal of Speech-Language Pathology, 6*(2), 85–94.

Global Parkinson's Disease Survey Steering Committee. (2002). Factors impacting on quality of life in Parkinson's disease: Results from an international survey. *Movement Disorders, 17*(1), 60–67.

Goetz, C. G., Stebbins, G. T., & Blasucci, L. M. (2000). Differential progression of motor impairment in levodopa-treated Parkinson's disease. *Movement Disorders, 15*(3), 479–484.

Goetz, C. G., Tilley, B. C., Shaftman, S. R., Stebbins, G. T., Fahn, S., Martinez-Martin, P., . . . Dodel, R. (2008). Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale presentation and clinimetric testing results. *Movement Disorders, 23*(15), 2129–2170.

Gu, J. W., Halpin, C. F., Nam, E.-C., Levine, R. A., & Melcher, J. R. (2010). Tinnitus, diminished sound-level tolerance, and elevated auditory activity in humans with clinically normal hearing sensitivity. *Journal of Neurophysiology, 104*(6), 3361–3370.

Hellman, R. P., & Zwislocki, J. (1961). Some factors affecting the estimation of loudness. *The Journal of the Acoustical Society of America, 33*(5), 687–694.

Ho, A. K., Bradshaw, J. L., & Iansek, R. (2000). Volume perception in parkinsonian speech. *Movement Disorders, 15*(6), 1125–1131.

Ho, A. K., Iansek, R., Marigliani, C., Bradshaw, J. L., & Gates, S. (1999). Speech impairment in a large sample of patients with Parkinson's disease. *Behavioural Neurology, 11*(3), 131–137.

Hoehn, M. M., & Yahr, M. D. (1998). Parkinsonism: Onset, progression, and mortality. *Neurology, 50*(2), 318.

Hogikyan, N. D., & Sethuraman, G. (1999). Validation of an instrument to measure Voice-Related Quality of Life (V-RQOL). *Journal of Voice, 13*(4), 557–569.

Holte, L., & Margolis, R. H. (1987). The relative loudness of third-octave bands of speech. *The Journal of the Acoustical Society of America, 81*(1), 186–190.

Jankovic, J. (2008). Parkinson's disease: Clinical features and diagnosis. *Journal of Neurology, Neurosurgery & Psychiatry, 79*(4), 368–376.

Launer, S. (1995). *Loudness perception in listeners with sensorineural hearing impairment* (Unpublished PhD thesis). Oldenburg University, Oldenburg, Germany.

Logemann, J. A., Fisher, H. B., Boshes, B., & Blonsky, E. R. (1978). Frequency and cooccurrence of vocal tract dysfunctions in the speech of a large sample of Parkinson patients. *Journal of Speech and Hearing Disorders, 43*(1), 47–57.

Marigliani, C., Gates, S., & Jacks, D. (1997). Speech pathology and Parkinson's disease. In M. S. Morris & R. N. Iansek (Eds.), *Parkinson's disease: A team approach* (pp. 127–149). Melbourne, Australia: Buscombe Vicprint.

Marks, L. E. (1978). Binaural summation of the loudness of pure tones. *The Journal of the Acoustical Society of America, 64*(1), 107–113.

Marsden, C. D. (1982). The mysterious motor function of the basal ganglia: The Robert Wartenberg lecture. *Neurology, 32*(5), 514–539.

Martinez-Martin, P., Rodriguez-Blazquez, C., Kurtis, M. M., & Chaudhuri, K. (2011). The impact of non-motor symptoms on health-related quality of life of patients with Parkinson's disease. *Movement Disorders, 26*(3), 399–406.

Mayeux, R. (2003). Epidemiology of neurodegeneration. *Annual Review of Neuroscience, 26*(1), 81–104.

McNamara, P., & Durso, R. (2003). Pragmatic communication skills in patients with Parkinson's disease. *Brain and Language, 84*(3), 414–423.

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

Oetting, D., Brand, T., & Ewert, S. D. (2014). Optimized loudness-function estimation for categorical loudness scaling data. *Hearing Research, 316,* 16–27.

Pringsheim, T., Jette, N., Frolkis, A., & Steeves, T. D. L. (2014). The prevalence of Parkinson's disease: A systematic review and meta-analysis. *Movement Disorders, 29*(13), 1583–1590.

Rasetshwane, D. M., Trevino, A. C., Gombert, J. N., Liebig-Trehearn, L., Kopun, J. G., Jesteadt, W., . . . Gorga, M. P. (2015). Categorical loudness scaling and equal-loudness contours in listeners with normal hearing and hearing loss. *The Journal of the Acoustical Society of America, 137*(4), 1899–1913.

Ricketts, T. A., & Bentler, R. A. (1996). The effect of test signal type and bandwidth on the categorical scaling of loudness. *The Journal of the Acoustical Society of America, 99*(4), 2281–2287.

Schow, R. L. (1991). Considerations in selecting and validating an adult/elderly hearing screening protocol. *Ear and Hearing, 12*(5), 337–348.

Stevens, S. S. (1956). The direct estimation of sensory magnitudes: Loudness. *The American Journal of Psychology, 69*(1), 1–25.

Tanner, C. M., & Aston, D. A. (2000). Epidemiology of Parkinson's disease and akinetic syndromes. *Current Opinion in Neurology, 13*(4), 427–430.

Verbaan, D., Marinus, J., Visser, M., Van Rooden, S. M., Stiggelbout, A. M., Middelkoop, H. A. M., & Van Hilten, J. J. (2007). Cognitive impairment in Parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry, 78*(11), 1182–1187.

Witjas, T., Kaphan, E., Azulay, J. P., Blin, O., Ceccaldi, M., Pouget, J., . . . Chérif, A. A. (2002). Nonmotor fluctuations in Parkinson's disease Frequent and disabling. *Neurology, 59*(3), 408–413.

Witte, R. S., & Witte, J. S. (2010). *Statistics.* Hoboken, NJ: Wiley.