ASHA American Speech-Language-Hearing Association

AJSLP

Research Article

The Effects of Speech Task on Lexical Stress in Parkinson's Disease

Andrew H. Exner,[a] Alexander L. Francis,[a] Megan K. MacPherson,[b] Meghan Darling-White,[c] and Jessica E. Huber[a]

[a] Department of Speech, Language, and Hearing Sciences, Purdue University, West Lafayette, IN [b] Department of Communication Sciences and Disorders, Central Michigan University, Mount Pleasant [c] Department of Speech, Language, and Hearing Sciences, The University of Arizona, Tucson

ARTICLE INFO

Article History:
Received June 9, 2022
Revision received September 19, 2022
Accepted October 23, 2022

Editor-in-Chief: Katherine C. Hustad
Editor: Yunjung Kim

https://doi.org/10.1044/2022_AJSLP-22-00185

ABSTRACT

Purpose: Hypokinetic dysarthria associated with Parkinson's disease (PD) is characterized by dysprosody, yet the literature is mixed with respect to how dysprosody affects the ability to mark lexical stress, possibly due to differences in speech tasks used to assess lexical stress. The purpose of this study was to compare how people with and without PD modulate acoustic dimensions of lexical stress—fundamental frequency, intensity, and duration—to mark lexical stress across three different speech tasks.
Method: Twelve individuals with mild-to-moderate idiopathic PD and 12 age- and sex-matched older adult controls completed three speech tasks: picture description, word production in isolation, and word production in lists. Outcome measures were the fundamental frequency, intensity, and duration of the vocalic segments of two trochees (initial stress) and two iambs (final stress) spoken in all three tasks.
Results: There were very few group differences. Both groups marked trochees by modulating intensity and fundamental frequency and iambs by modulating duration. Task had a significant impact on the stress patterns used by both groups. Stress patterns were most differentiated in words produced in isolation and least differentiated in lists of words.
Conclusions: People with PD did not demonstrate impairments in the production of lexical stress, suggesting that dysprosody associated with PD does not impact all types of prosody in the same way. However, there were reduced distinctions in stress marking that were more apparent in trochees than iambs. In addition, the task used to assess prosody has a significant effect on all acoustic measures. Future research should focus on the use of connected speech tasks to obtain more generalizable measures of prosody in PD.

One of the key features of the hypokinetic dysarthria associated with Parkinson's disease (PD) is dysprosody (Darley et al., 1969). The dysprosody associated with PD is characterized by monopitch (MacPherson et al., 2011; Rusz et al., 2011; Skodda et al., 2009; Tykalova et al., 2014), monoloudness (Rusz et al., 2011; Tykalova et al., 2014), and variable rate abnormalities (Tykalova et al., 2014; Watson & Munson, 2008). Dysprosody often leads to social withdrawal by people with PD (Miller et al., 2006), listener confusion (Pell et al., 2006), and reduced intelligibility (Lam & Tjaden, 2016).

While there has been ample research into the perceptual characterization of monopitch and monoloudness in the speech of people with PD, little is known about how they mark lexical stress. Lexical stress refers to the tendency for syllables within multisyllabic words to have different salience, typically realized by differential stress in pronunciation (Cutler, 2015). In two-syllable words, trochees have initial stress (with emphasis on the first syllable,

Correspondence to Jessica E. Huber: jhuber@purdue.edu. **Disclosure:** *Jessica E. Huber owns shares in SpeechVive, Inc., the company that manufactures and sells a device she has invented. Her interests have been reviewed and are managed by Purdue University according to the university's conflict-of-interest policy. She also serves as a member of the Medical Advisory Board for Rock Steady Boxing in Indianapolis, IN. All other authors have declared that no competing financial or non-financial interests existed at the time of publication.*

e.g., "COFfee"), and iambs have final stress (with emphasis on the second syllable, e.g., "balLOON"). Although there are some words in Standard American English that differ primarily by syllabic stress (e.g., the trochaic noun "FOREbear" and the iambic verb "forBEAR"), the primary purpose of lexical stress for listeners is less word recognition and more boundary recognition (Cutler et al., 1997). A word's stress pattern is part of its lexical representation, but most English words tend to have initial stress, such as trochees, and therefore, stress often indicates to listeners the presentation of a new word (Cutler, 2015). Word recognition can be delayed if words are stressed incorrectly, especially if the mis-stress occurs on the final syllable (Cutler & Clifton, 1984). For people with PD and their conversation partners, unexpected differences in stress patterns could cause communication breakdowns.

The prototypical acoustic pattern of lexical stress in healthy speakers of Standard American English is to increase fundamental frequency (f_o), intensity, and duration for the vocalic segment to be emphasized (Fry, 1955, 1958; Lieberman, 1960). People also tend to favor different cues to produce trochees than iambs, with increased f_o and intensity being prioritized for stress in trochees and durational contrasts being emphasized for stress in iambs (Walker et al., 2009). Furthermore, in a study of three- and four-syllable words, the magnitude of pitch contrasts across the first and second syllables was lowest while duration contrasts were largest (Kim et al., 2010). Thus, rather than using these cues equally in all contexts, cue trading is common for both healthy speakers (Lieberman, 1960) and speakers with dysarthria (Kim et al., 2010; Patel & Campellone, 2009). Cue trading refers to the process of using different combinations of cues, essentially trading one prosodic feature for another, which can be judged by a listener to be relatively equivalent (Howell, 1993; Lieberman, 1960). For instance, due to a motor bias of final lengthening, people tend to produce the unstressed final syllables of trochees with a longer duration than the stressed initial syllable (Snow, 1994). In this case, the durational cue is counter to what is expected of initial stress. Therefore, speakers may need to rely on a different cue, such as f_o or intensity, to clearly mark a trochee as a trochee (Goffman & Malin, 1999).

The use of these cues develops and changes with age and experience, although more is known about typical development than typical aging. Children show a strong bias toward final lengthening regardless of a word's stress pattern, which reduces but does not disappear with maturity (Goffman & Malin, 1999). As adults age, their syllables tend to be 20%–25% longer than those of young adults (Smith et al., 1987). Duration differences between syllables tend to be exaggerated in iambs as compared to trochees, and young adults rely on duration cues more than older adults when producing both iambs and trochees (Barnes, 2013). There may be a gender difference in the use of these cues as well, as older women are more likely than other groups to use f_o to differentiate between the syllables of trochees (Scukanec et al., 1996). In addition, lexical stress cues are often maintained from a speaker's native language when they learn a second language so that their spoken second-language stress patterns more closely match their spoken first-language patterns if there is a conflict (Zuraiq & Sereno, 2007).

There is ample evidence that stress-marking behavior is affected in neurological disabilities or injuries, from cerebral palsy (Kim et al., 2010) to traumatic brain injury (McHenry, 1998; Wang et al., 2005) and aphasia secondary to left-hemisphere cerebrovascular accident (Walker et al., 2009). This provides further credence to the current understanding that prosodic function is implicated in a variety of neural structures. Speakers with neurological diseases have generally been found to be able to mark lexical stress by modulating different acoustic cues, although the magnitude of the contrast across syllables is generally reduced compared to controls (Kim et al., 2010; McHenry, 1998; Wang et al., 2005).

As with other neurological conditions, PD may also change the use of cues in stress-marking behavior. Darkins et al. (1988) assessed the ability of people with PD to use f_o contour and pause times between syllables to differentiate between noun compounds (e.g., "BLACKboard," a trochee with initial stress) and noun phrases (e.g., "black BOARD," functionally an iamb with final stress) in a sentence-reading task. Whereas healthy controls marked noun compounds (trochaic stress) with a difference in the f_o contour and noun phrases (iambic stress) with a difference in pause duration, people with PD produced neither differences in f_o contour nor pauses between syllables to show a contrast between the noun compounds and phrases. This suggests a significant impairment in marking lexical stress related to PD. However, it is notable that, given known rate and pausing abnormalities in the speech of people with PD (Huber & Darling, 2011; Lowit et al., 2006; Skodda & Schlegel, 2008), an impaired pause time would be expected.

By contrast, Cheang and Pell (2007) found no significant difference between people with PD and healthy controls in their use of f_o to differentiate eight pairs of noun compounds (e.g., "HOTdog") from noun phrases (e.g., "hot DOG") following declarative carrier phrases (i.e., "this is a __") paired with pictures. However, people with PD produced smaller contrasts across syllables than controls for duration and intensity cues, suggesting a mild prosodic impairment. A listener study (Pell et al., 2006) conducted on the lexical stress data collected by Cheang and Pell showed that naïve listeners had a more challenging time distinguishing between the noun phrases and noun compounds produced by people with PD than those produced by healthy controls. This suggests that minor differences in stress patterns can lead to difficulties understanding stress produced by a person with PD. These limited studies showed mixed results about the impact of PD on a person's ability to use f_o, intensity, and duration to differentiate trochees and iambs.

Other studies have assessed contrastive stress-marking behavior in phrases (i.e., emphasizing different words, e.g., "SHE took it" it vs. "she TOOK it") and found mixed results. Darkins et al. (1988), Ma et al. (2015), and Rusz et al. (2011) found reduced use of f_o by people with PD to mark contrastive stress. However, Hertrich and Ackerman (1993), Lowit-Leuschel and Docherty (2001), and Cheang and Pell (2007) found that f_o marking remains intact in people with PD relative to controls. Similarly, Cheang and Pell and Rusz et al. showed reduced use of intensity to mark contrastive stress in people with PD, but Hertrich and Ackerman and Lowit-Leuschel and Docherty indicate that intensity marking is intact in people with PD relative to controls. All referenced studies have found that measures of duration (e.g., pause behavior and syllable duration) are altered (i.e., significantly longer or shorter) in people with PD relative to controls.

One possible reason for the discrepancies among studies is the variety of speech tasks used to assess lexical and contrastive stress patterns: sentence reading (Darkins et al., 1988), carrier phrases in response to a picture (Cheang & Pell, 2007), passage reading with typographical emphasis (Rusz et al., 2011) or without typographical emphasis (Lowit-Leuschel & Docherty, 2001; Ma et al., 2015), sentence repetition (Hertrich & Ackermann, 1993), monologue (Ma et al., 2015; Rusz et al., 2011), and conversation (Lowit-Leuschel & Docherty, 2001). Task differences, in general and for people with PD specifically, are common, and the nature of the task changes multiple acoustic correlates of phrasal prosody. For example, if a person is speaking a list of words, intonational stress may override the typical lexical pattern of individual words. That is, even though a trochaic word is canonically marked with stress on the initial syllable (which typically has a higher f_o), it is common for items in a list to have a rising (or flat) nuclear tone for all but the final item in the list (Couper-Kuhlen, 1986; Schubiger, 1958; Selting, 2007). This represents a violation of the expected lexical pattern due to phrasal intonation. In spontaneous speech, sentence and question intonation may also drive differences in the realization of canonical lexical stress patterns (Ladd, 2008).

Task may also affect how people with PD use duration to mark stress due to inherent speech rate differences across tasks. However, the literature has been mixed, with studies reporting similarities and differences in articulation rate across tasks in both people with PD and controls. Lowit et al. (2006) found that people with PD and controls had similar articulation rates during both passage- and sentence-reading tasks; however, speakers with both PD and cognitive decline had slower articulation rates. Skodda and Schlegel (2008) found no significant difference in the overall articulation rates of people with PD and controls in a reading passage but did find more rushes of speech for people with PD. Conversely, Huber and Darling (2011) found that both people with PD and controls tend to produce longer utterances and speak with a slower rate during monologue tasks versus in reading tasks, but that people with PD have a faster speaking rate than controls during spontaneous speech.

A third possible reason for discrepancies across studies is differences in the speech severity of participants in those studies. Two of the studies cited so far provided little to no information about the severity of the dysarthria of their participants (Darkins et al., 1988; Rusz et al., 2011), two studied a group of participants with no to mild dysarthria (Cheang & Pell, 2007; Hertrich & Ackermann, 1993), and two included participants along the continuum from mild to severe dysarthria (Lowit-Leuschel & Docherty, 2001; Ma et al., 2015). Both Lowit-Leuschel and Docherty (2001) and Ma et al. (2015), who included a wide range of dysarthria severity, found variable group performance within each severity group and concluded that overall speech severity is not a reliable predictor of prosodic impairment. Furthermore, Ma et al. found no interaction effect between speech severity and speech stimulus.

The first aim of this study was to compare how people with and without PD modulate acoustic dimensions of lexical stress—duration, f_o, and intensity—to mark iambs and trochees. We hypothesized that people with PD would have smaller between-syllables differences in f_o, intensity, and duration compared to age- and sex-matched controls. However, to better understand dysprosody in PD, we need to compare across tasks to understand the effects of task on prosody. Thus, the second aim of this study was to compare lexical stress marking across three different speech tasks: word production in isolation, word production in lists, and a connected speech task (map description). All tasks used pictures to support production; none of the tasks used reading for elicitation. We hypothesized that all participants would demonstrate smaller lexical stress differentiation on the map description speech task compared to the isolated word and list production tasks, with a significant Group × Task interaction effect showing that PD-specific prosodic deficits may be exacerbated by task-specific effects.

Method

Participants completed all procedures under a protocol approved by the Purdue University Institutional Review Board (Protocol #0705005388).

Participants

Twelve individuals (six women, six men) diagnosed with idiopathic PD and 12 age- and sex-matched controls participated in this study. Age matching was targeted to be within 4 years. Participants with PD were required to be diagnosed with PD by a neurologist. Participants were required to be over 60 years old. As shown in Table 1, participants with PD were 68–85 years old (M = 76.3, SD = 5.13), and controls were 66–85 years old (M = 77.1, SD = 5.40). A t test on the ages of the participants showed no significant differences in age between groups (t = −0.3587, p = .723). Participants were required to speak a North Midland dialect of American English. Exclusion criteria were a history of formal vocal training, respiratory illnesses other than allergies (which needed to be controlled at the time of testing), head and neck cancer or surgery, neurological disease other than PD, and smoking within the last 5 years. Participants were required to be free from infection on the day of testing. All participants passed a hearing screening at 40 dB at 500 and 1000 Hz, but three participants with PD did not pass at 2000 Hz. Their response thresholds ranged from 50 to 60 dB at 2000 Hz. Most participants with PD were taking anti-parkinsonian medications, and they participated in the study within 1–3 hr of taking their medications to account for potential on–off effects. One participant had a deep brain stimulation implant and was assessed while the implant was turned on. Several participants had enrolled in speech therapy prior to the study, though none were enrolled at the time of the study. All participants were given the Cognitive Linguistic Quick Test (CLQT) prior to data collection: Older controls needed to score within normal limits to continue the study, but people with PD could have mild-to-moderate impairments and continue in the study.

Two speech-language pathologists (SLPs), unaffiliated with the study, with experience diagnosing and treating adults with motor speech disorders rated the speech severity of all participants using a visual analog scale with anchors from *normal* to *very severe*. They listened to a 30-s speech sample from the middle of a monologue produced by the participant. Samples were clipped so that none of the samples started or ended in the middle of an utterance. Samples intensity-normalized normalized to 70 dBA and presented over headphones at a comfortable intensity (set prior to the start of the rating task). The

Table 1. Participant demographic information.

Participant	Age (years; months)	Years since diagnosis	CLQT	Speech severity (%)	Hx of speech Tx	PD-related [or depression] medication
Parkinson's disease						
F1	72;2	11	WNL (4.0)	43.0	Low volume	Eldepryl, Sinemet, [Wellbutrin, Zoloft]
F2	76;11	7	WNL (4.0)	11.4	No	Zelapar, Bromocriptine, Sinemet
F3	82;10	15	Mild (2.8)	38.2	No	None
F4	75;4	5	WNL (4.0)	11.2	No	None
F5	75;11	1	WNL (4.0)	10.0	No	None
F6	68;9	3	WNL (4.0)	4.0	No	None
M1	73;5	9	WNL (4.0)	82.0	Low volume, pitch breaks	Sinemet, Requip
M2	76;8	13	Moderate (1.8)	73.0	Speech clarity	Sinemet, Aricept, [Prozac]
M3	73;7	9	Mild (3.4)	8.9	Word finding	Sinemet, Mirtazapine, Donepezil, Aricept, [Remeron]
M4	85;2	7	WNL (3.8)	35.5	No	Amantadine, Lodosyn, Sinemet
M5	69;10	4	WNL (3.8)	20.6	No	None
M6	84;2	2	WNL (3.6)	31.4	No	None

Participant	Age (years; months)	CLQT	Speech severity (%)
Controls			
F1	78;6	WNL (4.0)	4.7
F2	77;1	WNL (4.0)	7.1
F3	69;7	WNL (4.0)	1.0
F4	77;2	WNL (3.6)	9.1
F5	76;9	WNL (4.0)	2.5
F6	83;3	WNL (4.0)	2.3
M1	74;1	WNL (4.0)	2.0
M2	74;0	WNL (4.0)	7.5
M3	85;6	WNL (4.0)	0.3
M4	84;7	WNL (4.0)	3.7
M5	77;3	WNL (4.0)	3.0
M6	66;8	WNL (4.0)	7.2

Note. Higher speech severity indicates more severe speech rating. CLQT = Cognitive Linguistic Quick Test; Hx = history; Tx = therapy; PD = Parkinson's disease; WNL = within normal limits.

SLPs were instructed to listen to the speech sample one time and then provide a rating on the visual analog scale. The raters were blinded to the disease status (PD or control). The distance from normal was measured and converted to a percentage of the entire scale so that higher numbers reflect more severe speech impairments. The two ratings were averaged to produce a final severity score. However, if there was a greater than 20% difference between the ratings, a third SLP (J. E. Huber) rated the sample. The average of this third rating and the closer of the original two ratings was used as the final severity score. Three individuals were rated by co-author J. E. Huber. Demographic and medication information, speech severity ratings, history of speech therapy, and CLQT scores are presented for each participant in Table 1.

Equipment

Speech samples were transduced using a condenser microphone with a flat-frequency response from 50 Hz to 20 kHz (Brüel & Kjær 4936). The microphone was held at a constant distance of 6 in. from the speaker's mouth, placed at a 45° angle. A sound-level meter (Quest 1700) was coupled to the microphone, but not in view of the participant. The acoustic signal from the microphone was recorded to a digital audio tape (TASCAM DA-P1) and then digitized for computer analysis using Praat (Boersma & Weenink, 2019). The audio signal was recorded at 44.1 kHz, resampled at 18 kHz, and then low-pass filtered at 9 kHz. The microphone was calibrated using a sound of known intensity and recorded to the digital audio tape before the speech tasks for each participant.

Speech Tasks

Participants produced lexical stress in three tasks. The first task was an isolated word production task, in which participants named a picture on a computer monitor that was presented with the item name labeled. The second task was a list production task (see Appendix A), in which four pictures were displayed together and the participant named all four pictures as a list. The same pictures were used for the isolated word production and list production tasks, and participants were familiarized to the pictures and the associated target words prior to the start of data collection. The target words selected for analysis from the list production were always in the second or third position in the list to avoid speech differences observed in list-initial and list-final words. For list production, participants were instructed to string the words together as a list, rather than producing each word individually, to help ensure list intonation. The researcher prompted participants to try again if pauses between words in the list were subjectively too long. In the third task, participants described a map of a town (see Appendix B) and a map of a zoo (see Appendix C), each displayed on a computer monitor. For both maps, the participant described what they saw as if they were walking around the location. The instructions were as follows:

> "Here is a map of a town. Please look at the map and tell me the place you could go and the things you could see in the town. Also, please tell me what you could do or purchase in each place."

> "Here is a map of a zoo. Pretend that you are going to spend a day at the zoo. Please tell me what you could do and the things you could see."

Participants were allowed time (as long as they wanted) to study the map before beginning, but most started in less than a minute. The map description task elicited connected, spontaneous speech. Task order was kept consistent for all participants. Participants completed additional speech tasks during the session that are not included in the present analysis.

Four target words were extracted as the focus of the study: two trochees (initial stress: "coffee" and "tiger") and two iambs (final stress: "balloons" and "giraffe"). In the reporting of results, the stressed initial syllables of the two trochees were the [ˈkɔ] of "coffee" and [ˈtɑɪ] of "tiger," while the unstressed final syllables of the two trochees were the [ˌfi] of "coffee" and [ˌgɚ] of "tiger." Similarly, the unstressed initial syllables of the two iambs were the [ˌbə] of "balloons" and [ˌdʒɚ] of "giraffe," while the stressed final syllables of the two iambs were the [ˈlunz] of "balloons" and [ˈɹæf] of "giraffe." These words were chosen because they were consistently produced in all three tasks by all subjects. One production of each word from each task was used in data analysis. One male participant with PD requested that he stop the list production task prior to completion, so only 25% of the data were collected from him for this task. There are no other missing data.

Measurements

Praat Version 6.1 (Boersma & Weenink, 2019) was used (with standard wideband spectrogram settings) to obtain three acoustic measures for the vocalic segment of each syllable in the target words: mean f_o, mean intensity, and vocalic segment duration. The onsets and offsets of vocalic segments in syllables were identified through the appearance of the voicing bar, along with other consonant-specific criteria. Specific vocalic segment marking criteria are identified below.

For "coffee," the onset of the first vocalic segment /ɔ/ was identified as the beginning of the voicing bar following the plosive, and the offset of the first vocalic segment was demarked by the onset of the fricative /f/, with low-intensity broad-spectrum noise. The onset of the second vocalic segment /i/ was the end of the fricative /f/, and the offset was the end of the voicing bar.

For "tiger," the onset of the first vocalic segment /ɑɪ/ was identified as the beginning of the voicing bar following the initial plosive, and the offset of the first vocalic segment was identified as the velar pinch for /g/ (the closing of the second [F2] and third [F3] formants). The onset of the second vocalic segment /ɚ/ was the opening of F2 and F3, or movement away from the velar pinch, and the offset was identified as the end of the voicing bar.

For "balloons," the onset of the first vocalic segment /ə/ was identified as the end of the burst of the plosive /b/. The offset of the first vocalic segment (and the onset of the second, [lu]) was the drop in the F3 for the /l/. The offset of the second vocalic segment [lu] was the /n/, which was identified by the appearance of antiformants in the spectrogram.

For "giraffe," the onset of the first vocalic segment /ə/ was identified as the end of the high spectral energy of the affricate /dʒ/. The offset of the first vocalic segment (and the onset of the second, /æ/) was identified by where the F2 and F3 both deflected upward in the spectrogram, signifying the transition from the /ɚ/ to the /ɹ/. The offset of the second vocalic segment /æ/ was the onset of the fricative /f/.

We do not currently know the precise nature of what stress-marking cues people attend to when decoding lexical stress. In order to promote comparison with past and future research on this subject, both the raw acoustic and calculated pairwise variability index (PVI) measures were used to demonstrate the differences in productions between speakers. Duration was measured as the length (in seconds) of the vocalic segment in Praat. Praat's "get pitch" and "get intensity" functions with default settings were used respectively to measure the f_o and intensity of vocalic segments produced with typical modal voice. Words produced with breathy voice, pressed voice, vocal fry, or other vocal timbres introduced errors into the f_o calculations. Changes in vocal quality are prevalent in all older adults but are especially common in people with PD. Rather than discarding these data, strategies were implemented to correct for these f_o-tracking errors. Most errors occurred at points where vocal fry was present. To correct the f_o-tracking errors, we measured the duration of each period, identified at the zero-crossing, and calculated the f_o as the inverse of the pitch period. In these cases, the final average f_o for the vocalic segment was the average across the correct (default Praat setting) and corrected segments, weighted by the duration of each segment. In cases where there were no f_o-tracking errors, we obtained the average f_o from the whole vocalic segment as identified by the default Praat settings. Out of 570 vocalic segments, 38 needed to be corrected.

PVIs (Grabe & Low, 2002) were calculated to quantify the differences in f_o, intensity, and duration across the two vocalic segments. Using duration (D) as an example, the formula used was

$$\left|D_{\text{Syllable1}} - D_{\text{Syllable2}}\right| / \left|D_{\text{Syllable1}} + D_{\text{Syllable2}}/2\right|. \quad (1)$$

The PVI values range from 0 (no difference across the syllables) to 2 (maximal difference across the syllables).

Statistical Analysis

Mixed-model analysis of variance with repeated measures was used to determine whether there were significant effects of group (PD, controls), task (isolated, list, map description), and syllable (trochee: stressed initial, trochee: unstressed final, iamb: unstressed initial, iamb: stressed final) on the dependent variables (duration, mean f_o, mean intensity). Stress and order were collapsed into a single factor "syllable" because syllable order is constant within these stress patterns. The analysis plan was the same for the three PVI measures except that stress pattern (trochee, iamb) was used instead of syllable because PVI is a word-level, not a syllable-level, measure. Interaction effects were also examined. For significant main effects and all significant interaction effects, the Tukey's honestly significant difference tests were used to test for significant comparisons. The alpha level was set at $p < .05$. Cohen's d are reported for significant group effects. Cohen's d for repeated measures (d_{RM}), with pooled standard deviations and accounting for intercorrelations between the outcomes, are reported for all significant comparisons involving within-subject factors or interactions (Lakens, 2013).

A second researcher (uninvolved in data collection or original measurements) selected a random subset of 25% of the data (all measures for three people with PD and three controls; three men and three women) to review and independently analyze. The mean differences between the measures of the primary researcher and the secondary researcher were small. Intraclass correlation estimates and their 95% confidence intervals (CIs) were calculated using SPSS 28 based on a single-rating, absolute-agreement, two-way random-effects model. The two raters had excellent agreement on duration and intensity and good agreement on f_o, demonstrating adequate interrater reliability (see Table 2).

Results

Table 3 contains a statistical summary of the results. There were no significant three-way interaction effects on any of the dependent variables. For all results, means are

Table 2. Interrater reliability.

Variable	Original (*M*)	Reliability (*M*)	Mean difference	ICC	95% CI	Agreement
Duration of vocalic segment (s)	0.186 (0.094)	0.184 (0.078)	−0.002	.933	[.912, .949]	Excellent
Mean f_o of vocalic segment (Hz)	131.74 (38.86)	129.49 (36.72)	−2.250	.890	[.856, .916]	Good
Mean intensity of vocalic segment (dB SPL)	58.52 (6.81)	58.50 (6.85)	−0.015	.994	[.992, .996]	Excellent

Note. ICC = intraclass correlation coefficient; CI = confidence interval; f_o = fundamental frequency.

presented with standard errors in parentheses (e.g., mean (standard deviation)). Cohen's d_{RM} are presented with 95% CIs.

Duration of Vocalic Segment

For duration of vocalic segment, there was a significant main effect of syllable. There was a statistically significant interaction effect for Task × Syllable. The Group × Task and Group × Syllable interaction effects were not significant. For all tasks, the stressed final syllable in iambs (isolated: 0.329 s (0.070), list: 0.270 s (0.084), map: 0.284 s (0.094)) was longer than the unstressed initial syllable in iambs (isolated: 0.087 s (0.027), list: 0.091 s (0.024), map: 0.089 s (0.024); $p < .001$ for all tasks, isolated: d_{RM} = 2.414, 95% CI [1.665, 3.173]; list: d_{RM} = 2.233, 95% CI [1.537, 2.930]; map: d_{RM} = 2.706, 95% CI [1.970, 3.441]). However, the stressed initial syllable in trochees (isolated: 0.180 s (0.055), list: 0.183 s (0.047), map: 0.177 s (0.043)) was not significantly longer than the unstressed final syllable in trochees (isolated: 0.160 s (0.046), list: 0.163 s (0.066), map: 0.181 s (0.077); p = .867 for isolated, p = .925 for list, and $p > .999$ for map). In addition, the stressed final syllable in iambs was longer than the stressed initial syllable in trochees ($p < .001$ for all tasks, isolated: d_{RM} = 4.378, 95% CI [3.561, 5.196]; list: d_{RM} = 1.837, 95% CI [1.211, 2.463], map: d_{RM} = 1.208, 95% CI [0.602, 1.814]), and the unstressed final syllable in trochees was longer than the unstressed initial syllable in iambs ($p < .001$ for all tasks, isolated: d_{RM} = 1.009, 95% CI [0.322, 1.697]; list: d_{RM} = 1.254, 95% CI [0.425, 2.083]; map: d_{RM} = 0.957, 95% CI [0.128, 1.786]). There was also a significant across-tasks interaction effect for Task × Syllable: The stressed final syllable in iambs was longer when produced in isolation than in either the list ($p < .001$, d_{RM} = 0.880, 95% 95% CI [0.285, 1.474]) or map description (p = .013, d_{RM} = 0.571, 95% CI [0.010, 1.153]; see Figure 1).

PVI of Duration

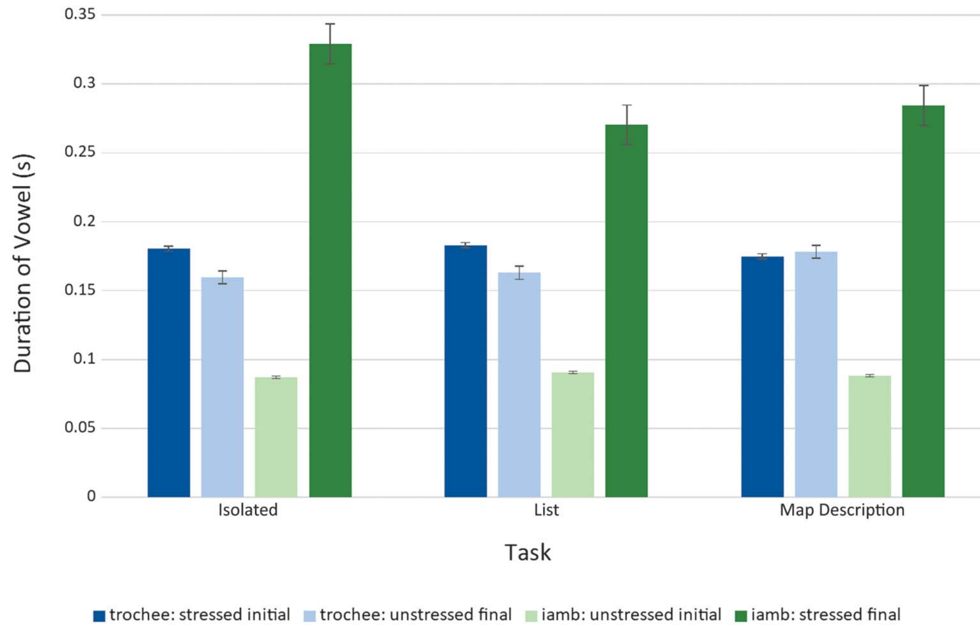
For the PVI of duration, there was a significant main effect of stress pattern and a significant interaction effect for Task × Stress Pattern. The Group × Task and Group × Stress Pattern interaction effects were not significant. For all tasks, the PVI of duration was higher in iambs (isolated: 1.155 (0.236), list: 0.965 (0.266), map: 1.023 (0.234)) than trochees (isolated: 0.262 (0.184), list: 0.328 (0.186), map: 0.276 (0.214); $p < .001$ for all, isolated: d_{RM} = 2.361, 95% CI [1.560, 3.162]; list: d_{RM} = 1.417, 95% CI [0.755, 2.078]; map: d_{RM} = 1.690, 95% CI [0.963, 2.418]), indicating

Table 3. Statistical summary.

Main/interaction effect	Duration		PVI of duration		Mean f_o		PVI of f_o		Mean intensity		PVI of intensity	
	F	*p*	*F*	*p*	*F*	*p*	*F*	*p*	*F*	*p*	*F*	*p*
Group	0.01	.910	0.05	.825	16.75	**< .001***	15.24	**< .001***	2.60	.107	0.21	.650
Task	2.01	.135	2.38	.095	0.78	.461	18.77	**< .001***	0.27	.767	3.60	**.029***
Syllable	290.07	**< .001***			0.91	.436			7.90	**< .001***		
Stress pattern			831.17	**< .001***			40.63	**< .001***			38.81	**< .001***
Group × Task	0.81	.444	1.04	.355	0.65	.520	2.65	.073	0.75	.475	1.47	.232
Group × Syllable	0.50	.684			0.38	.765			0.17	.916		
Group × Stress			1.41	.236			4.45	**.036***			2.16	.143
Task × Syllable	4.23	**< .001***			0.16	.987			0.56	.764		
Task × Stress Pattern			7.92	**< .001***			8.63	**< .001***			2.61	.076
Group × Task × Syllable	0.86	.523			0.33	.920			0.06	.999		
Group × Task × Stress Pattern			0.37	.694			2.34	.098			0.85	.429

Note. Bolded values are significant. PVI = pairwise variability index; f_o = fundamental frequency.
*Significance level is $p < .05$.

Figure 1. Mean duration of vocalic segment by task and syllable with standard error bars.

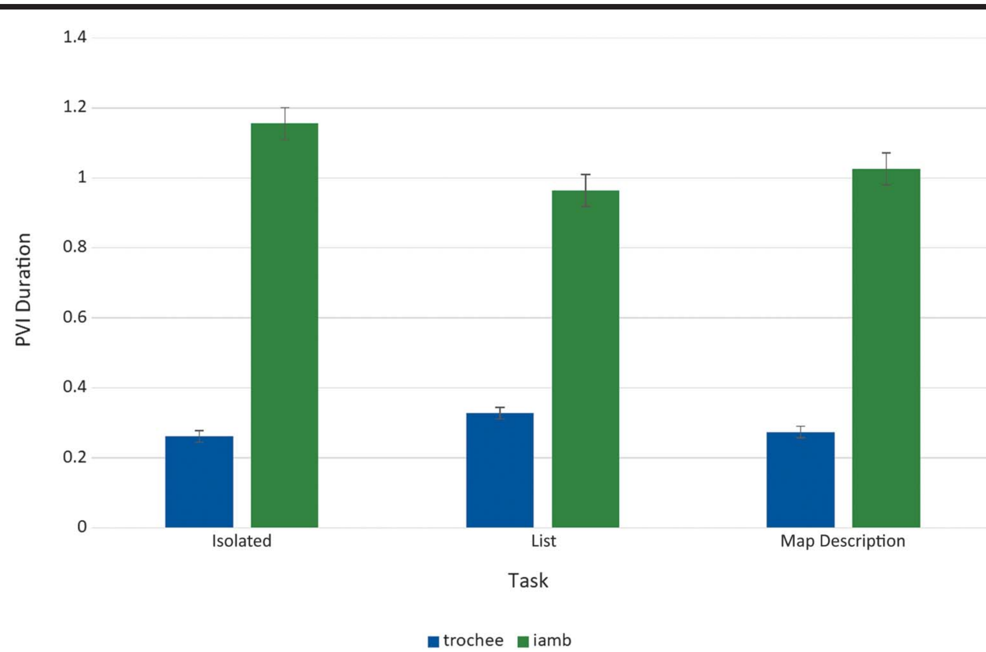


greater durational differences between syllables in iambs than trochees. There were two significant differences across tasks: The PVI of duration was higher for iambs produced in isolation than for iambs produced in both list intonation ($p < .001$, $d_{RM} = 0.473$, 95% CI [0.107, 1.053]) and the map task ($p = .043$, $d_{RM} = 0.510$, 95% CI [0.066, 1.086]; see Figure 2).

Mean f_o of Vocalic Segment

For mean f_o of vocalic segment, there was a significant main effect of group. The Group × Task, Group × Syllable, and Task × Syllable interaction effects were not significant. The mean f_o of vocalic segments was higher in people with PD (148.58 Hz (35.53)) than

Figure 2. Pairwise variability indices (PVIs) of duration of vocalic segment by task and stress pattern with standard error bars.



controls (135.40 Hz (40.31); $p < .001$, $d = 0.347$, 95% CI [0.460, 1.153]).

PVI of Mean f_o

For the PVI of mean f_o, there were significant main effects of group, task, and stress pattern. There were also significant interaction effects for Group × Stress Pattern and Task × Stress Pattern. The Group × Task interaction effect was not significant.

Regarding the Group × Stress Pattern interaction effect (see Figure 3), the PVI of f_o was higher for trochees than iambs for both people with PD (iamb: 0.065 (0.065), trochee: 0.135 (0.137); $p < .016$, $d_{RM} = 0.357$, 95% CI [0.231, 0.946]) and controls (iamb: 0.094 (0.093), trochee: 0.233 (0.245); $p < .001$, $d_{RM} = 0.581$, 95% CI [0.049, 1.212]). The PVI of f_o was higher in trochees produced by controls than in those produced by participants with PD ($p < .001$, $d = 0.484$, 95% CI [0.328, 1.295]), but there was no significant difference across groups in the PVI of f_o for iambs ($p = .582$).

Relative to the Task × Stress Pattern interaction effect (see Figure 4), the PVI of f_o was higher in trochees than iambs in both isolation (iamb: 0.081 (0.068), trochee: 0.276 (0.250); $p < .001$, $d_{RM} = 0.0819$, 95% CI [0.031, 1.606]) and map description (iamb: 0.111 (0.107), trochee: 0.198 (0.189); $p = .026$, $d_{RM} = 0.404$, 95% CI [0.179, 0.987]). However, there was no significant difference in the PVI of f_o in trochees and iambs produced with list intonation (iamb: 0.046 (0.043), trochee: 0.078 (0.083); $p = .892$). The PVI of f_o was greater in trochees produced in isolation and the map description than the syllables of trochees produced in list intonation ($p < .001$ for both; isolated vs. list: $d_{RM} = 0.641$, 95% CI [0.061, 1.22]; map vs. list: $d_{RM} = 0.531$, 95% CI [0.044, 1.107]), but there were no differences between the isolated and map tasks ($p = .067$). For iambs, there were no differences in PVI of f_o across tasks ($p = .823$ for isolated vs. list; $p = .900$ for isolated vs. map; $p = .208$ for list vs. map).

Mean Intensity of Vocalic Segment

For mean intensity of vocalic segment, there was a significant main effect of syllable. The Group × Task, Group × Syllable, and Task × Syllable interaction effects were not significant. The stressed initial syllable of trochees (59.89 dB SPL (6.46)) had a greater intensity than the unstressed final syllable of trochees (56.42 dB SPL (7.61); $p < .001$, $d_{RM} = 2.124$, 95% CI [1.513, 2.734]). However, there were no significant differences in the mean intensity of iambs between the unstressed initial (59.03 dB (7.02)) and stressed final (60.02 dB SPL (6.97)) syllables ($p = .639$). Mean intensity was higher in the unstressed initial syllable of iambs than the unstressed final syllable of trochees ($p = .011$, $d_{RM} = 0.745$, 95% CI [0.169, 1.321]). However, there were no significant differences in the mean intensity between the stressed initial syllable of trochees and the stressed final syllable of iambs ($p = .999$).

Figure 3. Pairwise variability indices (PVIs) of fundamental frequency (f_o) by group and stress pattern with standard error bars.

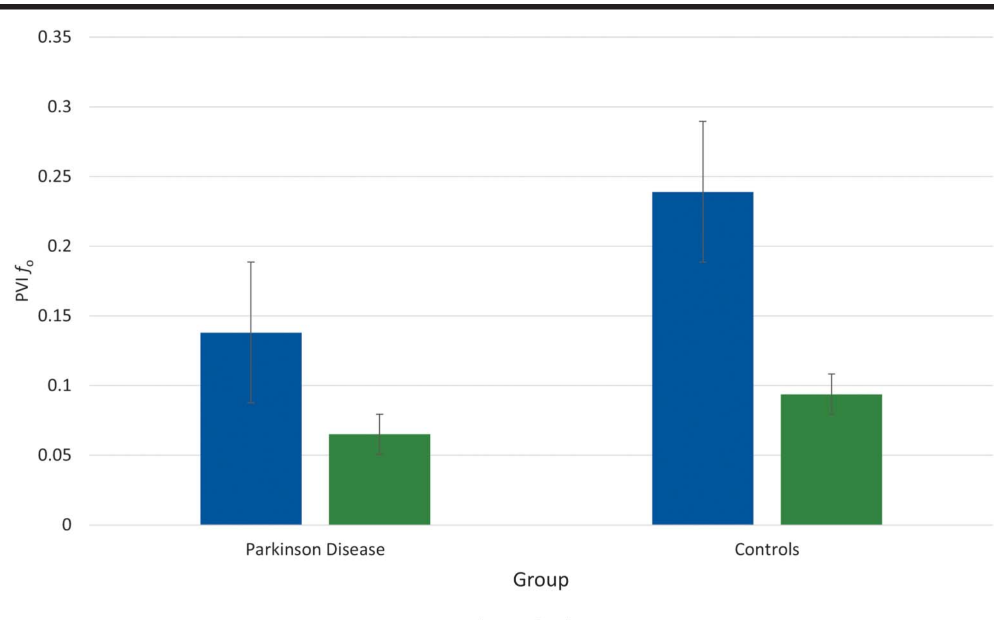
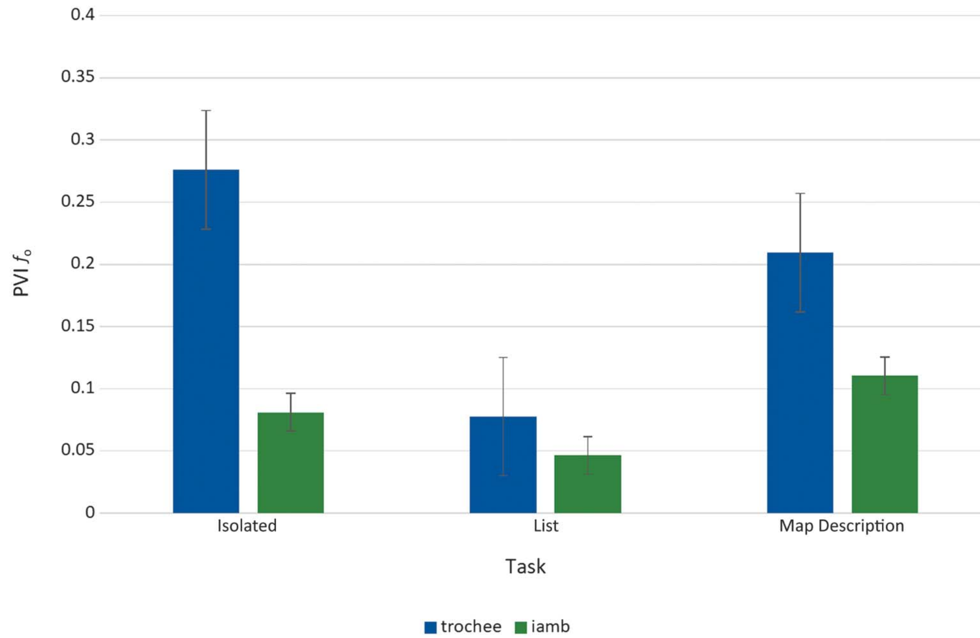


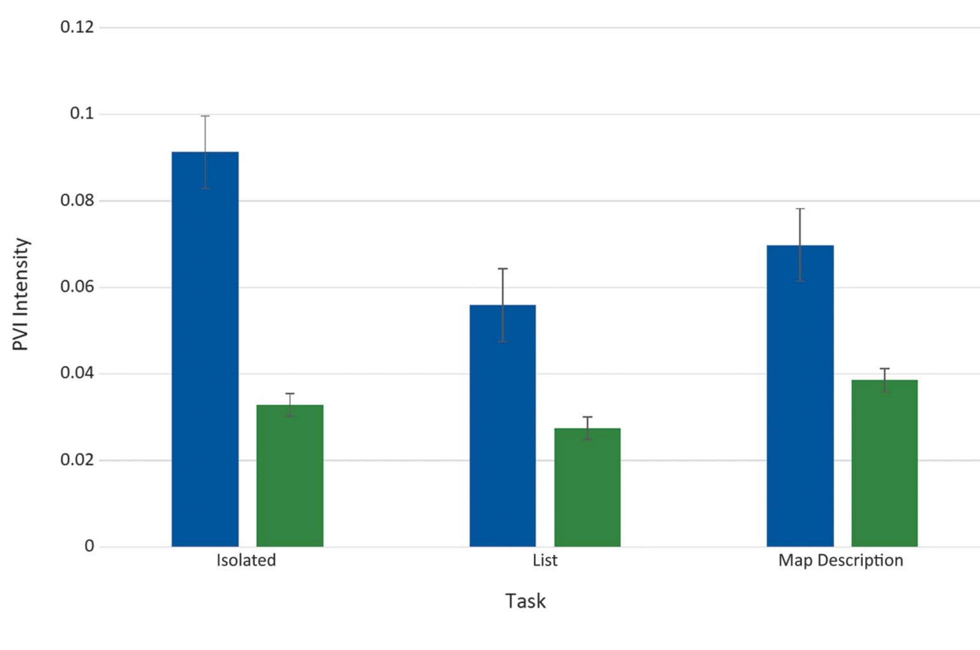
Figure 4. Pairwise variability indices (PVIs) of fundamental frequency (f_o) by task and stress pattern with standard error bars.



PVI of Mean Intensity

For the PVI of intensity, there were significant main effects of task and stress pattern. The Group × Task, Group × Stress Pattern, and Task × Stress Pattern interaction effects were not significant. For the main effect of task (see Figure 5), the PVI of intensity was higher in words produced in isolation (0.062 (0.070)) than in words produced in list intonation (0.042 (0.038); p = .021, d_{RM} = 0.227, 95% CI [0.340, 0.795]). However, there were no

Figure 5. Pairwise variability indices (PVIs) of intensity by task and stress pattern with standard error bars.



significant differences in the PVI of intensity between words produced in isolation and the map description (0.053 (0.054); p = .449) or between list intonation and the map description (p = .301). For the main effect of stress pattern (see Figure 5), the PVI of intensity was higher in trochees (0.072 (0.067)) than iambs (0.033 (0.033); $p < .001$, $d_{RM} = 0.0470$, 95% CI [0.104, 1.043]).

Discussion

The Effects of PD on the Marking of Lexical Stress

The first aim of this study was to examine whether PD affects a person's ability to adequately differentiate between the syllables of trochaic and iambic words using acoustic measures of f_o, intensity, and duration. Counter to our hypothesis, there were few differences in the way that people with PD marked lexical stress compared to age- and sex-matched controls. Both people with PD and controls marked trochees by modulating f_o, but the PVI of f_o (i.e., the magnitude of the difference between syllables) was significantly reduced in trochees produced by people with PD compared to controls; this result is supported by a medium effect size. These findings are consistent with most studies that have found that people with PD are able to mark lexical stress with f_o but that such distinctions are reduced (Darkins et al., 1988; Ma et al., 2015; Rusz et al., 2011). Our findings extend prior work by demonstrating that these reductions are more prominent in trochees than iambs. This suggests that a person with PD and dysprosody receiving speech therapy may benefit more from interventions targeting trochees than iambs. Although both iambs and trochees are automatic motor behaviors by the time PD symptoms manifest, it may be that iambs require, on some level, additional attention throughout the life span since trochees are the dominant metrical structure in Standard American English.

Both people with PD and adults without PD mark trochees differently than they mark iambs; this pattern does not appear to be significantly altered by PD at the group level. Both groups modulated duration more to mark differences between the syllables of iambs than they did for trochees, and this is supported by very large effect sizes. They modulated f_o—and to a lesser extent intensity—to differentiate between the syllables of trochees more than iambs, and this is supported by medium effect sizes. These findings are similar to prior results from typically developing children (Goffman & Malin, 1999), older adults (Barnes, 2013; Scukanec et al., 1996), people with aphasia (Walker et al., 2009), and adults with cerebral palsy (Kim et al., 2010) and follow an established pattern that native speakers of Standard American English produce trochees more with f_o and intensity contrasts and iambs more with durational contrasts. The PD-specific results are, however, counter to that of Cheang and Pell (2007), who found impaired use of both intensity and duration to mark lexical stress in people with PD.

This study cannot rule out a perceptual impairment of lexical stress as produced by people with PD. Pell et al. (2006) found that, despite similar lexical stress–marking behavior by people with PD and controls, naïve listeners still struggled to distinguish between the trochaic (noun compound) and iambic (noun phrase) productions of people with PD. It is noteworthy, however, that methods of analysis in listener studies are insufficient to determine which cues listeners attend to, as there are complex psychoacoustic interactions in the perception of f_o, intensity, and duration (Neuhoff et al., 2002; Tekman, 1995). For instance, changes to both f_o and intensity, whether converging or diverging in the direction of the change, influence the perception of the magnitude of the change of the other dimension (Neuhoff et al., 2002). Furthermore, there is evidence from music perception research that acoustic cues that converge make it easier for a listener to identify a pattern, especially when either duration or intensity changes support concurrent f_o changes (Tekman, 1995).

Even though there is some disagreement in the prior literature about the impact of PD on the use of duration to mark lexical stress, most studies agree that people with PD use duration in lexical stress in a similar manner to healthy controls. This finding is supported by research into other domains of relative timing, such as the shortening of the vowel with increased number of syllables in a word ("zip" – "zipper" – "zippering"; J. J. Sidtis & Sidtis, 2012). Data from people with right and left hemisphere damage associated with stroke demonstrate that speech timing is impaired with left prefrontal cortical damage, which is often intact in people with PD, especially those with mild or moderate progression of the disease (Braak et al., 2003; D. V. L. Sidtis et al., 2010; Walker et al., 2009). People with PD, however, appear to have intact relative timing of syllables, even though speech rate is abnormal (J. J. Sidtis & Sidtis, 2012). These findings are consistent with the lack of group differences in duration in this study. J. J. Sidtis and Sidtis (2012) conclude that the data from people with PD and those with left hemisphere stroke suggest that speech rate is controlled by subcortical structures, impaired in PD, and that relative syllable timing is controlled in the left cortex. The data reported here support that conclusion.

The Effects of Task on the Marking of Lexical Stress

The second aim of this study was to understand how task affects the marking of lexical stress. We found significant differences in lexical stress production across tasks, though the patterns were similar for both people with PD and controls. Words in isolation had higher PVIs than words produced in lists for duration (in iambs), mean f_o (in trochees), and mean intensity (in both iambs and trochees). These results are supported by a range of effect sizes, largest for duration. The differences in PVI across tasks mean that words produced in isolation had a greater magnitude of difference across the syllables of words. These PVI differences were, in general, the highest in words produced in isolation and the lowest in words produced in lists, with words produced in the map description tending to have PVI values between the other two tasks.

There are two potential reasons for these task differences. First, these differences may relate to suprasegmental patterns inherent to the tasks. Suprasegmental patterns are generally absent in words produced in isolation, and there is significant variability in suprasegmental patterns of words produced in spontaneous speech depending on where in an utterance the word is produced and what sentence type is being produced. However, suprasegmental patterns in lists tend to be more fixed (Couper-Kuhlen, 1986; Schubiger, 1958; Selting, 2007). One of the most common suprasegmental patterns for lists with greater than three items (such as the four-item lists used in this study) is for all items except for the final item to have a rising (or flat) nuclear tone (Couper-Kuhlen, 1986). Since all words analyzed in the list task were taken from the second or third item (out of four), this rising pattern could override the preferred lexical pattern (especially for trochees that may otherwise fall). If this were the case, we would expect to see smaller differences between the marking of iambs and trochees in lists than for the isolated and map description tasks, especially for f_o, as trochaic and iambic items would more closely resemble each other. Consistent with this explanation, the smallest PVI of f_o for both trochees and iambs occurred during lists. Furthermore, in lists, there was no significant difference in PVI of f_o between the stress patterns, suggesting that f_o differences between the two stress patterns have been collapsed in the list production task.

Alternatively, these task differences may also relate to the idea of articulatory undershoot: Speakers tend to achieve the canonical target of a given word more often when there is less coarticulation, such as in words produced in isolation. This has been observed both in people with PD (Ackermann & Ziegler, 1991) and in the general population (Lindblom, 1963). Since a word's stress pattern is a part of its lexical representation (Cutler, 2015), the words produced in isolation will be more likely to include canonical stress than coarticulated words in lists or spontaneous speech. Consistent with this explanation, we observed canonical stress (i.e., trochees marked with f_o and intensity, iambs marked with duration), and these intersyllabic differences were greatest in words produced in isolation. However, this does not explain why canonical stress was least present in words produced in lists. If this explanation were consistent with all the data, we would have expected to see the fewest syllabic differences in the map task.

Limitations and Future Directions

Due to the use of a map description task, the position of the target words could not be controlled in this task. While words in the list task were always taken from the middle rather than the beginning or the end of the list, words in the map description task were taken from wherever they were produced in an utterance. One concern with this paradigm was that words taken from the end of an utterance could show effects of phrase-final lengthening, possibly distorting duration measurements. However, we did not observe a significant effect of phrase-final lengthening in our data. In the map description task, 41.7% of the target words were produced in the middle of an utterance, with 46.9% of the target words produced at the end of an utterance. Most of the remaining target words were produced in an utterance-initial position, but four words were produced as single-word responses to researcher prompting, and four words were produced with longer pauses (greater than 150 ms) on either side. There was no significant difference in the duration of the vocalic segments produced in the middle or beginning of an utterance (M = 0.372 s, SE = 0.018 s) versus at the end of an utterance (M = 0.370 s, SE = 0.018 s), $t(44) = 0.097$, $p = .461$. Another limitation of the map task is related to the fact that some words appeared on the town map, but none of the target words were written down. Therefore, while the map description elicited connected speech, it was different from spontaneous speech in that some of the words they produced could have been read.

A second limitation of the study pertained to phonemic distribution. The target words were selected because they were the words consistently produced in all three tasks by all participants; however, phoneme selection was not controlled. There is a benefit to using nonheteronymic target words because they represent more naturalistic speech. However, with this sample, we cannot rule out the effects of voicing, place, and manner of coarticulated consonants on the acoustic correlates of lexical stress. Of these, the most likely effect in our data would be the impact of voiced versus unvoiced consonants on vocalic segment duration, as vowels have a longer duration prior to voiced consonants than unvoiced consonants (Kluender et al., 1988). Consistent with those predictions, the /ɑɪ/ in "tiger" had a significantly longer duration going into the voiced /g/ (M = 0.192 s, SD = 0.049 s) than the /ɔ/ of "coffee" going into the unvoiced /f/ (M = 0.167 s, SD = 0.044 s), $t(140) = 3.17$, $p = .009$. However, the length of the first vocalic segment for "tiger" as compared to "coffee" may be more attributable to the difference in length between a diphthong versus a monophthong. In contrast, for iambs, the /u/ of "balloon" had a significantly shorter duration going into the voiced /n/ (M = 0.281 s, SD = 0.090 s) than the /æ/ of "giraffe" going into the unvoiced /f/ (M = 0.308 s, SD = 0.081 s), $t(141) = 1.88$, $p = .031$. Thus, we did not see systematic voiced–voiceless differences across the four words in this study. The effects of these phoneme-specific influences can be taken as evidence to recommend the use of PVI measures in adjunct to raw acoustic measures in future research and clinical practice. PVI captures the relative salience of stress-marking features independent of global prosodic features and may be considered more sensitive or reliable than raw acoustic measures. The authors still recommend reporting raw acoustic measures to enable comparison with previous and future research on these subjects.

Last, the current study had a small sample size, and the speech of the participants with PD was mild-to-moderately impaired, although the range of severity was wide. The study included two participants with more severe speech impairments (M1 and M2; see Table 1). Disease progression almost certainly affects the results of this study, particularly since Braak's hypothesis asserts that cortical damage would be more apparent in patients in later stages of the disease (Braak et al., 2003). Prior research has shown that damage to the left prefrontal cortex and thalamocortical loops may underlie changes to relational timing and stress patterns (Schirmer, 2004; J. J. Sidtis & Sidtis, 2012). All participants were from the Midwest and were White. Future research should examine the generalizability of these findings to a larger, more diverse sample of people with PD. Although there is no evidence of dialectal differences in stress accuracy of typically developing children (Jarmulowicz et al., 2012), dialectal differences may emerge later in linguistic development, especially in nonnative speakers (Guo, 2022). Although Lowit-Leuschel and Docherty (2001) and Ma et al. (2015) argue that overall dysarthria severity cannot adequately predict prosodic impairment, it is likely that speech severity may impact these findings, and there may be dialectal differences in how speakers produce stress patterns. Intra-individual variability cannot be accounted for in the study because most participants only produced one production of the target word in the map description. Therefore, only one trial of each word in each task was analyzed.

It is possible that the group mean comparisons presented in this analysis obscure speaker-specific patterns and that there may be individual differences that could reveal subtypes based on factors such as an individual's speech severity. To assess this, we calculated the Pearson product–moment correlations of all participants with PD using severity as the independent variable and PVI as the dependent variable. For the PVI of duration, the correlation coefficient was .0037 ($p = .853$). For the PVI of f_o, the correlation coefficient was .0674 ($p = .415$). For the PVI of intensity, the correlation coefficient was .0242 ($p = .629$). The nonsignificant correlation coefficients suggest that there is not a strong relationship between the severity of a person with PD's speech impairment and their ability to use duration, f_o, and intensity to mark lexical stress.

We then looked more specifically at the two people with PD whose speech was rated the most severely impaired (M1, 82% and M2, 73%) and compared the PVI of each variable for their productions against the two people with PD whose speech was rated as the least impaired (F6, 4% and M3, 8.9%) using the subject means. In this subset, F6 produced the greatest average durational (0.830) and intensity (0.112) differences across syllables, while M2 produced the greatest f_o differences (0.201). M1 produced the smallest average durational differences across syllables (0.629), while M3 produced the smallest f_o (0.040) and intensity (0.037) differences. The differences in PVI for duration among M1, M2, and M3 are fairly small (all with average PVIs of duration in the 0.63–0.67 range). The PVIs of f_o and intensity present a similar mixed picture about the interaction of severity and stress marking. F6 (mild impairment) demonstrated the greatest differences in PVI of intensity out of these four participants. Compare this with M3 (mild impairment), who demonstrated very small average differences between syllables for both f_o (0.041) and intensity (0.037), and M2 (moderate–severe impairment), who demonstrated higher intersyllabic differences for both f_o (0.201) and intensity (0.080). Although M2 (moderate–severe) seemed to rely more on pitch and intensity than the other speakers, he still utilizes duration as well. Furthermore, M3 (mild) seemed to rely almost exclusively on durational cues, with relatively small differences in f_o and intensity across his measured speech sample. Taken together, these results suggest that the use of duration to mark lexical stress is not strongly associated with speech severity. Observed differences, where they exist, are more based on each individual's productions than on the severity of their speech. This is in agreement with previous research findings on the interaction of speech severity, lexical stress, and speech stimulus (Lowit-Leuschel & Docherty, 2001; Ma et al., 2015).

Future research should also examine the relationship between the acoustic and perceptual aspects of lexical stress patterns, as assessed previously by Pell et al. (2006). Since the current research demonstrates that reduced distinctions in stress marking are more prevalent in trochees than iambs, it will be important to determine whether naïve listeners experience perceptual differences in their ratings of speaker intelligibility in tasks designed to assess lexical stress.

Conclusions

This study suggests that, in general, people with PD use the same cues as age- and sex-matched controls in marking lexical stress. Trochees and iambs are marked with different acoustic cues, and these differences need to be considered when designing treatment for people with PD who demonstrate a prosodic impairment in lexical stress production. Furthermore, there are significant task differences between measures of lexical stress, with words in isolation generally displaying greater differences between syllables than words produced in lists or in the map description. Task differences are especially prominent for duration cues in iambs and f_o cues in trochees. On the basis of our results, we recommend that speech assessments, whether in research or clinical settings, use connected speech tasks in addition to more traditional controlled tasks to accurately describe a person's prosodic deficit.

Data Availability Statement

De-identified data can be requested from Jessica E. Huber. Availability of data is contingent on the execution of a data-sharing agreement between Purdue University and the receiving university.

Acknowledgments

This research was supported by a Research Support Incentive Grant from the Center on Aging and the Life Course at Purdue University, awarded to Jessica E. Huber. Andrew H. Exner was supported by an Institutional National Research Service Award T32 grant from the National Institute on Deafness and Other Communication Disorders, awarded to Elizabeth Strickland. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Center on Aging and the Life Course at Purdue University or the National Institute on Deafness and Other Communication Disorders. The authors thank Sandy Snyder for her help with data collection and management of data measurement. They also thank McKayla Schloemer for conducting interrater reliability of the measurements.

References

Ackermann, H., & Ziegler, W. (1991). Articulatory deficits in parkinsonian dysarthria: An acoustic analysis. *Journal of Neurology, Neurosurgery & Psychiatry, 54*(12), 1093–1098. https://doi.org/10.1136/jnnp.54.12.1093

Barnes, D. R. (2013). *Age-related changes to the production of linguistic prosody* [Master's thesis, Purdue University]. Open Access Theses, 17.

Boersma, P., & Weenink, D. (2019). *Praat: Doing phonetics by computer* [Computer program].

Braak, H., Del Tredici, K., Rüb, U., de Vos, R. A. I., Jansen Steur, E. N. H., & Braak, E. (2003). Staging of brain pathology related to sporadic Parkinson's disease. *Neurobiology of Aging, 24*(2), 197–211. https://doi.org/10.1016/s0197-4580(02)00065-9

Cheang, H. S., & Pell, M. D. (2007). An acoustic investigation of parkinsonian speech in linguistic and emotional contexts. *Journal of Neurolinguistics, 20*(3), 221–241. https://doi.org/10.1016/j.jneuroling.2006.07.001

Couper-Kuhlen, E. (1986). *An introduction to English prosody*. Max Niemeyer Verlag. https://doi.org/10.2307/415352

Cutler, A. (2015). Lexical stress in English pronunciation. In M. Reed & J. M. Lewis (Eds.), *The handbook of English pronunciation* (pp. 106–124). Wiley. https://doi.org/10.1002/9781118346952

Cutler, A., & Clifton, C. (1984). The use of prosodic information in word recognition. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance X: Control of language processes* (pp. 183–196). Erlbaum.

Cutler, A., Dahan, D., & van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech, 40*(2), 141–201. https://doi.org/10.1177/002383099704000203

Darkins, A. W., Fromkin, W. A., & Benson, D. R. (1988). A characterization of the prosodic loss in Parkinson's disease. *Brain and Language, 34*(2), 315–327. https://doi.org/10.1016/0093-934x(88)90142-3

Darley, F. L., Aronson, A. E., & Brown, J. R. (1969). Differential diagnostic patterns of dysarthria. *Journal of Speech and Hearing Research, 12*(2), 246–269. https://doi.org/10.1044/jshr.1202.246

Fry, D. B. (1955). Duration and intensity as physical correlates of linguistic stress. *The Journal of the Acoustical Society of America, 27*(4), 765–768. https://doi.org/10.1121/1.1908022

Fry, D. B. (1958). Experiments in the perception of stress. *Language and Speech, 1*(2), 126–152. https://doi.org/10.1177/002383095800100207

Goffman, L., & Malin, C. (1999). Metrical effects on speech movements in children and adults. *Journal of Speech, Language, and Hearing Research, 42*(4), 1003–1015. https://doi.org/10.1044/jslhr.4204.1003

Grabe, E., & Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. In C. Gussenhoven & N. Warner (Eds.), *Laboratory phonology VII* (pp. 515–546). Mouton de Gruyter. https://doi.org/10.1515/9783110197105.2.515

Guo, X. (2022). Acoustic correlates of English lexical stress produced by Chinese dialect speakers compared to native English speakers. *Frontiers in Psychology, 13,* 796252. https://doi.org/10.3389/fpsyg.2022.796252

Hertrich, I., & Ackermann, A. (1993). Acoustic analysis of speech prosody in Huntington's and Parkinson's disease: A preliminary report. *Clinical Linguistics & Phonetics, 7*(4), 285–297. https://doi.org/10.1080/02699209308985568

Howell, P. (1993). Cue trading in the production and perception of vowel stress. *The Journal of the Acoustical Society of America, 94*(4), 2063–2073. https://doi.org/10.1121/1.407479

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)

Jarmulowicz, L., Taran, V. L., & Seek, J. (2012). Metalinguistics, stress accuracy, and word reading: Does dialect matter? *Language, Speech, and Hearing Services in Schools, 43*(4), 410–423. https://doi.org/10.1044/0161-1461(2012/11-0060)

Kim, H., Hasegawa-Johnson, M., & Perlman, A. (2010, May). *Acoustic cues to lexical stress in spastic dysarthria*. Speech Prosody 2010, Chicago, IL, United States.

Kluender, K. R., Diehl, R. L., & Wright, B. A. (1988). Vowel-length differences before voiced and voiceless consonants: An auditory explanation. *Journal of Phonetics, 16*(2), 153–169. https://doi.org/10.1016/S0095-4470(19)30480-2

Ladd, D. (2008). *Intonational phonology*. Cambridge University Press. https://doi.org/10.1017/CBO9780511808814

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for *t*-tests and ANOVAs. *Frontiers in Psychology, 4,* 1–12. https://doi.org/10.3389/fpsyg.2013.00863

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

Lieberman, P. (1960). Some acoustic correlates of word stress in American English. *The Journal of the Acoustical Society of America, 32*(4), 451–454. https://doi.org/10.1121/1.1908095

Lindblom, B. (1963). Spectrographic study of vowel reduction. *The Journal of the Acoustical Society of America, 35*(11), 1773–1781. https://doi.org/10.1121/1.1918816

Lowit, A., Brendel, B., Dobinson, C., & Howell, P. (2006). An investigation into the influences of age, pathology and cognition on speech production. *Journal of Medical Speech-Language Pathology, 14*(4), 253–262. https://doi.org/10.1080/14015430127772

Lowit-Leuschel, A., & Docherty, G. J. (2001). Prosodic variation across sampling tasks in normal and dysarthric speakers. *Logopedics Phoniatrics Vocology, 26*(4), 151–164. https://doi.org/10.1080/14015430127772

Ma, J. K. Y., Schneider, C. B., Hoffman, R., & Storch, A. (2015). Speech prosody across stimulus types for individuals with Parkinson's disease. *Journal of Parkinson's Disease, 5*(2), 291–299. https://doi.org/10.3233/JPD-140451

MacPherson, M. K., Huber, J. E., & Snow, D. P. (2011). The intonation–syntax interface in the speech of individuals with Parkinson's disease. *Journal of Speech, Language, and Hearing Research, 54*(1), 19–32. https://doi.org/10.1044/1092-4388(2010/09-0079)

McHenry, M. (1998). The ability to effect intended stress following traumatic brain injury. *Brain Injury, 12*(6), 495–503. https://doi.org/10.1080/026990598122458

Miller, N., Noble, E., Jones, D., & Burn, D. (2006). Life with communication changes in Parkinson's disease. *Age and Ageing, 35*(3), 235–239. https://doi.org/10.1093/ageing/afj053

Neuhoff, J. G., Wayand, J., & Kramer, G. (2002). Pitch and loudness interact in auditory displays: Can the data get lost in the map? *Journal of Experimental Psychology: Applied, 8*(1), 17–25. https://doi.org/10.1037/1076-898X.8.1.17

Patel, R., & Campellone, P. (2009). Acoustic and perceptual cues to contrastive stress in dysarthria. *Journal of Speech, Language, and Hearing Research, 52*(1), 206–222. https://doi.org/10.1044/1092-4388(2008/07-0078)

Pell, M. D., Cheang, H. S., & Leonard, C. L. (2006). The impact of Parkinson's disease on vocal-prosodic communication from the perspective of listeners. *Brain and Language, 97*(2), 123–134. https://doi.org/10.1016/j.bandl.2005.08.010

Rusz, J., Cmejla, R., Ruzickova, H., & Ruzicka, E. (2011). Quantitative acoustic measurements for characterization of speech and voice disorders in early untreated Parkinson's disease. *The Journal of the Acoustical Society of America, 129*(1), 350–367. https://doi.org/10.1121/1.3514381

Schirmer, A. (2004). Timing speech: A review of lesion and neuroimaging findings. *Cognitive Brain Research, 21*(2), 269–287. https://doi.org/10.1016/j.cogbrainres.2004.04.003

Schubiger, M. (1958). *English intonation, its form and function*. Max Niemeyer Verlag. https://doi.org/10.2307/411069

Scukanec, G. P., Ptrosino, L., & Colcord, R. D. (1996). Age-related differences in acoustical aspects of contrastive stress in women. *Folia Phoniatrica et Logopaedica, 48*(5), 231–239. https://doi.org/10.1159/000266414

Selting, M. (2007). Lists as embedded structures and the prosody of list construction as an interactional resource. *Journal of Pragmatics, 39*(3), 483–526. https://doi.org/10.1016/j.pragma.2006.07.008

Sidtis, D. V. L., Kempler, D., Jackson, C., & Metter, E. J. (2010). Prosodic changes in aphasic speech: Timing. *Clinical Linguistics & Phonetics, 24*(2), 155–167. https://doi.org/10.3109/02699200903464439

Sidtis, J. J., & Sidtis, D. V. L. (2012). Preservation of relational timing in speech of persons with Parkinson's disease with and without deep brain stimulation. *Journal of Medical Speech-Language Pathology, 20*(4), 140–143.

Skodda, S., Rinsche, H., & Schlegel, U. (2009). Progression of dysprosody in Parkinson's disease over time—A longitudinal study. *Movement Disorders, 24*(5), 716–722. https://doi.org/10.1002/mds.22430

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

Smith, B. L., Wasowicz, J., & Preston, J. (1987). Temporal characteristics of the speech of normal elderly adults. *Journal of Speech and Hearing Research, 30*(4), 522–529. https://doi.org/10.1044/jshr.3004.522

Snow, D. (1994). Phrase-final syllable lengthening and intonation in early child speech. *Journal of Speech and Hearing Research, 37*(4), 831–840. https://doi.org/10.1044/jshr.3704.831

Tekman, H. G. (1995). Cue trading in the perception of rhythmic structure. *Music Perception, 13*(1), 17–38. https://doi.org/10.2307/40285683

Tykalova, T., Rusz, J., Cmejla, R., Ruzickova, J., & Ruzicka, E. (2014). Acoustic investigation of stress patterns in Parkinson's disease. *Journal of Voice, 28*(1), 129.e1–129.e8. https://doi.org/10.1016/j.jvoice.2013.07.001

Walker, J. P., Joseph, L., & Goodman, J. (2009). The production of linguistic prosody in subjects with aphasia. *Clinical Linguistics & Phonetics, 23*(7), 529–549. https://doi.org/10.1080/02699200902946944

Wang, Y. T., Kent, R. D., Duffy, J. R., & Thomas, J. E. (2005). Dysarthria associated with traumatic brain injury: Speaking rate and emphatic stress. *Journal of Communication Disorders, 38*(3), 231–260. https://doi.org/10.1016/j.jcomdis.2004.12.001

Watson, P. J., & Munson, B. (2008). Parkinson's disease and the effect of lexical factors on vowel articulation. *The Journal of the Acoustical Society of America, 124*(5), EL291–EL295. https://doi.org/10.1121/1.2987464

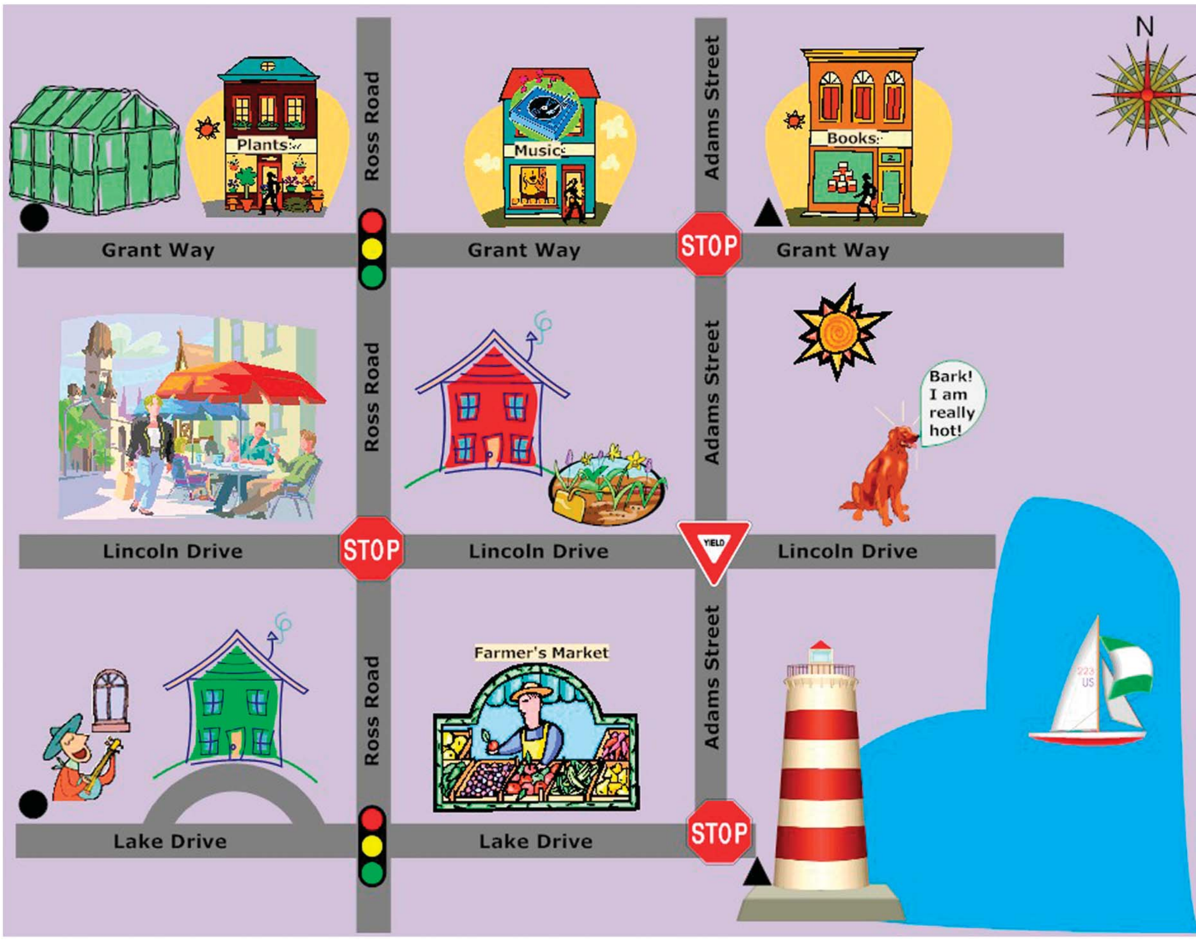
Zuraiq, W., & Sereno, J. (2007). English lexical stress cues in native English and non-native Arabic speakers. *Proceedings of the 16th International Congress of Phonetic Sciences, Saarbrücken, Germany*.

Appendix A

Picture Listing Task With Target Word Highlighted in Yellow

Appendix B

Map Description Speech Task: Streets



Appendix C

Map Description Speech Task: Zoo