JSLHR

Research Article

# A Comparison of Speech Amplification and Personal Communication Devices for Hypophonia

Thea Knowles,[a] Scott G. Adams,[b,c,d] Allyson Page,[b,c] Daryn Cushnie-Sparrow,[b,c] and Mandar Jog[d]

**Purpose:** This study compared the performance of three amplification devices hypothesized to improve speech communication in individuals with hypophonia (HP), as well as to identify individuals' device preferences.
**Method:** Twenty-two individuals with HP and their primary communication partners participated in a cross-over design study comparing three different speech amplification devices: a wired portable amplifier (Device A), a wireless stationary amplifier (Device B), and a one-way personal communication system (Device C). Participants attended one laboratory visit followed by 1-week trial periods with each device. At the first visit, HP participants completed speech tasks with and without the devices, in quiet and in noise. Following the in-laboratory test period, participants trialed each device at home for approximately 1 week per device. Following completion of the study, participants indicated whether or not they would like to continue using a device.
**Results:** Overall, in the presence of noise, all three devices demonstrated significant improvements in speech-to-noise levels and speech intelligibility compared to no device. A clear device hierarchy emerged such that the personal communication device (Device C) was associated with significantly better speech outcomes compared to the other two devices. The majority of participants elected to continue using a device at the completion of the study. Device preferences, however, did not clearly reflect the objective device hierarchy that was found for the objective speech measures. Each of the three devices was selected as a preferred device by at least three participants at the completion of the study.
**Conclusion:** Results from this study demonstrated clear differences in device performance in three distinct forms of amplification devices for individuals with HP. Findings suggest that amplification device use may be beneficial for this clinical population and underscore the potential to improve device availability and device selection criteria in future research.
**Supplemental Material:** https://doi.org/10.23641/asha.12735875

Hypophonia, an overall reduction in speech intensity, is one of the most prevalent and distinctive speech impairments associated with Parkinson's disease (PD; Darley et al., 1975). Hypophonia is often present even in early stages of the disease (Logemann et al., 1978) and has been reported to be perceptually detectable in up to 65.5% of individuals with PD (Gamboa et al., 1997; Ho et al., 1998; Ludlow & Bassich, 1984). Adams and Dykstra (2009) suggested that these proportions may underestimate the actual prevalence of individuals with PD and hypophonia, given that studies have shown that formal speech testing may be less sensitive to speech impairments in people with PD (Bunton, 2008; Ho et al., 1999; Sidtis et al., 2012).

Hypophonia hinders oral communication across acoustic, perceptual, and participatory domains (Adams & Dykstra, 2009). Acoustic speech studies have reported individuals with PD speak at intensity levels at an average of 2–4 dB SPL below age-matched healthy control participants in controlled laboratory-based speech tasks (Adams et al., 2010; Fox & Ramig, 1997) and up to 5 dB SPL lower in the context of background noise (Dykstra et al., 2012) and

[a]Department of Communicative Disorders and Sciences, University at Buffalo, NY
[b]School of Communication Sciences and Disorders, Western University, London, Ontario, Canada
[c]Health and Rehabilitation Sciences, Western University, London, Ontario, Canada
[d]Department of Clinical Neurological Sciences, University Hospital, London, Ontario, Canada
Correspondence to Thea Knowles: theaknow@buffalo.edu

Editor-in-Chief: Bharath Chandrasekaran
Editor: Kate Bunton
Received February 24, 2020
Revision received May 25, 2020
Accepted May 26, 2020
https://doi.org/10.1044/2020_JSLHR-20-00085

***Disclosure:*** *The authors have declared that no competing interests existed at the time of publication.*

in conversation (Dykstra et al., 2015). In daily life (i.e., not in a controlled laboratory setting), people with PD have been reported to speak approximately 7–8 dB SPL lower than healthy controls (Gustafsson et al., 2019). A metric of speech intensity that considers its relationship to the level of background noise is *speech-to-noise ratio* (SNR; Kryter, 1994).[1] SNR is an acoustic measure that may be more directly related to *speech intelligibility* (i.e., the extent to which a speech signal is recovered by a typical listener; Kent et al., 1989; Miller, 2013; Weismer, 2008). SNR is calculated as the background noise intensity level subtracted from the speech intensity level (DeBonis & Donohue, 2008). Small positive or negative SNR values indicate that the background noise is of near-equal levels to that of the speaker, making the task of the listener to detect, process, and understand the speech signal substantially more difficult. By definition, then, individuals with hypophonia with demonstrated reductions in speech intensity may demonstrate associated reductions in SNR compared to healthy talkers in the same situations. Adams et al. (2008) reported that SNR levels below 1.8 were associated with sentence intelligibility of 50% or less, whereas SNR levels between 5 and 7 dB SPL were associated with sentence intelligibility of approximately 80% for both individuals with and without hypophonia. Furthermore, there is evidence to suggest that individuals with PD and hypophonia demonstrate an impaired ability to perceive the reduced loudness in their own voice, possibly due to deficits in sensorimotor integration (Adams et al., 2010; Clark et al., 2014; Ho et al., 1999, 2000).

In the presence of background noise, lower speech intensity associated with hypophonia may be more pronounced. Most healthy individuals demonstrate a tendency to increase their speech intensity in the presence of background noise that exceeds 50 dB SPL, a phenomenon known as the Lombard effect (Lane & Tranel, 1971). The Lombard effect is a compensatory response that may help to maintain positive SNR levels in noisy environments (Pick et al., 1989). Individuals with PD also demonstrate the Lombard response in noise (Adams & Lang, 1992; Adams et al., 2005; Stathopoulos et al., 2014), though it is attenuated compared to that of healthy controls (Adams et al., 2005). While SNR and speech intelligibility tend to decrease even for healthy talkers as background noise levels increase, this effect is more detrimental for people with hypophonia who already demonstrate lower speech intensity and intelligibility baselines (Adams et al., 2008; Dykstra et al., 2012).

An estimated 65%–70% of people with PD demonstrate lower speech intelligibility compared to healthy age-matched control talkers (Coates & Bakheit, 1997; Miller et al., 2007). Individuals with PD also demonstrate further detriment in speech production with disease progression (Ash et al., 2017; Skodda et al., 2013). Reduced speech intelligibility in this population is likely affected by a combination of speech symptoms, rather than exclusively related to reduced speech intensity. Speech intelligibility is impaired at both sentence and conversational levels for individuals with hypophonia (Dykstra et al., 2015; Miller, 2012; Sidtis et al., 2012) and decreases as background noise levels increase (Adams et al., 2008). People with hypophonia also demonstrate reduced communicative effectiveness in various adverse communication settings (e.g., at a distance, in a noisy situation; Dykstra et al., 2015). Primary communication partners (CPs) of these individuals (e.g., family members) also report similar difficulties when asked to rate communication effectiveness of their partners with hypophonia (Dykstra et al., 2015).

### *Interventions for Hypophonia*

Evidence-based behavioral interventions exist that are specifically designed to increase speech intensity in individuals with PD (e.g., the Lee Silverman Voice Treatment program; Ramig et al., 2004). The predominant criticism of behavioral treatments for hypophonia is the concern that improvements in speech intensity may not transfer beyond the clinical setting because deficits related to cognition, sensorimotor integration, and motor learning may inhibit the incorporation of new speech strategies into habitual speech (Adams & Dykstra, 2009; Olson et al., 2019; Scott & Caird, 1983). Furthermore, individuals who show attenuated treatment responses or are less successful in maintaining their gains from therapy may be left without a viable option to remediate the consequences of their low speech intensity.

The use of an augmentative device to amplify the speech signal is a potential solution to this problem (Andreetta et al., 2016; Greene et al., 1972). Speech amplification devices have been recommended for individuals with reduced loudness and sufficient articulatory abilities, that is, not limited to individuals with hypophonia secondary to PD (Spencer et al., 2003). Despite the availability of a broad range of devices, however, only one published study to date has examined differences in device efficacy for hypophonia (Andreetta et al., 2016), and none has investigated long-term amplification device use (Bertrand, 2009).

Speech amplification devices designed for clinical use in treating or preventing speech and voice disorders may be broadly classified into two types: wired portable and wireless stationary. Both types minimally consist of a microphone and sound field speaker. Wired, portable devices typically involve a speaker system that is able to be worn on the body, for example, belted around the waist, clipped to a pocket, or worn on a lanyard. Attached to this portable amplifier is a headset or lavalier microphone that is worn by the user. Wireless amplifiers also include a microphone attached to a small, body-worn unit that transmits the speech signal wirelessly to an audio speaker located up to several meters away from the talker (i.e., like a portable public address system).

A third type of device similar to the amplification devices described above is a wireless personal communication

[1] "Speech-to-noise ratio" should be distinguished from the measure of signal-to-noise ratio, which is a measure of noise present in the speech signal itself (Kent et al., 1997).

system. These types of devices have been used by individuals with hearing impairments (Harkins & Tucker, 2007; Laplante-Lévesque et al., 2010) but have not previously been reported for use with speech disorders. Personal communication systems are typically wireless and transmit the audio signal over a frequency modulation (FM) or very high frequency (VHF) channel. These systems, such as the wireless amplifiers, typically include a small, body-worn transmitter. Unlike the amplifiers, personal communication systems are designed to transmit the signal to a small receiver designed to be worn with headphones, typically worn by the person with hearing loss.

Conceptually, personal communication systems could be a viable option for individuals with hypophonia, if their standard use were modified. For example, an individual, regardless of their own hearing status, who wished to communicate with a person with hypophonia could wear the receiver and headphones while the individual with the speech impairment wore the transmitter and microphone. While personal communication systems are often also able to transmit to loud speakers, similar to the wireless amplifiers described above, studies have shown a clear advantage of the use of headphones in terms of SNRs (Crandell et al., 2001; Náb lek et al., 1986).

Regardless of the type of device, the purpose of an amplification system is to increase the intensity of the speech signal. In the case of individuals with reduced speech intensity, this serves to compensate for their impaired communication without requiring any behavioral adjustments.

### *Previous Comparison of Amplification Devices*

Andreetta et al. (2016) compared the effectiveness of seven speech amplification devices for 11 people with PD and hypophonia. Efficacy measures included acoustic and perceptual measures as well as user experience ratings of the devices trialed. Overall, compared to unamplified speech, the authors found that the use of an amplification device was associated with an increase in SNR of approximately 1–5 dB SPL and an increase in conversational intelligibility scores ranging from 14% to 59%.

Across the seven devices, a series of hierarchical trends emerged. Regarding the user experience ratings, the Chattervox, a portable belt-pack amplifier, was rated by the users as having the best amplification power and sound quality. The Spokeman, the smallest of the portable amplifiers (both in dimensions and in weight), was rated as being the highest in physical comfort, appearance, and overall preference. Compared to the other device conditions, the BoomVox, a large, wireless stationary amplifier, was significantly associated with the highest intensity and SNR levels, as well as the highest conversational speech intelligibility visual analog scale ratings for most device comparisons.

Based on their results, Andreetta et al. (2016) put forth a tentative recommendation regarding speech amplifier device selection for individuals with PD and hypophonia. The BoomVox was the highest recommended device of the group of seven due to its performance on objective acoustic and perceptual measures of efficacy (SNR, intensity, intelligibility). The Chattervox was the second highest recommended device due to its overall performance in relation to the other devices as well as its portability. The Oticon Amigo, a wireless FM speaker system designed to be used by individuals with hearing loss, was the third highest recommended device due to its relative performance and portability. The Spokeman, the smallest of the portable wired amplification devices, was the fourth highest recommended. Although the Spokeman obtained the highest user preference ratings, likely due to its size and portability, it performed more poorly on the more objective speech outcome measures.

### *Purpose of This Study*

The primary purpose of this study was to conduct a more in-depth comparison of speech amplification devices similar to those recommended by Andreetta et al. (2016), as well as to evaluate the use of a wireless personal communication device for people with hypophonia. To this end, measures of device performance and user preference were of interest, and both highly controlled and more naturalistic settings were employed. Specifically, device performance was measured by examining SNR and speech intelligibility in quiet and in noise for each of the three devices as well as without any device in a series of laboratory speech tasks. Speech intelligibility scores were provided by both familiar and nonfamiliar listeners. Device preferences of individuals with hypophonia and their CPs were measured through the use of device-related questionnaires, both in the laboratory setting as well as following more naturalistic trial periods with each of the devices. The overarching goal was to provide more specific recommendations for the use of amplification devices for this population.

Four research questions (RQs) were of interest:

1. What differences in the acoustic metric of SNR exist across the device conditions, and how are these impacted by the presence of background noise?
2. How do familiar and naive listener sentence intelligibility scores differ across the device and noise conditions?
3. How do device user experience preferences differ across device conditions at two time points (immediately after and following the 1-week trial period), and how do these preferences differ for the individuals with hypophonia versus for their CPs?
4. Following a trial period, do individuals with hypophonia wish to continue using an amplification device and, if so, which one do they choose?

## Method

The proposed study employed a clinical cross-over design to compare three types of amplification devices that could be used by individuals with communication disorders:

(a) a wired belt-pack voice amplifier (Device A), (b) a wireless personal amplifier (Device B), and (c) a wireless personal communication system (Device C). Human Subjects Research Ethics Board Western University approved this study, and it was registered as a clinical trial (ClinicalTrials.gov Identifier: NCT02407067).

### Participants

Three groups participated in this study: a group of individuals with hypophonia, their CPs, and a group of naive listener participants. Participant demographics for the hypophonia group (HP) is presented in Table 1 below. The primary group included 22 individuals with hypophonia (four women, 18 men; age range: 54–88 years; herein referred to as the HP group) recruited from the Movement Disorders Clinic at University Hospital in London, Ontario, Canada. Of these individuals, 20 had a primary diagnosis of idiopathic PD confirmed by their primary neurologist (author M. J.). One individual had a primary diagnosis of multiple systems atrophy-predominant cerebellar ataxia, and one had a diagnosis of possible parkinsonism. All individuals presented with symptoms of hypophonia as identified by their primary neurologist.

Inclusion criteria for the HP group included that they (a) had received their neurological diagnosis at least 6 months prior to testing, (b) exhibited mild-to-moderate hypophonia (as rated by an experienced speech-language pathologist [SLP], S. A.), (c) were at least 50 years of age, (d) had no history of other neurological or voice disorders,[2] and (e) were otherwise in good general health. All participants with PD were stabilized on antiparkinsonian medication, with the exception of one participant (HP13), who had recently adjusted his medication regimen. Levodopa equivalent dose calculations were performed as per Tomlinson et al. (2010). Seven HP participants had received deep brain stimulation (DBS) surgery of the subthalamic nucleus as an adjunctive intervention to treat the symptoms of PD. Eight had previously received speech therapy to address speech concerns related to PD. While details of previous speech therapy were not elicited, most participants confirmed that their therapy had involved treatment involving loud speech production.[3] Hearing and cognitive status were screened but were not exclusion criteria. Hearing screenings were done at a 40–dB HL threshold at 500 Hz, 1 kHz, 2 kHz, and 4 kHz, and failing the screening was not an exclusion criterion. Eight HP participants passed the hearing screening, 10 failed at one or more frequencies, and hearing screenings were not available for four participants.[4] Cognitive status was not an exclusion criterion, though cognition was screened using the Montreal Cognitive Assessment (MoCA; Nasreddine et al., 2005).

The second group of participants involved individuals who identified as being a "CP" to the HP participants ($n$ = 22, 18 women, four men; age range: 54–79 years). The role of the CP participants was to act as a familiar listener during the speech production experiment and to provide device ratings alongside their partner throughout the trial periods. Prior to study enrolment, each potential HP participant was instructed to select someone in their daily life with whom they spoke regularly to accompany them to all study visits.[5] In 21 cases, this was a spouse. In one case (CP13), it was an adult child. In order not to restrict the CP selection appropriate for the HP participants, the only inclusion criteria for the CP group included that they were between 18 and 85 years of age. Hearing screenings for the CP group were completed at Visit 2. Five CP participants (CP04, CP07, CP10, CP17, CP21) did not pass the 40-dB hearing screening and did not wear hearing aids. Two CP participants (CP06 and CP14) did have hearing aids. CP hearing information is not available for the two participants (CP05 and CP15) who withdrew from the study after Visit 1.

The third group of participants included four (two female, two male) naive listeners recruited to listen and transcribe recordings made from the speech production study later on. These participants were graduate students in speech-language pathology or audiology and were not involved in any other aspect of the study. All were native speakers of English, and all but one passed a 20–dB HL hearing screening.[6] Previous studies of disordered speech perception have used comparably small listener groups (Adams et al., 2008; Dromey, 2003; Dykstra et al., 2012, 2015; Moreau et al., 2011; Tanaka et al., 2016; Tripoliti et al., 2014). Listener groups of this size (and smaller) have also been shown to offer reliable estimates of intelligibility (Abur et al., 2019).

### Devices

Four device conditions were included in the protocol: three devices consistent with those recommended from the results of Andreetta et al. (2016; referred to as Devices A, B, and C, described in more detail below) and a "no device"

[2]One participant, HP21, did report a history of developmental stuttering that had worsened since the onset of PD.

[3]Two participants, HP03 and HP12, reported that they had completed the Lee Silverman Voice Treatment program; most participants were unable to recall more specific details about the speech therapy they had received.

[4]In most cases, screenings were completed at Visit 2. Therefore, only participants who completed at least one device trial periods had their hearing screened. HP05 and HP15 did not complete the hearing screenings because they withdrew from the study after Visit 1 (HP20 withdrew after Visit 2; hence, hearing screening is available for him). HP09 and HP10 did not complete the hearing screening due to declining health at subsequent visits.

[5]During the recruitment period, participants with HP were informed that, if they could not identify a partner they felt comfortable bringing in, a graduate student researcher from the university would stand in as a conversation partner. This was proposed in order to ensure that all individuals who were contacted would have the opportunity to participate regardless of their social network. All HP participants, however, were able to bring in a primary CP from their daily life.

[6]One listener passed at 25 dB HL.

**Table 1.** Participant demographics for the hypophonia (HP) group.

| Participant | Sex | Age | LED | Diagnosis | Years since diagnosis | DBS | UPDRS | UPDRS-Speech | MoCA | Hearing screening | CP hearing screening |
|---|---|---|---|---|---|---|---|---|---|---|---|
| HP01 | m | 75 | 750 | PD | 9 | No | 40 | 3 | 16 | Fail | Pass |
| HP02 | m | 54 | 0 | PD | 7 | Yes | 31 | 3 | 22 | Fail | Pass |
| HP03 | m | 75 | 750 | PD | 8 | No | 49 | 2 | 23 | Fail | Pass |
| HP04 | f | 78 | 800 | PD | 14 | No | 35 | 2 | 20 | Fail | Fail |
| HP05 | m | 69 | 1250 | PD | 11 | No | 43 | 2 | 13 | NA | NA |
| HP06 | m | 67 | 550 | PD | 21 | Yes | 29 | 3 | 22 | Fail | Hearing aids |
| HP07 | f | 72 | 0 | PD | 16 | Yes | 30 | 1 | 26 | Pass | Fail |
| HP08 | m | 65 | 1200 | PD | 15 | No | 20 | 1 | 21 | Pass | Pass |
| HP09 | m | 79 | 800 | PD | 8 | No | 50 | 3 | 19 | NA | Pass |
| HP10 | m | 75 | 2000 | PD | 20 | No | 45 | 2 | NA | NA | Fail |
| HP11 | m | 72 | NA | PD | 11 | Yes | NA | NA | 20 | Pass | Pass |
| HP12 | m | 59 | 400 | PD | 10 | Yes | 37 | 2 | 24 | Pass | Pass |
| HP13 | m | 71 | 400 | PD | 0.5 | No | 31 | 1 | 22 | Fail | Pass |
| HP14 | f | 67 | 600 | PD | 31 | Yes | 43 | 2 | 19 | Fail | Hearing aids |
| HP15 | m | 88 | 1000 | PD | 18 | No | 23 | 2 | 14 | NA | NA |
| HP16 | m | 70 | 100 | PD | 17 | Yes | 18 | 2 | 23 | Fail | Pass |
| HP17 | m | 71 | 0 | MSA-C | 5 | No | 23 | 2 | 27 | Fail | Fail |
| HP18 | m | 72 | 820 | PD | 2 | No | 45 | 3 | 25 | Fail | Pass |
| HP19 | m | 59 | 1350 | MSA-P | 8 | No | 52 | 3 | 26 | Pass | Pass |
| HP20 | m | 65 | 1200 | PD | 15 | No | 4 | 2 | 18 | Pass | Pass |
| HP21 | m | 60 | 610 | PD | 12 | No | 17 | 2 | 29 | Pass | Fail |
| HP22 | f | 68 | 750 | PD | 15 | Yes | 36 | 1 | 25 | Pass | Pass |

*Note.* LED = levodopa equivalence dosage; DBS = deep brain stimulation; UPDRS = Unified Parkinson's Disease Rating Scale; UPDRS-Speech = speech item score from the UPDRS; MoCA = Montreal Cognitive Assessment; CP = communication partner; m = male; PD = Parkinson's disease; f = female; NA = not applicable; MSA-C = multiple systems atrophy cerebellar type; MSA-P = multiple systems atrophy Parkinsonian type.

(ND) condition. The three devices were expected to capture an array of device styles and capabilities and anticipated to appeal differently to each participant dyad depending on various factors including their speech symptoms, communication needs, and lifestyle.

Device A is a portable wired belt-pack speech amplifier (Chattervox). The intended setup for this amplification device is as follows. The talker wears a lightweight, headset microphone (included with the Chattervox device) connected to an external speaker worn as a belt back.

Device B is similar to the BoomVox in form and function (Nady WA120BT), consisting of a lightweight, wireless headset microphone (Nady HM20) that transmits wirelessly over a VHF channel to a larger, stationary speaker that projects the speech from up to several meters away from the talker. The external speaker is 8.3 × 10.6 × 5.4 in., weighs 2.4 kg, and has multiple possible audio adjustments, including volume, echo, treble, and bass.

Device C (Nady 351VR) is similar to the Phonic Ear Easy Listener body-worn FM system previously tested for use with individuals with hearing loss (Crandell et al., 2001). As mentioned earlier, personal communication systems have not previously been tested for people with hypophonia. A lightweight, headset microphone (Nady HM20) worn by the talker transmits the speech wirelessly to a pocket-sized VHF receiver, which is then amplified through headphones worn by the listener. Devices similar to Device C have been used with individuals with hearing loss (e.g., Easy Listener), but there are no reports of their use in PD. In this study, the HP participants wore the microphone, and the CP participants wore the headphones.

## *Protocol*

The study consisted of three stages: two stages completed by the HP and CP participant dyads (Stages 1 and 2) and a final stage completed by the naive listeners (Stage 3). HP and CP participants completed two stages of the study over four visits, described in greater detail below. Briefly, Stage 1 consisted of one visit at the laboratory (Visit 1), which involved administration and recording of the speech stimuli as well as a brief trial period with each of the three devices. Stage 2 involved three visits (Visits 2–4). Participant dyads were given each device for 1 week at a time to try at home. Following these longer home trial periods, they would either return to the laboratory or the researcher (T. K.) would visit them in their homes to discuss the trial period and to complete a battery of questionnaires related to its impact on their communication and their preferences.

### Stage 1: In-Laboratory Speech Tasks

At the first visit (Stage 1), the HP participants completed baseline speech tasks followed by experimental speech tasks below in each of the four device and two noise conditions, described below. The CP participants served as listeners for these tasks. All 22 participant dyads completed Stage 1.

*Physical setup.* In the laboratory, the HP and CP participants were seated 2 m apart and were facing each

other. HP participants wore a headset condenser microphone to record the speech stimuli (AKG c520). When a device was in use, the HP participants additionally wore the headset microphone associated with each device. A second recording microphone (Shure SM48) was positioned next to the CP participant at the 2-m distance from the talker. Two external audio speakers positioned at a 2-m distance from the talkers provided multitalker background noise for a subset of the conditions. The CP participants wore a lavalier microphone that was not used for acoustic analyses but was used to facilitate transcription of their utterances.

*Speech tasks.* For each condition, HP participants read aloud six randomly selected sentences (two each of five-, six-, and seven-word lengths) from the Sentence Intelligibility Test (SIT; Yorkston et al., 1996). Immediately following each sentence, the CP participants were instructed to repeat aloud what they heard their partner say. If they only heard part of the utterance, they were asked to repeat anything they did hear. If they were completely unsure, they were asked to verbally indicate this to facilitate transcription later. As part of a larger protocol, HP participants also verbally described pictures to their CPs. This picture description task will not be reported in this article.

HP09 and HP10 had difficulty completing the speech tasks due to attention and vision challenges, respectively. Data from HP09 were excluded from both the SNR and intelligibility analyses. HP10, who was able to complete a modified version of the sentence reading task in which he repeated the examiner, was included in the SNR analyses but excluded from the intelligibility analyses. HP05 and HP15 dropped out of the study after Visit 1 and were excluded for the intelligibility tasks.

*Conditions.* Eight conditions were included: four device conditions (ND, Device A, Device B, Device C) and two noise conditions (no noise [NN] and multitalker noise [MN]). The MN condition involved playing a recording of multitalker babble (Audiotech, four talker noise) at a level of 65 dB SPL measured at the 2-m distance from the talker and next to the CP participant. The ND condition was always elicited first, followed by a random ordering of the three device conditions. Within each device condition, the NN condition was always elicited first, and the sentence reading task was always elicited before the picture description task.

During the Device A and B conditions, the device volume was turned up as high as possible before device feedback was perceived by the investigators (i.e., gain-before-feedback level). While the final volumes likely varied slightly across participants, device and participant positioning were arranged as consistently as possible to mitigate this. Other possible settings (such as "echo" for Device B) were not activated. The belt-pack amplifier of Device A was positioned on the lap of the HP participant, and the stationary amplifier of Device B was situated on the floor next to their chair, facing the CP participant. During the Device C condition, the CP was instructed to adjust the volume of the receiver to a comfortable listening level while their partner counted from one to 20. A second receiver set to a constant volume was set up to record the speech in order to account for differences in CP participants' volume preferences. This constant volume was set by the first author (T. K.) before the study began and corresponded to a comfortable listening level of the speech of a healthy male talker (S. A.).

Three audio inputs were recorded, with recording input gain set to constant levels for the entire experiment: (a) the audio from the primary headset microphone worn by the HP participant; (b) the 2-m microphone; and (c) in the case of Device C, the audio from a secondary receiver on the same VHF channel, set to a constant volume for all participants. These latter two recordings were intended to capture the auditory experience of the CP participant and would be played to the naive listeners later. It is imperative to note that, while the CP controlled the listening volume of Device C, the actual recordings of the receiver input used a constant volume (which had been previously set to a comfortable listening level by the researchers) in order to standardize the eventual presentation of audio stimuli to the naive listeners.

*Device ratings.* Following each device condition, both the HP and the CP participants filled out a brief device user experience questionnaire (Andreetta et al., 2016; see Appendix) in which they were asked to rate various aspects of the device using a 100-mm visual analogue scale. This questionnaire contained five questions, including "Overall, is this a device you would prefer to use?" The HP participants were instructed to complete the questionnaire from their own perspective, whereas the CP participants were instructed to complete the questionnaires in terms of how they perceived their partner.

**Stage 2: At-Home Device Trials**

At the end of Visit 1, one of the three devices was randomly selected to be trialed first. Participants were told that they would be given the opportunity to try out each device. They were instructed on the basic elements of use for the device they would trial and were given a Device Diary to help them keep track of when they used the device, the context in which they used it, and any notes they would like to keep. This Device Diary was optional. Participants were instructed to try to use the device at least twice during the week and in more than one setting with more than one person, if possible.

Following the completion of the first 1-week trial period, the participant dyads met with the researcher for three subsequent visits (one visit following each 1-week device trial period), each approximately 1 week apart. These visits lasted approximately 1 hr in length and included an informal discussion of the trial period and a battery of questionnaires related to communication and device preferences. Both the HP participants and the CP participants completed their own set of questionnaires as described in Stage 1. Critically, the Device Experience Questionnaire was completed at each of these 1-week trial periods (i.e., for each device).

Upon the completion of all three device trial periods, the HP participants were given the option to continue using an amplification device. If they consented, their primary neurologist (M. J.) completed a prescription for the device of their choice and signed an application for funding through the Ontario Assistive Devices Program (OADP), which covers up to 75% of the cost of assistive communication devices up to $400 (CAD). An SLP with individual authorizer status (S. A.) completed the funding application, which included a recommendation for the amplification device the participant had selected. The total cost for each device after OADP funding was applied totaled approximately $75–$125 (CAD). Additional costs not covered by OADP included shipping and handling, as well as optional device add-ons. If the participant dyads did not want to or could not pay for the device, they were still given the opportunity to continue trialing the device of their choosing, thus removing a potential cost barrier (no participants chose this option). Completing the study was not a prerequisite for seeking a prescription or funding for a speech amplification device.

Nineteen participant dyads completed all device trial sessions. Of these 19, one HP participant (HP09) demonstrated a decline in functioning and was unable to actively participate in the device trials. HP10 was unable to fill out the questionnaires due to his vision difficulties mentioned above, so only his partner's ratings for this dyad are included in the device preference ratings.

**Stage 3: Naive Listener Intelligibility**

Utterances recorded at Visit 1 were later segmented and played to naive listeners via a custom Praat script (Boersma & Weenink, 2011) written by the first author. Utterances were presented at the natural sound pressure level (SPL) at which they were recorded (i.e., not scaled). Naive listeners attended approximately four self-paced visits usually lasting between 1 and 2 hr each and spaced approximately a week apart. During these visits, listeners were seated in a sound-treated booth (Industrial Acoustic Company) and heard stimuli presented on external speakers calibrated to 70 dB SPL.

Listeners heard four blocks (2 noise conditions and two speech tasks): NN-SIT, MN-SIT, NN-Picture description, MN-Picture description. Only the SIT tasks are described here. Within each block, utterance order presentation was fully randomized across participants and devices, and each utterance was presented 2 consecutive times. A random 10% of utterances were repeated for reliability. Listeners were instructed to transcribe (type) exactly what they heard the talker say. If they were unsure, they were instructed to write as much as they could understand or indicate that they could understand nothing by typing "NA."

## *Analyses*

Primary analyses included acoustic, perceptual, and device user experience metrics. To answer RQ 1, the acoustic SNR measure was calculated separately for each utterance as the difference between the average intensity of the speech signal and the intensity of the noise floor. Speech was continuous and excluded silent periods of 250 ms or greater, as was done by Andreetta et al. (2016). Average noise floor levels were calculated for each condition by measuring the intensity of the audio signal at approximately three periods of time not punctuated by speech or other background noise. SNR was calculated from the 2-m audio recordings for all conditions except for Device C, which used the receiver input recordings in order to more closely replicate the listener's experience (i.e., through headphones).

To answer RQ 2, two intelligibility metrics were analyzed: familiar listener intelligibility and naive listener intelligibility. Speech intelligibility was calculated as the percentage of words per sentence correctly repeated by the CP or correctly transcribed by the naive listener. Transcription of the CP's utterances to be used in the familiar listener intelligibility calculation was completed by the first author using recordings from the 2-m microphone (which was located next to the CP participant) or by consulting the recording from the supplemental lavalier microphone. All transcriptions were manually compared to the original utterances produced by the HP participants. For all listener transcriptions, contractions were counted as correct (e.g., no penalization if the target was "He is" and the CP said "He's"), and additions were not penalized so long as they did not alter the pronunciation of the target word (e.g., "doing" would be counted as correct even if "do" was the target, but "did" instead of "do" would be counted as incorrect).

RQ 3 was addressed using two device user experience rating time points for each participant group (HP and CP) and for each device: immediately following the devices in Stage 1 and following each 1-week trial period during Stage 2. User experience was analyzed as the percentage along the visual analogue scale for the question "Overall, is this a device you would prefer to use?" with anchors from *low preference* to *high preference*.

Finally, RQ 4 was answered at the conclusion of Stage 2 by the participant dyads' response to the question "Do you wish to continue using a device" (*yes*/*no*) and, if *yes*, state their device selection. In summary, the outcome measures are reported in Table 2.

## *Statistical Analyses*

The primary analyses for RQs 1–3 were modeled using linear mixed-effects regression with the `lmer()` function from the `lme4` package (Bates et al., 2015) in R (R Core Team, 2018). Post hoc pairwise comparisons were computed using the `emmeans` package (Lenth, 2018). To address RQs 1 and 2, each of the three outcome measures were modeled as a function of device and noise conditions. Each model is described below.

*SNR.* Fixed effects for the SNR model included device, noise, and their interaction. Device and noise were coded using treatment contrasts with the baseline condition as the reference level for each (i.e., ND and NN).

**Table 2.** Summary of outcome measures.

| Outcome | Metric | Stage | RQ |
|---|---|---|---|
| 1. SNR | dB | 1 | 1 |
| 2. Familiar listener intelligibility | % words correct | 1 | 2 |
| 3. Naive listener intelligibility | % words correct | 3 | 2 |
| 4. Device user experience ratings | % along VAS at two time points | 1 and 2 | 3 |
| 6. Decision to continue using device | Yes/no | 2 | 4 |
| 7. Device selection | Devices A, B, and C | 2 | 4 |

*Note.* RQ = Research question; SNR = speech-to-noise ratio; VAS = Visual Analog Scale.

Random effects included by-participant random intercepts and random slopes for device. This structure accounts for the individual variability across the HP participants and additionally accounts for differential effects of each device on individual talkers' SNR levels.

*Intelligibility.* Intelligibility was treated as a proportion from 0 to 1 for both familiar and naive listener groups and was logit-transformed. For both intelligibility models (familiar and naive), fixed effects included device, noise, and their interaction, as for the SNR model. Device and noise were coded as above. Random effects included by-participant random intercepts and random slopes for device, as well as by-sentence random intercepts.

Additionally, for the naive listener intelligibility model, a by-listener random intercept was included to account for baseline variability across listeners. This was not a relevant concern for the familiar listener model, as each HP talker had only one associated familiar listener (i.e., their spouse). Listener inter- and intrareliability was calculated using the intraclass correlation coefficient (ICC; Koo & Li, 2016). Interrater reliability across the four listeners was examined using average consistency in a two-way random model (ICC 2, *k*). Intrarater reliability for each listener was examined using average agreement in a two-way mixed model (ICC 3, *k*).

Primary analyses of interest included (a) comparisons between each device condition, within each noise condition, and (b) comparisons between noise conditions within each device. To this end, post hoc pairwise comparisons for all possible pairs were computed using estimated marginal means (i.e., least squares means) from the `emmeans` R package (Lenth, 2018). *P* values calculated with `emmeans` used the Kenward–Roger approximation and were adjusted using the Tukey post hoc method.

*Device preferences.* To address RQ 3, ratings were logit-transformed and modeled first with fixed effects of time, device, and group. The interaction between time and device was added to this base model and compared using a likelihood ratio test to determine whether the interaction improved the model fit. A random by-dyad intercept was included in the model. For example, HP01 and CP01 were participants from two distinct groups, but from the same dyad (i.e., CP01 was the spouse of HP01). Nineteen of the 22 dyads completed the at-home device trial periods. One HP participant (HP09) whose health deteriorated over the course of the study was excluded because he was unable to use the devices at all toward the end of the trials. The remaining 18 dyads were included in the device preference analysis. Final device decisions (RQ 4) were tabulated for the 18 dyads who completed the trials, and only qualitative remarks are included.

## Results

### *SNR*

SNR results are reported in Table 3 and Figure 1 for 21 participants. The pairwise comparisons for the SNR model are reported in the supplemental materials (Supplemental Material S1).

#### Effect of Background Noise

SNR for all device conditions was significantly lower in the presence of the MN compared to NN (ND: estimated difference [est. diff.] = 1.86, $p < .001$; Device A: est. diff. = 3.95, $p < .001$; Device B: est. diff. = 5.18, $p < .001$; Device C: est. diff. = 12.07, $p < .001$).

#### Compared to ND

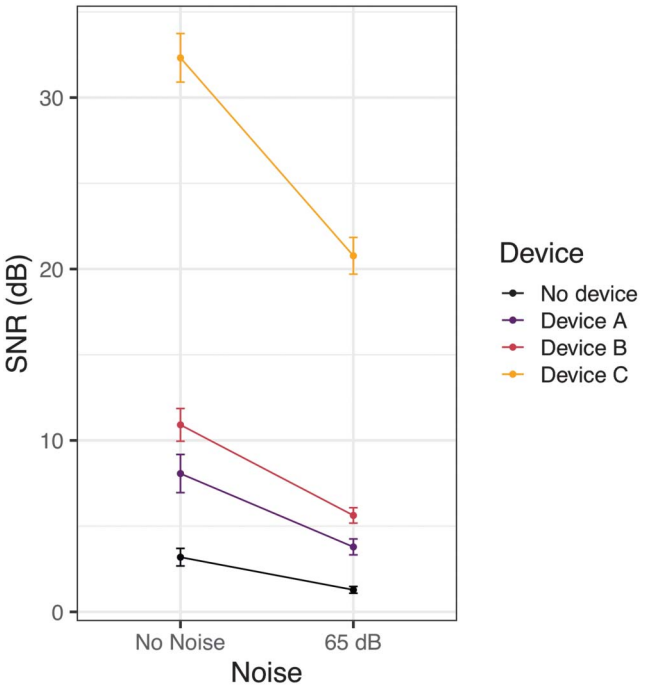
Each device was associated with greater SNR levels compared to not having any device at all, as is evident in Figure 1. SNR was significantly higher for each device compared to ND, both in quiet (Device A: est. diff. = −4.75, $p < .001$; Device B: est. diff. = −7.61, $p < .001$; Device C: est. diff. = −29.32, $p < .001$) and in noise (Device A: est. diff. = −2.66, $p = .002$; Device B: est. diff. = −4.3, $p < .001$; Device C: est. diff. = −19.11, $p < .001$). SNR levels were greatest for Device C (NN: 32.32 dB; noise: 20.77 dB),

**Table 3.** Speech-to-noise ratio (SNR; dB) means and standard deviations for all device and noise conditions.

| Device | Noise | SNR | *SD* |
|---|---|---|---|
| No device | NN | 3.189 | 2.349 |
| | MN | 1.283 | 0.881 |
| Device A | NN | 8.067 | 5.104 |
| | MN | 3.789 | 2.078 |
| Device B | NN | 10.908 | 4.367 |
| | MN | 5.625 | 2.006 |
| Device C | NN | 32.320 | 6.338 |
| | MN | 20.771 | 4.665 |

*Note.* NN = no noise; MN = multitalker noise.

**Figure 1.** Empirical data for speech-to-noise ratio (SNR) for each device condition in no noise and multitalker noise. Error bars represent standard error. Device A: belt-pack amplifier; Device B: wireless stationary amplifier; Device C: personal communication system.




followed by Device B (NN: 10.91 dB; noise: 5.63 dB), and then Device A (NN: 8.07 dB; noise: 3.79 dB).

### Device Comparison

Overall, Device C had the largest SNR, followed by Device B, followed by Device A. All pairwise device comparisons were significant in noise (A vs. B: est. diff. = −2.86, $p = .007$; A vs. C: est. diff. = −24.57, $p < .001$; B vs. C: est. diff. = −21.71, $p < .001$). In the presence of noise, both Devices A and B had lower SNR than C (A vs. C: est. diff. = −16.45, $p < .001$; B vs. C: est. diff. = −14.81, $p < .001$), but Devices A and B did not significantly differ from one another (A vs. B: est. diff. = −1.64, $p = .312$).

## *Intelligibility*

Intelligibility results are reported in Table 4 and Figure 2. The pairwise comparisons for the Intelligibility models are reported in the supplemental materials (Supplemental Material S2 and S3).

### Familiar Listeners

*Effect of background noise.* The downward slope for all device conditions visible in Figure 2 demonstrates that the presence of background noise was associated with a decrease in intelligibility within all but one of the device conditions, indicating that, generally, the familiar listeners

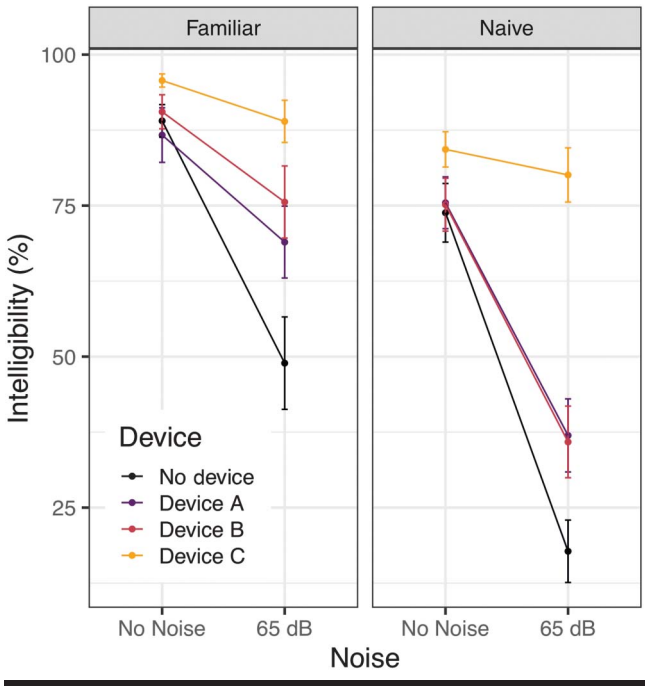
**Table 4.** Intelligibility (% words correct) means and standard deviations for all device and noise conditions.

| Device | Noise | Intelligibility | *SD* | Listener group |
|---|---|---|---|---|
| No device | NN | 89.032 | 12.048 | Familiar |
| | MN | 48.917 | 34.218 | |
| Device A | NN | 86.679 | 20.205 | |
| | MN | 68.960 | 26.611 | |
| Device B | NN | 90.536 | 12.571 | |
| | MN | 75.591 | 26.708 | |
| Device C | NN | 95.706 | 4.876 | |
| | MN | 88.941 | 15.629 | |
| No Device | NN | 73.811 | 21.697 | Naive |
| | MN | 17.770 | 23.101 | |
| Device A | NN | 75.467 | 19.187 | |
| | MN | 36.950 | 27.061 | |
| Device B | NN | 75.152 | 19.567 | |
| | MN | 35.875 | 26.529 | |
| Device C | NN | 84.307 | 12.745 | |
| | MN | 80.076 | 19.559 | |

*Note.* NN = no noise; MN = multitalker noise.

were less accurate in identifying what their partner said in a noisy environment. Specifically, all but Device C were associated with a significant decrease in intelligibility in the noise condition (ND: est. diff. = 2.23, $p < .001$; Device A: est. diff. = 1.02, $p < .001$; Device B: est. diff. = 0.88, $p < .001$; Device C: est. diff. = 0.41, $p = .346$). As can be

**Figure 2.** Empirical data for intelligibility ratings in each device and noise condition. Left panel indicates familiar listeners; right panel indicates naive listeners (means were first aggregated over listeners). Error bars represent standard errors. Device A: belt-pack amplifier; Device B: wireless stationary amplifier; Device C: personal communication system.



seen in Figure 2, intelligibility for Device C was also associated with a downward trend in noise, but this was a flatter slope than for the other device conditions and did not reach significance.

*Comparison to ND.* In quiet, having a device did not impact familiar listener intelligibility, as captured by a lack of any significant differences between each of the three device conditions (A, B, and C) and the ND condition ($p > .8$). In noise, however, all three amplification devices were associated with significantly higher intelligibility compared to ND (Device A: est. diff. = −1.07, $p < .001$; Device B: est. diff. = −1.47, $p < .001$; Device C: est. diff. = −2.25, $p < .001$).

*Device comparison.* In quiet, none of the devices demonstrated significant differences from one another ($p > .6$). In the presence of noise, Device C was associated with significantly higher intelligibility than Device A (est. diff. = −1.17, $p = .008$, but Devices A and B and Devices B and C did not differ from one another (A vs. B: est. diff. = −0.39, $p = .487$; B vs. C: est. diff. = −0.78, $p = .117$).

**Naive Listeners**

*Reliability.* Average interrater reliability for sentence intelligibility was .934 (95% confidence interval [.919, .946]). Average intrarater reliability was .934 (range: .908–.949). Both inter- and intrareliability may be interpreted as "excellent" reliability (i.e., > .90; Koo & Li, 2016).

*Effect of background noise.* As was the case for naive listeners, all device conditions except for Device C demonstrated significantly worse intelligibility in the presence of background noise (ND: est. diff. = 3.17, $p < .001$; Device A: est. diff. = 2.16, $p < .001$; Device B: est. diff. = 2.17, $p < .001$; Device C: est. diff. = 0.21, $p = .458$).

Figure 2 demonstrates a steeper slope for Devices A and B and the ND condition for the naive listeners compared to the familiar listeners, indicating that naive listener intelligibility suffered greater detriment when noise was introduced. This is also captured by the greater estimated pairwise differences reported for the naive listener model.

*Comparison to ND.* In the NN condition, none of the devices were associated with better naive listener intelligibility compared to ND except for Device C, which demonstrated a marginally significant intelligibility benefit (est. diff. = −0.56, $p = .049$). Devices A and B did not significantly differ from the ND condition (Device A: est. diff. = −0.09, $p = .999$; Device B: est. diff. = −0.04, $p > .999$). In the multitalker noise condition, all three devices were associated with significantly better intelligibility compared to ND (Device A: est. diff. = −1.1, $p < .001$; Device B: est. diff. = −1.04, $p < .001$; Device C: est. diff. = −3.52, $p < .001$). This effect was strongest for Device C (i.e., largest est. diff.).

*Device comparison.* In quiet, there was a nonsignificant trend for Device C to be associated with higher intelligibility than Device A (A vs. C: est. diff. = −0.46, $p = .095$). No other device comparisons were significant (A vs. B: est. diff. = 0.05, $p > .999$; B vs. C: est. diff. = −0.52, $p = .181$). In noise, Device C was associated with significantly higher intelligibility than Devices A and B (A vs. C: est. diff. = −2.42, $p < .001$; B vs. C: est. diff. = −2.48, $p < .001$). There was no significant difference between Devices A and B (est. diff. = 0.06, $p > .999$).

Overall, for speech intelligibility scores, Device C was associated with better intelligibility, followed by Device B and Device A, and finally by ND. Both listener groups (familiar and naive) demonstrated similar patterns, with more extreme effects for the naive listeners. While better speech intelligibility for both listener groups was found for Device C, naive listeners benefitted more from this device in both quiet and noise. Noise was associated with worse intelligibility for all device conditions except for Device C for both listener groups, but the negative effect of noise was lesser for the familiar listeners.

In summary, across all outcome measures for Stage 1, HP participants exhibited lower SNR and were less intelligible in the presence of noise in most device conditions. The exception to this pattern was for Device C, the personal communication system, which was associated with similar intelligibility levels both in quiet and in noise. Having a device was associated with greater SNR in both quiet and in noise. In general, having a device largely improved intelligibility in noise and, to a lesser extent, in quiet. Device C was associated with higher SNR levels than the other two devices and better intelligibility than Device A. Overall, device performance can be described as a hierarchy such that Device C > Device B ≥ Device A > ND.

## *Device Preferences*

Device preference ratings are shown in Figure 3. There was no main effect of time, device, nor group. Device preferences differed across time points though, evidenced by the fact that including the interaction between time and device improved the model fit ($\chi^2 = 3.368$, $p = .037$). Within each time point, pairwise comparisons revealed few device differences with the following notable exceptions: Immediately after the laboratory trials, Device C was rated as being more preferable than Devices A and B, as noted by a marginally significant difference between Devices A and C (est. diff. = −1.06, $p = .052$), and a trend toward significance between Devices B and C (est. diff. = −0.97, $p = .096$). Following the 1-week trial periods, however, these differences disappeared completely ($p > .9$ for all comparisons within the 1-week time point). In other words, while there was a slight (marginally significant) trend for a preference for Device C after approximately 5–10 min of controlled use with the devices, the opportunity to trial each device for longer periods of time in a more naturalistic environment led to the overall device preference hierarchy dissipating. While individuals did state explicit preferences for given devices, overall, there was no one-size-fits-all device that demonstrated a clear difference in preference ratings after the trials. Furthermore, device ratings did not change according to time period; that is, overall device ratings neither improved nor worsened following the opportunity to trial the device for longer.

**Figure 3.** Box plots showing device preference ratings for each device at each time point (left panel: immediately following device trials in Stage 1; right panel: following 1-week device trial periods in Stage 2. HP = hypophonia; CP = communication partner

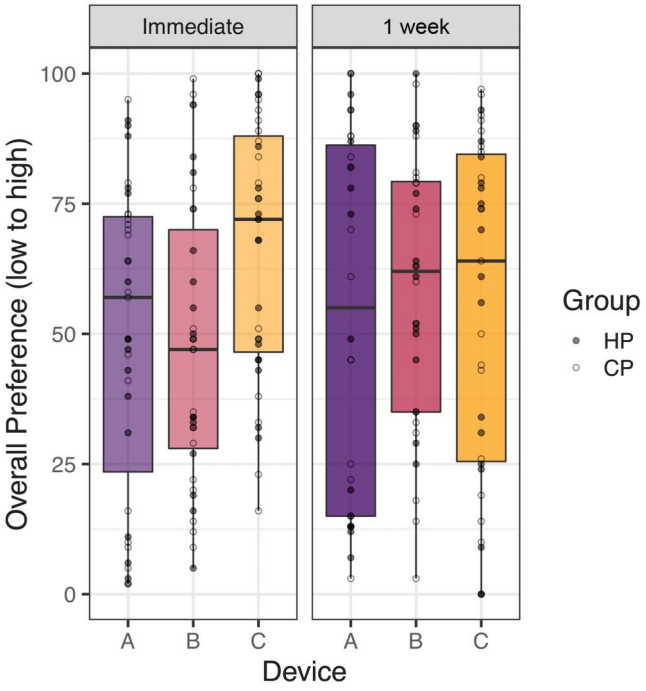
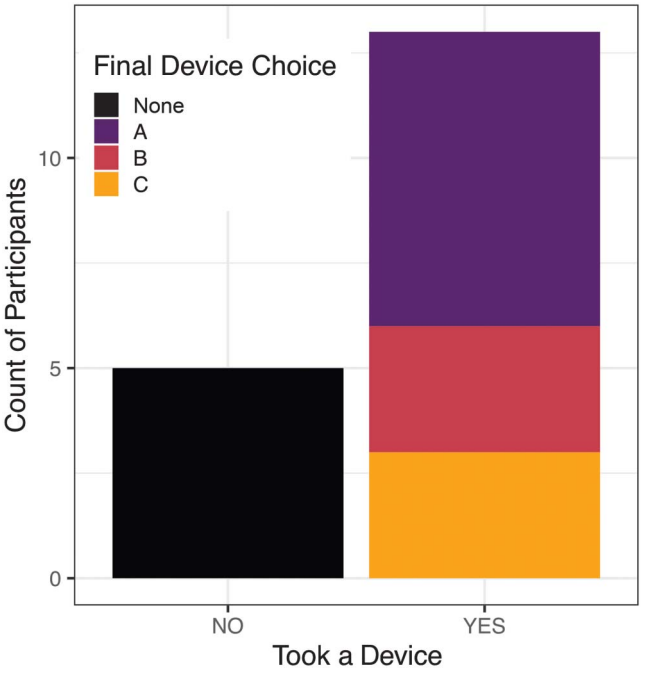



**Figure 4.** Count of final device choices across the dyads who completed the device trials. Device A: belt-pack amplifier; Device B: wireless stationary amplifier; Device C: personal communication system.




### *Final Device Decision*

Figure 4 reports the distribution of final device choices for the participants who completed the device trial periods. Of the 18 dyads who completed all at-home device trial periods (i.e., Stage 2 of the study procedures), 13 (72%) chose to continue using a device after the study had concluded. Seven (HP03, HP06, HP16, HP17, HP19, HP21, and HP22) chose Device A, three (HP04, HP10, and HP18) chose Device B, and three (HP01, HP02, and HP14) chose Device C. Five (HP07, HP08, HP11, HP12, and HP13) declined to take a device. Of the four (HP05, HP09, HP15, and HP20) dyads who did not complete the trials, two expressed that they were no longer interested in the study (indicating that they would not have been interested in taking a device), one (HP05) dropped out due to health concerns, and one (HP09) became unable to actively participate in the trials due to declining health.

Figure 5 reports SNR and familiar listener intelligibility without any device in the presence of background noise (i.e., ND, MN condition), ordered by individual participant SNR outcomes. HP10, who chose Device B, is not included in this figure because he was unable to do the sentence reading task.

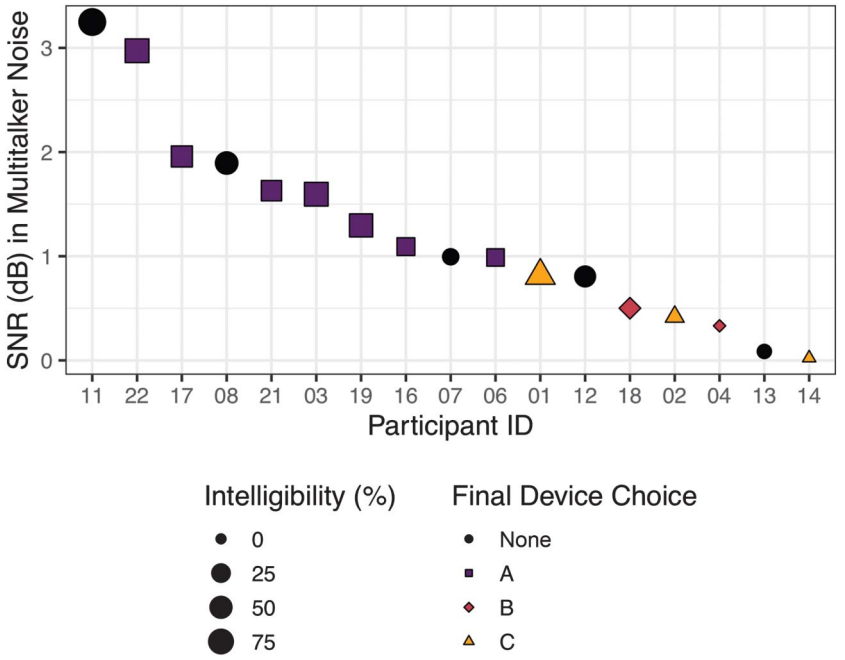
## Discussion

In summary, the results from this study provide evidence that speech amplification devices and personal communication systems may be a viable augmentative form of speech intervention for individuals with hypophonia. Results demonstrated that, compared to ND, each of the three devices tested were associated with gains in SNR and in speech intelligibility in adverse listening conditions (i.e., in noise). Following the opportunity to trial the devices in their natural living environments, the majority of the dyads in this study chose to continue using a device after the study concluded, indicating that they saw enough benefit in the device to incorporate it into their communication lifestyles.

A clear device hierarchy emerged from the objective speech tasks and can be summarized as Device C > Device B ≥ Device A > ND. Device C outperformed Devices A and B for both SNR and intelligibility, though the intelligibility benefit was most noticeable in adverse listening conditions. Device B outperformed Device A in SNR, but the two devices did not differ from one another in terms of how intelligible the speakers were to either familiar nor naive listeners, regardless of the amount of background noise.

This device hierarchy did not map on to device preferences or individuals' ultimate decisions of whether and which device to continue using. This is not entirely surprising, as user preferences are not necessarily driven by device performance alone. In the study conducted by Andreetta et al. (2016), the authors found that the Spokeman, a small, lightweight amplification device, received the highest ratings for dimensions of physical comfort, visual presentation,

**Figure 5.** Individual participant speech outcomes in the no device condition in multitalker noise. Participant IDs on the *x*-axis are ordered by their speech-to-noise ratio (SNR) values in the no device condition in the presence of multitalker noise. Larger point size represents higher familiar listener speech intelligibility in noise. Point color represents final device choice. Device A: belt-pack amplifier; Device B: wireless stationary amplifier; Device C: personal communication system.




and overall preference, despite the finding that it performed more poorly compared to other devices with respect to more objective speech measures (SNR and intelligibility). The findings from the study of Andreetta et al. suggested that user preference and user comfort likely do not predict device performance or effectiveness and cautioned that SLPs working with these individuals explore devices that can optimize performance without compromising a client's aesthetic preferences.

A number of anecdotal reports from participants in this study are of relevance to the patterns in device choice observed. Discussions with the participants of this study suggested that they would be disinclined to use devices they found to be too unsightly or uncomfortable. Therefore, it is important to consider these qualities in order to maximize not only speech performance but also user buy-in and likelihood of use. Participants were not required to elaborate on their final decisions; however, reasons for declining a device included feeling they could communicate effectively without it, declining health that made using a device too cumbersome, and dislike of the use of a device in general.

While more research is required to identify factors that would identify individuals who may be more likely to be interested in using one type of device over another, some preliminary observations may be made. Reasons given for choosing Device A (seven participants) included portability, independence, and being able to use it with more than one person at a time (i.e., in contrast to Device C).

Reasons given for choosing Device B (three participants) included being able to leave the speaker in a given location (e.g., by a chair in the living room for HP04 or in the kitchen for HP18). One participant (HP18) mentioned that he felt that hearing his own amplified speech made him more aware of his communication, and he reported that he spoke more clearly when using the amplifier as a result. HP10, who was residing in long-term care, reportedly liked being able to have the amplifier stationary in his living quarters or brought into the main social areas.

Reasons given for choosing Device C included greater discreteness (only the person wearing the receiver would hear the amplified speech) and greater amplification for an individual listener. Given that this is the first study to explore the use of a personal communication system for hypophonia, further reasons are speculated based on qualitative reports of individual participants.

For one dyad (HP/CP14), Device C was reportedly desirable because the spouse wore hearing aids, and the other amplification options were not loud and/or clear enough to be effective. This dyad expressed a desire to obtain a device like Device C but required greater customization to accommodate CP14's hearing aids. Following the conclusion of the study, this dyad was seen by an audiologist to discuss and trial similar options that could be paired to CP14's hearing aids, but a final solution was not reached. HP14 reported that she found the other devices to be effective when communicating with other individuals besides

her spouse but ultimately decided not to take these devices. CP06 also wore hearing aids, but both she and her HP partner expressed satisfaction with the amplified output of Device A.

It is worth noting that all three HP participants of the dyads who chose Device C had MoCA scores below the cutoff for mild cognitive impairment (i.e., < 26), and two of the three fell below the cutoff for dementia (i.e., < 21; Dalrymple-Alford et al., 2010). These two individuals (HP02 and HP14) were also DBS recipients, which is associated with cognitive declines in some individuals (Tröster et al., 2017). Although speculative, it is possible that final device preferences may be partially influenced by the CP and especially for Device C. The CP needs to recognize the benefits of a device like Device C in their communicative interactions with their partner with hypophonia and be willing to take on a primary role in managing the use of the device on a daily basis. The CP may need to adopt the perspective that the best way to improve communication in hypophonia is to place a new focus on "enhancing the ears of the listener" rather than only focusing on enhancing the speech of the talker. This may require a greater emphasis on developing listener-oriented verbal interaction strategies that address the specific effects of hypophonia on communication. The importance of including communicative partners in other aspects of speech evaluation and treatment has been suggested before as a means of facilitating the implementation of treatment goals that are of mutual benefit to both the person with PD and their primary partner (Donovan et al., 2008; Dykstra et al., 2015).

As can be seen in Figure 5, the six participants[7] who chose Devices B and C tended to demonstrate overall lower SNR values and were less intelligible to their CPs in adverse listening conditions at baseline (i.e., MN, ND). Participants who chose Device A tended to demonstrate higher SNR values (range: 0.988–2.973 dB) and were overall relatively more intelligible to their partners (range: 33.3%–97.6%). The five participants who chose not to take a device varied along these dimensions. For example, among the participants who chose not to take a device, SNR in the multitalker, ND conditions (i.e., baseline speech in adverse listening conditions) ranged from 0.086 dB (HP13) to 3.249 dB (HP11), and intelligibility ranged from 6.7% (HP13) to 97.6% (HP11).

Participants who chose not to take a device also varied in other demographic dimensions (see Table 1). Ages for these five participants ranged from 59 to 72 years and included one woman. One (HP13) participant had been diagnosed with PD for less than 1 year, whereas the others had been diagnosed more than 10 years prior. All but one (HP13) of the five HP participants and all but one (CP07) of the CP participants passed the hearing screening, and MoCA scores ranged from 20/30 (HP11) to 26/30 (HP07). Two (HP11 and HP12) reported receiving speech therapy in the past, and three (HP07, HP11, and HP12) had received DBS surgery.

People with PD also tend to demonstrate further reductions in loudness in their daily life (Gustafsson et al., 2019), though it is not known to what degree individual variability is present in the difference between sound levels in formal settings (such as a clinic or lab) versus in more naturalistic settings. Individual device preferences are likely more driven by the characteristics of an individual's speech in their daily life that may not be fully captured by the formal speech tasks.

Another consideration is the degree to which hearing one's own voice amplified affects user experience. For the participants who did not like Device C after the trial periods, it is possible that they found it frustrating to use because they had no feedback that their speech was being amplified. With Devices A and B, which amplified the speech via an external speaker, all persons within hearing range of the device are able to hear the amplified signal. With Device C, only the person wearing the headphones is able to hear the signal. Some HP participants also indicated frustration that Device C did not function more like a two-way device would, in which they would be able to hear their spouse's voice amplified and be able to communicate over larger distances.

Conversely, one participant (HP18, as mentioned above) mentioned that hearing his voice amplified made him more aware of his speech, and he and his wife noticed that he would speak more loudly when he wore the device (and, in particular, Device B). In this case, Device B was his final preference because he liked the external amplification it afforded but also benefitted from the smaller wearable transmitter (compared to the portable speaker worn with Device A).

More research is needed to determine specific factors that predict user preferences and long-term device use. Overall, the findings of this study point to a need to advocate for the exploration of speech amplification as a form of intervention for this population and to continue to develop devices that better suit their needs.

### *Methodological Considerations and Limitations*

Several limitations in this study warrant discussion. One limitation is the heterogeneity and relatively small sample size of the HP and CP participants. While all participants were recruited on the basis of the presence of hypophonia as a predominant speech symptom, the severity of both their hypophonia, other speech symptoms, and other parkinsonian features varied. Furthermore, the CP participants varied in their hearing status and likely also varied in other, undocumented ways, such as cognitive status.

It is likely that factors unrelated to speech contributed to device experience for both user groups. For example, physical factors such as mobility or the presence of dyskinesias may have impacted the comfort of wearing the devices. The social networks of the dyads also likely impacted their experience with the devices but were not controlled

[7]Only five of the six participants who chose Device B or C are pictured in Figure 5 because HP10 who was unable to do the speech task is not shown.

for or characterized in detail. For example, couples that mainly communicated with one another may have been more inclined to be satisfied with a device such as Device C, whereas couples that often participated in social outings may have been more likely to find the same device limiting.

Another possible limitation is that all participants trialed the devices in different orders. While device trials were counterbalanced, it is possible that trialing one device before another could have impacted the experience ratings of subsequent devices and could have interacted with individual baseline device preferences as well. Given the relatively small sample size, this may not be fully controlled for, despite the counterbalanced presentation.

Individual experiences with the device also varied considerably. Though all participants were instructed to use each device on at least a few different occasions and amounting to at least 2 hr, adherence to these guidelines varied. Many participants acknowledged using devices less than this, especially when they did not perceive them to be particularly useful to their individual lifestyle and circumstances. A possible outcome of this may be that individuals who had initial poor impressions of a given device did not give ample time to trial each one, potentially resulting in decreased overall preference scores in some cases. Even so, the purpose of the trial periods was for individuals to assess their likelihood and ability to use the devices in their day-to-day lives; if they did not perceive them to be useful or did not enjoy using them, perhaps an extended, more rigorous trial period (i.e., longer in duration and/or greater requirements of the frequency of device usage during that period) would not have impacted this.

Disease-related factors also pose limitations to the formal laboratory testing procedures. As mentioned in the introduction, previous research suggests that formal speech contexts, such as read speech, may be less sensitive to speech impairments in people with PD (Bunton, 2008; Ho et al., 1999; Sidtis et al., 2012). While this study attempted to combine acoustic and perceptual measures subject to these drawbacks with person-centered device preference decisions, future studies should consider comparisons of speech production in more ecologically valid contexts, such as in the home and during conversation. Future planned analyses from this study will report on spontaneous speech production from the picture description task not presented here.

The familiar listener intelligibility task was subject to task-specific limitations. For example, the CP participants were required to repeat back what they heard their partner say, which requires attentional and working memory cognitive resources (Poll et al., 2016). Sentence repetition ability also declines with healthy aging (Schum & Sivan, 1997). While precautions were put in place to limit these loads (short sentences, word order/contractions not penalized in the intelligibility metric), it is not possible to definitively know the effect individual differences may have had. Future studies would also perhaps benefit from the inclusion of practice sessions for CP participants.

On the other hand, the naive listener task was lengthy (approximately 10 hr over four to five sessions). It is possible that the naive listeners' attention and fatigue varied across this task. Multiple sessions, a self-paced schedule, and frequent breaks as needed were employed to reduce listener fatigue. It is also possible that listeners became more proficient at the task over time as they adapted to hearing dysarthric speech (Lansford et al., 2019; McAuliffe et al., 2017). Perceptual learning is an acknowledged drawback of studies of this nature (McAuliffe et al., 2017). Full randomization of the recordings was implemented to attempt to account for this.

Another limitation related to the naive listener task is the small number of listeners used ($n = 4$). However, other similar studies of disordered speech perception have used comparably small listener groups (Adams et al., 2008; Dromey, 2003; Dykstra et al., 2012, 2015; Moreau et al., 2011; Tanaka et al., 2016; Tripoliti et al., 2014). Furthermore, a recent study found that reliability between different intelligibility metrics (i.e., transcription and visual analog scale ratings) were highly related even with only one or two listeners, indicating that larger numbers of listeners do not necessarily contribute to more reliable estimates (Abur et al., 2019). Inter- and intrarater reliability was found to be excellent, which further suggests that the small listener group in this case was mitigated.

### *Clinical Implications and Limitations*

While this study points to the promising potential of speech amplification devices for people with hypophonia, a number of clinical limitations warrant further discussion and caution. Limitations should also be considered in the context of the American Speech-Language-Hearing Association's (ASHA) discussion of SLPs' roles and responsibilities in the screening, assessment, and treatment of potential candidates for augmentative devices (ASHA, 2020). First, to reiterate a methodological limitation, this study was conducted on a small, heterogeneous group of individuals with hypophonia and thus cannot provide recommendations regarding the impact of individual characteristics including (but not limited to) hearing status, cognition, and presence of additional speech symptoms. Further research with a larger participant group is needed to have a better understanding of the precise factors that drive device preferences and device usage in order to make appropriate and functional comprehensive clinical recommendations.

Another limitation is related to the time period of testing in this study. Participants were given approximately 1 week with each device. While this trial period represents a relative strength, 1 week is likely not a sufficiently representative time period to truly gauge whether a device will be effective long term. Future studies should consider longer trial periods (e.g., multiple weeks or even months) that facilitate and encourage more frequent device use in more diverse settings and situations. Another timing consideration is related to long-term follow-ups. While the majority of participants who completed this study (13/19) did choose to take a device, it is not yet known to what extent these individuals actively used the device at home

after the study. In a clinical setting, SLPs who prescribe amplification devices should do so with caution and consider routine check-ins over a period of several months to a year or more in order to ensure devices continue to function as intended outside the context of the clinic and to identify any barriers that may impede effective use on a case-by-case basis. PD and parkinsonian disorders are progressive, and most individuals will experience declines in speech symptoms as time goes on, including hypophonia (Ash et al., 2017; Skodda et al., 2013). These changes in speech may happen in addition to cognitive and motor changes but are overall independent of them (Ash et al., 2017). Longer term follow-ups for individuals who seek speech amplification devices are thus especially important, as symptom severity may evolve in a way that impacts device effectiveness over time. These considerations are consistent with ASHA's recommendations that SLPs who prescribe augmentative devices such as amplifiers "document progress" and "determine appropriate AAC modifications…if indicated" (ASHA, 2020).

Another major finding of this study was the potential effectiveness of a personal communication system, which has not been previously investigated for use in hypophonia. This too warrants careful consideration, as effective use additionally relies on the active participation of one or more CPs who would wear the headphones. While the communication system outranked the other, more traditional speech amplification devices in its ability to overcome background noise (SNR) and transmit a spoken message (intelligibility), it was only selected by three participants as their final preferred device. Having at least 1 week with each device separately likely allowed the participants to identify other factors that either increased or decreased their preference for each of the different devices, as suggested by Figure 4. Clinical recommendations surrounding any form of augmentative device, but especially one that requires compliance of another CP, should heed ASHA's recommendations regarding consideration of counseling and involvement of family members in order to prevent device abandonment (ASHA, 2020).

In summary, from a clinical implementation standpoint, this study points to the potential benefits of speech amplification devices and personal communication systems but does not provide evidence on the impact of individual characteristics on device preference or usage, nor does it provide indications of long-term adoption of devices. Clinically, in addition to evidenced-based recommendations pending future research, SLPs should, whenever possible, consider the benefits of home trial periods that allow clients to gain sufficient experience with devices to make informed decisions about the kind that is best for them.

## Summary

Overall, this study points to the potential for three distinct types of speech amplification devices for individuals with hypophonia, including a personal communication system never before tested for this population. The majority (13/19) of individuals enrolled in the study chose to continue using a device permanently after the study period ended. While an objective device performance hierarchy emerged, user preferences varied extensively. Each of the three devices was selected as a preferred device by at least three participants, pointing to both the potential for improving the availability of device options for this population and for a need for further research in this area.

## Author Contributions

**Scott G. Adams:** Conceptualization (Lead), Data curation (Supporting), Formal analysis (Supporting), Investigation (Supporting), Methodology (Equal), Resources (Lead), Writing - Review & Editing (Supporting). **Allyson Page:** Investigation (Supporting), Writing - Review & Editing (Supporting). **Daryn Cushnie-Sparrow:** Data curation (Supporting), Investigation (Supporting), Writing - Review & Editing (Supporting). **Mandar Jog:** Data curation (Supporting), Resources (Supporting).

## Acknowledgments

This study was supported by Graduate Student Research Funding awarded to Thea Knowles by the Parkinson Society Southwestern Ontario. The authors thank the participants for their time and contributions to this work.

## References

**Abur, D., Enos, N. M., & Stepp, C. E.** (2019). Visual analog scale ratings and orthographic transcription measures of sentence intelligibility in Parkinson's disease with variable listener exposure. *American Journal of Speech-Language Pathology, 28*(3), 1222–1232. https://doi.org/10.1044/2019_AJSLP-18-0275

**Adams, S. G., & Dykstra, A.** (2009). Hypokinetic dysarthria. In M. R. McNeil (Ed.), *Clinical management of sensorimotor speech disorders*. Thieme.

**Adams, S. G., Dykstra, A., Jenkins, M., & Jog, M.** (2008). Speech-to-noise levels and conversational intelligibility in hypophonia and Parkinson's disease. *Journal of Medical Speech-Language Pathology, 16*(4), 165–172.

**Adams, S. G., Haralabous, O., Dykstra, A., Abrams, K., & Jog, M.** (2005). Effects of multi-talker background noise on the intensity of spoken sentences in Parkinson's disease. *Canadian Acoustics, 33*(3), 94–95.

**Adams, S. G., & Lang, A. E.** (1992). Can the Lombard effect be used to improve low voice intensity in Parkinson's disease. *European Journal of Disorders of Communication, 27*(2), 121–127. https://doi.org/10.3109/13682829209012034

**Adams, S. G., Winnell, J., & 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), 1–9.

**American Speech-Language-Hearing Association.** (2020, May 21). *Augmentative and alternative communication* [Professional issues]. https://www.asha.org/Practice-Portal/Professional-Issues/Augmentative-and-Alternative-Communication/

**Andreetta, M. D., Adams, S. G., Dykstra, A. D., & Jog, M.** (2016). Evaluation of speech amplification devices in Parkinson's

disease. *American Journal of Speech-Language Pathology, 25*(1), 29–45. https://doi.org/10.1044/2015_ajslp-15-0008

**Ash, S., Jester, C., York, C., Kofman, O. L., Langey, R., Halpin, A., Firn, K., Perez, S. D., Chahine, L., Spindler, M., Dahodwala, N., Irwin, D. J., McMillan, C., Weintraub, D., & Grossman, M.** (2017). Longitudinal decline in speech production in Parkinson's disease spectrum disorders. *Brain and Language, 171,* 42–51. https://doi.org/10.1016/j.bandl.2017.05.001

**Bates, D., Mächler, M., Bolker, B., & Walker, S.** (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

**Bertrand, A. N.** (2009). *Speech-language pathologists decision making pertaining to speech amplification devices: A survey* [Master's thesis]. Texas A&M University-Kingsville.

**Boersma, P., & Weenink, D**. (2011). *Praat: Doing phonetics by computer* (Version 5.3) [Computer program]. http://www.praat.org/

**Bunton, K.** (2008). Speech versus nonspeech: Different tasks, different neural organization. *Seminars in Speech and Language, 29*(4), 267–275. https://doi.org/10.1055/s-0028-1103390

**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. https://doi.org/10.1016/j.jcomdis.2014.08.001

**Coates, C., & Bakheit, A. M.** (1997). The prevalence of verbal communication disability in patients with Parkinson's disease. *Disability and Rehabilitation, 19*(3), 104–107. https://doi.org/10.3109/09638289709166834

**Crandell, C., Charlton, M., Kinder, M., & Kreisman, B.** (2001). Effects of portable sound field FM systems on speech perception in noise. *Journal of Educational Audiology, 9,* 8–12.

**Dalrymple-Alford, J. C., MacAskill, M. R., Nakas, C. T., Livingston, L., Graham, C., Crucian, G. P., Melzer, T. R., Kirwan, J., Keenan, R., Wells, S., Porter, R. J., Watts, R., & Anderson, T. J.** (2010). The MoCA well-suited screen for cognitive impairment in Parkinson disease. *Neurology, 75*(19), 1717–1725. https://doi.org/10.1212/WNL.0b013e3181fc29c9

**Darley, F. L., Aronson, A. E., & Brown, J. R.** (1975). *Motor speech disorders* (pp. 2–297). Saunders.

**DeBonis, D. A., & Donohue, C. L.** (2008). *Survey of audiology: Fundamental for audiologists and health professionals* (2nd ed.). Pearson Education.

**Donovan, N. J., Kendall, D. L., Young, M. E., & Rosenbek, J. C.** (2008). The communicative effectiveness survey: Preliminary evidence of construct validity. *American Journal of Speech-Language Pathology, 17*(4), 335–347. https://doi.org/10.1044/1058-0360(2008/07-0010)

**Dromey, C.** (2003). Spectral measures and perceptual ratings of hypokinetic dysarthria. *Journal of Medical Speech-Language Pathology, 11*(2), 85–94.

**Dykstra, A. D., Adams, S. G., & Jog, M.** (2012). The effect of background noise on the speech intensity of individuals with hypophonia associated with Parkinson's disease. *Journal of Medical Speech-Language Pathology, 20*(3), 19–30.

**Dykstra, A. D., Adams, S. G., & Jog, M.** (2015). Examining the relationship between speech intensity and self-rated communicative effectiveness in individuals with Parkinson's disease and hypophonia. *Journal of Communication Disorders, 56,* 103–112. https://doi.org/10.1016/j.jcomdis.2015.06.012

**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's disease. *American Journal of Speech-Language Pathology, 6*(2), 85–94. https://doi.org/10.1044/1058-0360.0602.85

**Gamboa, J., Jiménez-Jiménez, F. J., Nieto, A., Montojo, J., Ortı -Pareja, M., Molina, J. A., García-Albea, E., Cobeta, I., & Cobeta, I.** (1997). Acoustic voice analysis in patients with Parkinson's disease treated with dopaminergic drugs. *Journal of Voice, 11*(3), 314–320. https://doi.org/10.1016/s0892-1997(97)80010-0

**Greene, M. C., Watson, B., Gay, P., & Townsend, D.** (1972). A therapeutic speech amplifier and its use in speech therapy. *The Journal of Laryngology & Otology, 86*(6), 595–605. https://doi.org/10.1017/S0022215100075629

**Gustafsson, J. K., Södersten, M., Ternström, S., & Schalling, E.** (2019). Voice use in daily life studied with a portable voice accumulator in individuals with Parkinson's disease and matched healthy controls. *Journal of Speech, Language, and Hearing Research, 62*(12), 4324–4334. https://doi.org/10.1044/2019_JSLHR-19-00037

**Harkins, J., & Tucker, P.** (2007). An Internet survey of individuals with hearing loss regarding assistive listening devices. *Trends in Amplification, 11*(2), 91–100. https://doi.org/10.1177/1084713807301322

**Ho, A. K., Bradshaw, J. L., & Iansek, T.** (2000). Volume perception in parkinsonian speech. *Movement Disorders, 15*(6), 1125–1131. https://doi.org/10.1002/1531-8257(200011)15:6<1125::AID-MDS1010>3.0.CO;2-R

**Ho, A. K., Bradshaw, J. L., Iansek, R., & Alfredson, R.** (1999). Speech volume regulation in Parkinson's disease: Effects of implicit cues and explicit instructions. *Neuropsychologia, 37*(13), 1453–1460. https://doi.org/10.1016/S0028-3932(99)00067-6

**Ho, A. K., Iansek, R., Marigliani, C., Bradshaw, J. L., & Gates, S.** (1998). Speech impairment in a large sample of patients with Parkinson'ss disease. *Behavioural Neurology, 11*(3), 131–137. https://doi.org/10.1155/1999/327643

**Kent, R. D., Kent, J. F., Rosenbek, J. C., Vorperian, H. K., & Weismer, G.** (1997). A speaking task analysis of the dysarthria in cerebellar disease. *Folia Phoniatrica et Logopaedica, 49*(2), 63–82. https://doi.org/10.1159/000266440

**Kent, R. D., Weismer, G., Kent, J. F., & Rosenbek, J. C.** (1989). Toward phonetic intelligibility testing in dysarthria. *The Journal of Speech and Hearing Disorders, 54*(4), 482–499. https://doi.org/10.1044/jshd.5404.482

**Koo, T. K., & Li, M. Y.** (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

**Kryter, K. D.** (1994). *The handbook of hearing and the effects of noise*. Academic Press.

**Lane, H., & Tranel, B.** (1971). The Lombard sign and the role of hearing in speech. *Journal of Speech and Hearing Research, 14*(4), 677–709. https://doi.org/10.1044/jshr.1404.677

**Lansford, K. L., Borrie, S. A., & Barrett, T. S.** (2019). Regularity matters: Unpredictable speech degradation inhibits adaptation to dysarthric speech. *Journal of Speech, Language, and Hearing Research, 62*(12), 4282–4290. https://doi.org/10.1044/2019_JSLHR-19-00055

**Laplante-Lévesque, A., Hickson, L., & Worrall, L.** (2010). Rehabilitation of older adults with hearing impairment: A critical review. *Journal of Aging and Health, 22*(2), 143–153. https://doi.org/10.1177/0898264309352731

**Lenth, R**. (2018). *Emmeans: Estimated marginal means, aka least-squares means*. https://CRAN.R-project.org/package=emmeans

**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.

*The Journal of Speech and Hearing Disorders, 43*(1), 47–57. https://doi.org/10.1044/jshd.4301.47

**Ludlow, C. L., & Bassich, C. J.** (1984). Relationships between perceptual ratings and acoustic measures of hypokinetic speech. In M. R. McNcil, J. C. Rosenbek, & A. E. Aronson (Eds.), *The dysarthrias: Physiology, acoustics, perception, management* (pp. 163–195). College-Hill Press.

**McAuliffe, M. J., Fletcher, A. R., Kerr, S. E., O'Beirne, G. A., & Anderson, T.** (2017). Effect of dysarthria type, speaking condition, and listener age on speech intelligibility. *American Journal of Speech-Language Pathology, 26*(1), 113–123. https://doi.org/10.1044/2016_AJSLP-15-0182

**Miller, N.** (2012). Speech, voice and language in Parkinson's disease: Changes and interventions. *Neurodegenerative Disease Management, 2*(3), 279–289. https://doi.org/10.2217/nmt.12.15

**Miller, N.** (2013). Measuring up to speech intelligibility. *International Journal of Language & Communication Disorders, 48*(6), 601–612. https://doi.org/10.1111/1460-6984.12061

**Miller, N., Allcock, L., Jones, D., Noble, E., Hildreth, A. J., & Burn, D. J.** (2007). Prevalence and pattern of perceived intelligibility changes in Parkinson's disease. *Journal of Neurology, Neurosurgery, & Psychiatry, 78*(11), 1188–1190. https://doi.org/10.1136/jnnp.2006.110171

**Moreau, C., Pennel-Ployart, O., Pinto, S., Plachez, A., Annic, A., Viallet, F., Destée, A., & Defebvre, L.** (2011). Modulation of dysarthropneumophonia by low-frequency STN DBS in advanced Parkinson's disease. *Movement Disorders, 26*(4), 659–663. https://doi.org/10.1002/mds.23538

**Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H.** (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society, 53*(4), 695–699. https://doi.org/10.1111/j.1532-5415.2005.53221.x

**Nábělek, A., Donahue, A., & Letowski, T.** (1986). Comparison of amplification systems in a classroom. *Journal of Rehabilitation Research and Development, 23*(1), 41–52.

**Olson, M., Lockhart, T. E., & Lieberman, A.** (2019). Motor learning deficits in Parkinson's disease (PD) and their effect on training response in gait and balance: A narrative review. *Frontiers in Neurology, 10,* 62. https://doi.org/10.3389/fneur.2019.00062

**Pick, H. L. J., Siegel, G. M., Fox, P. W., Garber, S. R., & Kearney, J. K.** (1989). Inhibiting the Lombard effect. *The Journal of the Acoustical Society of America, 85*(2), 894–900. https://doi.org/10.1121/1.397561

**Poll, G. H., Miller, C. A., & van Hell, J. G.** (2016). Sentence repetition accuracy in adults with developmental language impairment: Interactions of participant capacities and sentence structures. *Journal of Speech, Language, and Hearing Research, 59*(2), 302–316. https://doi.org/10.1044/2015_JSLHR-L-15-0020

**Ramig, L. O., Fox, C., & Sapir, S.** (2004). Parkinson's disease: Speech and voice disorders and their treatment with the Lee Silverman voice treatment. *Seminars in speech and language*, *25*(2), 169–180. Thieme. https://doi.org/10.1055/s-2004-825653

**R Core Team.** (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

**Schum, R. L., & Sivan, A. B.** (1997). Verbal abilities in healthy elderly adults. *Applied Neuropsychology, 4*(2), 130–134. https://doi.org/10.1207/s15324826an0402_6

**Scott, S., & Caird, F.** (1983). Speech therapy for Parkinson's disease. *Journal of Neurology, Neurosurgery, & Psychiatry, 46*(2), 140–144. https://doi.org/10.1136/jnnp.46.2.140

**Sidtis, D., Cameron, K., Bonura, L., & Sidtis, J.** (2012). Speech intelligibility by listening in Parkinson speech with and without deep brain stimulation: Task effects. *Journal of Neurolinguistics, 25*(2), 121–132. https://doi.org/10.1016/j.jneuroling.2011.08.004

**Skodda, S., Grönheit, W., Mancinelli, N., & Schlegel, U.** (2013). Progression of voice and speech impairment in the course of Parkinson's disease: A longitudinal study. *Parkinson's Disease, 2013,* 389195. https://doi.org/10.1155/2013/389195

**Spencer, K. A., Yorkston, K. M., & 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), xxxix–xxxix.

**Stathopoulos, E. T., Huber, J. E., Richardson, K., Kamphaus, J., DeCicco, D., Darling, M., Fulcher, K., & Sussman, J. E.** (2014). Increased vocal intensity due to the Lombard effect in speakers with Parkinson's disease: Simultaneous laryngeal and respiratory strategies. *Journal of Communication Disorders, 48,* 1–17. https://doi.org/10.1016/j.jcomdis.2013.12.001

**Tanaka, Y., Tsuboi, T., Watanabe, H., Kajita, Y., Nakatsubo, D., Fujimoto, Y., Ohdake, R., Ito, M., Atsuta, N., Yamamoto, M., Wakabayashi, T., Katsuno, M., & Sobue, G.** (2016). Articulation features of Parkinson's disease patients with subthalamic nucleus deep brain stimulation. *Journal of Parkinson's Disease, 6*(4), 811–819. https://doi.org/10.3233/jpd-160838

**Tomlinson, C. L., Stowe, R., Patel, S., Rick, C., Gray, R., & Clarke, C. E.** (2010). Systematic review of levodopa dose equivalency reporting in Parkinson's disease. *Movement Disorders, 25*(15), 2649–2653. https://doi.org/10.1002/mds.23429

**Tripoliti, E., Limousin, P., Foltynie, T., Candelario, J., Aviles-Olmos, I., Hariz, M. I., & Zrinzo, L.** (2014). Predictive factors of speech intelligibility following subthalamic nucleus stimulation in consecutive patients with Parkinson's disease. *Movement Disorders, 29*(4), 532–538. https://doi.org/10.1002/mds.25816

**Tröster, A. I., Jankovic, J., Tagliati, M., Peichel, D., & Okun, M. S.** (2017). Neuropsychological outcomes from constant current deep brain stimulation for Parkinson's disease. *Movement Disorders, 32*(3), 433–440. https://doi.org/10.1002/mds.26827

**Weismer, G.** (2008). Speech intelligibility. In M. J. Ball, M. R. Perkins, N. Muller, & S. Howard (Eds.), *The handbook of clinical linguistics* (pp. 568–582). Blackwell. https://doi.org/10.1002/9781444301007.ch35

**Yorkston, K. M., Beukelman, D. R., & Tice, R.** (1996). *Sentence Intelligibility Test* [Computer software]. Institute for Rehabilitation Science and Engineering at Madonna Rehabilitation Hospital.

**Appendix**

User Experience Questionnaire

## User Experience Questionnaire

Unique ID: _____ Device: _____ Date: _____

Please evaluate your experience with each communication device. Please indicate your answers to the following questions by placing a cross ("X") on the line where it best represents your answer.
A cross towards the left side indicates a **poorer** rating of the speech amplification device, while a cross towards the right side indicates a **better** rating of the device.

1. ***Physical Comfort:*** How comfortable is this device to wear?

*Uncomfortable* _____ *Comfortable*

2. ***Visual Presentation:*** How acceptable is this device to wear in public?

*Unacceptable* _____ *Acceptable*

3. ***Sound Quality:*** What is the sound quality of the speech output from the device?

*Poor sound quality* _____ *Good sound quality*

4. ***Amplification Power:*** How well does the output from the device overcome the background noise?

*Poor amplification* _____ *Good amplification*

5. ***Overall Preference:*** Overall, is this a device that you would prefer to use?

*Low preference* _____ *High preference*
(prefer not to use) (prefer to use)