JSLHR

Research Article

Cognitive–Perceptual Examination of Remediation Approaches to Hypokinetic Dysarthria

Megan J. McAuliffe,[a] Sarah E. Kerr,[a] Elizabeth M. R. Gibson,[a] Tim Anderson,[b] and Patrick J. LaShell[a]

Purpose: To determine how increased vocal loudness and reduced speech rate affect listeners' cognitive–perceptual processing of hypokinetic dysarthric speech associated with Parkinson's disease.
Method: Fifty-one healthy listener participants completed a speech perception experiment. Listeners repeated phrases produced by 5 individuals with dysarthria across habitual, loud, and slow speaking modes. Listeners were allocated to habitual (n = 17), loud (n = 17), or slow (n = 17) experimental conditions. Transcripts derived from the phrase repetition task were coded for overall accuracy (i.e., intelligibility), and perceptual error analyses examined how these conditions affected listeners' phonemic mapping (i.e., syllable resemblance) and lexical segmentation (i.e., lexical boundary error analysis).
Results: Both speech conditions provided obvious perceptual benefits to listeners. Overall, transcript accuracy was highest in the slow condition. In the loud condition, however, improvement was evidenced across the experiment. An error analysis suggested that listeners in the loud condition prioritized acoustic–phonetic cues in their attempts to resolve the degraded signal, whereas those in the slow condition appeared to preferentially weight lexical stress cues.
Conclusions: Increased loudness and reduced rate exhibited differential effects on listeners' perceptual processing of dysarthric speech. The current study highlights the insights that may be gained from a cognitive–perceptual approach.

Key Words: intelligibility, dysarthria, Parkinson's disease, speech perception

Parkinson's disease (PD) is a degenerative neurological condition that affects approximately 1% of people above age 60 years, rising to about 4% in elderly individuals (de Lau & Breteler, 2006). Of those with PD, it is estimated that 50% to 89% will develop the speech disorder *hypokinetic dysarthria* (Hartelius & Svensson, 1994; Johnson & Pring, 1990). Hypokinetic dysarthria affects all speech motor control subsystems, with salient speech features including reduced vocal loudness; a breathy or harsh voice; monopitch and monoloudness; imprecision of consonant and vowel production; and, in some cases, a fast rate of speech and associated dysfluencies (Duffy, 2005). Combined, PD and hypokinetic dysarthria have considerable negative effects on quality of life. Individuals with PD have expressed concern regarding how the speech changes affect both their ability to communicate and their self-image (Miller, Noble, Jones, & Burn, 2006). It would seem that these concerns are not unfounded, given that speakers with PD have been rated as sounding significantly less interested, happy, friendly, and involved compared with healthy speakers (Jaywant & Pell, 2010). It is clear that the challenges that arise from their speech disorder, combined with its effect on the perceptions of others, negatively influence both social participation and well-being in individuals with PD.

Currently, behavioral strategies that focus on increased vocal loudness—for example, the Lee Silverman Voice Treatment (Ramig, Sapir, Countryman, et al., 2001)—and reduced speech rate (Lowit, Dobinson, Timmins, Howell, & Kröger, 2010) are regularly implemented in the treatment of hypokinetic dysarthria (Miller, Deane, Jones, Noble, & Gibb, 2011). Although these approaches are commonly used in clinical settings, a recent Cochrane review concluded that "there is insufficient evidence to conclusively support or refute the efficacy of [speech-language therapy] for speech problems in Parkinson's disease" (Herd et al., 2012, p. 2). Indeed, much of what is known regarding treatment outcomes for hypokinetic dysarthria has been directed at the level of the speaker, in particular how specific treatments, or treatment simulations, influence

[a]University of Canterbury, Christchurch, New Zealand
[b]New Zealand Brain Research Institute, Christchurch, New Zealand
Correspondence to Megan J. McAuliffe: megan.mcauliffe@canterbury.ac.nz

Editor: Jody Kreiman
Associate Editor: Kate Bunton
Received November 4, 2012
Revision received May 24, 2013
Accepted January 7, 2014
DOI: 10.1044/2014_JSLHR-S-12-0349

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

speaker-based dependent variables such as intensity, fundamental frequency, vowel space, or rate of speech (e.g., Hammen & Yorkston, 1996; Johnson & Pring, 1990; Ramig, Sapir, Countryman, et al., 2001; Ramig, Sapir, Fox, & Countryman, 2001; Tjaden & Wilding, 2011a). However, for many individuals with dysarthria the primary aim of participating in speech treatment is to be better understood—that is, to obtain improvements to overall speech intelligibility. Yet empirical evidence highlighting the success, or otherwise, of loudness- and rate-based techniques in improving speech intelligibility for persons with PD is limited (Lansford, Liss, Caviness, & Utianski, 2011). Furthermore, only limited attention has been given to the effect of these behavioral interventions on listener perception, that is, a listener's ability to perceive, process, and comprehend speech that has been modified.

Borrowing theoretical perspectives from speech perception, we can hypothesize as to the effect of common remediation approaches on listeners' perceptual processing and measure outcomes using listener-based indices. Hence, in this study, we examined the phrase transcriptions of listeners who were exposed to hypokinetic dysarthric speech produced across habitual, loud, and slow speech conditions. From these transcripts, listeners' accuracy comprehending speech across conditions was determined and a detailed analysis of error patterns conducted. Transcription error analysis provides insight into listeners' perceptual strategies as they attempt to comprehend dysarthric speech that is reduced in intelligibility and has been modified via common intervention techniques. Conceptualizing remediation from this perspective therefore provides a theoretically motivated means of determining how improvements to intelligibility are achieved. From this point forward we refer to this approach as the *cognitive–perceptual* approach to remediation.

Cognitive–Perceptual Approach

The cognitive–perceptual approach to conceptualizing assessment and treatment for dysarthria advocates bidirectional consideration of the acoustic degradation present in dysarthric speech and its effect on listeners' perceptual processing (Lansford et al., 2011). The theoretical underpinnings for this approach were derived from theories of speech perception, in particular what is known of the basic perceptual processes that occur when a listener attempts to comprehend speech. These processes include lexical segmentation, lexical activation, and lexical competition. At any point, cue degradation associated with dysarthric speech may interrupt these processes, resulting in communication failure (Lansford et al., 2011; Liss, 2007).

Lexical segmentation refers to the processes by which the speech stream is broken into word units (Jusczyk & Luce, 2002). Once this occurs, possible word candidates are activated (i.e., *lexical activation*) and competition between lexical candidates results. The best match between a speaker's acoustic output and a listener's mental lexical representations of words is ultimately selected and recognized as the target. Lexical segmentation proceeds relatively automatically in good listening conditions (e.g., when conversing with a healthy talker in a quiet environment). The listener, aided by a clear signal and the presence of semantic and contextual knowledge, hears individual words, allowing him or her to reliably parse the speech stream. However, when signal degradation reduces word salience (e.g., in the case of dysarthric speech), listeners are forced to rely on lower level cues (e.g., coarticulatory, allophonic, and rhythmic) in their attempts to determine word boundaries (Mattys, White, & Melhorn, 2005). In English, lexical segmentation in adverse listening conditions is aided by the relatively predictable patterning of strong and weak syllables, with over 70% of English words beginning with a stressed syllable (Cutler & Butterfield, 1992). When signal quality is poor, listeners may exploit this patterning to assist in determining word boundaries. This exploitation of stressed syllables to inform word boundary decisions is known as the *metrical segmentation strategy* (MSS; Cutler & Butterfield, 1992).

In a practical sense, this strategy is evidenced through listener transcription errors that demonstrate selection of word boundaries based on the presence of a strong syllable. For example in the word *amend*, syllabic patterning is weak–strong—a weak syllable ("a") followed by a strong syllable ("mend"). In degraded listening conditions, listeners may transcribe *amend* as "a man," which is evidence of a lexical segmentation strategy that preferences strong syllables through the insertion of a word boundary before the strong syllable. A number of means of analysis are available to quantify this perceptual effect, including counts of error types (e.g., number of insertions or deletion of a word boundary before or after a strong or weak syllable) and calculation of an *MSS* ratio,[1] which quantifies a listeners' overall strength of adherence to the MSS strategy (Spitzer, Liss, & Mattys, 2007). Although not an exhaustive list, these indices highlight the possibility for examination of listeners' attention to syllabic stress as they parse the speech stream.

Using this methodological approach, a number of studies have shown that, when faced with dysarthric speech, listeners exploit syllabic stress cues as they attempt to identify word boundaries (Borrie et al., 2012; Choe, Liss, Azuma, & Mathy, 2012; Liss, Spitzer, Caviness, Adler, & Edwards, 1998; Liss, Spitzer, Caviness, Adler, & Edwards, 2000). What is not known is whether the behavioral speech strategies commonly used assist listeners in their attempts to identify word boundaries. Given that the identification of word boundaries is a fundamental component of perception, any speech strategy that facilitates this process will likely prove beneficial to intelligibility.

Once the speech stream has been broken up into appropriate word-sized units, the acoustic–phonetic information available within these word-sized "packets" is used

[1]The calculation procedure for the MSS ratio (also termed *proportion of predictable errors*) is provided in the Method section.

to identify available possible lexical candidates; that is, a listener matches the available packet of acoustic–phonetic information with stored words within his or her lexicon. Once possible lexical candidates are identified, the best fit between the acoustic input and the words available in a listener's mental representation is recognized as the target word (Lansford et al., 2011). When conversing with a healthy talker in an optimal listening environment, this process occurs with minimal error. However, as articulatory imprecision associated with dysarthric speech renders this process more difficult, the phonemic characteristics of the word ultimately recognized by a listener may not resemble those of the target. For example, a dysarthric speaker's articulation deficits may result in complete failure to activate the lexical target—because there was no relationship between the dysarthric speaker's acoustic output and a listener's mental representation of the phonemes of that spoken word (Liss, 2007). Alternatively, phoneme distortion may result in a listener selecting a phonemically similar, but incorrect, word (e.g., *sip* instead of *tip*). A number of analysis techniques are available to quantify the link between speech produced and a listener's mental representation of the signal. For example, single-word intelligibility tests enable listeners to select a word from a constrained set of targets and, through this, gain some insight into the acoustic–phonetic origins of listeners' errors (e.g., if a listener selected the word *need* instead of *bead*, nasality is likely a contributing factor to the perceptual error; e.g., Kent, Weismer, Kent, & Rosenbek, 1989).

However, more recent work has attempted to characterize perception errors by examining the degree of phonemic relatedness of a listener's transcription error to the intended target. Spitzer, Liss, Caviness, and Adler (2000) showed that the transcription errors of listeners who were exposed to dysarthric speech prior to completing a transcription task exhibited greater phonemic resemblance to the target compared with those who did not complete a familiarization procedure. Borrie et al. (2012) subsequently developed a measure called *syllable resemblance*, which quantified the degree of phonemic relatedness of an erroneous transcription to the target. Borrie et al. demonstrated that when a listener was familiarized with dysarthric speech, erroneously transcribed syllables were more phonemically like the intended target than when listeners had not received familiarization. The results were interpreted to suggest that familiarization with dysarthric speech "enabled listeners to better map acoustic–phonetic aspects of the disordered signal onto existing mental representations of speech sounds" (Borrie et al., 2012, p. 1052).

In summary, a cognitive–perceptual approach goes beyond intelligibility to examine the perceptual mechanisms invoked as listeners attempt to resolve a degraded signal—in this case, dysarthric speech. Through examination of transcription error patterns, evidence of listeners' cognitive–perceptual processing strategies can be observed. To date, studies that have used a cognitive–perceptual processing approach to dysarthric speech have characterized speech spoken in habitual conditions (Choe et al., 2012; Liss et al., 1998, 2000) or have investigated the effects of familiarization (i.e., listener training) on perception (Borrie et al., 2012; Liss, Spitzer, Caviness, & Adler, 2002; Spitzer et al., 2000). To the best of our knowledge, researchers have yet to take this one step further and pit remediation techniques against each other to directly examine how listeners' processing of dysarthric speech is enhanced, or otherwise, by common behavioral speech interventions. Given that current approaches to the treatment of hypokinetic dysarthria commonly focus on two primary strategies—reducing speech rate and increasing vocal loudness—we posited that early research examining a cognitive–perceptual approach to remediation is best placed to begin at this point. On the basis of current knowledge of the effect of both strategies on acoustic and kinematic parameters of dysarthric speech, we formulated hypotheses about the effects of these remediation techniques on listeners' cognitive–perceptual processing.

Predicted Effects of Reduced Rate and Increased Loudness on Listeners' Cognitive–Perceptual Processing

The cue to speak loudly generally results in significant positive changes to intelligibility (e.g., Neel, 2009; Tjaden & Wilding, 2004). Common acoustic findings include increased vocal intensity (Neel, 2009; Tjaden & Wilding, 2004, 2011a), greater fundamental frequency variation across the phrase (Tjaden & Wilding, 2011a), improved acoustic distinctiveness for stops (Tjaden & Wilding, 2004), and changes to vowel production (Neel, 2009; Sapir, Spielman, Ramig, Story, & Fox, 2007; Tjaden & Wilding, 2004). Although the cognitive–perceptual source of this intelligibility gain is not clear, it appears likely that it results from an improved ability, on the part of the listener, to complete the processes of lexical segmentation and lexical access.

If, as Lansford et al. (2011) suggested, it is presumed that this combination of acoustic cue changes in loud speech results in a general improvement in the production of syllabic stress (i.e., to pitch variation and vowel production), then it follows that it should result in improved lexical segmentation; that is, the increased salience of syllabic stress cues available to listeners will assist them in their attempts to parse the speech signal. Accurate placement of word boundaries also has the flow-on effect of increasing the possibility of activation of the intended lexical target. In addition, co-occurring improvements to articulatory precision (e.g., Tjaden & Wilding, 2004) should reduce phonemic ambiguity, further facilitating the processes of lexical activation and competition. In sum, it is expected that the cue to speak loudly would facilitate listeners' processing of the dysarthric speech signal and that this will be observed through improvements to overall intelligibility, greater phonemic resemblance of syllable errors to the target, and changes to the number and pattern of lexical boundary errors (LBEs).

In regard to speech rate, studies that have modified rate in speakers with dysarthria, using a variety of elicitation methods, have reported both positive (Adams, 1994; Hammen, Yorkston, & Minifie, 1994; Yorkston, Hammen, Beukelman, & Traynor, 1990) and negligible or negative effects on intelligibility (Van Nuffelen, De Bodt, Vanderwegen, & Van de Heyning, 2010; Van Nuffelen, De Bodt, Wuyts, & Van de Heyning, 2009). Although the intelligibility findings have been variable, and dependent on the type of rate cue used, a number of outcomes appear generally consistent across studies.

Reduced speech rate generally results in an increased number and duration of pauses (Hammen & Yorkston, 1996; Tjaden & Wilding, 2011b). Provided these occur at syntactically appropriate boundaries, they should serve to delineate word boundaries for listeners, in turn facilitating the process of lexical segmentation (Lansford et al., 2011). This effect should be evidenced through a reduction in the number of LBEs. For speakers with hypokinetic dysarthria and a faster speech rate, this effect is likely to be particularly salient. Slower speech rates have also been associated with larger acoustic working spaces for both vowels and consonants (McRae, Tjaden, & Schoonings, 2002; Tjaden & Wilding, 2004) and reductions in coarticulation (Tjaden & Wilding, 2004) in dysarthric speakers. The resultant enhanced articulatory precision should further help to lessen phonemic ambiguity experienced by listeners, and error analysis is predicted to show enhanced phonemic resemblance of syllable errors to the target. The combined effects of these and other changes would result in improved speech intelligibility. There are, however, two caveats to the above assumptions. First, a reduced fundamental frequency (F0) range has also been reported to co-occur with reductions in speech rate (Tjaden & Wilding, 2011a). This has the potential to reduce the salience of stress-based cues to segmentation (Spitzer et al., 2007), which may mitigate potential perceptual benefits. Second, it is possible that speech rate may slow speech to such an extent that longer perceptual processing times are required. If this were to occur, then working memory resources would be taxed, which may also negatively affect the accuracy of listeners' transcriptions (Liss, 2007). This remains to be tested.

Research Aims

There is much to be learned regarding listeners' processing of dysarthric speech, including whether various treatments enhance or reduce the ability to make accurate lexical boundary decisions and whether these techniques improve acoustic–phonetic aspects of the signal to such a degree that lexical access is enhanced. The purpose of the current investigation was to examine the effect of habitual, loud, and slow speech—produced by speakers with hypokinetic dysarthria—on healthy listeners' perceptual processing. The dependent variables included listeners' overall accuracy at transcribing the signal, the phonemic resemblance of any syllable errors to the target syllable, and the location and type of word boundary errors. We addressed three primary questions. First, do increased loudness and rate reduction result in significant improvements to listeners' transcription accuracy across the experiment (i.e., intelligibility)? Second, do these behavioral modifications result in an improved ability on the part of the listener to map phonemic aspects of the signal (as measured by syllable resemblance scores)? Third and last, how do these modifications affect listeners' lexical segmentation (as measured by LBE analysis)?

Method

This study received ethical approval from the Upper South A Regional Ethics Committee, Ministry of Health, New Zealand. All individuals provided written consent to participate.

Listener (Experimental) Participants

A total of 51 listeners (45 women and six men) participated in the study. The average age of the group was 20.83 years (*SD* = 2.94, range: 18–30 years). All were native speakers of New Zealand English and reported no significant history of contact with persons having a motor speech disorder or a history of language, learning, or cognitive disabilities. All listeners passed a pure-tone hearing screen at 20 dB HL for 1000, 2000, and 4000 Hz and at 30 dB HL for 500 Hz bilaterally. Listeners received a $10 (≈8.46 USD) voucher as compensation for their participation.

Speech Stimuli Acquisition and Selection

Speaker Participants

Speech recordings were collected from five individuals with hypokinetic dysarthria associated with PD. All individuals were native speakers of New Zealand English, and their biographical details are shown in Table 1. The speakers exhibited mild to mild–moderate dysarthria and were selected for participation if they met an operational definition of hypokinetic dysarthria as described by Liss et al. (1998). This included a perceptual impression of a fast rate of speech, monopitch, monoloudness, consonant imprecision, and a weak and/or breathy voice. We acknowledge that not all speakers with PD exhibit a fast rate of speech; however, we selected this operational definition in order to obtain as homogeneous a speech stimuli sample as possible and to ensure comparability with the few studies that have examined similar perceptual error patterns in speakers with hypokinetic dysarthria (Borrie et al., 2012; Liss et al., 1998). Perceptual impressions of deviant speech characteristics and an informal categorization of severity (as shown in Table 1) were determined by two speech-language pathologists (the first and third authors) via a consensus rating procedure and based on speakers' readings of the Grandfather Passage (Darley, Aronson, & Brown, 1975).

Speech Recording Procedure

Speaker participants attended a single assessment session. Digital audio recordings were made in a sound-treated room. During recordings, speakers wore an Audix HT2 headset condenser microphone positioned approximately

Table 1. Biographical details of the five participants with hypokinetic dysarthria.

Participant	Age	Sex	YPD	SIT (%)	Dysarthria severity	Perceptual impression
1	75	M	NA	81.8	Mild–moderate	Fast rate, imprecise consonants, repeated phonemes, monoloudness, and monopitch
2	79	M	10	80.6	Mild–moderate	Fast rate, short rapid rushes of speech, instances of pallilalia, imprecise consonants, breathiness, monoloudness, and monopitch
3	73	M	25	95.2	Mild	Fast rate, short rapid rushes of speech, imprecise consonants, reduced volume
4	69	M	5	90.5	Mild–moderate	Fast rate, imprecise consonants, reduced volume, breathiness, monoloudness, and monopitch
5	65	M	11	90.6	Mild	Fast rate, imprecise articulation, reduced volume, monoloudness, and monopitch

Note. Dysarthria severity, based on perceptual impressions, was rated by two experienced speech-language pathologists. YPD = years postdiagnosis; SIT = Sentence Intelligibility Test; M = male; NA = not available.

5 cm from the mouth. Stimuli were recorded directly to a laptop computer using Sony SoundForge Version 9.0 at a sampling rate of 48 kHz with 16 bits of quantization. Before each assessment commenced, we used a calibration procedure to ensure that the sound pressure level of the recordings could be compared across speaking conditions, as per the procedure detailed by Tjaden and Wilding (2004). In brief, a 1000-Hz calibration tone was played at 90 dB and recorded and saved for later calculation of sound pressure level from the acoustic trace. In doing so, the head-mounted microphone and sound level meter were positioned side by side in a position that corresponded to the mouth-to-microphone distance. A sound level meter (Reed ST-805 Compact Digital) verified the amplitude of the tone, and this value was saved to a file for analysis. During the assessment, speaker participants were recorded completing the Sentence Intelligibility Test (Yorkston, Beukelman, & Hakel, 1996), reading the Grandfather Passage and reading a list of 80 experimental phrases (which we describe in detail in subsequent sections).

The 80 experimental phrases were read in three conditions: (a) habitual, (b) loud, and (c) slow. For the habitual condition, speakers were asked to read the experimental phrase list in their everyday speaking voice, as if talking with a family member or friend. A magnitude scaling procedure was used to elicit the slow and loud conditions. We acknowledge that a number of methods are available, both clinically and in research studies, to achieve such modifications (in particular that of rate) and that these methods are likely to have differential effects on cognitive–perceptual processing. However, we aimed to achieve a degree of methodological consistency with existing studies that have examined the effect of rate and loudness manipulations on speech production (e.g., Dromey & Ramig, 1998; Kleinow, Smith, & Ramig, 2001; Rosen et al., 2011; Tjaden & Wilding, 2004, 2011a, 2011b) and to ensure procedural consistency across tasks for the speakers with PD. Therefore, a magnitude estimation procedure was deemed optimal. For the slow condition, speakers were asked to read each phrase at "what feels like half your normal speed." For the loud condition, speakers were asked to say each phrase "at a level that feels like twice as loud as normal" (Tjaden & Wilding, 2004). Stimuli were produced in the habitual condition first, followed by the slow condition and then the loud condition. The set phrase collection order was established after pilot testing that had revealed pervasive carry-over effects of condition. Specifically, if either of the speech manipulation conditions preceded the habitual condition, carryover of one or the other condition was noted. Furthermore, if the loud condition preceded the slow condition, carryover of increased loudness was persistent. Hence, in keeping with Dromey and Ramig's (1998) seminal paper, we decided that speech should be produced in a consistent order. Any stimuli that contained errors were repeated.

Experimental Stimuli

Composition. The experimental phrases were modeled on those established by Cutler and Butterfield (1992) and further developed by Liss et al. (1998). All were semantically anomalous, to limit the contribution of semantic information to listeners' processing, and syntactically correct (although the syntactic structure differed across phrases). Each phrase contained six syllables and was designed to ensure an alternating phrasal stress pattern. Half of the stimuli were iambic (exhibited a weak–strong phrasal stress pattern), and the remaining half were trochaic (exhibited a strong–weak stress pattern). Although in the English language trochaic words are more frequent than iambic, with approximately 73% of English words exhibiting strong initial syllables (Cutler & Carter, 1987), the inclusion of balanced numbers of stressed and unstressed syllables allows us to elucidate whether lexical stress is indeed exploited to parse the signal. It addition, the phrase stimuli allowed for comparisons with prior studies that have used the MSS ratio (Borrie et al., 2012; Liss et al., 1998, 2000). Given the primary importance of phrasal stress patterning, phrase length was allowed to vary from three to five words in length and words contained within the phrases were either mono- or bisyllabic.

Perceptual selection and acoustic verification. As is usual in studies of this kind, a phrase selection and verification process was undertaken (e.g., Borrie et al., 2012; Liss et al., 1998, 2000; McAuliffe, Wilding, Rickard, & O'Beirne, 2012). The intent of this process was to select phrases from each speaker that most strongly conformed to our perceptual criteria (as judged by two expert judges, the first and third authors) and subsequently to verify acoustically these perceptual impressions. Such a procedure was necessary to ensure a relatively homogeneous set of experimental stimuli from which interpretation of perceptual error patterns could be conducted. For the first step, perceptual selection, 12 phrases spoken in the habitual condition were selected from each of the five speakers with hypokinetic dysarthria, for a total of 60 experimental stimuli. Each speaker's corresponding loud and slow tokens then underwent further perceptual screening to ensure that, perceptually, they exhibited evidence of louder and more effortful speech production (loud condition) and a reduced rate of speech (slow condition), respectively, relative to the habitual condition. Through this process we noted that two phrases produced in the slow condition, which corresponded with the selected habitual recording, were produced in error. These were removed from the stimulus set and the final slow condition comprised 58 phrases only. The final list of experimental phrases used in the perception experiment is provided in Appendix A.

Step 2, acoustic analysis, was undertaken for two reasons: (a) to confirm our perceptual impression that the behavioral modifications used elicited the required differences in speaking rate and vocal loudness and (b) to provide further information on the general effects of these manipulations on speech production, to later assist in the interpretation of listeners' syllable resemblance and lexical boundary error (LBE) patterns. To complete the acoustic analysis, each of the 178 experimental stimuli (60 from the habitual condition, 60 from the loud condition, and 58 from the slow condition) were transcribed and then automatically segmented at the phoneme level using the Hidden Markov Model Toolkit (Young et al., 2002). All automatically derived phoneme boundaries were then visually checked for accuracy by the second author, using standard criteria (Peterson & Lehiste, 1960). If any uncertainty arose in discriminating boundaries for consecutive consonants (e.g., /t/ and /s/) the boundary derived from automatic segmentation was preferred.

After the manual checking, custom Praat (Boersma & Weenink, 2009) scripts extracted the following nine acoustic data for each phrase: (a) intensity (calibrated mean intensity across the phrase); (b) total phrase duration (in seconds); (c) total pause duration (in milliseconds), with a pause defined as a silent interval greater than 50 ms (Robb, Maclagan, & Chen, 2004); (d) articulation rate (syllables/s; Robb et al., 2004); (e) total vowel duration across the phrase (in seconds); (f) total consonant duration across the phrase (in seconds); (g) F0 (mean F0 across the phrase, noting that F0 was checked, and manual adjustment of pitch traces completed if required); (h) F0 variation (F0 standard deviation across the phrase); and (i) intensity variation (intensity standard deviation across the phrase). We also calculated a measure of vowel space area using the temporal midpoint of the first and second formants from the vowels /i/, /a/, and /ɔ/ across the phrases. Mean formant values across each of the vowels were used to produce a measure of vowel space. In keeping with earlier work (Borrie et al., 2012), six tokens of each speaker's /a/, /i/, and /ɔ/ vowels (point vowels in New Zealand English), located within strong syllables, were averaged and used to calculate a vowel space area (VSA). The formula used in this calculation was as follows:

$$\text{VSA in Hz}^2 = 0.5 \times \text{ABS}\big[\text{F1/i/} \times (\text{F2/a/–F2/ɔ/}) + \text{F1/ɔ/} \times (\text{F2/i/–F2/a/}) + \text{F1/a/} \times (\text{F2/ɔ/–F2/i/})\big],$$

where ABS is equal to the absolute value.

To ensure the reliability of the acoustic findings, 20% of the original segmented phrases were randomly selected and again manually rechecked by the original judge (intrarater reliability) and a second judge (interrater reliability). Cronbach's alpha was used as a measure of agreement, with a value above .93 for intrarater reliability across all measures and above .89 for interrater reliability. The reliability of the acoustic measures was deemed acceptable.

The results of the acoustic analysis are presented in Table 2. To determine whether significant differences in acoustic parameters existed across the three conditions, we conducted repeated measures analyses of variance (ANOVAs) followed by post hoc pairwise comparisons with Bonferroni correction. Significant main effects of condition were found across all acoustic parameters (see Table 2). Post hoc tests confirmed that intensity was significantly higher in the loud condition compared with both the habitual and slow conditions ($p < .001$) and that intensity was significantly reduced in the slow condition compared with the habitual condition ($p = .007$). Significantly longer phrase durations were present in the slow condition compared with the habitual and loud conditions ($p < .05$), with the habitual and loud conditions exhibiting similar phrase durations ($p > .05$). Hence, our behavioral modifications elicited the required changes to speech rate and intensity.

Additional acoustic measures provided some indication of the locus of speech change across conditions (see Table 2). As expected, the slow condition exhibited significantly greater pause durations than either the habitual or loud conditions ($p < .001$); however, the habitual condition exhibited significantly longer pause durations than the loud conditions ($p < .05$). Articulation rate was significantly reduced in the slow condition compared with both the habitual and loud conditions ($p < .001$), and the loud condition exhibited a significantly reduced articulation rate relative to the habitual condition ($p < .05$). For vowel duration, all post hoc comparisons were significant (at either $p < .01$ or $p < .001$) indicating that vowel duration was greatest in the slow condition, followed by the loud condition and then the habitual condition. Consonant durations

Table 2. Mean and standard deviation values across the habitual, loud, and slow speaking conditions for selected acoustic measures.

Acoustic measure	Habitual ($M \pm SD$)	Loud ($M \pm SD$)	Slow ($M \pm SD$)	F
Speech intensity (dB)	87.20 ± 5.61	94.21 ± 4.52	85.38 ± 5.34	180.34*
Phrase duration (s)	1.33 ± 0.21	1.37 ± 0.23	1.85 ± 0.30	201.52*
Pause duration (ms)	16.35 ± 46.75	2.10 ± 16.30	136.61 ± 185.42	28.78*
Articulation rate (syll/s)	4.70 ± 0.71	4.52 ± 0.78	3.61 ± 0.65	160.84*
Total vowel duration (s)	0.62 ± 0.12	0.67 ± 0.13	0.80 ± 0.19	69.29*
Total consonant duration (s)	0.69 ± 0.17	0.70 ± 0.16	0.92 ± 0.22	76.32*
F0 (Hz)	139.57 ± 17.52	165.83 ± 12.52	140.10 ± 17.25	58.49*
F0 *SD* (Hz)	18.88 ± 8.76	28.36 ± 12.95	17.84 ± 7.68	32.44*
Speech intensity *SD* (dB)	6.97 ± 1.95	8.15 ± 2.17	9.92 ± 3.20	45.98*
Vowel space area (Hz^2)	142,714	134,131	160,399	

Note. Statisical analysis across conditions was completed with repeated measures analysis of variance, and resultant *F* statistics are reported. syll = syllables.

*$p < .001$.

were significantly longer in the slow condition compared with both the habitual and loud conditions (both $ps < .001$), although consonant durations in the habitual and loud conditions were similar ($p > .05$). Mean F0 was significantly higher in the loud condition compared with both the habitual and slow conditions ($p < .001$), which exhibited similar F0s ($p > .05$). In addition, F0 variation was significantly greater in the loud condition relative to both the habitual and slow conditions ($p < .001$), whereas the habitual and slow conditions exhibited similar levels of F0 variation ($p > .05$). Intensity variation was greatest in the slow condition, followed by the loud condition and then the habitual condition, respectively, with all post hoc tests significant at $p < .001$. Last, vowel space area was reduced by approximately 6% in the loud condition relative to the habitual condition and increased by approximately 12% in the slow condition relative to the habitual condition.

Procedure

Each of the 51 listener participants was assigned to one of three experimental conditions: (a) habitual ($n = 17$), (b) loud ($n = 17$), or (c) slow ($n = 17$). Data from the habitual condition were collected as part of an earlier study (McAuliffe, Gibson, Kerr, Anderson, & LaShell, 2013); hence, the additional participants were allocated consecutively to either the loud or slow conditions. All listeners underwent hearing screening and then completed the speech perception experiment.

The experiment was conducted in a quiet room and took approximately 25 min to complete. It was programmed using DirectRT (Jarvis, 2010). During the perception experiment, participants were seated in front of a laptop computer and the experimental phrases were presented through Sennheiser HD280 Pro circumaural headphones at a volume of approximately 65 dB. We chose to present the stimuli to listeners at the same intensity across all experimental conditions. Faint speech induces greater attention to lexical stress cues (Cutler & Butterfield, 1992), and the possible variation introduced by not controlling for intensity was deemed unwise for this initial work. Before the perception experiment commenced, all listeners were asked whether the volume of the practice phrases was appropriate to complete the task, and none requested a change to the presentation levels.

All participants were presented with task instructions on a laptop screen. They were told that they would hear a series of phrases and that each phrase would be composed of real English words, although the phrase itself may not make sense. Listeners were asked to verbally repeat each of the phrases exactly as they heard them (using real English words) and to respond with their best guess if they did not understand what was said. In cases where participants did not feel they were able to make a reasonable guess, they were to report the word *something*. A trained research assistant transcribed listener responses in real time. Responses were also audio-recorded, and the transcriptions were cross-checked prior to data coding and analysis (as per McAuliffe et al., 2012, 2013). We chose this process over a standard computer-based transcription task because it enables younger listener transcriptions to be compared with older listener transcriptions (e.g., McAuliffe et al., 2013).

Participants began the perception task with a short practice phase. In this phase, participants were presented with four example phrases, similar in construction to the test stimuli. This practice phase allowed participants to gain experience with the perception task and provided an informal evaluation of participants' short-term memory (Davis, Johnsrude, Hervais-Adelman, Taylor, & McGettigan, 2005). All participants successfully repeated the practice phrases, and thus no one was excluded from the experiment at this stage. After the practice phase, participants completed the perception task. Phrase stimuli were presented in six blocks of 10 stimuli, with the stimuli randomized within each block (because the slow condition consisted of 58 stimuli only, the final block in this condition comprised eight phrases). The six stimuli blocks were matched for number of phrases per speaker, stress pattern (iambic vs. trochaic). and number of possible LBEs. Participants were given a 2-min rest break after the first three blocks.

Data Analysis: Error Coding and Statistical Analysis

The final data set consisted of a series of 51 phrase transcriptions, 17 from each of the three experimental conditions. A rater blinded to listener group allocation analyzed the transcripts and coded for three dependent variables: (a) number of words correct (termed *accuracy*), (b) syllable resemblance and (c) the occurrence of LBEs. In the calculation of accuracy, words were deemed correct if they matched the target exactly, differed in the suffix by the tense "ed" or the plural "s," or involved a substitution of "a" and "the" (Borrie et al., 2012; Liss et al., 1998). The two further measures—syllable resemblance and LBEs—provide an indication of the error patterns that underlie the success, or otherwise, of listeners' interpretation attempts.

Syllable resemblance provides a measure of the phonemic resemblance of incorrectly reported syllables to the correct syllable. A syllable is scored as resembling the target if at least half of its phonemes match those of the target syllable. Thus, for a syllable to be counted as resembling the target syllable, syllables with two phonemes required one phoneme to match the target, syllables with three phonemes required two correct phonemes, syllables with four phonemes required at least two correct phonemes, and so on. A higher syllable resemblance score indicates that although a syllable may have been heard in error, the phonemic composition of the transcription bears some resemblance to the target. It is interpreted to indicate that the acoustic–phonetic representation of the signal was preserved enough for a listener to map this to an existing phonemic representation. In contrast, a lower syllable resemblance score is interpreted to indicate that the acoustic–phonetic aspects of that signal were less intact; hence, the listener was less successful in mapping this to existing phonemic representations. For further examples of the use and interpretation of this index, see Borrie et al. (2012). A worked example of coding of syllable resemblance and accuracy is provided in Appendix B.

The type and number of LBEs were also coded. LBEs were classified according to either the incorrect deletion or insertion of a lexical boundary and for the occurrence of this error before a strong or weak syllable. Hence, there were four possible error types: (a) insertion of a lexical boundary before a strong syllable (IS), (b) deletion of a lexical boundary before a weak syllable (DW), (c) insertion of a lexical boundary before a weak syllable, and (d) deletion of a lexical boundary before a strong syllable. Both IS and DW error types are considered predictable in the English language: If listeners are making segmentation decisions based on the presence of strong syllables, misperceptions in which a strong syllable is taken to be word initial would be predicted (Cutler & Butterfield, 1992). If the number of predictable error types outweighs the number of unpredictable errors, a listener is said to be using a stress-based approach to lexical segmentation (Spitzer et al., 2007). This proportion of predicted error types is also called the *MSS ratio*, and it is calculated as (IS + DW)/total number of LBEs. See Appendix C for a worked example of the coding process, and for further examples see Liss et al. (1998).

We conducted reliability analyses for all three dependent measures. Twenty percent of the transcripts were randomly selected and rechecked by the original judge and a second judge. Agreement across all measures was high, with Spearman's rho scores of above .95 for interrater reliability and above .98 for intrarater reliability across all measures.

We conducted statistical analysis of the transcription data using a variety of measures. To analyze accuracy of phrase transcription and syllable resemblance scores, we used mixed-effects modeling. Mixed-effects models are advantageous for repeated measures data such as these because they enable simultaneous consideration of multiple sources of individual variance (e.g., the effect of subject and items) while evaluating fixed factors (e.g., experimental condition). This is in contrast to traditional ANOVA analyses, which require different analyses for multiple sources of individual variance (Baayen, Davidson, & Bates, 2008). For analysis of LBEs, we used a combination of a one-way ANOVA with post hoc tests using Bonferroni correction, *t* tests, and chi-square goodness of fit.

Results

Accuracy of Phrase Transcription

Slow and loud speech modification strategies demonstrated clear positive effects on intelligibility. The average transcript accuracy score for participants in the habitual condition was 45.23% (SD = 5.21%), for participants in the loud condition it was 60.45% (SD = 4.03%), and for participants in the slow condition it was 69.28% (SD = 3.65%). Further analyses were performed using binomial mixed-effects models to characterize the data more completely.

The analysis began with a full model consisting of the fixed effects of condition and phrase block and participant descriptors of listener age and sex. We created a maximized random effects structure for participants (with individual slopes for phrase block), and phrase number was nested within speaker. Model evaluation proceeded in a backward-stepwise iterative fashion seeking to reduce the full model to a reduced model containing only significant fixed effects (with alpha set at .05). Model fitting was independently supported by fitness comparisons. The final model is detailed in Table 3.

The final model revealed that transcript accuracy was significantly increased in both the loud (p = .01) and slow (p < .001) conditions relative to the habitual condition. Although the main effect of phrase block was not significant (p > .05), it interacted significantly with condition (see Figure 1). For participants in the loud condition, transcript accuracy was found to increase significantly across blocks when compared with participants in the habitual condition (p < .001). Performance across blocks was similar for participants in the slow and habitual conditions (p > .05).

To compare the loud and slow conditions against each other, we re-leveled the model, with the slow condition

Table 3. Coefficients of a binomial mixed-effects model for transcript accuracy, with the habitual condition mapped to the intercept.

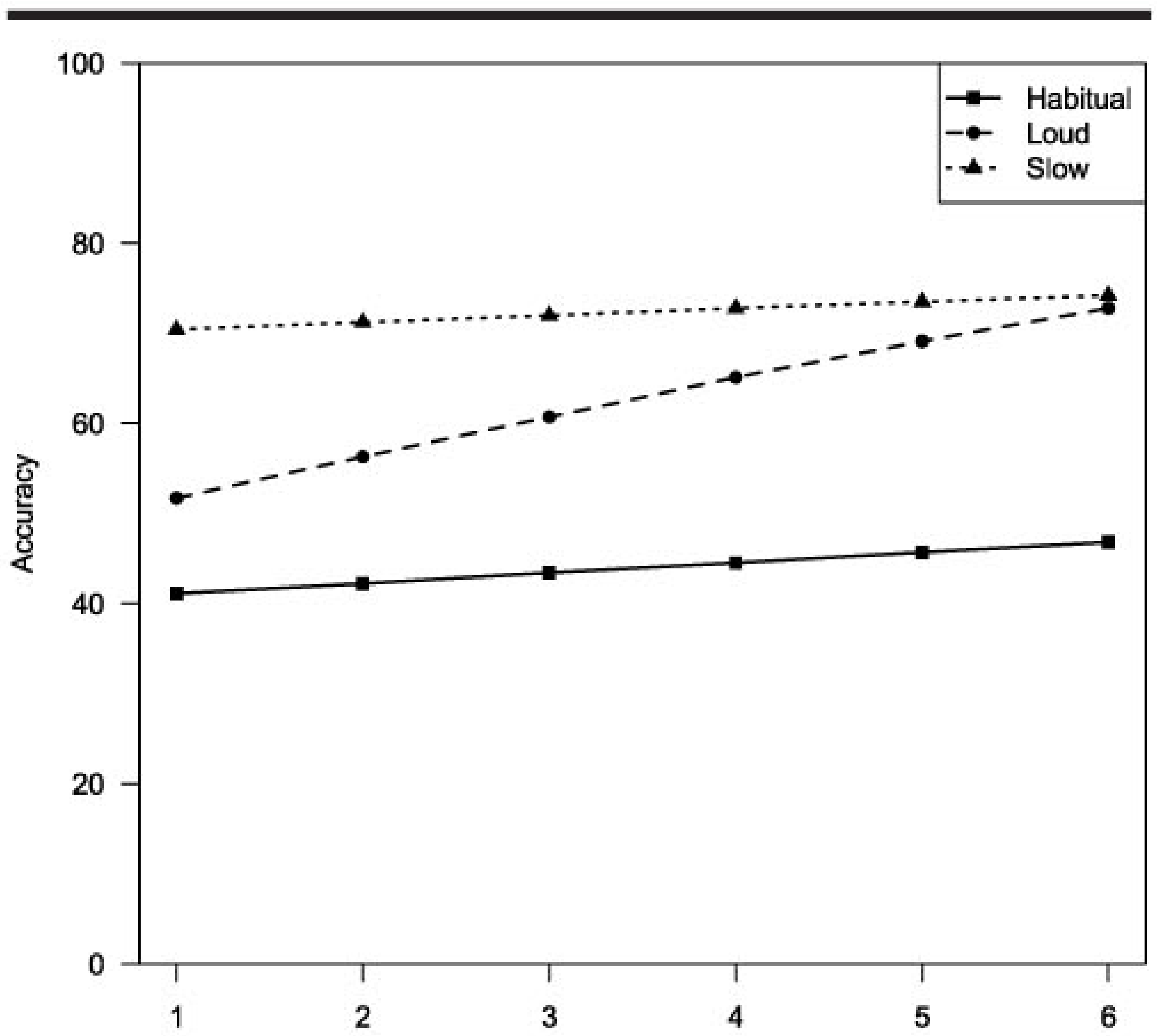
Fixed effects	Estimate	*SE*	*z*	Pr (> |z|)
Intercept	−0.406	0.364	−1.116	.265
Loud condition	0.290	0.113	2.579	.010
Slow condition	1.237	0.117	10.542	< .001
Phrase block	0.046	0.064	0.720	.471
Loud × Phrase Block	0.137	0.030	4.648	< .001
Slow × Phrase Block	−0.009	0.031	−0.292	.770

mapped to the intercept. In comparing the slow and loud conditions, our analysis indicated that participants' transcript accuracy was less accurate in the loud condition than in the slow condition, $\beta = -.947$, $SE = .117$, $p < .001$. There was also an interaction between condition and phrase block. Participant accuracy in the loud condition was found to increase significantly across block compared with participant accuracy in the slow condition, $\beta = .146$, $SE = .031$, $p < .001$. The significance of the remaining results remained unchanged from the original model.

Syllable Resemblance

The measure of syllable resemblance indicated that slow and loud speech modification strategies also demonstrated clear positive effects on listeners' ability to map the acoustic–phonetic aspects of the signal to existing phonemic representations. The average syllable resemblance score for participants in the habitual condition was 28.98% ($SD = 7.89\%$), for participants in the loud condition it was 42.04% ($SD = 6.26\%$), and for participants in the slow condition it was 43.19% ($SD = 12.67\%$). To characterize the data more completely, we conducted further analyses using binomial mixed-effects models. These analyses used the same process and independent variables as in the analysis of transcript accuracy. The final model for syllable resemblance is detailed in Table 4.

Figure 1. Interaction between listener condition and phrase block for the dependent variable of transcript accuracy.



The final model confirmed that the probability of an erroneous syllable resembling the target was significantly increased in the loud ($p < .001$) and slow ($p < .001$) conditions relative to the habitual condition. Phrase block was also a significant predictor of syllable resemblance ($p < .01$). In addition, a significant interaction (see Figure 2) was present between the slow condition and phrase block ($p < .01$), such that in the slow condition the probability of syllable resemblance increasing across blocks was significantly reduced compared with the habitual condition. There was no interaction between the loud condition and phrase block ($p > .05$).

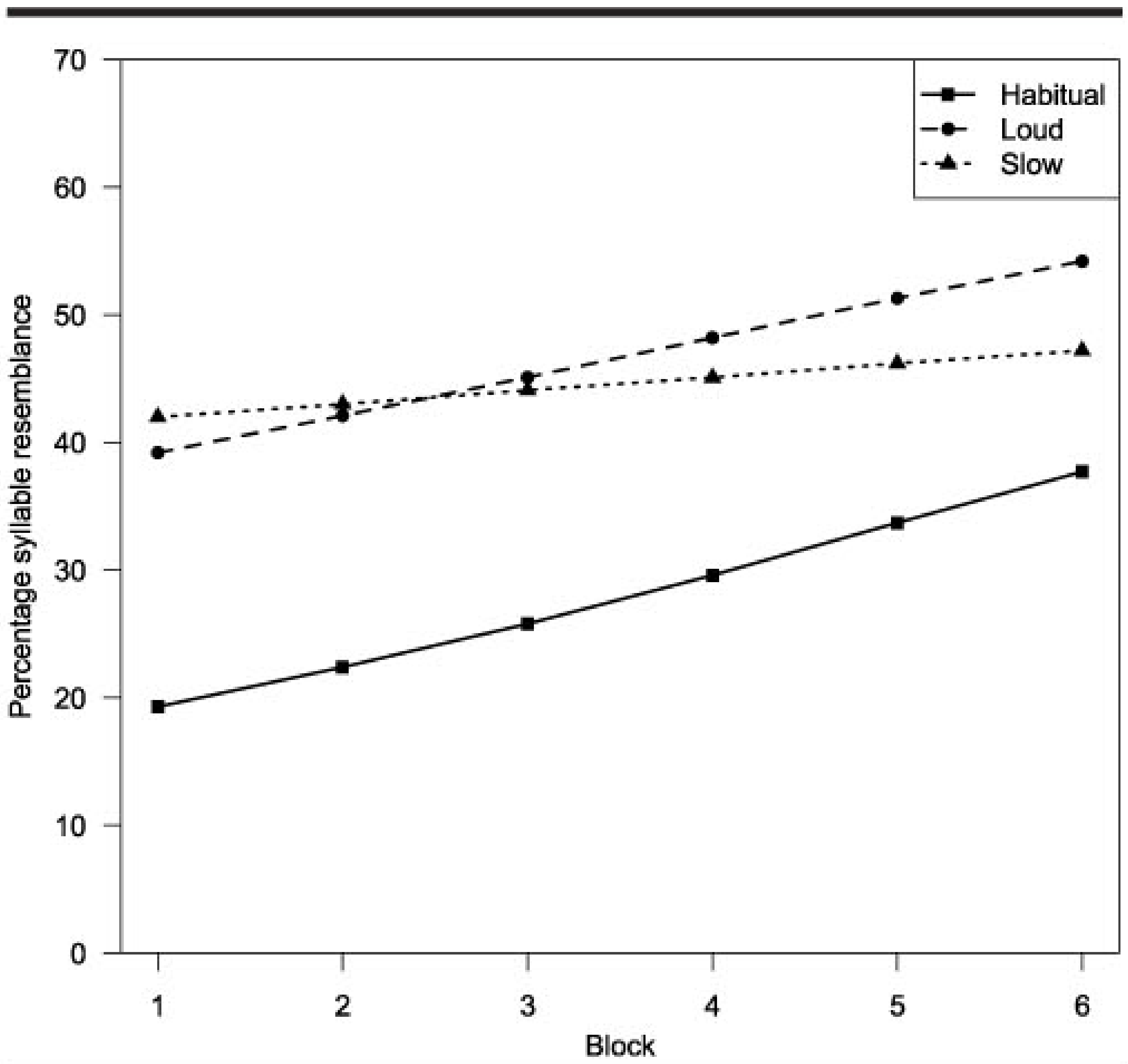
To compare the loud and slow conditions against each other, we re-leveled the model, with the slow condition mapped to the intercept. This analysis revealed no significant difference between the loud and slow conditions for syllable resemblance ($p > .05$). Furthermore, there was no significant main effect of block ($p > .05$) and no interaction between the loud condition and phrase block ($p > .05$); that is, listeners in the loud and slow conditions exhibited similar rates of syllable resemblance, and furthermore this did not change significantly across blocks. The significance of the remaining results remained unchanged from the original model.

The lack of a Phrase Block × Condition interaction for participants in the loud condition contradicts the previous model. To evaluate the discrepancy, we performed two additional analyses. In the first analysis, the model was refit to the data by reordering loud as the comparison level of condition. This model revealed a positive trend of block, $\beta = .122$, $SE = .065$, $p = .061$, but this was not significant at the alpha = .05 level. In the second analysis, participants in the loud condition were analyzed separately from the participants in the other conditions. In this model, block was again found to have a positive trend, $\beta = .147$, $SE = .085$, $p = .084$. Combined, all models support a trend in the loud condition for an increase in the proportion of syllable resemblance errors across block, although this was not statistically significant in our current data.

Table 4. Coefficients of a binomial mixed-effects model for syllable resemblance, with the habitual condition mapped to the intercept.

Fixed effects	Estimate	*SE*	*z*	Pr (> |z|)
Intercept	−1.615	0.339	−4.761	< .001
Loud condition	1.053	0.211	4.996	< .001
Slow condition	1.250	0.225	5.713	< .001
Phrase block	0.186	0.063	2.938	.003
Loud × Phrase Block	−0.065	0.043	−1.485	.138
Slow × Phrase Block	−0.145	0.047	−3.076	.002

Figure 2. Interaction between listener condition and phrase block for the dependent variable of syllable resemblance.



Lexical Boundary Error Patterns

The analysis of LBEs is summarized in Table 5. A one-way ANOVA revealed a significant difference in average number of LBEs across conditions, $F(2, 48) = 11.09$, $p < .001$, $\omega = .53$. Post hoc *t* tests with Bonferroni correction showed that individuals in the slow condition exhibited significantly fewer LBEs than those in the habitual ($p < .001$) and loud ($p < .01$) conditions. There was no significant difference in the average number of errors in the loud and habitual conditions ($p > .05$). With chi-square analyses, we examined whether predicted versus unpredicted errors were evenly distributed within the three listening conditions. Within each condition, error distribution followed the predicted pattern, with significantly more errors of the predicted type versus the unpredicted type: habitual condition, $\chi^2(1) = 78.04$, $p < .001$; loud condition, $\chi^2(1) = 22.51$, $p < .001$; slow condition, $\chi^2(1) = 49.71$, $p < .001$. Although all groups adhered to the predicted error pattern, they did so with differing strengths of adherence. A one-way ANOVA revealed a significant difference in proportion of predicted errors across conditions (i.e., MSS ratio), $F(2, 48) = 7.18$, $p < .01$, $\omega = .44$. Post hoc *t* tests with Bonferroni correction showed that individuals in the loud condition exhibited a significantly reduced proportion of predicted errors compared with those in both the habitual condition ($p = .01$) and the slow condition ($p < .01$). The slow and habitual conditions exhibited a similar ratio of predicted errors ($p > .05$).

Given the significant effects of phrase block in the preceding accuracy and syllable resemblance analyses, we conducted a further descriptive examination of the LBE data across the first versus second half of the study to gain insight into whether some degree of shift in attention to lexical stress cues had occurred in the first versus second half of the experiment (see Table 6). For participants in the habitual group, the average number of LBEs increased significantly in the second half of the listening experiment, $t(16) = 4.37$, $p < .001$, and more predicted errors were made, $t(16) = 5.96$, $p < .001$. Listeners in the loud conditions exhibited a significant reduction in the number of LBEs in the second half of the experiment, $t(16) = -5.82$, $p < .001$, although their attention to syllabic stress remained constant ($p > .05$). Finally, for listeners in the slow condition the number of errors remained constant across the two halves of the experiment ($p > .05$); however, they produced significantly more predicted errors in the second half, $t(16) = 2.27$, $p < .05$.

Discussion

In the current investigation we examined how two commonly used behavioral intervention strategies—increased vocal loudness and reduced speech rate—affected listeners' perceptual processing of hypokinetic dysarthric speech. The intent was to provide a window into the cognitive–perceptual strategies that listeners used on encountering dysarthric speech modified in such ways. First, the results of the study indicated that both behavioral speech modifications provided an obvious perceptual benefit to listeners, as demonstrated through significant gains to transcript accuracy relative to the habitual speaking condition. Overall, for the speakers with hypokinetic dysarthria sampled here, the slow speaking condition afforded the greatest intelligibility benefit. However, variable learning effects were observed within

Table 5. Lexical boundary errors across conditions and the proportion of those errors that were of a predictable or unpredictable type.

		Lexical boundary errors			
Group	**Total no.**	**Average no. (*M* ± *SD*)**	**No. predict.**	**No. unpredict.**	**Prop. predict. (*M* ± *SD*)**
Habitual	406	23.88 ± 7.00	292	114	0.72 ± 0.07
Loud	375	22.06 ± 7.55	234	142	0.62 ± 0.09
Slow	239	14.06 ± 4.42	174	65	0.74 ± 0.12

Note. Predicted errors are of IS (insertion of a word boundary before a strong syllable) and DW (deletion of a word boundary before a weak syllable) type. Unpredicted errors are of DS (deletion of a word boundary before a strong syllable) and IW (insertion of a word boundary before a weak syllable) type. The ratio of predicted to unpredicted errors is also termed the *metrical segmentation strategy* ratio. prop. = proportion; predict. = predictable error type; unpredict. = unpredicted error type.

Table 6. Details of the average number of lexical boundary errors and proportion of predicted errors in the first versus last 30 phrases of the perception experiment.

	Average no. LBEs ($M \pm SD$)		Proportion predict. ($M \pm SD$)	
Condition	First 30	Last 30	First 30	Last 30
Habitual	10.12 ± 3.46	13.76 ± 4.29	0.61 ± 0.12	0.81 ± 0.07
Loud	13.76 ± 4.98	8.29 ± 3.35	0.60 ± 0.13	0.67 ± 0.14
Slow	7.35 ± 2.76	6.71 ± 2.62	0.69 ± 0.16	0.81 ± 0.15

the conditions, suggesting that the different acoustic profiles of loud and slow speech acted on perceptual processing in different ways. Indeed, for the loud condition in particular it appeared that listeners had learned, and successfully applied, some component of the speech signal to enhance transcription accuracy across the course of the experiment. Second, the mechanism by which improved comprehension of dysarthric speech occurred depended on the speaking condition. We suggest that, when exposed to loud speech, listeners prioritized acoustic–phonetic cues in their speech processing, whereas for listeners in the slow condition, lexical stress cues appeared paramount. In the following paragraphs we discuss the study findings in relation to the existing literature on cognitive–perceptual processing of distorted speech, the acoustic changes that may have underpinned the perceptual processing observations, and models of lexical access and segmentation.

Overall Accuracy

Both the loud and slow speech conditions resulted in significant improvements to transcript accuracy relative to the habitual speaking mode. Hence, we can conclude that both forms of speech modification resulted in demonstrable benefit to listeners' comprehension of hypokinetic dysarthric speech. Overall, listeners exposed to slow speech exhibited the highest transcript accuracy scores, significantly outperforming those in the habitual and loud speech conditions. Although previous authors have cautioned that slow speech may interfere with perceptual processing through the taxing of memory resources (Liss, 2007), for the individuals with hypokinetic dysarthria included in the current investigation, this strategy was in fact advantageous. It appeared likely that this finding was related to the fast speech rate exhibited by the dysarthric speakers in this investigation. When rate reduction strategies were enacted, speech rate increased in the direction of the normal range; hence, concerns regarding increased processing times and taxing of memory resources were mitigated. Whether similar results would be obtained if the speech rate of the dysarthric speakers was within the normal range (or slow) in their habitual speaking mode or, alternatively, if the phrase stimuli were longer, included greater syntactic complexity, or were elicited using different cuing paradigms, remains to be seen.

It is interesting that although slow speech produced an intelligibility advantage overall compared with the loud condition, listeners allocated to the loud condition exhibited an improved ability to decipher dysarthric speech across phrase blocks—an effect that was not present in the slow condition. Testable explanations include that listeners in the loud condition had either learned something about the acoustic–phonetic properties of loud speech, and were applying this in subsequent attempts to decipher the signal, or had evidenced adaptation to its rhythmic properties in some way, or a combination of the two. The source of this perceptual change across phrase blocks cannot be determined from a simple analysis of intelligibility. However, an error analysis could provide insight into these potential explanations by examining differential processing strategies by condition.

Syllable Resemblance

Syllable resemblance scores were significantly higher in both the loud and slow speaking conditions relative to the habitual condition. This indicated that both conditions improved the acoustic–phonetic realization of the speech signal to such an extent that phonemic ambiguity was reduced, and listeners were better able to map the signal to existing phonemic representations. Although similar improvements to syllable resemblance were made in both the slow and loud conditions, it is likely that different underlying acoustic properties stimulated these perceptual improvements.

We posit that, for listeners in the loud condition, the reduced phonemic ambiguity evidenced was due primarily to enhanced consonant precision on the part of the dysarthric speakers. We take this position for two reasons. First, previous studies have noted that loud speech is associated with improved acoustic distinctiveness for stops (Tjaden & Wilding, 2004) and increased articulatory velocity and excursion (Dromey & Ramig, 1998). Although a measure of consonant integrity was not included in the current study, we assume that similar changes may have occurred within the current speaker group. Second, vowel space was reduced in the loud condition relative to the habitual condition. It is unlikely that speech produced with a reduced vowel space would result in improvements to phonemic mapping, leading us to conclude that improved consonant precision was the most obvious source of the higher syllable resemblance scores in the loud condition. Of particular interest is that listeners in the loud condition exhibited a trend toward increased syllable resemblance across phrase blocks. It appears that listeners had, by some perceptual mechanism, expanded or updated their existing phoneme representations to include the distortion present in dysarthric speech. Such an interpretation is supported by previous studies that have observed listeners' ability to update their phonemic categories to include ambiguous sounds (Kraljic & Samuel, 2005; Norris, McQueen, & Cutler, 2003).

Analysis of syllable resemblance errors also highlighted a reduction in phonemic ambiguity for listeners in the slow condition relative to those in the habitual condition. It appears likely that the increased vowel space associated with slower speech played at least some part in the reduced phonemic ambiguity experienced by listeners. The presence of a concurrent increase in phoneme durations in this condition may also have been a factor in this finding. It is interesting that, unlike both the loud and habitual conditions, individuals in the slow condition exhibited no change to their syllable resemblance scores across the duration of the experiment. Either the listeners in this condition did not modify their phonemic representations when speech was produced in this manner, or modification did occur but was not captured by the syllable resemblance measure included in this study. Analysis of the LBE findings shed further light on these findings.

LBE Analysis

Our LBE analysis demonstrated that all three listener groups attended to lexical stress in their attempts to segment speech; however, those in the habitual and slow conditions exhibited significantly greater strength of adherence to this strategy compared with listeners in the loud condition. Listeners in the slow condition also exhibited significantly fewer LBEs than those in either the habitual or loud groups, and differences were noted in the groups' performances across the first and second halves of the experiment.

We speculated earlier in this article that improved integrity of consonant production may have played a role in the syllable resemblance findings for listeners in the loud condition. This interpretation also fits with the pattern of LBEs: While listeners in the loud condition attended to stress contrasts in determining word boundaries, they paid the least attention of any of the three groups. Acoustic analyses provided some insight into why this might have occurred. First, the loud condition exhibited the smallest vowel space of the three conditions, hence the contrast between weak and strong syllables may not have been as prominent for listeners in this group. Second, speech rate was fast in the loud condition—similar to that of the habitual condition—without obvious pause boundaries between words. Third, F0 variation was significantly higher in this condition, perhaps unnaturally or unusually high in relation to the speech of healthy speakers of New Zealand English (Borrie et al., 2012; McAuliffe et al., 2013). This may have interfered with prosody, thereby reducing rhythmic expectancy and limiting listeners in their ability to use rhythm as a cue to word segmentation. Combined, these acoustic changes likely led to a reduced ability, on the part of listeners, to confidently apply knowledge of predictable lexical stress patterns to segment the speech stream. This, plus the finding of improved syllable resemblance, suggests that listeners in the loud condition may have given higher weighting to acoustic–phonetic cues in completing lexical segmentation.

Further support for this interpretation is shown through comparison of LBE findings in the first and second halves of the experiment. Listeners in the loud condition maintained their proportion of predicted error types across the course of the experiment (i.e., they maintained their level of attention to lexical stress). However, they exhibited a significant drop in their average number of LBEs in the latter half of the experiment. This indicates that a shift in perceptual processing had occurred that enabled listeners in this group to better parse the speech stream. When we consider this in tandem with the syllable resemblance findings, we suggest that listeners use of acoustic–phonetic properties of the signal as cues to segmentation produced the greatest magnitude of benefit in the later stages of the experiment. It is quite possible that by the latter half of the study listeners had updated their phonemic representations (as evidenced by the syllable resemblance findings) and were successfully applying a segmentation strategy that prioritized acoustic–phonetic cues. This strategy resulted in more accurate detection of word boundaries and a concomitant increase in transcript accuracy. Such an interpretation also appears consistent with models of perceptual learning in which listeners turn their attention toward the most high-yield cue (e.g., Sebastián-Gallés, Dupoux, Costa, & Mehler, 2000)—in this case, the acoustic–phonetic aspects of the signal.

LBE analysis of data from listeners in the slow condition revealed differences in approaches to, and success of, lexical segmentation. As expected, listeners in this condition were significantly more accurate in their detection of word boundaries compared with listeners in the habitual and loud conditions. The existence of between-word pauses for at least some words in the phrase resulted in a clearer delineation of word boundaries. Of particular interest, however, was the use of lexical stress as a cue to segmentation. Listeners in the slow condition exhibited the strongest adherence to metrically based segmentation, significantly greater than listeners in the loud condition. Lansford and colleagues (2011) hypothesized that increased vowel space present in a slow speaking mode would likely facilitate listeners' use of rhythm cues in detecting word boundaries, and it appeared that this occurred in the current study. Furthermore, listeners' use of this cue increased significantly in the later stages of the experiment, indicating a greater weighting given to these rhythm cues as the experiment progressed. It appears that listeners had learned the rhythmic qualities of slow speech and were applying this strategy to aid parsing. This is not to say that listeners did not also use acoustic–segmental cues to segmentation; they simply prioritized lexical stress to a greater extent. This approach led to more successful segmentation that, combined with reduced phonemic ambiguity, resulted in improved lexical access and higher transcript accuracy overall.

One further issue that we should address is that of learning in the slow condition. It was somewhat surprising that individuals in this condition did not exhibit the same degree of learning as those in the loud condition, and this was not due to ceiling effects (see Figure 1). From the outset, however, slow speech was inherently more understandable than the loud and habitual speech conditions. Perhaps listeners in the slow condition, having been provided with relatively salient word boundaries, more distinctive and natural cues to rhythm, and additional processing time, did not demonstrate the same level of need to search out optimal cues, or learn from these cues. Alternatively, it is possible that some form of perceptual learning ceiling existed that participants in neither the slow group nor the loud group could overcome without additional perceptual training directed toward specific cues in the signal. Whether loud and slow speech conditions promote different speeds and magnitudes of learning requires further study.

Summary, Limitations, and Directions for Future Research

This investigation has provided insight into the cognitive–perceptual processing strategies used by listeners when attempting to comprehend dysarthric speech modified by increased loudness and reduced rate. Our findings indicate that slow speech provided a quick and relatively effective means of enhancing listeners' accuracy at deciphering the speech signal; in contrast, loud speech required listeners to undergo a period of familiarization or learning before obtaining a similar level of perceptual benefit. Although both techniques resulted in improved transcription accuracy, the mechanisms by which these gains occurred differed.

We theorized that loud speech drove listeners to prioritize acoustic–phonetic cues in their attempts to resolve the degraded signal, whereas slow speech occasioned listeners to prioritize lexical stress cues. These conclusions must be considered in context. First, loud and slow speech were achieved using a magnitude-estimation paradigm; it is likely that other elicitation techniques may produce different results. Second, the loud condition was undertaken without any additional benefit that may have been derived by maintaining the amplitude difference between this and other conditions. Third, speech stimuli were collected in a set order (habitual, slow, loud). Finally, the study results relate to the current speaker group only; alternative types of dysarthria or perceptual selection processes may result in different findings.

In spite of these limitations, the findings highlight the utility of the cognitive–perceptual approach in improving our knowledge base related to the treatment of dysarthria. This evolving research paradigm, and the results of the current study, provide a pathway for future investigations. Researchers should consider whether different types of speech cues have the potential to provide even greater benefit to listeners (e.g., clear speech—which commonly results in slower and more clearly articulated speech—may provide greater benefit); whether different speech characteristics, variations in severity, and more complex or realistic speech stimuli provide different magnitudes of benefit or elicit changes in processing; whether different listener characteristics influence processing (e.g., older listeners and those who have familiarity with dysarthric speech); and whether similar findings would be achieved if listeners were presented with audio-visual cues in processing of dysarthric speech.

It is premature to provide direct clinical implications based on the present findings; however, the results do speak to the need for consideration of the listener/communicative partner in intervention planning for individuals with hypokinetic dysarthria, namely, consideration of the possible effects of intervention choices on specific cognitive–perceptual processing strategies of listeners and the possibility of incorporating listener-based variables as outcome measures. The findings of this and subsequent investigations will ultimately provide an improved understanding of how the signal characteristics of dysarthric speech, modified through intervention, act on listener processing. This information can only positively augment clinical decision making and provide a better understanding of the effectiveness and suitability of treatment choices.

Acknowledgments

We sincerely thank the participants with Parkinson's disease, their families, and the individual listeners who participated in this research. Support from the New Zealand Neurological Foundation (Project Grant 0827-PG) is gratefully acknowledged. We also thank Julie Liss for advice and suggestions on various aspects of this research, Rene Utianski for comments on an earlier version of the article, and Maryam Ghaleh for research assistance.

References

Adams, S. G. (1994). Accelerating speech in a case of hypokinetic dysarthria: Descriptions and treatment. In J. A. Till, K. M. Yorkston, & D. R. Beukelman (Eds.), *Motor speech disorders: Advances in assessment and treatment* (pp. 213–228). Baltimore, MD: Brookes.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59,* 390–412.

Boersma, P., & Weenink, D. (2009). Praat: Doing phonetics by computer (Version 5.1.05) [Computer software]. Retrieved from http://www.praat.org/

Borrie, S. A., McAuliffe, M. J., Liss, J. M., Kirk, C., O'Beirne, G., & Anderson, T. (2012). Familiarization conditions and the mechanisms that underlie improved recognition of dysarthric speech. *Language and Cognitive Processes, 27,* 1039–1055.

Choe, Y.-K., Liss, J. M., Azuma, T., & Mathy, P. (2012). Evidence of cue use and performance differences in deciphering dysarthric speech. *The Journal of the Acoustical Society of America, 131,* EL112–EL118.

Cutler, A., & Butterfield, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory and Language, 31,* 218–236.

Cutler, A., & Carter, D. M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language, 2,* 133–142.

Darley, F. L., Aronson, A. E., & Brown, J. R. (1975). *Motor speech disorders*. Philadelphia, PA: W.B. Saunders.

Davis, M. H., Johnsrude, I. S., Hervais-Adelman, A., Taylor, K., & McGettigan, C. (2005). Lexical information drives perceptual learning of distorted speech: Evidence from the comprehension of noise-vocoded sentences. *Journal of Experimental Psychology: General, 134,* 222–241.

de Lau, L. M. L., & Breteler, M. M. B. (2006). Epidemiology of Parkinson's disease. *The Lancet Neurology, 5,* 525–535.

Dromey, C., & Ramig, L. O. (1998). Intentional changes in sound pressure level and rate: Their impact on measures of respiration, phonation and articulation. *Journal of Speech, Language, and Hearing Research, 41,* 1003–1118.

Duffy, J. R. (2005). *Motor speech disorders: Substrates, differential diagnosis, and management* (2nd ed.). St. Louis, MO: Elsevier Mosby.

Hammen, V. L., & Yorkston, K. M. (1996). Speech and pause characteristics following speech rate reduction in hypokinetic dysarthria. *Journal of Communication Disorders, 29,* 429–445.

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,* 244–253.

Hartelius, L., & Svensson, P. (1994). Speech and swallowing symptoms associated with Parkinson's disease and multiple sclerosis: A survey. *Folia Phoniatrica et Logopaedica, 46,* 9–17.

Herd, C. P., Tomlinson, C. L., Deane, K. H. O., Brady, M. C., Smith, C. H., Sackley, C., & Clarke, C. E. (2012). Speech and language therapy versus placebo or no intervention for speech problems in Parkinson's disease. *Cochrane Database of Systematic Reviews, Issue 8, Article No. CD002812*. doi:10.1002/14651858.CD002812.pub2

Jarvis, B. G. (2010). DirectRT (Version 2010) [Computer software]. New York, NY: Empirisoft Corporation.

Jaywant, A., & Pell, M. D. (2010). Listener impressions of speakers with Parkinson's disease. *Journal of the International Neuropsychological Society, 16,* 49–57.

Johnson, J. A., & Pring, T. R. (1990). Speech therapy and Parkinson's disease: A review and further data. *British Journal of Disorders of Communication, 25,* 183–194.

Jusczyk, P. W., & Luce, P. A. (2002). Speech perception and spoken word recognition: Past and present. *Ear and Hearing, 23,* 2–40.

Kent, R. D., Weismer, G., Kent, J. F., & Rosenbek, J. C. (1989). Toward phonetic intelligibility testing in dysarthria. *Journal of Speech and Hearing Disorders, 54,* 482–499.

Kleinow, J., Smith, A., & Ramig, L. O. (2001). Speech motor stability in IPD: Effects of rate and loudness manipulations. *Journal of Speech, Language, and Hearing Research, 44,* 1041–1051.

Kraljic, T., & Samuel, A. G. (2005). Perceptual learning for speech: Is there a return to normal? *Cognitive Psychology, 511,* 141–178.

Lansford, K. L., Liss, J. M., Caviness, J. N., & Utianski, R. L. (2011). A cognitive–perceptual approach to conceptualizing speech intelligibility deficits and remediation practice in hypokinetic dysarthria. *Parkinson's Disease, 2011.* Retrieved from http://www.hindawi.com/journals/pd/2011/150962/

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. 195–231). San Diego, CA: Plural.

Liss, J. M., Spitzer, S., Caviness, J. N., & Adler, C. (2002). The effects of familiarization on intelligibility and lexical segmentation in hypokinetic and ataxic dysarthria. *The Journal of the Acoustical Society of America, 112,* 3022–3030.

Liss, J. M., Spitzer, S., Caviness, J. N., Adler, C., & Edwards, B. (1998). Syllabic strength and lexical boundary decisions in the perception of hypokinetic dysarthric speech. *The Journal of the Acoustical Society of America, 104,* 2457–2466.

Liss, J. M., Spitzer, S. M., Caviness, J. N., Adler, C., & Edwards, B. W. (2000). Lexical boundary error analysis in hypokinetic and ataxic dysarthria. *The Journal of the Acoustical Society of America, 107,* 3415–3424.

Lowit, A., Dobinson, C., Timmins, C., Howell, P., & Kröger, B. (2010). The effectiveness of traditional methods and altered auditory feedback in improving speech rate and intelligibility in speakers with Parkinson's disease. *International Journal of Speech-Language Pathology, 12,* 426–436.

Mattys, S. L., White, L., & Melhorn, J. F. (2005). Integration of multiple speech segmentation cues: A hierarchical framework. *Journal of Experimental Psychology: General, 134,* 477–500.

McAuliffe, M. J., Gibson, E. M. R., Kerr, S. E., Anderson, T., & LaShell, P. J. (2013). Vocabulary influences older and younger listeners' processing of dysarthric speech. *The Journal of the Acoustical Society of America, 134,* 1358–1368.

McAuliffe, M. J., Wilding, P. J., Rickard, N. A., & O'Beirne, G. A. (2012). Effect of speaker age on speech recognition and perceived listening effort in older adults with hearing loss. *Journal of Speech, Language, and Hearing Research, 55,* 838–847.

McRae, P. A., Tjaden, K., & Schoonings, B. (2002). Acoustic and perceptual consequences of articulatory rate change in Parkinson disease. *Journal of Speech, Language, and Hearing Research, 45,* 35–50.

Miller, N., Deane, K. H. O., Jones, D., Noble, E., & Gibb, C. (2011). National survey of speech and language therapy provision for people with Parkinson's disease in the United Kingdom: Therapists' practices. *International Journal of Language & Communication Disorders, 46,* 189–201.

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

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

Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology, 47,* 204–238.

Peterson, G. E., & Lehiste, I. (1960). Duration of syllable nuclei in English. *The Journal of the Acoustical Society of America, 32,* 693–703.

Ramig, L. O., Sapir, S., Countryman, S., Pawlas, A. A., O'Brien, C., Hoehn, M. M., & Thompson, L. (2001). Intensive voice treatment (LSVT) for patients with Parkinson's disease: A 2 year follow up. *Journal of Neurology, Neurosurgery, & Psychiatry, 71,* 493–498.

Ramig, L. O., Sapir, S., Fox, C. M., & Countryman, S. (2001). Changes in vocal loudness following intensive voice treatment (LSVT) in individuals with Parkinson's disease: A comparison with untreated patients and normal aged-matched controls. *Movement Disorders, 16,* 79–83.

Robb, M. P., Maclagan, M. A., & Chen, Y. (2004). Speaking rates of American and New Zealand varieties of English. *Clinical Linguistics & Phonetics, 18,* 1–15.

Rosen, K. M., Folker, J. E., Murdoch, B. E., Vogel, A. P., Cahill, L. M., Delatycki, M. B., & Corben, L. A. (2011). Measures of spectral change and their application to habitual, slow, and clear speaking modes. *International Journal of Speech-Language Pathology, 13,* 165–173.

Sapir, S., Spielman, J. L., Ramig, L. O., Story, B. H., & Fox, C. (2007). Effects of intensive voice treatment (the Lee Silverman Voice Treatment [LSVT]) on vowel articulation in dysarthric individuals with idiopathic Parkinson disease: Acoustic and

perceptual findings. *Journal of Speech, Language, and Hearing Research, 50,* 899–912.
Sebastián-Gallés, N., Dupoux, E., Costa, A., & Mehler, J. (2000). Adaptation to time-compressed speech: Phonological determinants. *Perception & Psychophysics, 62,* 834–842.
Spitzer, S. M., Liss, J. M., Caviness, J. N., & Adler, C. (2000). An exploration of familiarization effects in the perception of ataxic and hypokinetic dysarthric speech. *Journal of Medical Speech-Language Pathology, 8,* 285–293.
Spitzer, S. M., Liss, J. M., & Mattys, S. L. (2007). Acoustic cues to lexical segmentation: A study of resynthesised speech. *The Journal of the Acoustical Society of America, 122,* 3678–3687.
Tjaden, K., & Wilding, G. E. (2004). Rate and loudness manipulations in dysarthria: Acoustic and perceptual findings. *Journal of Speech, Language, and Hearing Research, 47,* 766–783.
Tjaden, K., & Wilding, G. E. (2011a). The impact of rate reduction and increased loudness on fundamental frequency characteristics in dysarthria. *Folia Phoniatrica et Logopaedica, 63,* 178–186.
Tjaden, K., & Wilding, G. E. (2011b). Speech and pause characteristics associated with voluntary rate reduction in Parkinson's disease and multiple sclerosis. *Journal of Communication Disorders, 44,* 655–665.
Van Nuffelen, G., De Bodt, M., Vanderwegen, J., & Van de Heyning, P. (2010). Effect of rate control on speech production and intelligibility in dysarthria. *Folia Phoniatrica et Logopaedica, 62,* 110–119.
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,* 69–75.
Yorkston, K. M., Beukelman, D. R., & Hakel, M. (1996). Sentence Intelligibility Test for Windows [Computer software]. Lincoln, NE: Communication Disorders Software.
Yorkston, K. M., Hammen, V. L., Beukelman, D. R., & Traynor, C. D. (1990). The effect of rate control on the intelligibility and naturalness of dysarthric speech. *Journal of Speech and Hearing Disorders, 55,* 550–560.
Young, S., Evermann, G., Hain, T., Kershaw, D., Moore, G., Odell, J., . . . Woodland, P. (2002). *The HTK book: Revised for HTK Version 3.2.* Retrieved from http://htk.eng.cam.ac.uk/

Appendix A

Experimental Phrase Stimuli Used in the Study

account for who could knock*	kick a tad above them
address her meeting time	mark a single ladder
admit the gear beyond	mate denotes a judgement
advance but sat appeal	may the same pursued it
afraid beneath demand	measure fame with legal
amend escape approach	mistake delight for heat
appear to wait then turn	mode campaign for budget
assume to catch control	narrow seated member
attend the trend success	or spent sincere aside
avoid or beat command	pain can follow agents
award his drain away	pooling pill or cattle
balance clamp and bottle	push her equal culture
beside a sunken bat	remove and name for stake
bolder ground from justice	resting older earring*
bush is chosen after	rocking modern poster
butcher in the middle	rode the lamp for testing
confused but roared again	round and bad for carpet
connect the beer device	rowing father matters
constant willing walker	secure but lease apart
darker painted baskets	sinking rather tundra
define respect instead	sparkle enter broken
for coke a great defeat	stable wrist and load it
forget the joke below	target keeping season
frame her seed to answer	technique but sent result
functions aim his acid	thinking for the hearing
had eaten junk and train	to sort but fear inside
hold a page of fortune	transcend almost betrayed
increase a grade sedate	unless escape can learn
indeed a tax ascent	unseen machines agree
it's harmful note abounds	vital seats with wonder

Note. All phrases listed above were included in the habitual and loud conditions; however, those marked with an asterisk were excluded from the slow condition because of a recording error.

Appendix B

Worked Examples of the Coding of Accuracy and Syllable Resemblance (SR)

Target	Listener response	Accuracy (bold)	SR (underlined)
unless escape can learn	*a list* a *shape* **can learn**	50% (2/4)	100% (3/3)
measured fame with legal	**measure fame** *il*legal	50% (2/4)	0% (0/1)
mark a single ladder	*like* **a single** *mother*	50% (2/4)	33% (1/3)
increase a grade sedate	**increased a** *great today*	50% (2/4)	100% (3/3)
avoid or beat command	**avoid** *all beach* **commands**	50% (2/4)	100% (2/2)
roam the lamp for testing	**roam the** *lab* **for testing**	80% (4/5)	100% (1/1)
round and bad for carpet	**round** *in bed* **for carpet**	60% (3/5)	50% (1/2)
target keeping season	*take it* ki*wi* **season**	33% (1/3)	33% (1/3)

Note. *Accuracy* refers to the number of words correct out of the total number of words, with correct words in boldface type. For SR, italicized syllables are those produced in error. Syllables that are both italicized and underlined resemble, phonemically, the target; that is, at least half of the phonemes in the transcription match those in the target syllable (Borrie et al., 2012). For both accuracy and SR measures, the information in parentheses relates to the number correct out of the possible total. For example, "(1/4)" for accuracy means one word correct out of a total of four words. For SR, this means one syllable resembled the target out of a total of four erroneous syllables within the phrase.

Appendix C

Examples of Coding of LBEs

Target	Listener response	LBE type
unseen machines agree	unseen machine in green	IS
define respect instead	to find respect instead	IS
mistake delight for heat	a fake delight for heat	IS
hold a page of fortune	older page of fortune	DW
address her meeting time	adjacent meeting time	DW
may the same pursued it	may the same computer	DW
target keeping season	the good keeping season	IW
sparkle enter broken	sparkle eat a broken	IW
butcher in the middle	watch her in the middle	IW
for coke a great defeat	forgo a great defeat	DS
or spent sincere aside	expense sincere aside	DS
admit the gear beyond	admit adhere beyond	DS

Note. IS = insertion of a word boundary before a strong syllable; DW = deletion of a word boundary before a weak syllable; IW = insertion of a word boundary before a weak syllable; DS = deletion of a word boundary before a strong syllable.