ASHA American Speech-Language-Hearing Association

AJSLP

Research Article

Intelligibility Across a Reading Passage: The Effect of Dysarthria and Cued Speaking Styles

Frits van Brenk,[a,b] Kaila Stipancic,[a] Alexander Kain,[c] and Kris Tjaden[a]

[a] Department of Communicative Disorders and Sciences, University at Buffalo, NY [b] Utrecht Institute of Linguistics OTS, Utrecht University, the Netherlands [c] Department of Pediatrics, Oregon Health & Science University, Portland

ARTICLE INFO

Article History:
Received May 24, 2021
Revision received August 10, 2021
Accepted October 8, 2021

Editor-in-Chief: Katherine C. Hustad
Editor: Jessica E. Huber

https://doi.org/10.1044/2021_AJSLP-21-00151

ABSTRACT

Objective: Reading a passage out loud is a commonly used task in the perceptual assessment of dysarthria. The extent to which perceptual characteristics remain unchanged or stable over the time course of a passage is largely unknown. This study investigated crowdsourced visual analogue scale (VAS) judgments of intelligibility across a reading passage as a function of cued speaking styles commonly used in treatment to maximize intelligibility.
Patients and Method: The Hunter passage was read aloud in habitual, slow, loud, and clear speaking styles by 16 speakers with Parkinson's disease (PD), 30 speakers with multiple sclerosis (MS), and 32 control speakers. VAS judgments of intelligibility from three fragments representing the beginning, middle, and end of the reading passage were obtained from 540 crowdsourced online listeners.
Results: Overall passage intelligibility was reduced for the two clinical groups relative to the control group. All speaker groups exhibited intelligibility variation across the reading passage, with trends of increased intelligibility toward the end of the reading passage. For control speakers and speakers with PD, patterns of intelligibility variation across passage reading did not differ with speaking style. For the MS group, intelligibility variation across the passage was dependent on speaking style.
Conclusions: The presence of intelligibility variation within a reading passage warrants careful selection of speech materials in research and clinical practice. Results further indicate that the crowdsourced VAS rating paradigm is useful to document intelligibility in a reading passage for different cued speaking styles commonly used in treatment for dysarthria.

A primary goal of a motor speech evaluation is establishing the presence and severity of impairment (Duffy, 2019). A comprehensive motor speech evaluation may also identify specific components (i.e., subsystems, features) of speech production that are affected. Once identified, aspects of deviant speech production become a primary target for behavioral treatment, with an overall aim of maximizing speech intelligibility (Yorkston, 2010). An important and widely used task in the battery of speech tasks used for the assessment of motor speech disorders is the reading aloud of a written passage. Passage reading is a contextual speech task used to obtain samples of connected speech, and allows for holistic evaluation of the respiratory, phonatory, articulatory, resonatory, and prosodic aspects of speech production. Arguably, passage reading is a more ecologically valid task than tasks such as word or sentence reading, for the evaluation of speech intelligibility (Weismer et al., 2001). Compared to reading single words or sentences, passage reading is considered to more closely approximate the linguistic and cognitive demands of spontaneous speech, while still providing the structure of a controlled and repeatable speaking task (Duffy, 2019; Patel et al., 2013).

Passage Reading in Research

A variety of reading passages have found their way into research and clinical practice. Among the most

Correspondence to Frits van Brenk: f.j.vanbrenk@uu.nl. ***Disclosure:*** *The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.*

prominently used are the Rainbow Passage (Fairbanks, 1960), the Grandfather Passage (Van Riper, 1963), the Hunter Script (Crystal & House, 1982), and the Caterpillar Story (Patel et al., 2013). A number of studies have employed reading passages to assess acoustic and perceptual aspects of speech produced by individuals with dysarthria. The majority of these studies predominantly considered the passage as a single "unit" and, thus, examined speech features within the passage as a whole (i.e., rather than examining variation in speech features across the course of the passage). For example, Nishio and Niimi (2001, 2006) analyzed speaking rate and articulation rate in the 223 morae Japanese standard passage "The North Wind and the Sun" read by speakers with various types of dysarthria and healthy control speakers. Averaged across the entire passage, findings of slower speaking rates were reported for all speakers with dysarthria compared to control speakers. Slower articulation rates also were found for all speakers with dysarthria except those with flaccid and hypokinetic dysarthria. To evaluate the effect of different cueing techniques on intelligibility and speech production, as inferred from the acoustic signal, Tjaden and Wilding (2004) employed the John Passage. Speakers with dysarthria due to Parkinson's disease (PD) and multiple sclerosis (MS), and neurotypical speakers read the passage aloud in habitual, loud, and slow speaking styles. Acoustic measures of interest included articulation rate, sound pressure level (SPL), vowel space area, first moment difference measures, and second formant (F2) trajectory characteristics. Listeners estimated intelligibility for an excerpt taken from the middle of the reading passage using a free-modulus magnitude estimation paradigm. Control speakers and speakers with MS demonstrated larger vowel space areas in the slow speaking style compared to the habitual speaking style, while such an effect was not found for the speakers with PD. Speech intelligibility was found to be higher in the loud versus habitual and slow speaking styles for the speakers with PD, whereas between-style differences were absent for the other two groups, indicating notable disparities in acoustic and perceptual findings across speaker groups. The Grandfather Passage was employed by Eijk et al. (2020), who investigated the effects of word frequency and probability on word duration in speakers with various subtypes of dysarthria. Across the reading passage, measurements of word duration, frequency, transitional word probability, and visual analog scale (VAS) ratings of speech precision were obtained. Results indicated that with increasing dysarthria severity, word durations were less affected by probabilistic variables. The authors suggested that this result might be due to reductions in the control and execution of muscle movement exhibited by speakers with more severe dysarthria (Eijk et al., 2020). Taken together, these studies demonstrate the utility of reading passages for investigating a variety of speech production and perceptual characteristics in dysarthria.

In the studies discussed above, the manner in which read passages were analyzed implicitly assumes that speakers maintain consistency in speech production across the length of a reading passage, such that a measure derived from the beginning of a passage would be comparable to the same measure derived from the middle or end of the passage. Given that these studies included measures of either single passage reading excerpts or the reading passage as a whole, such an approach does not account for within-task change or variation in speech production along the course of a reading passage, something that may be expected to occur both in neurotypical speakers and speakers with dysarthria (Kent et al., 2000; Smiljanić & Bradlow, 2008). In particular, factors including neuromuscular fatigue, planning, and coordination difficulties in speech formulation and respiratory support, and self-monitoring deficits may differentially impact intelligibility as they vary across the course of a reading passage (Huber & Darling, 2011; Pinto et al., 2017; Spencer & Rogers, 2005; Yorkston, 2010).

Only a few dysarthria studies have investigated variation in speech production during passage reading. Skodda and Schlegel (2008) analyzed articulation rate and speech-to-pause ratios on the first and last sentence of a 170-syllable standard reading passage produced by speakers with PD and control speakers. The results indicated that both speaker groups produced an accelerated speech rate at the end of the reading passage, but speakers with PD increased their rate in the last sentence to a larger extent compared to the control speakers, indicating an impaired speech timing organization unfolding toward the end of the reading passage. Kuo and Tjaden (2016) employed the John Passage to investigate variation in acoustic measures of global speech timing, vocal intensity, and segmental articulation. Speakers with MS and PD, and control speakers produced the reading passage in habitual, slow, and loud speaking styles. Acoustic measures were obtained from three segments located at the beginning, middle, and end of the reading passage. Acoustic measures showing between-fragment, between-task, and between-group variation included F2 frequency interquartile range (IQR), articulation rate, and pause duration. For F2 IQR, all groups showed a decrease in the middle segment compared to the beginning and end segments for the loud speaking style, while the slow and habitual speaking styles showed a decrease in the end segments compared to the beginning and middle segments. For articulation rate, the end segment for all groups was associated with increased rates compared to the beginning segment across speaking styles. In addition, a Group × Speaking Style interaction indicated that for the slow speaking style, healthy speakers and speakers with MS decreased articulation rate to a larger extent compared to speakers with PD. For pause duration, a Significant Group × Segment interaction was found, indicating that for healthy speakers, mean pause durations increased throughout the passage. In contrast, for speakers with MS, pause durations were shortest for the middle segment, whereas for speakers with PD, pause durations were shortest for the beginning segment, with comparable durations for the middle and end segments. Overall, results indicate significant within-task variation for a variety of acoustic measures during passage reading. The studies by Skodda and Schlegel (2008) and Kuo and Tjaden (2016) are of particular interest because they revealed important findings that would have been missed had the reading passage been considered as a whole.

The studies discussed above indicate the presence of within-task variation in speech acoustic characteristics during passage reading. It is feasible that such within-task variation is also present in perceptual measures of intelligibility, and, moreover, the amount of within-task intelligibility variation may be influenced by speaking style. Global treatment techniques incorporating nonhabitual speaking styles play an increasingly important role in maximizing intelligibility in speakers with dysarthria and are aimed at producing holistic improvements of the speech subsystems (Duffy, 2019; Lee et al., 2014; Yorkston et al., 2007). Treatment techniques commonly used in clinical practice due to ease of trainability and high stimulability are rate reduction, increased vocal intensity, and clear speech (Duffy, 2019; Freed, 2018; Yorkston, 2010). The impact of these techniques on perceptual and acoustic aspects of speech production are considered in the following section.

Perceptual and Acoustic Findings of Global Treatment Techniques

Rate reduction is a commonly practiced behavioral management technique for improving intelligibility in dysarthria (Yorkston et al., 2007). Mechanisms that have been proposed to explain the improved intelligibility associated with rate reduction include increased articulatory precision by enabling the achievement of more extreme vocal tract configurations (Blanchet & Snyder, 2010; Yorkston et al., 2007), allowing speakers to simultaneously manipulate multiple speech subsystems, and enabling speakers to produce more appropriate breath group units (Hardcastle & Tjaden, 2008; Yorkston, 2010). Listeners may also take advantage of a slowed speech rate to improve accuracy of lexical segmentation by using the extra time to better decode speech (Liss, 2007; McAuliffe et al., 2014). Rate control as a behavioral management technique has shown treatment effectiveness across many neurological diagnoses and clinical dysarthria subtypes (Duffy, 2019; Yorkston et al., 2007). However, the effect of rate reduction on intelligibility has been found to vary, even within relatively homogeneous populations of speakers with the same neurological diagnosis and pathophysiology, or comparable severity.

Increasing vocal intensity is another commonly used treatment technique in the management of dysarthria. A variety of training programs are aimed at increasing vocal loudness by increasing respiratory and phonatory effort. Increased vocal intensity has been associated with increased SPL, increased fundamental frequency range, and improved segmental productions (Tjaden & Martel-Sauvageau, 2017; Yorkston et al., 2007). Improvements in intelligibility have been associated with increased vocal loudness across a number of assessment tasks and patient populations (Neel, 2008; Stipancic et al., 2016; Tjaden et al., 2014; Yorkston et al., 2007).

A clear speaking style is characterized by exaggerated or hyperarticulated speech production. Clear speech has been associated with a number of segmental and suprasegmental acoustic adjustments including a reduced rate of speech, increased vocal intensity, increased tense and lax vowel space areas, and steeper F2 slopes (Ferguson, 2004; Kuo et al., 2014; Lam & Tjaden, 2016; Smiljanić & Bradlow, 2009). These acoustic adjustments have been associated with increased intelligibility, as measured by transcription accuracy and VAS scores (Cannito et al., 2011; Kuo et al., 2014; Stipancic et al., 2016; Tjaden et al., 2014).

Although the three behavioral strategies discussed above have been shown to produce some positive effects on speech intelligibility in speakers with dysarthria, the durability of such an intelligibility benefit across longer, more extended speech material has not been thoroughly and systematically studied (e.g., see the work of Tjaden & Wilding, 2004). This knowledge is important for strengthening the validity of passage reading in evaluating global treatment techniques. Improved understanding of intelligibility variation as a function of particular behavioral strategies over an extended connected speech task like passage reading could have implications for the selection of treatment strategies used in clinical practice. For example, both clear and loud speaking styles may improve intelligibility for a given speaker. However, one style may yield a decreasing benefit over the length of speech output due to putative difficulties in maintaining the production alterations necessitated by the style, while the other strategy may yield preserved intelligibility benefits. Especially for speakers with dysarthria with impaired neuromuscular control, a given speaking strategy may be feasible and beneficial for the production of a single sentence, but a connected speech task may present challenges for sustaining the new mode of production. As within-task variation in speech production may be evident in a passage read in a habitual or typical speaking style (Kuo & Tjaden, 2016; Skodda & Schlegel, 2008), it is important to establish to what extent such variation is present in speaking styles commonly used in therapy. Such knowledge may aid in the optimal use of passage reading for dysarthria evaluation and treatment selection.

A further clinical challenge in the assessment of intelligibility of dysarthria is the use of and access to reliable measurement tools (Abur et al., 2019). While orthographic transcription is arguably the most ecologically valid method of assessing intelligibility (Duffy, 2019; Miller, 2013), it is time and labor consuming, requiring significant effort on the part of the listener and the scorer, which may not always be feasible. Listener familiarity with both speakers and stimuli is another critical element to consider when employing this methodology. Studies have shown that listeners previously exposed to the speaker or the type of dysarthria show higher transcription performance compared to unfamiliar listeners (Beukelman & Yorkston, 1980; Hustad & Cahill, 2003; Utianski et al., 2011). VAS judgments of intelligibility are an alternative to time-intensive orthographic transcription (Miller, 2013; Stipancic et al., 2016). The VAS paradigm involves rating a speech stimulus on a particular perceptual dimension such as intelligibility or naturalness using a scale of predefined length. Compared to orthographic transcription, VAS is less time consuming for both listener and scorer. In addition, previous research has shown that VAS scores and transcription scores are strongly correlated when employing naïve listeners (Abur et al., 2019; Stipancic et al., 2016). Crowdsourcing websites offer a platform to recruit large numbers of naïve listeners as well as the opportunity to sample the broader population of possible listeners compared to laboratory-based recruitment practices. Previous studies have demonstrated the feasibility of using crowdsourced transcription paradigms to quantify intelligibility in dysarthria (Lansford et al., 2016; Nightingale et al., 2020; van Brenk et al., 2021). However, the feasibility of using a crowdsourced VAS paradigm to assess intelligibility in dysarthria has not yet been demonstrated, although this knowledge would strengthen use of VAS as a measurement tool in both research and clinical practice.

Purpose

As suggested in the preceding review, prior acoustic studies have reported within-task variation across a reading passage. Moreover, the within-task acoustic variation has been found to depend on the neurological status of the speaker (Kuo & Tjaden, 2016; Nishio & Niimi, 2006; Skodda & Schlegel, 2008). This investigation extends this line of inquiry to intelligibility. Following previous research (e.g., Kuo & Tjaden, 2016), this study also evaluated intelligibility variation for a number of speaking styles commonly used therapeutically to maximize intelligibility in dysarthria. Capturing robust differences in acoustic measures across a reading passage attributable to speaking style in prior research has been challenging. It is, therefore, important to evaluate whether differences in variation across a reading passage as function of speaking style may be revealed using perceptual measures.

Against this background, this study evaluated scaled intelligibility over the course of a reading passage read in habitual, slow, loud, and clear speaking styles by speakers with dysarthria secondary to MS and PD as compared to neurologically healthy speakers. We hypothesized that a crowdsourced VAS paradigm would reveal intelligibility differences associated with cued speaking styles employed in clinical practice. We also hypothesized that the crowdsourced VAS paradigm would reveal intelligibility variation across three fragments of a reading passage (i.e., passage beginning, middle, and end) produced in habitual, slow, loud, and clear speaking styles by individuals with PD, MS, and control speakers. Comparing intelligibility variation across a reading passage for different speaking styles affords the opportunity to identify favorable speaking styles with respect to cueing and trainability, or ease of maintenance across a longer stretch of speech. Acoustic measures of duration and SPL were obtained to validate that speakers produced the cued slow, loud, and clear speaking styles. As lexical properties of speech stimuli may impact intelligibility (Chiu & Forrest, 2018; A. Fletcher & McAuliffe, 2021), word frequency and neighborhood density values also were calculated for each word in the three fragments. Findings have the potential to inform clinical use of reading passage materials and choice of speaking style in determining optimal behavioral strategies in the management of dysarthria, as well as to inform the use of crowdsourced VAS paradigms for research purposes.

Method

Speakers

The 78 speakers of interest for this study have been described in greater detail in Tjaden et al. (2014) and Stipancic et al. (2016). Control speakers (n = 32) were 10 men (25–70 years old, M = 56) and 22 women (27–77 years old, M = 57) who reported the absence of neurological disease. Speakers with PD (n = 16) were eight men (55–78 years, M = 67) and eight women (48–78 years, M = 69) who reported being diagnosed with idiopathic PD by their neurologist. Speakers with MS (n = 30) were 10 men (29–60 years, M = 51) and 20 women (27–66 years, M = 50) who reported a diagnosis of MS from their neurologist. All participants were native speakers of American English, had achieved at least a high school diploma, denied use of hearing aids, reported no other history of neurologic disease, and had cognitive ability adequate for the purposes of the experiment as indicated by score of at least 26 points on the Standardized Mini-Mental State Examination (Molloy, 1999). Pure-tone thresholds were obtained by an audiologist at the University Speech-Language and Hearing Clinic for the purpose of providing a global indication of auditory status. However, no speaker was excluded on the basis of pure-tone thresholds. Pure-tone thresholds averaged across 500, 1000, and 4000 Hz were 25 dB HL or better in both ears, except one speaker who displayed elevated average pure-tone threshold of 37 dB HL in the left ear.

Clinical metrics of sentence intelligibility and scaled estimates of speech severity for the Grandfather Passage (Duffy, 2019) and procedures to obtain these measures were reported in a previous study by Sussman and Tjaden (2012), and are summarized below to characterize participants' speech. Stimuli were presented in quiet through headphones at the same SPL at which they were naturally produced by the speakers. Sentence intelligibility scores were obtained using the Sentence Intelligibility Test (SIT; Yorkston et al., 1996) and scored by 42 naïve listeners. The mean sentence intelligibility for control speakers was 94% (SD = 2.7). Average sentence intelligibility was 93% (SD = 4.5) for speakers with MS and 85% (SD = 10) for speakers with PD. Speech severity, an operationally defined perceptual construct intended to assess speech naturalness and prosodic adequacy (Feenaughty et al., 2014; Kuo et al., 2014), was judged by 10 naïve and three expert speech-language pathologists who were blinded to the speakers' neurological diagnoses and identities. Listeners used a computerized VAS to judge speech severity, with scale end points of 0 (*no impairment*) and 1.0 (*severe impairment*). The mean scaled speech severity value for the Grandfather Passage was 0.18 (SD = .08) for control speakers. Mean scaled speech severity was 0.44 (SD = .25) for speakers with MS and 0.46 (SD = .21) for speakers with PD. In addition, Sussman and Tjaden (2012) reported anecdotal perceptual observations made by three expert listeners of the speech of the clinical groups. Speakers with MS had predominantly reduced segmental precision and some prosodic (e.g., slow speech rate, excess stress) and voice deficits (e.g., harshness, hoarseness). Speakers with PD had mostly reduced segmental precision and a breathy, monotonous voice. While sentence intelligibility was relatively spared in the speakers with MS and PD (e.g., high mean SIT scores: MS = 93%, PD = 85%), these speaker groups had a noticeable speech impairment, as reflected in the higher scaled speech severity scores relative to control speakers. The combination of the clinical measures of intelligibility, the scaled severity for the Grandfather Passage, and anecdotal perceptual judgments suggest mild dysarthria for many speakers (Yorkston, 2010).

Speech Production Task

Each speaker was audio-recorded while reading a modified version of the Hunter Script, a short 17-sentence story (Crystal & House, 1982). A small number of words were replaced or inserted to ensure an adequate number of phoneme occurrences for segmental acoustic analyses not of interest to this study. The modified Hunter Script consisted of 240 monosyllabic and 39 polysyllabic words, with a total of 328 syllables (see Appendix). The passage was read in habitual, slow, fast, loud, and clear speaking styles. Speaking styles used clinically to maximize intelligibility (i.e., slow, loud, clear) were of interest to this study. For the slow speaking style, participants were instructed to read the passage at a rate half as fast as their regular rate. Speakers were encouraged to stretch out words rather than solely insert pauses and to say each sentence on a single breath, compare McHenry (2003). This instruction was intended to discourage speakers from only using pauses to reduce speech rate, as only adjusting pause characteristics to reduce rate would likely not enhance intelligibility (van Brenk et al., 2021). For the loud speaking style, participants were asked to read the passage twice as loud as their regular talking voice. For the clear speaking style, clear speech instructions were modeled after other clear speech studies and were intended to maximize the likelihood that speakers would not only exaggerate articulation but would also increase vocal intensity and reduce rate (Smiljanić & Bradlow, 2009). Specifically, speakers were instructed to say each sentence twice as clearly as their typical speech. Speakers were told to exaggerate the movements of their mouth as how they might speak to someone in a noisy environment or to someone with a hearing loss. All participants first read the passage in their habitual speaking style to establish a baseline performance. Other speaking styles were implemented in a randomized order.

Audio recordings were collected using an AKG C410 head-mounted microphone with a constant mouth-microphone distance, positioned 10 cm and 45°–50° from the left oral angle. The acoustic signal was preamplified, low-pass filtered at 9.8 kHz, and sampled at 22 kHz. A 1000-Hz calibration tone was recorded to permit off-line measurement of SPL of the acoustic signal.

Stimulus Preparation and Dependent Variables

Three fragments were segmented from each of the Hunter Script recordings for presentation to listeners for judgment of scaled intelligibility. As shown in the Appendix, one fragment was located near the beginning of the passage, one near the middle of the passage, and one near the end of the passage. Fragments were carefully selected to avoid potential idiosyncrasies associated with initial and final sentences of a reading passage. In addition, fragments were selected to include a comparable number of words, coincided with sentence boundaries, and were separated by at least one sentence (c.f. Allison et al., 2019; Kuo & Tjaden, 2016). Fragment 1 was composed of 21 words and 21 syllables. Fragment 2 was composed of 18 words and 20 syllables. Fragment 3 was composed of 17 words and 20 syllables. To minimize sentence-level loudness differences, each of the stimuli were intensity normalized using the following procedure: Speech signal waveforms were first filtered with an A-weighted filter, and levels were calculated by averaging frame-by-frame root-mean-squared (rmsA) sample values. Each waveform was then multiplied by an appropriate gain factor so that the resulting waveforms all had the same average rmsA value (Kain et al., 2008). Speakers had mild dysarthria based on transcription scores obtained using the SIT (Yorkston et al., 1996), as reported in the work of Sussman and Tjaden (2012). To avoid ceiling effects (i.e., VAS ratings averaging 100% regardless of fragment) in the perceptual task described below, sentence variants were mixed with 10-talker babble sampled at 22 kHz and low-pass filtered at 11 kHz. Pilot testing with nine crowdsourced listeners evaluating a total of 216 fragments indicated that a signal-to-noise ratio (SNR) of 0 dB (c.f. Abur et al., 2019; Moya-Galé et al., 2018) minimized floor and/or ceiling effects and yielded a mean VAS score of 56.7 (SD = 30.5) demonstrating listeners' use of the full scale.

Previous research has shown that lexical properties of speech stimuli may impact intelligibility. Specifically, transcription accuracy for speech mixed with noise produced by individuals with dysarthria has been found to be higher for high-frequency words than low-frequency words, and for words with low phonological neighborhood density compared to words with high phonological neighborhood density (Chiu & Forrest, 2018; A. Fletcher & McAuliffe, 2021). It was, therefore, important to document the presence of any differences in lexical characteristics among fragments. Word frequency and neighborhood density values for each word in the three distinct operationally defined fragments were calculated using the Irvine Phonotactic Online Dictionary (Vaden et al., 2009).

To verify whether speakers successfully increased speech intensity during the loud speaking style, measures of SPL of the original nonintensity normalized fragments were performed with TF32 (Milenkovic, 2005). To verify whether speakers increased durations in the slow speaking style, measures of fragment durations were performed with Praat (Boersma & Weenink, 2020). Both SPL and durational measures were performed to document the implementation of a clear speaking style, as a reduction in rate and increased vocal intensity are expected to characterize clear speech (Smiljanić & Bradlow, 2009; Uchanski, 2005).

Listeners and Perceptual Methodology

A total of 415 adults (212 women, 194 men, nine not specified), 18–71 years of age (M = 31.4, SD = 10.6), judged intelligibility in the context of a crowdsourced experimental paradigm. All listeners self-reported to be native speakers of American English, living in the United States, and without a history of speech, language, or hearing problems. Extended demographic information of participating listeners is displayed in Table 1. The study was programmed and executed in jsPsych (De Leeuw, 2015) and hosted on Pavlovia.org (Peirce & MacAskill, 2018). Participants were recruited using the crowdsourcing website Prolific (prolific.co; Palan & Schitter, 2018).

Listeners were allowed to participate after fulfilling a number of prerequisites, including an 80% approval record for previously completed studies on the Prolific platform, and confirmed status of U.S. residence. Participants were limited to using a personal computer or laptop only, excluding the use of mobile devices or tablets. Participants were provided a brief description of the task including the requirement of wearing headphones or earphones and

Table 1. Demographic details of listener participants.

Demographic details	*N*	%
Gender		
Males	194	46.7
Females	212	51.1
Other/blank	9	2.2
Age		
≥ 50	31	7.5
40–49	46	11.1
30–39	133	32.0
≤ 29	205	49.4
Education completed		
Doctorate	13	3.1
Master's	61	14.7
Bachelor's	173	41.7
Associate	43	10.4
High school/GED	125	30.1
Region		
Midwest	80	19.3
Northeast	96	23.1
Southeast	119	28.7
Southwest	32	7.7
Rocky mountains	8	1.9
Pacific	72	17.3
Noncontiguous	4	1.0
Not specified	3	0.7
Race		
American Indian/Alaska Native	2	0.5
Asian	39	9.4
White	294	70.8
Black or African American	39	9.4
More than one race	31	7.5
Other/prefer not to say	10	2.4

being in a quiet environment when participating. After consenting to participate using the institutional review board–approved consent form, listeners were asked to complete a demographic questionnaire. Then, participants performed a sound check by playing an example sentence to allow them to adjust their listening volume to a comfortable level. Listeners practiced using the computerized VAS interface by evaluating three practice fragments produced by speakers who were not part of this study.

Written instructions for the intelligibility task directed listeners to judge how well a sentence was understood using a computerized, continuous VAS with scale end points labeled *Cannot Understand Anything* (0) to *Understand Everything* (100). The continuous scale contained no tick marks and was oriented horizontally on a computer monitor. Participants were instructed as follows: "Judge how well you can understand each utterance, using the scale from CANNOT UNDERSTAND ANYTHING (0) to UNDERSTAND EVERYTHING (100). Slide the indicator positioned in the middle of the line using a mouse, pen, or touch screen to indicate the intelligibility judgment. Wait until the utterance is finished before making your judgment. Evaluate the entire utterance. You are encouraged to use the entire scale when judging the utterances." Listeners judged each fragment without knowledge of the speaker's neurological diagnosis. Listeners were allowed to listen to each fragment once and were not able to alter a previously submitted score after moving on to the next fragment.

Each listener judged 18 fragments, randomly selected from the list of stimuli. To assess intrarater reliability, a random selection of nine fragments was presented a second time to each listener. Stimuli were presented in a blocked and randomized fashion, ensuring that fragments presented the second time for the purpose of intrarater reliability measures were separated by at least two other stimuli. The fast rate condition was not of interest to this study, but was used as an experimental foil. Thus, perceptual judgments for the fast rate condition were not included in the statistical analyses discussed below.

The experiment lasted 12–16 min with associated remunerations between \$1.60 and \$2.20. Listeners were allowed to participate once. For each listener, intrarater reliability was determined by calculating a Pearson correlation for the VAS scores of the nine stimuli that were presented twice. Listeners with nonsignificant Pearson correlations ($\alpha > .05$, equivalent to Pearson $r < .583$) were replaced (c.f. Sussman & Tjaden, 2012). Overall, about 30% of listeners were replaced based on low intrarater reliability values. For each stimulus, a minimum of three valid VAS scores were obtained to permit calculation of standard deviations as well as to meet minimum effect size targets (c.f. Abur et al., 2019; Tjaden et al., 2014). The average number of observations for each stimulus was 4.85 (SD = 1.65), equally distributed across groups, fragments, and speaking styles. The average intrarater reliability expressed as Pearson r across all listeners meeting intrarater reliability criteria was .790 (range: .583–.994). Overall intrarater reliability and interrater reliability were assessed by means of two-way mixed (with fixed measures and random listener effects) intraclass correlation coefficient (ICC) models. ICCs computed for absolute agreement among listeners were used to determine intrarater reliability. The single measures ICC was .76 (95% CI [.69–.81]; $p < .0001$), and the average measures ICC was .86 (95% CI [.81–.90]; $p < .0001$), both indicating good intrarater reliability. An ICC model computing consistency aggregated over all listeners was used to determine interrater reliability (Neel, 2009; Weismer et al., 2012). Aggregated interrater reliability as assessed by average measures ICC was .85 (95% CI [.84–.86]; $p < .0001$), indicating good reliability. Listener reliability is considered further in the Discussion section.

Data Analysis

Statistical analyses were carried out using R software (R Core Team, 2019). The averaged logarithmically transformed word frequency and raw neighborhood density values of each word were compared across fragments using a one-way analysis of variance (ANOVA). The outcome measures fragment duration (in seconds), SPL (in dB), and intelligibility (in VAS score) were analyzed by means of linear mixed model analyses. To reduce model complexity and to be able to focus on interaction effects separately for each group, separate models were used to analyze between-groups effects and within-group effects. Between-groups effects were analyzed with a model containing group, speaking style, and fragment as fixed factors, and speaker as the independent random factor. In these models, the main effects of group, speaking style, and fragment were of interest. The factor speaker sex was included as covariate in consideration of different proportions of males and females in the three speaker groups. Speaker sex was a significant contributing factor for intelligibility only, and excluded from the models analyzing fragment duration and SPL. Given the need for hypothesis-driven analyses, planned comparisons were performed for the factor speaking style, with post hoc comparisons limited to slow, loud, and clear speaking styles compared to the habitual speaking style.

Within-group effects for each of the three speaker groups were analyzed using linear mixed models containing speaking style and fragment as fixed factors, and speaker as the independent random factor. For intelligibility, listener was an additional independent random factor. As the covariate speaker sex was not a statistically significant factor in any of the within-speaker group analyses, this variable was excluded from the within-speaker group models. In these models, the main and interaction effects of speaking style and fragment were of interest. Significant post hoc differences of estimated marginal means were further explored using Tukey's method to correct for multiple comparisons, and Satterthwaite's method to estimate the degrees of freedom (Lenth et al., 2020). Again, planned comparisons were performed for the factor speaking style, with pairwise post hoc analyses limited to comparisons of slow, loud, and clear speaking styles versus the habitual speaking style. Standardized effect sizes, expressed as Cohen's *d,* were derived from the estimated marginal means and population standard deviations. A significance level of .05 was used for all hypothesis testing. With a few exceptions for completeness, the Results section focuses on statistically significant findings.

Results

Lexical Characteristics

Results from the one-way ANOVA indicated that log-transformed word frequency values were not statistically different across the three fragments, $F(2, 53) = 0.458$, $p = .635$. The average word neighborhood densities were also not statistically different across the fragments, $F(2, 53) = 1.647$, $p = .202$. Thus, any differences in intelligibility across the three fragments could not be attributable to differences in word frequency or neighborhood density.

Duration

Between-Speaker Groups

The results of the statistical analyses of the fragment durations indicated a main effect of group: $F(2, 75.03) = 8.18$, $p < .001$. Post hoc pairwise testing indicated longer fragment durations for the control speakers ($M = 9.56$ s; $t(74.9) = 3.66$, $p = .001$, $d = 1.08$) as well as the speakers with MS ($M = 9.62$ s; $t(75.1) = 3.71$, $p = .001$, $d = 1.11$) compared to the speakers with PD ($M = 7.55$ s). Fragment durations were not significantly different between the speakers with MS and the speakers of the control group. There was a main effect of speaking style: $F(3, 814.6) = 288.1$, $p < .001$. Durations were longer in the slow ($M = 11.38$ s; $t(814) = 24.59$, $p = .001$, $d = 2.39$) and clear ($M = 10.09$ s; $t(815) = 17.35$, $p = .001$, $d = 1.70$) speaking styles compared to the habitual speaking style ($M = 6.93$ seconds). The habitual speaking style was not significantly different from the loud speaking style ($M = 7.24$ s). There was a main effect of fragment: $F(2, 814.1) = 10.4$, $p < .001$. Durations were longer in the beginning fragment ($M = 9.29$ s) compared to the middle ($M = 8.84$ s; $t(814) = 2.86$, $p = .012$, $d = 0.24$) and end fragment ($M = 8.59$ s; $t(814) = 4.50$, $p = .001$, $d = 0.38$). The middle and end fragment did not differ in duration.

Within-Speaker Groups

Speakers in control group. Within-group analyses of fragment durations for the control speakers indicated a significant effect of speaking style: $F(3, 340.1) = 255.3$, $p < .001$, and fragment: $F(2, 340.1) = 4.18$, $p = .016$. Post hoc comparisons across fragments indicated that the longest fragment durations were produced during the slow speaking style. Both the slow speaking style ($M = 13.23$ s; $t(340) = 23.3$, $p < .001$, $d = 3.36$) and the clear speaking style ($M = 11.36$ s; $t(340) = 16.7$, $p < .001$, $d = 2.42$) yielded significantly longer fragment durations, compared to the habitual speaking style ($M = 6.52$ s). The loud speaking style ($M = 7.14$ s) was not significantly different from the habitual speaking style. Across speaking styles, the beginning fragment ($M = 9.96$ s) had longer durations compared to the end fragment ($M = 9.25$ s; $t(340) = 2.784$ $p = .013$, $d = 0.36$). Comparisons involving the middle fragment ($M = 9.49$ s) were not significant. The interaction effect of speaking style by fragment was not significant.

Speakers with MS. The statistical analyses for the speakers with MS indicated significant main effects of speaking style: $F(3, 310.5) = 100.9$, $p < .001$, and fragment: $F(2, 310.1) = 7.26$, $p < .001$. The post hoc results indicated that, across fragments, speakers increased durations to the greatest extent in the slow speaking style ($M = 11.84$ s; $t(310) = 14.1$, $p < .001$, $d = 2.11$), followed by the clear speaking style ($M = 11.03$ s; $t(311) = 11.2$, $p < .001$, $d = 1.71$), compared to the habitual speaking style ($M = 7.57$ s). The loud speaking style ($M = 7.98$ s) was not significantly different from the habitual speaking style. Across speaking styles, durations were longest in the beginning fragment ($M = 10.15$ s) compared to the middle ($M = 9.54$ s; $t(310) = 2.36$, $p = .0496$, $d = 0.31$) and end fragments ($M = 9.17$ s; $t(310) = 3.77$, $p < .001$, $d = 0.49$). The middle and end fragments did not differ in duration. The interaction effect of speaking style by fragment was not significant.

Speakers with PD. The results of the statistical analyses for the speakers with PD also indicated a significant main effect of speaking style: $F(3, 164.0) = 47.8$, $p < .001$. Parallel to the results for other speaker groups, the results of the post hoc comparisons indicated the longest fragment durations for the slow speaking style ($M = 9.06$ s; $t(164) = 10.1$, $p < .001$, $d = 2.06$), followed by the clear speaking style ($M = 7.85$ s; $t(164) = 5.0$, $p < .001$, $d = 1.03$), compared to the habitual speech style ($M = 6.66$ s). Durations of the loud speaking style ($M = 6.61$ s) was not significantly different from the habitual speaking style. The main effect of fragment and the interaction effect of speaking style by fragment were not significant.

Sound Pressure Level

Between-Speaker Groups

The results of the statistical analyses of SPL indicated a main effect of group: $F(2, 72.95) = 4.96$, $p = .001$. Post hoc pairwise testing comparing the three speaker groups indicated significantly higher SPLs for the control speakers ($M = 78.7$ dB) compared to the speakers with PD ($M = 75.4$ dB; $t(72.9) = 3.08$, $p = .008$, $d = 1.45$). SPLs were not significantly different for between-group comparisons involving speakers with MS ($M = 77.1$ dB). The main effect of speaking style was significant: $F(3, 794.1) = 325.6$, $p < .001$. SPLs were significantly higher in the loud ($M = 81.0$ dB; $t(794) = 26.4$, $p < .001$, $d = 2.60$) and clear ($M = 77.4$ dB; $t(794) = 10.7$, $p < .001$, $d = 1.06$) speaking styles compared to the habitual speaking style ($M = 75.0$ dB). SPL for the slow speaking style ($M = 74.8$ dB) was not different from the habitual speaking style. The main effect of fragment was also significant: $F(2, 794.0) = 3.79$, $p = .023$. SPLs were higher in the middle fragment ($M = 77.3$ dB; $t(794) = 2.75$, $p = .017$, $d = 0.23$) compared to the beginning fragment ($M = 76.8$ dB). Comparisons involving the end fragment ($M = 77.1$ dB) were not significant.

Within-Speaker Groups

Speakers in control group. Control speakers demonstrated a significant effect of speaking style: $F(3, 318.0) = 174.4$, $p < .001$. The results of the post hoc comparisons indicated that, across speaking styles, the highest mean SPL values were produced during the loud speaking style ($M = 83.5$ dB; $t(318) = 20.1$, $p < .001$, $d = 3.00$), followed by the clear speaking style ($M = 79.2$ dB; $t(318) = 8.68$, $p < .001$, $d = 1.30$), when compared to habitually produced speech ($M = 75.9$ dB). No significant differences were found between the habitual and slow speaking styles ($M = 76.2$ dB). The main effect of fragment and the interaction effect of speaking style by fragment were not significant.

Speakers with MS. The results of the speakers with MS indicated a significant effect of speaking style: $F(3, 312.0) = 182.9$, $p < .001$. Similar to the results of the control speakers, the post hoc comparisons indicated highest mean SPL values during the loud speaking style ($M = 81.2$ dB; $t(312) = 20.9$, $p < .001$, $d = 3.15$), followed by the clear speaking style ($M = 77.4$ dB; $t(312) = 8.08$, $p < .001$, $d = 1.23$), compared to habitually produced speech ($M = 74.4$ dB). No significant differences were found between the habitual and slow speaking style ($M = 75.1$ dB). The effect of fragment was significant: $F(2, 312.0) = 6.84$, $p = .001$. Across speaking styles, SPLs were higher in the middle ($M = 77.4$ dB; $t(312) = 3.52$, $p = .002$, $d = 0.46$) and end fragments ($M = 77.2$ dB; $t(312) = 2.74$, $p = .018$, $d = 0.36$), compared to the beginning fragment ($M = 76.5$ dB). SPLs of the middle and end fragments were not significantly different. The interaction effect of speaking style by fragment was not significant.

Speakers with PD. A significant effect of speaking style was also found for the speakers with PD: $F(3, 164.0) = 47.5$, $p < .001$. Parallel to the results of the other two speaker groups, mean SPL values were highest during the loud speaking style ($M = 78.2$; $t(164) = 8.34$, $p < .001$, $d = 1.71$), followed by the clear speaking style ($M = 75.9$ dB; $t(164) = 3.21$, $p = .009$, $d = 0.66$), when compared to habitually produced speech ($M = 74.4$ dB). In addition, higher SPL values were found during the habitual speaking style compared to the slow speaking style ($M = 73.0$ dB). The main effect of fragment and the interaction effect of speaking style by fragment were not significant.

In summary, no significant differences were found in lexical variables between the three fragments. When comparing speaking styles, all three speaker groups produced the longest fragment durations during the slow speaking styles. Fragment durations were also significantly longer in the clear speaking style compared to habitually produced speech. All three speaker groups produced the highest SPL during the loud speaking style. Compared to the habitual speaking style, SPLs were also significantly higher in the clear speaking style. When comparing fragments, findings across groups varied. The control speakers showed longer duration in the beginning compared to the end fragment, with SPLs equal between fragments. The speakers with MS showed longer durations and lower SPLs in the beginning versus the middle and end fragments. Durations and SPLs did not differ between fragments for the speakers with PD.

Intelligibility

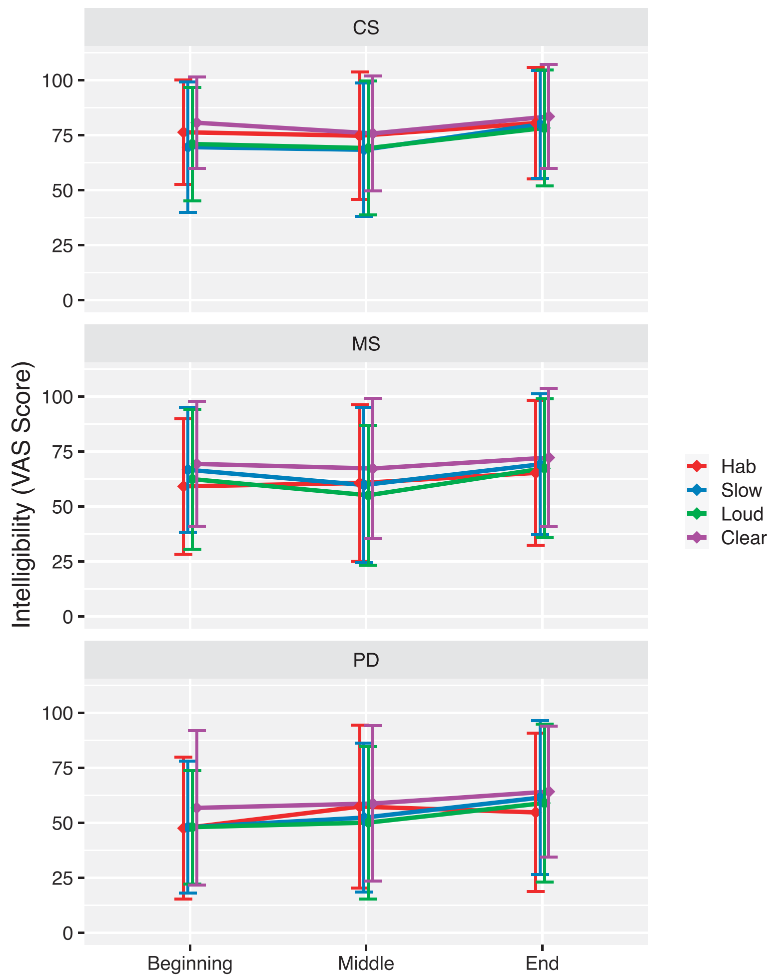
Between-Speaker Groups

VAS intelligibility scores across fragments and speaking styles are plotted in Figure 1. Intelligibility scores in Figure 1 are reported separately for each speaker group. Summative statistics of VAS scores across groups and sex, separately by speaking style, are reported in Table 2. Averaged differences in intelligibility scores between nonhabitual and habitual speaking styles across fragments and speaking styles are plotted in Figure 2.

When comparing overall intelligibility between groups, the statistical analysis indicated a significant main effect of group: $F(2, 73.7) = 9.26$, $p < .001$. Post hoc pairwise testing indicated higher intelligibility for the control group compared to the MS group (mean difference = 11.4, $t(73.8) = 2.50$, $p = .038$, $d = 0.56$) and the PD group (mean difference = 23.4, $t(73.7) = 4.21$, $p < .001$, $d = 1.14$). Intelligibility was not significantly different between the MS and PD groups.

The main effect of speaking style was also significant, $F(3, 4173) = 27.2$, $p < .001$. Post hoc analyses

Figure 1. Intelligibility scores across fragments and speaking styles, separately by speaker group. Error bars indicate ± 1 *SD*. VAS = visual analogue scale; Hab = habitual; CS = control speakers; MS = multiple sclerosis; PD = Parkinson's disease.



indicated that when pooled over groups and fragments, intelligibility was significantly higher in the clear speaking style compared to the habitual speaking style (mean difference = 6.4, $t(4170) = 6.5$, $p < .001$, $d = 0.31$). There also was a significant effect of fragment: $F(2, 4038) = 43.6$, $p < .001$. Post hoc comparisons indicated that the end fragment was significantly more intelligible compared to the beginning (mean difference = 6.3, $t(4028) = 7.8$, $p < .001$, $d = 0.31$) and middle fragments (mean difference = 6.9, $t(4045) = 8.4$, $p < .001$, $d = 0.34$). It should further be noted that the covariate speaker sex significantly contributed to the between-speaker model: $F(1, 73.8) = 4.53$, $p = .037$. Overall, male speakers were more intelligible than female speakers.

Within-Speaker Groups

Findings for within-group analyses of intelligibility are detailed below for each of the three speaker groups. *Speakers in control group*. Within-group statistics of the control speakers indicated significant main effects of speaking style: $F(3, 1749.9) = 13.6$, $p < .001$, and fragment: $F(2, 1664.4) = 24.1$, $p < .001$. The interaction of speaking style by fragment was not significant. The results of the post hoc comparisons are reported in Table 3. In summary, pooled over fragments, intelligibility was significantly lower in the slow and loud speaking styles compared to the habitual speaking style. Intelligibility of the clear and habitual speaking styles were not significantly different. Pooled over speaking styles, the end fragment was more intelligible compared to the beginning and middle fragments, and the beginning fragment was more intelligible compared to the middle fragment.

Speakers with MS. The results of the statistical analyses for the speakers with MS indicated significant main effects of speaking style: $F(3, 1643.6) = 12.3$, $p < .001$, and fragment: $F(2, 1583.9) = 18.0$, $p < .001$. In addition, the interaction of speaking style and fragment was significant, $F(6, 1641.5) = 3.23$, $p = .004$. The results of the post

Table 2. VAS scores (mean [standard deviation]) as a function of group and sex.

Group	Sex	Habitual	Slow	Loud	Clear
CS	Both	77.4 (26.1)	72.7 (28.7)	72.7 (27.9)	79.8 (24.1)
	Female	75.1 (26.4)	66.4 (31.5)	67.4 (29.5)	75.8 (25.7)
	Male	79.1 (23.4)	75.4 (27.3)	77.3 (24.4)	81.0 (22.7)
MS	Both	61.8 (33.2)	65.1 (32.2)	61.8 (32.1)	69.6 (30.7)
	Female	59.8 (34.0)	60.3 (34.2)	59.8 (31.9)	64.4 (32.4)
	Male	62.0 (31.5)	70.1 (28.6)	65.6 (30.6)	71.9 (29.6)
PD	Both	52.8 (35.0)	53.8 (33.3)	52.1 (32.2)	59.9 (33.5)
	Female	46.3 (36.4)	40.6 (33.7)	42.4 (30.5)	50.1 (34.6)
	Male	51.5 (33.7)	63.2 (29.2)	56.4 (34.4)	64.1 (31.6)

Note. VAS = visual analogue scale; CS = control speakers; MS = multiple sclerosis; PD = Parkinson's disease.

hoc comparisons are reported in Table 4. The significant interaction indicated that variation in intelligibility across the beginning, middle, and end fragments varied with speaking style. In particular, with respect to the habitual speaking style, the end fragment was significantly more intelligible than the beginning fragment. For the slow and clear speaking styles, the end fragment was significantly more intelligible than the middle fragment. For

Figure 2. Averaged differences in intelligibility scores between the habitual speaking style and the nonhabitual speaking styles. Results are displayed across fragments and speaking styles, separately by speaker group. VAS = visual analogue scale; Hab = habitual; CS = control speakers; MS = multiple sclerosis; PD = Parkinson's disease.

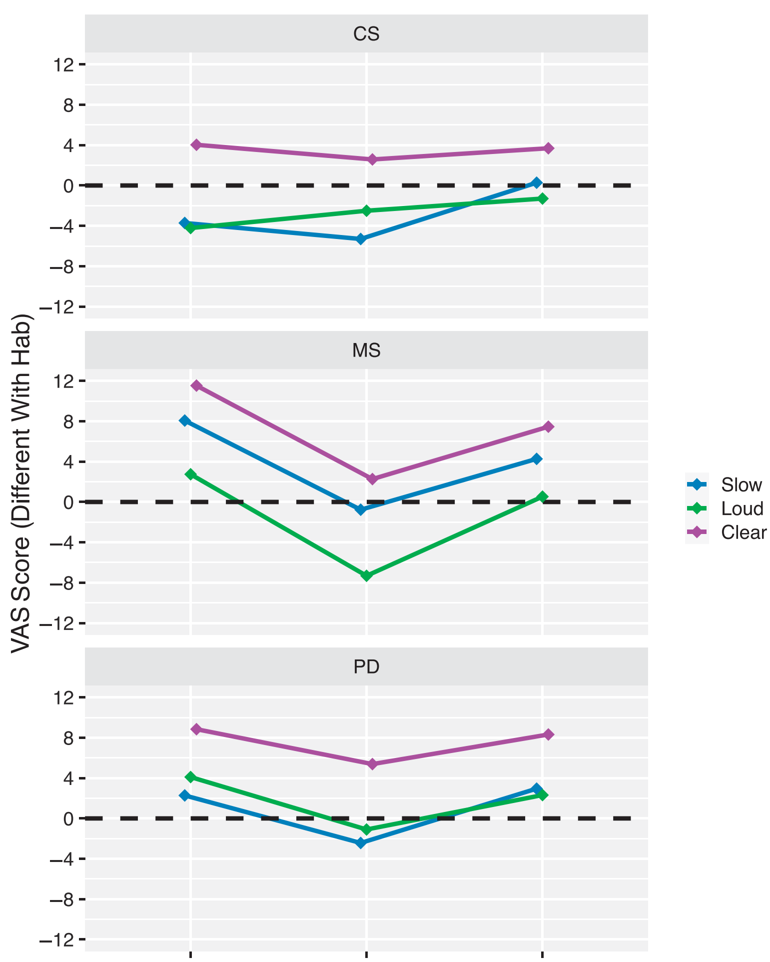


Table 3. Post hoc comparisons of fragment intelligibility in control speakers.

Speaking style	Difference (VAS score)	*t*	*p*	*d*
Contrasts: Main effect				
Habitual–slow	4.52	3.06	.012*	0.24
Habitual–loud	4.59	3.12	.001*	0.24
Habitual–clear	−3.18	−2.15	.137	−0.17
Contrasts: beginning fragment				
Habitual–slow	4.88	1.96	.205	0.26
Habitual–loud	4.49	1.77	.290	0.24
Habitual–clear	−6.22	−2.36	.084	−0.33
Contrasts: middle fragment				
Habitual–slow	5.55	2.12	.149	0.29
Habitual–loud	5.22	2.07	.164	0.28
Habitual–clear	−2.95	−1.16	.651	−0.16
Contrasts: end fragment				
Habitual–slow	3.13	1.28	.577	0.17
Habitual–loud	4.06	1.64	.357	0.21
Habitual–clear	−0.37	−0.15	.999	−0.02
Fragment ID				
Beginning–middle	2.94	2.45	.039*	0.16
Beginning–end	−5.22	−4.39	< .001*	−0.28
Middle–end	−8.17	−6.84	< .001*	−0.43
Contrasts: habitual				
Beginning–middle	1.78	0.69	.771	0.09
Beginning–end	−6.14	−2.50	.033*	−0.32
Middle–end	−7.92	−3.19	.004*	−0.42
Contrasts: slow				
Beginning–middle	2.45	0.99	.584	0.13
Beginning–end	−7.89	−3.31	.003*	−0.42
Middle–end	−10.3	−4.15	< .001*	−0.55
Contrasts: loud				
Beginning–middle	2.50	1.04	.593	0.13
Beginning–end	−6.57	−2.70	.019*	−0.35
Middle–end	−9.07	−3.75	.001*	−0.48
Contrasts: clear				
Beginning–middle	5.05	2.02	.108	0.27
Beginning–end	−0.29	−0.12	.993	−0.02
Middle–end	−5.34	−2.21	.071	−0.28

Note. Asterisk indicates statistically significant differences. VAS = visual analogue scale.

Table 4. Post hoc comparisons of fragment intelligibility in speakers with multiple sclerosis.

Speaking style	Difference (VAS score)	*t*	*p*	*d*
Contrasts: Main effect				
Habitual–slow	−3.89	−2.45	.068	−0.19
Habitual–loud	1.85	1.14	.664	0.09
Habitual–clear	−6.96	−4.35	< .001*	−0.34
Contrasts: beginning fragment				
Habitual–slow	−7.32	−2.68	.037*	−0.36
Hab–loud	−3.78	−1.34	.538	−0.19
Hab–clear	−11.92	−4.28	< .001*	−0.58
Contrasts: middle fragment				
Hab–slow	2.26	0.82	.845	0.11
Habitual–loud	11.44	4.04	< .001*	0.56
Habitual–clear	−1.26	−0.45	.969	−0.06
Contrasts: end fragment				
Habitual–slow	−6.60	−2.38	.082	−0.32
Habitual–loud	−2.10	−0.78	.864	−0.10
Habitual–clear	−7.71	−2.83	.025*	−0.38
Fragment ID				
Beginning–middle	2.54	1.88	.14	0.13
Beginning–end	−5.24	−3.94	< .001*	−0.26
Middle–end	−7.78	−5.87	< .001*	−0.38
Contrasts: habitual				
Beginning–middle	−6.32	−2.26	.063	−0.31
Beginning–end	−6.89	−2.55	.030*	−0.34
Middle–end	−0.57	−0.21	.977	−0.03
Contrasts: slow				
Beginning–middle	3.26	1.22	.444	0.16
Beginning–end	−6.17	−2.25	.064	−0.30
Middle–end	−9.43	−3.44	.002*	−0.46
Contrasts: loud				
Beginning–middle	8.90	3.12	.005*	0.44
Beginning–end	−5.21	−1.90	.138	−0.26
Middle–end	−14.1	−5.14	< .001*	−0.69
Contrasts: clear				
Beginning–middle	4.34	1.60	.246	0.21
Beginning–end	−2.68	−0.98	.590	−0.13
Middle–end	−7.02	−2.60	.025*	−0.34

Note. Asterisk indicates statistically significant differences. VAS = visual analogue scale.

the loud speaking style, the beginning and end fragments were significantly more intelligible compared to the middle fragment.

Speakers with PD. The statistical analyses for the speakers with PD indicated significant main effects of speaking style: $F(3, 869.6) = 5.74$, $p = .001$, and fragment: $F(2, 854.0) = 11.6$, $p < .001$. The interaction of speaking style and fragment was not statistically significant. The results of the post hoc comparisons are reported in Table 5. The speakers with PD demonstrated increased intelligibility in the clear versus the habitual speaking style, when data were pooled over fragments. No significant differences were found between the slow or loud speaking styles, compared to the habitual speaking style. Pooled over speaking styles, the end fragment was more intelligible compared to the beginning and middle fragments.

Discussion

Validating Speaking Style Implementation, Lexical Characteristics, and Listener Reliability

This study evaluated the effect of slow, loud, and clear speaking styles on crowdsourced scaled intelligibility across a

Table 5. Post hoc comparisons of fragment intelligibility in speakers with Parkinson's disease.

Speaking style	Difference (VAS score)	*t*	*p*	*d*
Contrasts: Main effect				
Habitual–slow	−2.35	−1.00	.748	−0.12
Habitual–loud	−1.75	−0.73	.886	−0.09
Habitual–clear	−9.42	−3.89	.001*	−0.47
Contrasts: beginning fragment				
Habitual–slow	−0.50	−0.13	.999	−0.03
Habitual–loud	−2.20	−0.55	.946	−0.11
Habitual–clear	−9.83	−2.34	.091	−0.49
Contrasts: middle fragment				
Habitual–slow	2.03	0.49	.961	0.10
Habitual–loud	3.65	0.84	.834	0.18
Habitual–clear	−7.12	−1.73	.309	−0.36
Contrasts: end fragment				
Habitual–slow	−8.57	−2.11	.151	−0.43
Habitual–loud	−6.70	−1.55	.407	−0.33
Habitual–clear	−11.32	−2.67	.039*	−0.57
Fragment ID				
Beginning–middle	−3.96	−2.00	.114	−0.20
Beginning–end	−9.53	−4.81	< .001*	−0.48
Middle–end	−5.58	−2.68	.020*	−0.28
Contrasts: habitual				
Beginning–middle	−6.73	−1.68	.213	−0.34
Beginning–end	−6.02	−1.49	.297	−0.30
Middle–end	0.71	0.17	.984	0.04
Contrasts: slow				
Beginning–middle	−4.20	−1.08	.530	−0.21
Beginning–end	−14.1	−3.63	.001*	−0.70
Middle–end	−9.88	−2.45	.039*	−0.49
Contrasts: loud				
Beginning–middle	−0.88	−0.21	.977	−0.04
Beginning–end	−10.5	−2.53	.031*	−0.53
Middle–end	−9.64	−2.19	.073	−0.48
Contrasts: clear				
Beginning–middle	−4.02	−0.95	.611	−0.20
Beginning–end	−7.51	−1.75	.188	−0.38
Middle–end	−3.49	−0.83	.683	−0.17

Note. Asterisk indicates statistically significant differences. VAS = visual analogue scale.

reading passage produced by speakers with dysarthria due to MS and PD as compared to neurologically healthy control speakers. Pertinent experimental considerations are discussed briefly prior to addressing the hypotheses.

First, all three speaker groups significantly increased fragment durations during the slow and clear speaking styles and significantly increased SPL during the loud and clear speaking styles, compared to the habitual speaking style. These results indicate that when cued, speakers altered their speech output in ways consistent with slow, loud, and clear speaking styles (Lam et al., 2012; Picheny et al., 1986; Smiljanić & Bradlow, 2005). Second, neither log-transformed word frequency or neighborhood density values differed across fragments. The implication is that intelligibility variation among the three fragments cannot be attributed to variation in these lexical characteristics. Third, when implementing an intrajudge reliability criterion for listener participants, using crowdsourced VAS judgments may yield sufficiently acceptable intrarater and interrater reliability measures. In this study, about 30% of initially recruited listeners were replaced based on low intrarater reliability values. Following this procedure, intrarater reliability was characterized by a single measure ICC of .76 and an average measure ICC of .86. These metrics are comparable to those reported in previous literature using a VAS to obtain judgments of intelligibility in a controlled laboratory setting for individuals with dysarthria. Intrarater reliability values in these prior studies ranged from .76 to .85 (Abur et al., 2019; Tjaden et al., 2014; Van Nuffelen et al., 2009). Similarly, for interrater reliability, the consistency aggregated over all listeners as assessed by average ICC was .85. These findings are also comparable to previously reported studies wherein average ICC values ranged from .83 to .84 (Abur et al., 2019; A. R. Fletcher et al., 2017; Tjaden et al., 2014).

Effect of Speaking Styles on Intelligibility

It was hypothesized that a crowdsourced VAS paradigm would reveal intelligibility differences for a passage read by speakers with PD, speakers with MS, and healthy controls in habitual, slow, loud, and clear speaking styles. When pooled over the three groups, intelligibility was significantly higher for the clear speaking style compared to the habitual speaking style, while intelligibility in the slow and loud speaking styles did not differ from the habitual speaking style. The results for the within-group analyses were more varied. For the control group (see Table 3), intelligibility was significantly poorer in the slow and loud speaking styles compared to the habitual speaking style, while intelligibility for the clear speaking style did not differ from habitual. This suggests that extending global dysarthria treatment strategies to passage reading for neurotypical speakers is not effective for enhancing intelligibility and may even be detrimental. In contrast, the clear speaking style yielded improved intelligibility for the MS and PD groups relative to the habitual speaking style, although slow and loud speaking styles were not effective in enhancing intelligibility (see Tables 4 and 5). When descriptively comparing these findings to previous studies examining cueing strategies, the following observations can be made. Tjaden et al. (2014) and Stipancic et al. (2016) reported lab-sourced judgments of sentence intelligibility for the same speakers and speaking conditions of interest to this study. These prior studies found that the speakers with MS as well as neurotypical control speakers had higher intelligibility scores for Harvard sentences produced in both the clear and loud speaking styles relative to the habitual speaking style, while no intelligibility differences were found for the slow and habitual speaking styles. In this study, given that both clinical groups successfully increased SPL for passage reading, the fact that neither group showed an intelligibility improvement in the loud speaking style may be unexpected. While speech with a more favorable SNR should be more audible and, therefore, result in improved speech intelligibility (Kim & Kuo, 2012), speech fragments in this study were equated for peak intensity prior to mixing with multitalker babble. The implication is that audibility would not contribute to any intelligibility differences among speaking styles. However, other adjustments to segmental and suprasegmental speech production characteristics including enhanced vowel and consonant segmental contrasts, a larger fundamental frequency range, and long-term spectral change have been found to be associated with speaking styles of interest to this study, including the loud speaking style (Neel, 2009; Tjaden & Wilding, 2004). These acoustic changes may be presumed to be especially beneficial to speakers with PD, whose speech is often characterized by reduced loudness, monoloudness, and monopitch. In this study, speakers with PD increased SPL by about 3.8 dB and speakers with MS by about 6.4 dB from the habitual to the loud speaking style. In comparison, Tjaden et al. (2014) and Stipancic et al. (2016) reported larger SPL changes, with an average increase of 7 dB for speakers with PD and 8 dB for speakers with MS during the loud speaking style. One might speculate that the intelligibility benefit of loud speech is related to the intensity gain achieved during this speaking style. Thus, the absence of intelligibility improvement for the loud speaking style might possibly be explained by the smaller increase in SPL and corresponding speech production adjustments realized during passage reading task (Huber & Darling, 2011; Kempler & Van Lancker, 2002). The absence of intelligibility improvements in the slow speaking style in speakers with PD and MS was not unexpected, as the varying impact of rate reduction on intelligibility has been previously documented (Stipancic et al., 2016; Tjaden et al., 2014).

The finding that both lab-sourced (previous studies discussed above) and crowdsourced (current study) experimental paradigms documented intelligibility improvements for clear speech suggests that speech improvements to neurologically impaired speech associated with this speaking style are robust and may be perceived in a variety of contexts. Furthermore, these findings are affirmation of the potential effectiveness and stimulability of this global behavioral treatment technique. On average, intelligibility for the clear speaking style improved intelligibility in the MS group by 7.0% and in the PD group by 9.4% compared to the habitual speaking style. Previous research has suggested this magnitude of intelligibility difference is clinically meaningful, especially for mildly impaired speakers in a perceptually challenging environment of multitalker babble (Lansford et al., 2019; Stipancic et al., 2018; Tjaden et al., 2014). In addition, the mild speech impairment of most speakers in this study may have contributed to the relatively modest intelligibility gains for the clear speaking style. Future research including speakers with dysarthria spanning a wider severity continuum may shed light on whether notable intelligibility gains for passage reading are limited to clear speech only, or whether loud and slow speaking styles might also lead to improved intelligibility.

Although reliability for crowdsourced and lab-sourced listeners was comparable, overall sensitivity to production contrasts of different speaking styles might be lower in the crowdsourced listener group, resulting in less pronounced intelligibility differences between speaking styles. This is in spite of the differences in signal-to-noise listening conditions across studies. Whereas Tjaden et al. (2014) and Stipancic et al. (2016) employed an SNR of –3 dB for their sentence materials, this study used a more favorable SNR of 0 dB in the speech fragments, theoretically leading to improved sensitivity to production differences across conditions. Despite this apparent loss of sensitivity, these findings indicate the feasibility of using crowdsourced VAS judgments of intelligibility for speakers with dysarthria.

Variation in Intelligibility across the Reading Passage

It also was hypothesized that the crowdsourced VAS paradigm would reveal intelligibility variation across the course of the reading passage as a factor of speaking style. Pooled over the three speaker groups, the end fragment was significantly more intelligible compared to the beginning and middle fragments. The within-group results were more varied. With respect to the control group (see Table 3), variation in intelligibility was noted, with overall higher intelligibility in the beginning and end fragments. The absence of a Speaking Style × Fragment interaction effect indicated that this variation was not influenced by speaking style. The overall high intelligibility scores in combination with a notable absence of a significant clear speech intelligibility benefit across all three fragments may be an indication that speakers already produced the reading passage at maximum performance during the habitual speaking style, without room for further improvement.

The results for the speakers with MS (see Table 4), pooled over speaking styles, indicated increased intelligibility for the end fragment compared to other fragments. Speaking style contributed with varying degree to the intelligibility gain for the end fragment. When comparing intelligibility in the nonhabitual speaking styles to the habitual speaking style across the three fragments for speakers with MS (see Figure 2), a trend emerged in which all three nonhabitual speaking styles were characterized by an intelligibility decline from the beginning to the middle fragment, followed by notable gains from the middle to the end fragment. Since the speakers successfully produced the speaking styles as evidenced by acoustic measures of duration and SPL, there is no evidence that speakers became less effective in maintaining the nonhabitual speaking styles during the middle part of the reading passage, making it difficult to explain the nature of the intelligibility variation across fragments, at least based on the acoustic measures that were obtained.

The results for the speakers with PD (see Table 5) showed that, pooled over speaking styles, intelligibility was significantly higher for the end fragment compared to the beginning and middle fragments, but this variation was not influenced by speaking style. The intelligibility differences between the nonhabitual and habitual speaking styles showed a similar but less pronounced pattern compared to the speakers with MS: Intelligibility declined from the beginning to the middle fragment, followed by gains from the middle to the end fragment. Parallel to the results for the speakers with MS, acoustic measures of duration, and SPL did not provide clues that may explain this trend in intelligibility variation. This suggests that other factors such as enhanced segmental articulation or fundamental frequency modulation may be at play, which could be explored in future studies.

Drawing on the results of Kuo and Tjaden (2016), who reported on acoustic variation across segments in the shorter John Reading passage (192 vs. 279 words in the Hunter Script), some notable parallels with this study may be discerned. Acoustic variation in segments was mostly independent of speaking style, similar to these intelligibility findings of the control speakers and speakers with PD. With respect to the loud speaking style, an overall trend of F2 IQR decreasing in the middle segment compared to the beginning and end segments was found. As reduction in F2 IQR has been associated with decreases in intelligibility (Feenaughty et al., 2014), this trend aligns well with the intelligibility patterns of the loud speaking style for the speakers with MS in this study. Furthermore, Kuo and Tjaden (2016) reported increased articulation rates for the end segment compared to earlier segments. Although the association between articulation rates and intelligibility is not unequivocal, some studies have indicated an increased intelligibility with decreasing articulation rate (A. R. Fletcher et al., 2017; Van Nuffelen et al., 2009). If such a trend were present in this study, one would predict a decrease, not an increase, in intelligibility for the end fragment. These findings suggest that assessing articulation rates might be warranted in future studies.

Although the finding of enhanced intelligibility of nonhabitual speaking styles during the beginning and end fragments for the speakers with PD and MS needs to be replicated in other studies, this points to methodological considerations when using passage reading for estimating intelligibility in clinical populations. First, there might be an interplay between physical (e.g., an increased fatigue from beginning toward the end) and psychological factors (e.g., speakers are emotionally more motivated toward the end of the passage) leading to the observed variation in intelligibility for the different fragments. Alternatively, the finding that the clinical speaker groups appeared to more effectively implement the nonhabitual speaking styles during the beginning and end of the reading passage relative to their habitual productions (see Figure 2) suggests that these speakers were immediately able to implement nonhabitual speaking styles on cueing, as evidenced by the intelligibility gains in the beginning fragment for the nonhabitual speaking styles. During the middle fragment, the intelligibility gains for the slow and clear speaking styles disappeared for the speakers with MS, with the loud speaking style resulting in a relative intelligibility loss. Since the interaction effect of speaking style by fragment was absent for the speakers with PD, limited conclusions can be drawn. A trend can be discerned in which gains for the slow and loud speaking styles seem to diminish during the middle fragment, while intelligibility gains for the clear speaking style were largely maintained. Given the significant interaction effect of segment by speaking style for the speakers with MS in combination with the trends displayed in Figure 2, the two clinical groups seem to differ in the relative intelligibility variation for the nonhabitual speaking styles. Overall, the magnitude of variation was descriptively larger for the speakers with MS compared to the speakers with PD, suggesting that speakers of the former group are more sensitive to implementing nonhabitual speaking styles compared to the speakers with PD. A number of factors may underlie these observations. First, speakers with PD were more severely impaired compared to speakers with MS (see Table 2). A neuromuscular system with diminished neurological capabilities and/or increased neuromuscular constraints might result in a reduced ability to change speech production styles when cued to do so (Tsao et al., 2006; Tsao & Weismer, 1997). Second, rigidity and bradykinesia are common clinical signs of parkinsonism. Bearing in mind that the manifestation of these symptoms to the axial speech musculature is not fully understood, rigidity may limit flexibility and intentional adaptation to the speech subsystems for speakers with PD compared to speakers with MS (Pinto et al., 2017; Skodda et al., 2012). Third, deficits in perceptual processing of speech intensity, timing, fundamental frequency, and formant frequencies and the effects on speech production have been reported in speakers with PD (De Keyser et al., 2016; Mollaei et al., 2019). Thus, perceptual deficits may also have contributed to the limited changes in speech production across speaking styles for the speakers with PD.

These findings have several clinical implications. The presence of intelligibility differences between fragments suggest that passage reading may supplement standard metrics of average sentence intelligibility and provide a more nuanced and complete assessment of intelligibility by yielding a range of intelligibility values not captured by average or overall sentence intelligibility. While intelligibility variation might be present to a similar degree among a series of sentences, to date such analyses at the level of individual sentences has not been carried out. For example, for the SIT, intelligibility is calculated by averaging across the 11 sentences and does not consider any variation in intelligibility across sentences (Yorkston & Beukelman, 1981). In contrast, this study elucidated the presence of distinct intelligibility variation patterns at group level across fragments of a reading passage as a factor of different speaking styles. These patterns may serve as a baseline and inform intervention decisions. For example, although, on average, intelligibility was the highest at the end of the passage reading, there may be speakers for whom this pattern does not hold throughout the passage, due to their inability to maintain effective implementation of the chosen speaking style. In this case, an alternative technique that achieves durable intelligibility benefits would be advantageous. Therefore, the use of a reading passage may be appropriate to test stimulability and trainability of treatment methods in an ecologically valid manner. Furthermore, the findings suggest the need to consider task or session length for their potential impact on intelligibility. Specifically, these results indicate that some minimum length of stimulus material may be required to fully evaluate whether a specific cued speaking style has been incorporated and applied in a consistent manner. Similar to previous work reporting the intelligibility benefit of sentences with fewer words (see the work of Allison et al., 2017), this type of research could have implications for determining speaking and practice time with the goal of optimizing intelligibility.

Limitations and Other Considerations

This study confirms the feasibility of crowdsourced VAS judgments to assess intelligibility variation over longer speech materials. However, a number of limitations were present that may be addressed in future work using crowdsourcing paradigms for intelligibility assessment. Overall dysarthria severity for the clinical groups was relatively mild in quiet listening conditions. To avoid ceiling effects during the assessment of intelligibility in all three speaker groups and to maximize between-speaker differences, stimuli were mixed with multitalker babble before presentation to listeners. The addition of noise might impact intelligibility differently for dysarthria and neurologically healthy speech. Previous studies reported mixed results, with Chiu and Forrest (2018) reporting a significant decrease in intelligibility in noise for speakers with dysarthria secondary to PD compared to neurotypical speakers, whereas Yoho and Borrie (2018) reported the decrease in intelligibility due to increasing levels of noise to be similar for both speakers with dysarthria and neurologically healthy controls. The potential impact of added noise to degraded speech is a complex topic that requires further investigation.

In addition, overall speech severity might affect an individual's response to a cue or instruction to produce different speaking styles. Individuals with more severe dysarthria than those included in this study may have benefitted more from nonhabitual speaking styles, resulting in significantly larger differences across conditions (A. R. Fletcher et al., 2017; Hammen et al., 1994). The relatively mild severity of participating speakers may not only have prevented more pronounced differences but may also have impacted the extent and patterning of intelligibility variation across the reading passage. Neuromuscular fatigue and difficulty sustaining respiratory support may start to play a role in more severely affected speakers, possibly resulting in patterns of intelligibility decline toward the end of the reading passage. Future research including speakers with a range of dysarthria severities is needed to further explore this possibility.

Listeners' assessment of intelligibility might potentially be influenced by having access to linguistically predictive information (Miller, 2013). The average predictability of words for each fragment was not considered and might have varied across the three fragments. In addition, the emotionality of the end fragment (aiming to shoot a bird) may have driven speakers to be more intelligible compared to the other fragments. These topics could be pursued in future research. Finally, results may not reflect spontaneous or conversational speech on intelligibility variation over time, as it is widely accepted that spontaneous speech requires additional cognitive resources that impact intelligibility (Kempler & Van Lancker, 2002; Tjaden & Wilding, 2011). However, we argue that a reading passage more closely approximates spontaneous speech versus isolated, single sentences and, therefore, may be more useful for understanding intelligibility variation that may be encountered in everyday speaking situations.

Acknowledgments

This study was funded by the National Institute on Deafness and Other Communication Disorders (Grant R01DC004689 awarded to Kris Tjaden). The authors gratefully thank all speaker and listener participants as well as Nicole Feeley, Grace Galvin, Laura Saitta, Noora

Somersalo, and Mengyang Qiu for their help with stimuli preparation, data collection, and analysis. Preliminary results of this study were presented at the 18th biennial conference of the International Clinical Phonetics and Linguistics Association.

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

Allison, K. M., Annear, L., Policicchio, M., & Hustad, K. C. (2017). Range and precision of formant movement in pediatric dysarthria. *Journal of Speech, Language, and Hearing Research, 60*(7), 1864–1876. https://doi.org/10.1044/2017_JSLHR-S-15-0438

Allison, K. M., Yunusova, Y., & Green, J. R. (2019). Shorter sentence length maximizes intelligibility and speech motor performance in persons with dysarthria due to amyotrophic lateral sclerosis. *American Journal of Speech-Language Pathology, 28*(1), 96–107. https://doi.org/10.1044/2018_AJSLP-18-0049

Beukelman, D. R., & Yorkston, K. M. (1980). Influence of passage familiarity on intelligibility estimates of dysarthric speech. *Journal of Communication Disorders, 13*(1), 33–41. https://doi.org/10.1016/0021-9924(80)90019-2

Blanchet, P. G., & Snyder, G. J. (2010). Speech rate treatments for individuals with dysarthria: A tutorial. *Perceptual and Motor Skills, 110*(3), 965–982. https://doi.org/10.2466/pms.110.3.965-982

Boersma, P., & Weenink, D. (2020). *Praat: Doing phonetics by computer* (6.1.09) [Computer software]. Phonetic Sciences, University of Amsterdam. http://praat.org

Cannito, M. P., Suiter, D. M., Beverly, D., Chorna, L., Wolf, T., & Pfeiffer, R. M. (2011). Sentence intelligibility before and after voice treatment in speakers with idiopathic Parkinson's disease. *Journal of Voice, 26*(2), 214–219. https://doi.org/10.1016/j.jvoice.2011.08.014

Chiu, Y.-F., & Forrest, K. (2018). The impact of lexical characteristics and noise on intelligibility of Parkinsonian speech. *Journal of Speech, Language, and Hearing Research, 61*(4), 837–846. https://doi.org/10.1044/2017_JSLHR-S-17-0205

Crystal, T. H., & House, A. S. (1982). Segmental durations in connected speech signals: Preliminary results. *The Journal of the Acoustical Society of America, 72*(3), 705–716. https://doi.org/10.1121/1.388251

De Keyser, K., Santens, P., Bockstael, A., Botteldooren, D., Talsma, D., De Vos, S., Van Cauwenberghe, M., Verheugen, F., Corthals, P., & De Letter, M. (2016). The relationship between speech production and speech perception deficits in Parkinson's disease. *Journal of Speech, Language, and Hearing Research, 59*(5), 915–931. https://doi.org/10.1044/2016_JSLHR-S-15-0197

De Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a web browser. *Behavior Research Methods, 47*(1), 1–12. https://doi.org/10.3758/s13428-014-0458-y

Duffy, J. R. (2019). *Motor speech disorders: Substrates, differential diagnosis, and management* (4th ed.). Elsevier.

Eijk, L., Fletcher, A., McAuliffe, M., & Janse, E. (2020). The effects of word frequency and word probability on speech rhythm in dysarthria. *Journal of Speech, Language, and Hearing Research, 63*(9), 2833–2845. https://doi.org/10.1044/2020_JSLHR-19-00389

Fairbanks, G. (1960). *Voice and articulation drillbook* (2nd ed.). Harper and Brothers.

Feenaughty, L., Tjaden, K., & Sussman, J. (2014). Relationship between acoustic measures and judgments of intelligibility in Parkinson's disease: A within-speaker approach. *Clinical Linguistics & Phonetics, 28*(11), 857–878. https://doi.org/10.3109/02699206.2014.921839

Ferguson, S. H. (2004). Talker differences in clear and conversational speech: Vowel intelligibility for normal-hearing listeners. *The Journal of the Acoustical Society of America, 116*(4), 2365–2373. https://doi.org/10.1121/1.1788730

Fletcher, A., & McAuliffe, M. (2021). Comparing lexical cues in listener processing of dysarthria and speech in noise. *American Journal of Speech-Language Pathology, 30*(3S), 1572–1579. https://doi.org/10.1044/2020_AJSLP-20-00137

Fletcher, A. R., McAuliffe, M. J., Lansford, K. L., Sinex, D. G., & Liss, J. M. (2017). Predicting intelligibility gains in individuals with dysarthria from baseline speech features. *Journal of Speech, Language, and Hearing Research, 60*(11), 3043–3057. https://doi.org/10.1044/2016_JSLHR-S-16-0218

Freed, D. (2018). *Motor speech disorders: Diagnosis and treatment* (3rd ed.). Plural.

Hammen, V. L., Yorkston, K. M., & Minifie, F. D. (1994). Effects of temporal alterations on speech intelligibility in parkinsonian dysarthria. *Journal of Speech and Hearing Research, 37*(2), 244–253. https://doi.org/10.1044/jshr.3702.244

Hardcastle, W. J., & Tjaden, K. (2008). Coarticulation and speech impairment. In M. J. Ball, M. R. Perkins, N. Muller, & S. Howard (Eds.), *The handbook of clinical linguistics* (pp. 506–524). Blackwell. https://doi.org/10.1002/9781444301007.ch32

Huber, J. E., & Darling, M. (2011). Effect of Parkinson's disease on the production of structured and unstructured speaking tasks: Respiratory physiologic and linguistic considerations. *Journal of Speech, Language, and Hearing Research, 54*(1), 33–46. https://doi.org/10.1044/1092-4388(2010/09-0184)

Hustad, K. C., & Cahill, M. A. (2003). Effects of presentation mode and repeated familiarization on intelligibility of dysarthric speech. *American Journal of Speech-Language Pathology, 12*(2), 198–208. https://doi.org/10.1044/1058-0360(2003/066)

Kain, A., Amano-Kusumoto, A., & Hosom, J.-P. (2008). Hybridizing conversational and clear speech to determine the degree of contribution of acoustic features to intelligibility. *The Journal of the Acoustical Society of America, 124*(4), 2308–2319. https://doi.org/10.1121/1.2967844

Kempler, D., & Van Lancker, D. (2002). Effect of speech task on intelligibility in dysarthria: A case study of Parkinson's disease. *Brain and Language, 80*(3), 449–464. https://doi.org/10.1006/brln.2001.2602

Kent, R. D., Kent, J. F., Weismer, G., & Duffy, J. R. (2000). What dysarthrias can tell us about the neural control of speech. *Journal of Phonetics, 28*(3), 273–302. https://doi.org/10.1006/jpho.2000.0122

Kim, Y., & Kuo, C. (2012). Effect of level of presentation to listeners on scaled speech intelligibility of speakers with dysarthria. *Folia Phoniatrica et Logopaedica, 64*(1), 26–33. https://doi.org/10.1159/000328642

Kuo, C., & Tjaden, K. (2016). Acoustic variation during passage reading for speakers with dysarthria and healthy controls. *Journal of Communication Disorders, 62,* 30–44. https://doi.org/10.1016/j.jcomdis.2016.05.003

Kuo, C., Tjaden, K., & Sussman, J. E. (2014). Acoustic and perceptual correlates of faster-than-habitual speech produced by speakers with Parkinson's disease and multiple sclerosis.

Journal of Communication Disorders, 52, 156–169. https://doi.org/10.1016/j.jcomdis.2014.09.002
Lam, J., & Tjaden, K. (2016). Clear speech variants: An acoustic study in Parkinson's disease. *Journal of Speech, Language, and Hearing Research, 59*(4), 631–646. https://doi.org/10.1044/2015_JSLHR-S-15-0216
Lam, J., Tjaden, K., & Wilding, G. (2012). Acoustics of clear speech: Effect of instruction. *Journal of Speech, Language, and Hearing Research, 55*(6), 1807–1821. https://doi.org/10.1044/1092-4388(2012/11-0154)
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
Lansford, K. L., Borrie, S. A., & Bystricky, L. (2016). Use of crowdsourcing to assess the ecological validity of perceptual-training paradigms in dysarthria. *American Journal of Speech-Language Pathology, 25*(2), 233–239. https://doi.org/10.1044/2015_AJSLP-15-0059
Lee, J., Hustad, K. C., & Weismer, G. (2014). Predicting speech intelligibility with a multiple speech subsystems approach in children with cerebral palsy. *Journal of Speech, Language, and Hearing Research, 57*(5), 1666–1678. https://doi.org/10.1044/2014_JSLHR-S-13-0292
Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2020). *Emmeans: Estimated marginal means, aka least-squares means* (R package version 1.5.1.) [Computer software]. https://github.com/rvlenth/emmeans
Liss, J. M. (2007). The role of speech perception in motor speech disorders. In G. Weismer (Ed.), *Motor speech disorders: Essays for Ray Kent* (pp. 187–219).
McAuliffe, M. J., Kerr, S. E., Gibson, E. M. R., Anderson, T., & LaShell, P. J. (2014). Cognitive–perceptual examination of remediation approaches to hypokinetic dysarthria. *Journal of Speech, Language, and Hearing Research, 57*(4), 1268–1283. https://doi.org/10.1044/2014_JSLHR-S-12-0349
McHenry, M. A. (2003). The effect of pacing strategies on the variability of speech movement sequences in dysarthria. *Journal of Speech, Language, and Hearing Research, 46*(3), 702–710. https://doi.org/10.1044/1092-4388(2003/055)
Milenkovic, P. (2005). *TF32* [Computer software]. University of Winsconsin–Madison.
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
Mollaei, F., Shiller, D. M., Baum, S. R., & Gracco, V. L. (2019). The relationship between speech perceptual discrimination and speech production in Parkinson's disease. *Journal of Speech, Language, and Hearing Research, 62*(12), 4256–4268. https://doi.org/10.1044/2019_JSLHR-S-18-0425
Molloy, D. W. (1999). *Standardized Mini-Mental State Examination*. New Grange Press.
Moya-Galé, G., Goudarzi, A., Bayés, À., McAuliffe, M., Bulté, B., & Levy, E. S. (2018). The effects of intensive speech treatment on conversational intelligibility in Spanish speakers with Parkinson's disease. *American Journal of Speech-Language Pathology, 27*(1), 154–165. https://doi.org/10.1044/2017_AJSLP-17-0032
Neel, A. T. (2008). Vowel space characteristics and vowel identification accuracy. *Journal of Speech, Language, and Hearing Research, 51*(3), 574–585. https://doi.org/10.1044/1092-4388(2008/041)
Neel, A. T. (2009). Effects of loud and amplified speech on sentence and word intelligibility in Parkinson disease. *Journal of Speech, Language, and Hearing Research, 52*(4), 1021–1033. https://doi.org/10.1044/1092-4388(2008/08-0119)
Nightingale, C., Swartz, M., Ramig, L. O., & McAllister, T. (2020). Using crowdsourced listeners' ratings to measure speech changes in hypokinetic dysarthria: A proof-of-concept study. *American Journal of Speech-Language Pathology, 29*(2), 873–882. https://doi.org/10.1044/2019_AJSLP-19-00162
Nishio, M., & Niimi, S. (2001). Speaking rate and its components in dysarthric speakers. *Clinical Linguistics & Phonetics, 15*(4), 309–317. https://doi.org/10.1080/02699200010024456
Nishio, M., & Niimi, S. (2006). Comparison of speaking rate, articulation rate and alternating motion rate in dysarthric speakers. *Folia Phoniatrica et Logopaedica, 58*(2), 114–131. https://doi.org/10.1159/000089612
Palan, S., & Schitter, C. (2018). Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance, 17,* 22–27. https://doi.org/10.1016/j.jbef.2017.12.004
Patel, R., Connaghan, K., Franco, D., Edsall, E., Forgit, D., Olsen, L., Ramage, L., Tyler, E., & Russell, S. (2013). "The caterpillar": A novel reading passage for assessment of motor speech disorders. *American Journal of Speech-Language Pathology, 22*(1), 1–9. https://doi.org/10.1044/1058-0360(2012/11-0134)
Peirce, J., & MacAskill, M. (2018). *Building experiments in PsychoPy*. SAGE.
Picheny, M. A., Durlach, N. I., & Braida, L. D. (1986). Speaking clearly for the hard of hearing II. *Journal of Speech and Hearing Research, 29*(4), 434–446. https://doi.org/10.1044/jshr.2904.434
Pinto, S., Chan, A., Guimarães, I., Rothe-Neves, R., & Sadat, J. (2017). A cross-linguistic perspective to the study of dysarthria in Parkinson's disease. *Journal of Phonetics, 64,* 156–167. https://doi.org/10.1016/j.wocn.2017.01.009
R Core Team. (2019). *R: A language and environment for statistical computing* (3.6.2) [Computer software]. R Foundation for Statistical Computing. https://www.r-project.org/
Skodda, S., Grönheit, W., & Schlegel, U. (2012). Impairment of vowel articulation as a possible marker of disease progression in Parkinson's disease. *PLOS ONE, 7*(2), e32132. https://doi.org/10.1371/journal.pone.0032132
Skodda, S., & Schlegel, U. (2008). Speech rate and rhythm in Parkinson's disease. *Movement Disorders, 23*(7), 985–992. https://doi.org/10.1002/mds.21996
Smiljanić, R., & Bradlow, A. R. (2005). Production and perception of clear speech in Croatian and English. *The Journal of the Acoustical Society of America, 118*(3), 1677–1688. https://doi.org/10.1121/1.2000788
Smiljanić, R., & Bradlow, A. R. (2008). Temporal organization of English clear and conversational speech. *The Journal of the Acoustical Society of America, 124*(5), 3171–3182. https://doi.org/10.1121/1.2990712
Smiljanić, R., & Bradlow, A. R. (2009). Speaking and hearing clearly: Talker and listener factors in speaking style changes. *Language and Linguistics Compass, 3*(1), 236–264. https://doi.org/10.1111/j.1749-818X.2008.00112.x
Spencer, K. A., & Rogers, M. A. (2005). Speech motor programming in hypokinetic and ataxic dysarthria. *Brain and Language, 94*(3), 347–366. https://doi.org/10.1016/j.bandl.2005.01.008
Stipancic, K. L., Tjaden, K., & Wilding, G. (2016). Comparison of intelligibility measures for adults with Parkinson's disease, adults with multiple sclerosis, and healthy controls. *Journal of Speech, Language, and Hearing Research, 59*(2), 230–238. https://doi.org/10.1044/2015_JSLHR-S-15-0271

Stipancic, K. L., Yunusova, Y., Berry, J. D., & Green, J. R. (2018). Minimally detectable change and minimal clinically important difference of a decline in sentence intelligibility and speaking rate for individuals with amyotrophic lateral sclerosis. *Journal of Speech, Language, and Hearing Research, 61*(11), 2757–2771. https://doi.org/10.1044/2018_JSLHR-S-17-0366

Sussman, J. E., & Tjaden, K. (2012). Perceptual measures of speech from individuals with Parkinson's disease and multiple sclerosis: Intelligibility and beyond. *Journal of Speech, Language, and Hearing Research, 55*(4), 1208–1219. https://doi.org/10.1044/1092-4388(2011/11-0048)

Tjaden, K., & Martel-Sauvageau, V. (2017). Consonant acoustics in Parkinson's disease and multiple sclerosis: Comparison of clear and loud speaking conditions. *American Journal of Speech-Language Pathology, 26*(2S), 569–582. https://doi.org/10.1044/2017_AJSLP-16-0090

Tjaden, K., Sussman, J. E., & Wilding, G. E. (2014). Impact of clear, loud, and slow speech on scaled intelligibility and speech severity in Parkinson's disease and multiple sclerosis. *Journal of Speech, Language, and Hearing Research, 57*(3), 779–792. https://doi.org/10.1044/2014_JSLHR-S-12-0372

Tjaden, K., & Wilding, G. (2011). Effects of speaking task on intelligibility in Parkinson's disease. *Clinical Linguistics & Phonetics, 25*(2), 155–168. https://doi.org/10.3109/02699206.2010.520185

Tjaden, K., & Wilding, G. E. (2004). Rate and loudness manipulations in dysarthria. *Journal of Speech, Language, and Hearing Research, 47*(4), 766–783. https://doi.org/10.1044/1092-4388(2004/058)

Tsao, Y. C., & Weismer, G. (1997). Interspeaker variation in habitual speaking rate. *Journal of Speech, Language, and Hearing Research, 40*(4), 858–866. https://doi.org/10.1044/jslhr.4004.858

Tsao, Y. C., Weismer, G., & Iqbal, K. (2006). Interspeaker variation in habitual speaking rate: Additional evidence. *Journal of Speech, Language, and Hearing Research, 49*(5), 1156–1164. https://doi.org/10.1044/1092-4388(2006/083)

Uchanski, R. M. (2005). Clear speech. In D. B. Pisoni & R. E. Remez (Eds.), *The handbook of speech perception* (pp. 207–235). Blackwell. https://doi.org/10.1002/9780470757024.ch9

Utianski, R. L., Lansford, K. L., Liss, J. M., & Azuma, T. (2011). The effects of topic knowledge on intelligibility and lexical segmentation in hypokinetic and ataxic dysarthria. *Journal of Medical Speech-Language Pathology, 19*(4), 25–36.

Vaden, K. I., Halpin, H. R., & Hickok, G. S. (2009). *The Irvine Phonotactic Online Dictionary, Version 2.0*. [Data file]. http://www.iphod.com

van Brenk, F., Kain, A., & Tjaden, K. (2021). Investigating acoustic correlates of intelligibility gains and losses during slowed speech: A hybridization approach. *American Journal of Speech-Language Pathology, 30*(3S), 1343–1360. https://doi.org/10.1044/2021_AJSLP-20-00172

Van Nuffelen, G., De Bodt, M., Wuyts, F., & Van de Heyning, P. (2009). The effect of rate control on speech rate and intelligibility of dysarthric speech. *Folia Phoniatrica et Logopaedica, 61*(2), 69–75. https://doi.org/10.1159/000208805

Van Riper, C. (1963). *Speech correction: Principles and methods*. Prentice-Hall.

Weismer, G., Jeng, J., Laures, J. S., Kent, R., & Kent, J. F. (2001). Acoustic and intelligibility characteristics of sentence production in neurogenic speech disorders. *Folia Phoniatrica et Logopaedica, 53*(1), 1–18. https://doi.org/10.1159/000052649

Weismer, G., Yunusova, Y., & Bunton, K. (2012). Measures to evaluate the effects of DBS on speech production. *Journal of Neurolinguistics, 25*(2), 74–94. https://doi.org/10.1016/j.jneuroling.2011.08.006

Yoho, S. E., & Borrie, S. A. (2018). Combining degradations: The effect of background noise on intelligibility of disordered speech. *The Journal of the Acoustical Society of America, 143*(1), 281–286. https://doi.org/10.1121/1.5021254

Yorkston, K. M. (2010). *Management of motor speech disorders in children and adults* (3rd ed.). Pro-Ed.

Yorkston, K. M., & Beukelman, D. R. (1981). Communication efficiency of dysarthric speakers as measured by sentence intelligibility and speaking rate. *Journal of Speech and Hearing Disorders, 46*(3), 296–301. https://doi.org/10.1044/jshd.4603.296

Yorkston, K. M., Beukelman, D. R., & Tice, R. (1996). *Sentence Intelligibility Test* [Measurement instrument]. Tice Technology Services.

Yorkston, K. M., Hakel, M., Beukelman, D. R., & Fager, S. (2007). Evidence for effectiveness of treatment of loudness, rate, or prosody in dysarthria: A systematic review. *Journal of Medical Speech-Language Pathology, 15*(2), 11–36.

Appendix

The Hunter Script (Crystal & House, 1982) With the Three Operationally Defined Fragments (Indicated in Bold)

In late spring and early June, short rays of the sun call a true son out-of-doors back to the places of his childhood. Tom Brooks was such a man**. Each year, his desk seemed like a stone whose weight made him wish for the life he knew as a boy (Fragment 1; beginning)**. In the years since leaving college, he had not revisited his past haunts before. But this March, Tom found himself by a stream with an apple, some cash, and a gun at rest in the crook of his arm. The steady desk that had needed to be set apart was gone and his one thought was for quail. **He had been jogging on the foggy trail since dawn, but not one bird had crossed his path (Fragment 2; middle).** It seemed as though five years without hunting had made him lose touch with signs that he once knew – signs that would tell for sure if an animal was near or not. Once he thought he saw birds feeding, but it was just garbage and a leaf from a rosebush up above that had failed to drop to the ground during winter. Tom stopped to rest on jagged rocks. Soon after he put down his gun, he heard wings making sounds from across the stream, and saw pale birds dash out of the brush. They flew to the edge of the stream, seemingly appalled at the hunter. **Tom placed his hand on his gun quietly. He raised it to his shoulder and took aim (Fragment 3; end).** Seconds ticked off abruptly, but the birds drank like piggies. Quick shots rang out. Years of waiting seemed to disappear with the successful culmination of the hunt.