Vocabulary influences older and younger listeners' processing of dysarthric speech

Megan J. McAuliffe, Elizabeth M. R. Gibson, and Sarah E. Kerr
Department of Communication Disorders and New Zealand Institute of Language, Brain and Behaviour, University of Canterbury, Private Bag 4800, Christchurch 8140, New Zealand

Tim Anderson
New Zealand Brain Research Institute, 66 Stewart Street, Christchurch 8011, New Zealand

Patrick J. LaShell
New Zealand Institute of Language, Brain and Behaviour, University of Canterbury, Private Bag 4800, Christchurch 8140, New Zealand

(Received 13 August 2012; revised 31 May 2013; accepted 12 June 2013)

This study examined younger ($n = 16$) and older ($n = 16$) listeners' processing of dysarthric speech—a naturally occurring form of signal degradation. It aimed to determine how age, hearing acuity, memory, and vocabulary knowledge interacted in speech recognition and lexical segmentation. Listener transcripts were coded for accuracy and pattern of lexical boundary errors. For younger listeners, transcription accuracy was predicted by receptive vocabulary. For older listeners, this same effect existed but was moderated by pure-tone hearing thresholds. While both groups employed syllabic stress cues to inform lexical segmentation, older listeners were less reliant on this perceptual strategy. The results were interpreted to suggest that individuals with larger receptive vocabularies, with their presumed greater language familiarity, were better able to leverage cue redundancies within the speech signal to form lexical hypothesis—leading to an improved ability to comprehend dysarthric speech. This advantage was minimized as hearing thresholds increased. While the differing levels of reliance on stress cues across the listener groups could not be attributed to specific individual differences, it was hypothesized that some combination of larger vocabularies and reduced hearing thresholds in the older participant group led to them prioritize lexical cues as a segmentation frame.
[http://dx.doi.org/10.1121/1.4812764]

PACS number(s): 43.71.Lz, 43.71.An, 43.71.Bp [BHS] Pages: 1358–1368

I. INTRODUCTION

Older listeners, even those with relatively preserved peripheral hearing, commonly exhibit difficulty understanding speech relative to younger adults when a signal is degraded. It has been shown that age-related deteriorations in hearing acuity, cognitive processing (e.g., Committee on Hearing and Bioacoustics and Biomechanics (CHABA), 1988; Wingfield *et al.*, 2005), and the ability to resolve phonological distortion (e.g., Schvartz *et al.*, 2008) all affect processing. It has also been observed that older listeners tend to derive considerable benefit from semantic-contextual information in their speech processing. While this is the case, there has been only limited consideration of vocabulary knowledge in the comprehension of degraded speech—and the results have been equivocal. Vocabulary knowledge appears particularly important as it has the potential to influence listeners' relative weighting of cues to speech segmentation (e.g., Mattys *et al.*, 2005), a fundamental process in speech recognition. The current study presents data that examine younger and older listeners' processing of dysarthric speech—a naturally degraded speech signal. It aims to determine if the two groups exhibit differences in the ability to understand dysarthric speech and whether age, hearing acuity, cognitive processing, and vocabulary knowledge influence listeners' speech recognition and lexical segmentation strategies.

Studies have commonly noted that older listeners demonstrate greater reliance on semantic-contextual information to inform speech comprehension in challenging listening conditions, particularly through the use of supportive sentential context in word identification (e.g., Pichora-Fuller *et al.*, 1995; Sheldon *et al.*, 2008; Sommers and Danielson, 1999) (though see Dubno *et al.*, 2000, for an opposing finding). It has been suggested that frequent listening in poor signal-to-noise conditions has enabled older listeners to develop expertise in using semantic context as a facilitative strategy (Pichora-Fuller *et al.*, 1995). However, more recently it has also been suggested that older listeners' lexical *knowledge* may play a role in this common finding—perhaps as a result of their many years of language experience or accumulated lexical knowledge (Sheldon *et al.*, 2008).

Two recent studies have investigated the role of linguistic/vocabulary knowledge in older listeners' speech processing, and they reported conflicting findings. Benichov *et al.* (2012) found that verbal ability, as measured by a composite score from word definition and word knowledge tasks, was not a significant predictor of word recognition in a group of participants aged 19 to 89 yr. Instead, hearing acuity and general cognitive ability played significant roles in word

 0001-4966/2013/134(2)/1358/11/$30.00

recognition. In contrast, Janse and Adank (2012) demonstrated that the rate at which older listeners ($M = 74$ yr) adapted to a novel accent was influenced by vocabulary knowledge. In this study vocabulary was measured via a receptive multiple-choice test, and participants with higher receptive vocabularies exhibited greater improvement in recognition accuracy over blocks of trials. It was postulated that individuals with larger lexicons may have been better able to retune category boundaries or learn something of the systematic qualities of the novel accent (Janse and Adank, 2012).

Based on the findings of these studies, it is unclear whether a listener's lexical knowledge is indeed a factor in their ability to resolve a degraded signal—over and above that contributed by age, hearing acuity, and cognitive ability. However, it seems possible that if enough lexical information could be glimpsed from the degraded signal (Cooke, 2006), then those with greater language experience or knowledge may be better able to leverage glimpses to form speech hypotheses and reconstruct the degraded signal (Cooke *et al.*, 2008). The first component of this study therefore examines the relative contributions of age, hearing, cognitive ability, and vocabulary knowledge to the task of comprehending degraded speech signals. However, one fundamental component of speech perception that has yet, to our knowledge, been investigated in older adults is lexical segmentation. Lexical segmentation is the process by which the continuous speech stream is segmented into word units (Jusczyk and Luce, 2002). In undertaking segmentation, lexical knowledge is thought to play a key role.

Mattys *et al.* (2005) outlined a hierarchical model of segmentation, which posits that listeners' reliance on various cues to segmentation is graded, and dependent on interpretive conditions. According to the model, when conditions are good (e.g., listening to speech in quiet) listeners rely on sentential context to facilitate segmentation. As interpretive conditions deteriorate, listeners rely on lexical knowledge followed by segmental cues, in turn, to facilitate segmentation. Of particular interest, when these cues are diminished or unavailable, Mattys *et al.* (2005) demonstrated that word stress (i.e., suprasegmental/prosodic cues) provided a strong cue to segmentation. In such cases, lexical segmentation is thought to be informed by the relatively predictable occurrence of strong and weak syllables in English, with listeners predisposed to inserting word boundaries before strong syllables and deleting them before weak syllables (Cutler and Butterfield, 1992; Cutler and Norris, 1988).

Given this, we suggest that if older listeners are indeed more reliant on lexical knowledge in their speech processing, then they should evidence that reliance in their segmentation decisions. Considered within the framework of Mattys *et al.* (2005), it seems reasonable to predict that, when faced with a degraded speech signal, older listeners—or simply those with higher levels of vocabulary knowledge—will be less likely to defer to, or use, sublexical cues to segmentation compared to younger listeners. First older listeners who partake in speech perception research tend to exhibit higher levels of lexical knowledge relative to younger participants (e.g., Sheldon *et al.*, 2008; Taler *et al.*, 2010). If, as previously stated, older listeners are able to make efficient use of lexical information in the speech stream to construct better word matches, then it follows that they may favor this tier of the segmentation hierarchy and preference a lexically driven parsing solution. Such a strategy would have significantly greater communicative value, increasing their chances of communicative success (Mattys *et al.*, 2005). Moreover, it is thought that impoverished auditory input associated with aging places stress on cognitive processing (e.g., Pichora-Fuller, 2008; Wingfield *et al.*, 2005), and listeners have been shown to demonstrate increased reliance on lexical-semantic cues to segmentation in the presence of cognitive load (Mattys *et al.*, 2009). While it is acknowledged that Mattys *et al.* (2009) define cognitive load as "the cost associated with actively processing a competing source of information" (p. 226), it seems possible that the generalized taxing of cognitive resources that occurs for older listeners, when listening to degraded signals, may occasion them to rely more heavily on lexical-semantic cues to inform segmentation.

To examine our predictions, we invoked adverse listening conditions through the use of speech stimuli from individuals with dysarthria associated with Parkinson's disease (PD). The stimuli were also semantically anomalous—to reduce the effects of contextual support on any findings. Parkinson's disease is a degenerative neurological condition that affects between 1% and 4% of those aged above 60 yr (de Lau and Breteler, 2006). The resultant speech disorder, hypokinetic dysarthria, affects between 50% and 89% of individuals with the disease (Hartelius and Svensson, 1994; Johnson and Pring, 1990); hence its occurrence is non-trivial in the population. Hypokinetic dysarthria is classically described as having a fast rate of speech, monopitch, monoloudness, and phoneme imprecision (Darley *et al.*, 1969). This constellation of speech features results in a signal that varies from mildly through to profoundly unintelligible.

Prior studies that have examined older listeners' processing using laboratory-induced methods of signal distortion have reported that older listeners have greater difficulty than younger listeners in the comprehension of speech that is produced at a fast rate (e.g., Vaughan and Letowski, 1997; Wingfield *et al.*, 2003), has varied or reduced segment durations (e.g., Gordon-Salant *et al.*, 2006; Souza, 2000), and unusual or diminished prosodic patterns (e.g., Baum, 2003; Wingfield *et al.*, 1992). Therefore, it seems logical to infer—generally—that given these key features are also present in hypokinetic dysarthric speech, that this form of signal distortion will pose a particular challenge to older listeners relative to younger listeners, particularly in the absence of contextual support. Given that this is a somewhat new line of research, Table I outlines possible predictions for this relationship, pairing specific features of hypokinetic dysarthria with characteristics of normal speech processing—and extrapolating to potential perceptual consequences of dysarthric speech for older listeners.

The use of dysarthric speech to tax the perceptual system is beneficial for a number of reasons. First, the natural variation that exists within the dysarthric speech signal is perhaps more representative of natural speech processing than constrained samples generated in the laboratory (Mattys and Liss, 2008)—enabling research to validate existing

TABLE I. Key characteristics of hypokinetic dysarthria and possible perceptual consequences for older listeners.

Hypokinetic dysarthria	General consequences for listener processing[a]	Potential general perceptual consequences for older listeners relative to younger listeners
Fast rate of speech	Increased processing time required. Blurring of word boundaries reduces cues to segmentation. Lexical and phonemic uncertainty may result in incorrect phoneme/word access.	Associated memory and sensory[b] decline may result in (1) Increased processing time and occasions of non-response on speech perception tasks; (2) increased listening effort; and (3) greater proportion of misperceptions as evidenced by incorrect lexical selection.
Reduced loudness	Lexical and phonemic uncertainty may result in incorrect phoneme/word access. Misperceptions of lexical boundaries reflect an increased reliance on stress cues. Increased listening effort.	Reduced audibility of the signal, coupled with sensory decline may result in heightened lexical and phonemic uncertainty. Consequences include: (1) Overall reductions in intelligibility; (2) increased listening effort; and (3) greater proportion of misperceptions as evidenced by incorrect lexical selection.
Monopitch	Reduced contrastivity between strong and weak syllables results in reduced cues to syllabic stress.	Older listeners exhibit greater reliance on prosody in speech perception (Baum, 2003; Wingfield *et al.*, 1992). The attenuation of this cue will therefore have heightened perceptual consequences, resulting in greater reductions in intelligibility.
Reduced vowel space	Reduced contrastivity between strong and weak syllables results in attenuation of cues to syllabic stress. Phonemic uncertainty leads to a larger number of possible lexical candidates and possibly incorrect lexical selection.	Predictions as discussed above (in relation to reliance on prosody). In addition, sensory decline may result in heightened phonemic uncertainty relative to younger listeners resulting in misidentification of consonants, inappropriate word selection etc.
Consonant distortion	Phonemic uncertainty may lead to larger number of possible lexical candidates and—possibly—incorrect lexical selection.	Sensory decline results in heightened phonemic uncertainty relative to younger listeners and subsequent misidentification of consonants, inappropriate word selection etc.

[a]General consequences for listener processing are modeled on those detailed in Liss (2007) and Lansford *et al.* (2011). All features of hypokinetic dysarthria are acknowledged to result in reduced intelligibility of speech both in isolated cases and in combination.
[b]Sensory, in this case, refers to hearing acuity as well as temporal and frequency resolution.

theoretical models with "real world" distortion. Second, our aging population ensures that the number of individuals living with speech disorders of neurological origin is increasing. As older listeners are more likely to be the most common communication partners of other older individuals, a comprehensive understanding of how older people comprehend such naturally occurring speech distortion is important for both speech and audiological habilitation and rehabilitation.

The current study therefore aims to determine if younger and older listeners exhibit significant differences in their ability to comprehend dysarthric speech. However, we quantify listeners' hearing acuity, memory, and vocabulary knowledge to determine how these factors interact first, in the recognition of dysarthric speech and second, in the use of syllabic stress cues in informing segmentation decisions. Given that hearing acuity, working memory, and short-term memory have all been shown to be associated with speech recognition (e.g., CHABA, 1988; Benichov *et al.*, 2012; Wingfield *et al.*, 2005), it is predicted that older listeners will have greater difficulty comprehending dysarthric speech than younger listeners, and that reduced hearing acuity and memory scores will be associated with poorer performance on the task. Yet, given recent findings (Janse and Adank, 2012), we also suggest that higher levels of vocabulary knowledge may be associated with improved performance on the task. Concerning lexical segmentation, we ask: What is the relative weight of reliance on syllabic stress cues to inform segmentation in younger and older listeners? Given that older listeners are expected to perform more poorly, overall, than younger listeners at the task, it is expected that they will exhibit a greater number of segmentation errors. However, if our hypothesis that older listeners weight more heavily toward lexical cues in their speech segmentation is correct, we expect that they will evidence reduced reliance on syllabic stress cues to inform speech segmentation, compared with younger listeners, in the challenging listening conditions imposed by dysarthric speech.

II. METHOD

A. Participants

1. Listeners

Thirty-two individuals, allocated to a younger ($n = 16$; 14 females, two males) and older ($n = 16$; 12 females, four males) group, participated in the study. All were native speakers of New Zealand English and reported no significant history of contact with persons having motor speech disorders. They also reported no history of language, learning, or cognitive disabilities. Hearing thresholds of all participants were determined using behavioral pure-tone audiometry. The younger listeners exhibited pure-tone thresholds no greater than 20 dB hearing level (HL) at intervals from 500 to 4000 Hz in both the right and left ears. The older listeners exhibited good hearing for their age—as determined by pure-tone thresholds of no greater than 25 dB HL at intervals from 500 to 4000 Hz in both the right and left ears (with the exception of one participant who exhibited a pure-tone threshold of 30 dB HL at 4000 Hz in the right ear only). While the older listeners' pure-tone thresholds were higher relative to those of the younger listeners they would, in general, have been classified as non-impaired clinically. The participants also completed a variety of cognitive and linguistic assessments. All participants scored within the normal range (i.e., $\geq$ 26) on the Montreal Cognitive Assessment, a screening tool that identifies individuals with mild cognitive impairment (Nasreddine *et al.*, 2005). Short-term (forward and backward digit span) and working memory were also assessed using the Wechsler Memory Scale (third edition) (Wechsler, 1997) and receptive vocabulary was examined with the Peabody Picture Vocabulary Test (PPVT, fourth edition) (Dunn and Dunn, 2007). Details of the listeners, and group averages of performance on the hearing, cognitive, and linguistic measures, are provided in Table II.

2. Speakers

Speech stimuli were collected from five individuals with hypokinetic dysarthria associated with PD and five age- and gender-matched controls. All individuals were native speakers of New Zealand English. The speakers with PD were aged between 65 and 79 yr ($M = 72.2$ yr, $SD = 5.4$ yr) and exhibited intelligibility scores of between 80% and 95% on the Sentence Intelligibility Test (Yorkston *et al.*, 1996). In addition, all speakers with PD met the operational definition of hypokinetic dysarthria as per previous studies—a perceptual impression of a fast rate of speech, monopitch, monoloudness, consonant imprecision, and a weak and/or breathy voice (e.g., Borrie *et al.*, 2012; Liss *et al.*, 1998).

B. Speech stimuli

The selected speakers attended a single speech recording session. Digital audio recordings of the speech stimuli were made in a sound treated room. During recordings, speakers wore an Audix HT2 Headset Condenser Microphone placed approximately 5 cm from the mouth. Stimuli were recorded directly to a laptop computer using Sony Sound Forge Version 9.0 at a sampling rate of 48 kHz with 16 bits of quantization. Speakers were asked to read the phrases in their everyday speaking voice, as if talking with a family member or friend, and stimuli were repeated if any errors in reading were noted.

Each speaker read a list of 80 phrases. Each phrase was six syllables in length, and ranged from three to five words. Each word contained within the phrases was either mono- or bi-syllabic. The stimuli alternated phrasal stress patterns, such that half of the stimuli were iambic and exhibited a weak-strong phrasal stress pattern and the remaining half were trochaic and exhibited a strong-weak stress pattern. All phrases were syntactically correct but semantically anomalous. Following recording, a phrase selection procedure was conducted as per previous studies (Borrie *et al.*, 2012; Liss *et al.*, 1998), and a total of 60 phrases selected for the perception experiment.

First two speech-language pathologists, with significant experience in motor speech disorders (MM and EG), rated individual phrases and only those phrases that most strongly conformed to the perceptual criteria were selected. A short pilot investigation was conducted in which five young healthy participants transcribed selected phrases that conformed to our perceptual criteria. The final 60 phrases selected from this cohort exhibited intelligibility levels of approximately 50% on pilot testing, with a total of 12 phrases chosen from each of the five speakers with hypokinetic dysarthria (see Table III). Of the 12 phrases from each speaker, six were trochaic and six iambic in phrasal stress pattern. Once perceptual screening was conducted and the experimental phrases selected, each phrase was subject to acoustic analysis to verify the judges' perceptual impressions and to ensure that the phrase productions of the individuals with hypokinetic dysarthria differed significantly from healthy control speakers on a range of parameters.

To conduct the acoustic analysis, the recorded phrases were transcribed then automatically segmented at phoneme

TABLE II. Summary description of the two listener groups (with standard deviations in parenthesis).[a]

	Younger Listeners	Older Listeners	t-statistic ($df = 30$)
No. participants	16	16	—
Age	20.1 years	64.8 years	—
	(3.0 years)	*(3.4 years)*	
Years of education	14.44 years	14.19 years	$t = 0.27$, *n.s.*
	(2.34 years)	*(2.97 years)*	
Pure-tone average	3.32 dB HL	12.78 dB HL	$t = -7.02$, $p < .001$
	(3.76 dB HL)	*(3.86 dB HL)*	
Forward digit span (raw)	10.38	11.00	$t = -0.78$, *n.s.*
	(2.36)	*(2.31)*	
Backward digit span (raw)	7.06	7.63	$t = -0.77$, *n.s.*
	(2.11)	*(2.00)*	
Letter-number sequencing (raw)	11.38	11.00	$t = 0.38$, *n.s.*
	(3.20)	*(2.31)*	
Receptive vocabulary (raw)	208.13	221.25	$t = -4.62$, $p < .001$
	(10.63)	*(4.01)*	

[a]*n.s.* = not significant at $p \leq .05$. Digit span and letter-number sequencing measures derived from Weschler Memory Scale third edition (Wechsler, 1997), receptive vocabulary measured using the Peabody Picture Vocabulary Test fourth edition (Dunn and Dunn, 2007), raw = average raw score of the group on the selected measure.

TABLE III. Phrase stimuli employed 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 judgment
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

level using the Hidden Markov Model Toolkit (HTK) (Young *et al.*, 2002). The segments were phonemically labeled based on the ONZE Miner orthographic-phonemic dictionary (Fromont and Hay, 2008), constructed from CELEX (Baayen *et al.*, 1996), and additional hand labeled entries. The third author, based on common criteria, manually checked all automatic segment boundaries. The primary indicators for vowels were steady formant structures, regular voicing, and regular waveform amplitude. Consonants were primarily determined by the following criteria: Measured from the closure or low energy waveforms preceding the release of a plosive until the burst (voiced plosives) or until voicing of the next vowel (voiceless plosives), aperiodic waveform with high frequency energy in the spectrogram (fricatives), low energy waveform accompanied by lightening of the spectrogram (nasals), and sudden changes in second and third formants (liquids and glides). As the HTK segmentation was completed at the individual phoneme level, if there was uncertainty in discriminating boundaries for consecutive consonants (e.g., /t/ and /s/) the boundary derived from automatic segmentation was preferred. Following manual checking, custom Praat (Boersma and Weenink, 2009) scripts extracted the following acoustic data for each phrase: (1) Total phrase duration (in seconds); (2) total vowel duration across the phrase (in seconds); (3) total consonant duration across the phrase (in seconds); (4) speech rate (syllables per second); (5) articulation rate (syllables per second); (6) normalized vocalic Pairwise Variability Index for vowel duration, nPVI(V) (calculated as per Grabe and Low, 2002); (7) pitch variation (fundamental frequency standard deviation across the phrase); and (8) intensity variation (intensity standard deviation across the phrase). A measure of vowel space area was also calculated 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 employed to produce a measure of vowel space.

The results of the acoustic analysis are detailed in Table IV. As can be seen, the phrases selected from the speakers with hypokinetic dysarthria were spoken at a significantly faster rate overall compared to control speech, with the faster rate of production evidenced in both consonant and vowel productions. Phrases selected from speakers with dysarthria also exhibited significantly reduced values on nPVI(V), highlighting reduced variation in the duration of subsequent vowels for hypokinetic compared with control phrases. The presence of monoloudness was also clear. Surprisingly, there was no significant difference between the two groups for F0 in phrases; however, review of our control speakers indicated a certain degree of monopitch to their speech. The value of 18.89 Hz as a F0 standard deviation for the phrases from speakers with hypokinetic dysarthria in the current study compares favorably with an earlier investigation from our laboratory in which a pitch variation of 17.67 Hz was achieved from speakers with hypokinetic dysarthria from the same phrase list relative to 25.96 for controls (Borrie *et al.*, 2012). Hence, it is presumed that the phrases employed in the current experimentation exhibited reduced F0 variation. Finally, vowel space area generated from the phrases of the participants with hypokinetic dysarthria was reduced by approximately 24% relative to that from the control speakers (dysarthric vowel space area = 142 714 Hz^2; control vowel space area = 186 873 Hz^2). To ensure the reliability of the acoustic findings, 20% of the original automatically segmented phrases were randomly selected and manually rechecked by the original judge (intra-rater reliability) and a second judge (inter-rater reliability). Cronbach's alpha was used as a measure of agreement, with a value above 0.95 for intra-rater reliability across all measures, and above 0.93 for inter-rater reliability with the exception of the nPVI, which was 0.86. The reliability of the acoustic measures was deemed acceptable.

TABLE IV. Comparison of the acoustic characteristics of phrases produced by speakers with hypokinetic dysarthria (n = 60) and matched controls (n = 60) on a range of parameters.[a]

Parameter	Hypokinetic	Control	*t*-statistic (*df* = 118)
Phrase duration (s)	1.33 *(0.21)*	1.86 *(0.30)*	11.21, $p < 0.001$
Speech rate (syll/s)	4.64 *(0.72)*	3.31 *(0.60)*	−10.98, $p < 0.001$
Articulation rate (syll/s)	4.70 *(0.71)*	3.40 *(0.58)*	−10.92, $p < 0.001$
Total vowel duration (s)	0.62 *(0.12)*	0.76 *(0.17)*	5.38, $p < 0.001$
Total consonant duration (s)	0.69 *(0.17)*	1.04 *(0.20)*	10.30, $p < 0.001$
nPVI(V)	58.15 *(21.11)*	66.13 *(22.76)*	1.99, $p < 0.05$
F0 sd (Hz)	18.89 *(8.76)*	17.41 *(4.21)*	−1.18, $p = 0.24$
SPL sd (dB)	6.97 *(1.95)*	10.16 *(2.32)*	8.15, $p < 0.001$

[a]Standard deviation in parentheses, nPVI(V) = normalised vocalic pairwise variability index for vowel duration, F0 = fundamental frequency, sd = standard deviation, SPL = sound pressure level and dB = decibels. Speech rate includes pause time whereas articulation rate does not.

C. Procedure

All listeners completed the study in the following order: (1) Hearing test, (2) perception experiment, and (3) ancillary assessments (i.e., memory and vocabulary testing). The order of the ancillary testing was randomized across participants. All experimental testing was completed in a sound treated booth and approximately 25 min was required for the completion of the listening experiment. During the experiment, participants were seated in front of a laptop computer, with the experiment programmed in DirectRT (Jarvis, 2010). The experimental phrases were presented through Sennheiser HD280 Pro circumaural headphones at a set volume of approximately 65 dB sound pressure level (SPL). Prior to the experiment commencing, all participants were asked whether the volume of the practice phrases was appropriate to complete the task—none of the participants requested a change to the presentation levels. All participants received identical instructions, presented on the laptop screen via the experimental software. Participants were told that they would hear a series of phrases, and that each of the words within the phrases were real English words; however, the phrases themselves may not make sense. Participants were asked to repeat each of the phrases exactly as they heard them and to take a guess even if they did not understand what was said. Participants were also asked to say the word "something" instead of the missing word if they were unable to understand a word within the phrase. The spoken repetition paradigm was implemented to reduce any disparities in cognitive load across listeners, or groups, that may have arisen due to difficulties typing or using the keyboard—as was common to a number of older participants. The participant's responses were transcribed by a research assistant, in real time, during experiment completion. In addition, the session was audio recorded, and transcriptions were cross-checked prior to data analysis. This paradigm was used successfully in a previous study from our laboratory that included older listeners (McAuliffe *et al.*, 2012).

Prior to commencing the experiment, participants completed a short practice phase in which four phrases, similar in construction to the test stimuli, were presented. The participant was asked to repeat each of the four phrases and only if all four phrases were repeated successfully was the participant able to complete the experiment. No participant was excluded from the study at this point. The phrase stimuli were subsequently presented in blocks of six in a consistent order (order of blocks was kept consistent due to the requirements of an on-going investigation). The 10 stimuli within each block were presented in random order. The six stimuli blocks were matched for phrase per speaker, stress pattern, and number of possible lexical boundary errors. After the presentation of the first three blocks, participants were given a 2 min rest break before continuing.

D. Data analysis

The final data set consisted of 32 sets of phrase transcriptions, 16 from younger listeners, and 16 from older listeners. The second author, blind to the group allocation of the listeners, completed transcript coding for words correct and for the occurrence of lexical boundary errors (LBEs). Words correct were calculated according to accepted procedures (Borrie *et al.*, 2012; Liss *et al.*, 1998). Words were scored as correct if they matched the target exactly, or differed only by the tense "ed," the plural "s," or possessed a substitution of "a" and "the." Homonyms and obvious misspellings were also counted as correct (Borrie *et al.*, 2012; Liss *et al.*, 1998).

The type and number of lexical boundary errors were also coded. An LBE can be classified as an incorrect deletion or insertion of a word or lexical boundary and was further coded for its occurrence before either a strong or weak syllable. Four types of possible errors result, two that are considered *predictable* error types and the remaining two that are considered *unpredictable* error types. These error predictions are language-specific and arise from the predictable pattern of English, in which strong syllables are common to word onsets, the occurrence of which have been shown to trigger lexical segmentation. This is commonly referred to as the metrical segmentation strategy (MSS) (Cutler and Norris, 1988). In English, predictable error types include the following: (a) Insertion of a lexical boundary before a strong syllable (IS) and (b) deletion of a lexical boundary before a weak syllable (DW). Unpredictable error types include the following: (a) Insertion of a lexical boundary before a weak syllable (IW) and (b) deletion of a lexical boundary before a strong syllable (DS). To measure whether listeners do indeed employ this form of stress-based segmentation, we also calculated an MSS ratio for each listener. This was defined as the number of IS errors, plus the number of DW errors, divided by the total number of LBEs [i.e., (IS + DW)/total number of LBEs]. If listeners exhibited an MSS ratio of greater than 0.50, listeners were said to employ a stress-based approach to lexical segmentation (Spitzer *et al.*, 2007).

Across our phrase list, there were 300 potential LBE opportunities. If in a trial a participant failed to respond with a complete phrase or responded with "something," that portion of the response was not used in classifying LBEs. The remainder of the response, however, was still used in assigning LBEs. Accuracy scoring was performed on all trials regardless of the LBE classification. Twenty-five percent of the listeners' transcripts were randomly selected for examination of inter- and intra-coder reliability. Agreement across all measures was high with Spearman's ρ scores of above 0.930 for inter-rater and above 0.987 for intra-rater reliability across all measures.

To analyze accuracy of phrase transcription mixed effects modeling was employed. Mixed-effects models are advantageous for repeated measures data, as 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, participant age). This is in contrast to traditional analysis of variance (ANOVA) analyses which require different analyses for multiple sources of individual variance (Baayen *et al.*, 2008). For analysis of LBEs, a combination of *t*-tests, χ^2 goodness of fit and correlation analysis were employed.

III. RESULTS

A. Accuracy

Overall accuracy in comprehending dysarthric speech did not vary between the two groups; older listeners ($M = 46.24\%$, $SD = 6.03\%$) performed similarly to younger listeners ($M = 45.28\%$, $SD = 5.39\%$), $t(30) = 0.47$, $p = 0.64$. A series of binomial mixed effects models were used to analyze listeners' accuracy of responses. The analysis examined the fixed effects of group (young or older), pure-tone average (PTA) threshold, receptive vocabulary (PPVT), short-term and working memory (forward digit span, backward digit span, letter-number sequencing, and a composite of all three measures), and number of words per phrase. Random effects for subject, speaker, and phrase were also included. Furthermore, review of the data indicated that receptive vocabulary and PTA were highly correlated with group. Older adults possessed worse hearing and better vocabularies than younger adults (see Fig. 1). Specifically, age group was correlated with PTA ($r = 0.79$, $p < 0.001$) and vocabulary knowledge ($r = 0.64$, $p < 0.001$), and PTA was correlated with receptive vocabulary ($r = 0.58$, $p < 0.001$). Hence, both PTA and receptive vocabulary were residualized against age group and these residualized scores were used in the analysis.

The resulting binomial mixed effects model fit on the overall dataset revealed a significant main effect of receptive vocabulary [$\beta = 0.022$ (0.006), $p < 0.001$], indicating that as receptive vocabulary knowledge increased, so too did the probability of an accurate response. The main effects of group [$\beta = 0.083$ (0.079), not significant (*n.s.*)] and PTA [$\beta = -0.025$ (0.018), *n.s.*] were non-significant. No significant two-way interactions were present, although the interaction of PTA and receptive vocabulary approached significance [$\beta = 0.004$ (0.002), $p = 0.070$]. Additionally, there was a significant three-way interaction of group, PTA, and receptive vocabulary [$\beta = -0.016$ (0.005), $p < 0.001$]. This suggested that younger listener's performance was influenced by receptive vocabulary only, whereas both PTA and receptive vocabulary influenced older listener's performance. To confirm this interpretation and examine the near significant two-way interaction of PTA with receptive vocabulary, this model was followed up by separate analyses of the younger and older groups.

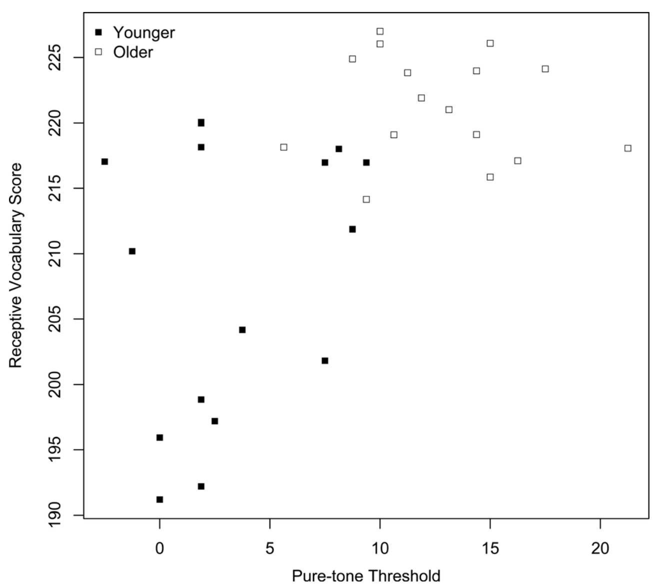

FIG. 1. Relationship between receptive vocabulary score and pure-tone average for individuals in the younger and older listener groups, respectively.

A second set of model fitting—again using binomial mixed effects modeling—was conducted to analyze the accuracy of response separately for each age group. In the final model for younger participants, only receptive vocabulary exhibited a significant effect on accuracy of response [$\beta = 0.016$ (0.006), $p < 0.01$]. Therefore for the younger listeners, their receptive vocabulary score was predictive of ability to decipher dysarthric speech—as receptive vocabulary score increased, so too did response accuracy. The final model for older participants revealed no main effect of PTA [$\beta = -0.023$ (0.017), *n.s.*] or receptive vocabulary, though the effect of receptive vocabulary approached significance [$\beta = 0.030$ (0.015), $p = 0.060$]. However, there was a significant interaction between receptive vocabulary and hearing [$\beta = -0.012$ (0.005), $p < 0.01$], such that while higher vocabulary scores facilitated recognition with low PTA, an increased PTA resulted in impaired recognition.

To exclude the possibility that the lexical frequency of items in the phrase stimulus list was responsible for the significant findings related to receptive vocabulary, we conducted further analysis. The lexical frequency of each content word contained within the phrase stimuli list was extracted from the CELEX database (Baayen *et al.*, 1996) and converted to a log value. The average log of each phrase and the minimum log value across a phrase were calculated. These values were then considered, in turn, in replications of the original statistical models. In all instances, the variables were included as separate factors and in interactions. Results indicated that neither measure of lexical frequency was a significant predictor of word accuracy, nor did they qualitatively change the final models. We therefore concluded that frequency of words within the phrase stimuli list did not explain the current findings.

B. Lexical boundary errors

Summary data describing the LBE performance of younger and older participants is provided in Table V. As can be seen, the younger and older participants exhibited a similar number of LBEs across the transcription task. No significant differences were found between the two groups for average number of LBEs [$t(30) = 0.05$, $p = 0.96$]. Further examination of the pattern of predicted and unpredicted errors revealed it was not uniform across the groups [X^2 (1) $= 6.65$, $p < 0.01$]. For both groups, the error distribution followed the expected pattern, with significantly more errors of the predicted type versus the unpredicted type [younger group: X^2 (1) $= 82.01$, $p < 0.001$; older group: X^2 (1) $= 31.15$, $p < 0.001$]. While both groups adhered to the predicted error patterns, they did so with differing strengths of adherence—as evidenced by the differences in their MSS ratios. Both groups exhibited MSS ratios of greater than

TABLE V. Pattern of lexical boundary errors by group. Where group averages are displayed, the standard deviation is included in parenthesis.

Lexical Boundary Errors	Younger Listeners	Older Listeners
Total number of errors	336	334
Average number of errors	21.00 (*6.45*)	20.88 (*7.11*)
Total number of predicted errors	251	218
Total number of unpredicted errors	85	116
MSS ratio	0.76 (0.*10*)	0.66 (0.*09*)
Breakdown of errors (total number)		
Insert before strong (IS)[a]	152	144
Delete before weak (DW)[a]	99	74
Delete before strong (DS)	23	29
Insert before weak (IW)	62	87

[a]Predicted error types in the English language. MSS = metrical segmentation strategy.

0.50, indicating relative adherence to stress-based segmentation in their attempts to decipher dysarthric speech (Spitzer *et al.*, 2007). However, the MSS ratio of the younger participant group was significantly higher than that of the older participant group [$t(30) = 3.22$, $p < 0.01$, $d = 1.14$]. Therefore, while both groups relied on syllabic strength to inform lexical boundary decisions, the younger group relied on this strategy to a significantly greater extent than the older listener group.

Further examination of LBEs revealed more detail about the pattern of errors delineating the two groups. Younger participants tended toward deleting word boundaries before weak syllables more often than older participants [X^2 (1) = 3.61, $p = 0.06$]. In contrast, older participants inserted word boundaries before weak syllables significantly more often than younger participants [X^2 (1) = 4.19, $p < 0.05$]. No differences were found between younger and older participants in their frequency of inserting word boundaries before strong syllables [X^2 (1) = 0.22, $p = 0.64$], nor in their frequency of deleting word boundaries before strong syllables [X^2 (1) = 0.69, $p = 0.41$].

A regression analysis was conducted to examine whether listeners' error patterns, specifically their MSS ratios, were predicted by their performance on the various cognitive and linguistic measures. Because of the previously described correlations between PTA, vocabulary, and age group, PTA and vocabulary were regressed upon age group to create residual measures. These residualized measures were used in the analysis. Results of this analysis indicated group membership was a significant predictor of MSS ratio, t (1) = 3.49, $p = 0.002$. There were no other significant relationships or interactions with working memory, residualized vocabulary, or residualized hearing. The finding of a significant relationship between age group and MSS ratio must be interpreted within the constraints of this dataset—both PTA and receptive vocabulary were highly correlated with age group. Therefore, it is unable to be determined by the current investigation whether a reduced reliance on syllabic stress is indeed related to PTA or receptive vocabulary across listeners or whether this is a by-product of the previously demonstrated effect of age group.

IV. DISCUSSION

This study investigated younger and older listeners' processing of dysarthric speech—a naturally degraded speech signal. It aimed to determine the relative contribution of age, hearing acuity, cognitive ability and vocabulary knowledge to both listeners' accuracy comprehending dysarthric speech and the weight given to syllabic stress cues in segmentation of the speech stream. We found that younger and older listeners exhibited similar levels of accuracy comprehending hypokinetic dysarthric speech. For younger listeners receptive vocabulary was the only factor tested that was predictive of the ability to understand dysarthric speech. However, for older listeners an interaction between hearing and receptive vocabulary indicated that this same vocabulary effect was present, but was moderated by hearing thresholds. Subsequent analysis confirmed that the vocabulary effects were not related to lexical frequency. The study also found that while younger and older listeners both attended to syllabic stress cues when segmenting the degraded speech stream, they did so with differing strengths of adherence to this strategy—in keeping with our predictions the older listener group exhibited significantly less reliance on syllabic stress to parse the speech stream compared to the younger group. The study findings are discussed in relation to the existing literature on speech understanding in older listeners, and models of lexical access and segmentation.

Regarding accuracy in comprehending dysarthric speech, it appears that the following conclusion can be proposed: The greater the size of an individual's receptive vocabulary the more accurate their ability to comprehend the dysarthric speakers included in the current study, except in instances of elevated hearing thresholds. It seems possible that a critical threshold existed in which a larger vocabulary could no longer compensate for the difficulties in resolving a degraded signal that arose from elevated hearing thresholds. Ours is not the first study to highlight a link between listeners' receptive vocabulary and their ability to comprehend a degraded signal. Prior investigations have noted similar effects in older adults (Janse and Adank, 2012), children (Munson, 2001) and in non-native speakers of English (Alamsaputra *et al.*, 2006). Combined, these findings provide growing evidence that receptive vocabulary plays a role in the ability, or otherwise, to decipher a degraded signal.

To an extent, our findings conflict with those of Benichov *et al.* (2012) who reported that verbal ability was not a significant contributor to speech recognition in their group of participants aged 19 to 89 yr. However, the two studies differed considerably in their methodologies, which may have underpinned these differences. While we employed a forced-choice receptive vocabulary task, participants in Benichov *et al.* (2012) completed two expressive linguistic tasks involving defining and reading various words. Furthermore, our experimental task required participants to repeat back an entire phrase, whereas participants in Benichov *et al.* (2012) were required to state the final word following a carrier phrase. Interestingly, the current study and two others that reported a link between vocabulary knowledge and speech perception (Alamsaputra *et al.*, 2006;

Munson, 2001) measured receptive vocabulary using the PPVT (Dunn and Dunn, 2007), and Janse and Adank (2012) also employed a force-choice receptive vocabulary task (Janse and Adank, 2012). It seems possible that the different mechanisms targeted by these tests may, at least partially, explain the pattern of findings. While the current study found that working memory did not play a role in listeners' accuracy at comprehending dysarthric speech, the short length of phrases employed and minimal variance in memory scores across listeners may have played a role in this finding.

The reason why a larger vocabulary may facilitate the recognition of degraded speech is not immediately obvious. However, dysarthric speech is produced from a baseline of disordered speech motor control and is therefore inherently variable in its production. Even within utterances a speaker with dysarthria may exhibit variation in intelligibility. This is particularly apparent in hypokinetic dysarthric speech, which is characterized by "short rushes" of reduced intelligibility (Duffy, 2005). In the current study, speech recognition averages of approximately 45% suggest that either certain words were intelligible within the speech stream, or that enough intelligible speech fragments were available within the signal for listeners to piece together and select, on occasion, the appropriate word. Given the utterances were semantically anomalous, contextual support was limited. Therefore, we suggest that listeners with a higher vocabulary score—and hence, we assume, greater levels of general experience/familiarity with speech—were better able to make use of redundancies within the acoustic signal and, ultimately, leverage this prior experience to draw accurate lexical hypotheses (Cooke *et al.*, 2008). Such an interpretation appears consistent with a glimpsing model of speech perception in noise (Cooke, 2006). Additionally, Heaps' law (Ángeles Serrano *et al.*, 2009; Heaps, 1978) suggests that a listener with a larger receptive vocabulary would exhibit more examples, within their lexicon, of the relatively familiar words employed in this study. The increased number of examples would afford a broader experience with language, and cues, which would also advantage a high vocabulary listener in the task of piecing together glimpses of intelligible fragments. The interaction between hearing and receptive vocabulary within the older age group implies that a larger receptive vocabulary advantaged listeners only to the point at which elevated hearing thresholds reduced its effectiveness. For those participants, we suggest that the diminishment or distortion of cues within the signal led to the activation of too many lexical candidates, or simply that they were unable to make use of cue redundancies to the same effect.

The LBE error analysis appears to provide some support for our interpretation of the vocabulary findings. Consistent with prior studies of predominately younger listeners processing of dysarthric speech (Borrie *et al.*, 2012; Choe *et al.*, 2012; Liss *et al.*, 1998), both the younger and older listeners in our cohort used syllabic stress as a cue to segmentation—as evidenced by their MSS ratios. However, novel to the current study was the finding that older listeners demonstrated significantly less reliance on stress-based segmentation compared to younger listeners. Given that both PTA and receptive vocabulary were highly correlated with age group, we were unable to ascertain whether individual differences on these measures was related to listeners' choice of cue to lexical segmentation. However, we can conclude that age was indeed a mediating factor in listeners' lexical segmentation strategies.

From this study and others (Choe *et al.*, 2012; Weiss *et al.*, 2010), there is growing evidence that it is not simply the strength of a particular cue that affects segmentation, but that listener-specific correlates may also influence the weight given to various cues in segmentation decisions. While we were unable to determine which listener-specific correlates may have influenced segmentation, it appears reasonable to suggest that this group of older listeners—with their larger receptive vocabularies and increased hearing thresholds relative to the younger group—may have prioritized lexical cues as a segmentation frame. The lack of contextual support within the speech stream, and availability of lexical "glimpses," meant that attempts at segmentation within the lexical tier were most efficient, and resulted in the greatest communicative gain. This finding appears consistent with the hierarchical segmentation model of Mattys *et al.* (2005)—and provides preliminary evidence that behavioral characteristics of listeners may also play some part in segmentation decisions. Given the similarity in cognitive ability between the two groups, we assume that this did not play a role in the current findings. However, follow-up studies including participants with wider ranges of hearing, vocabulary, and cognitive ability are required to determine whether these individual differences do indeed inform listeners' segmentation strategies.

V. CONCLUSION

These findings should be interpreted relative to the study design and its limitations. First, the younger and older groups differed in terms of hearing and vocabulary. While the older listeners demonstrated hearing that would generally be considered within the normal range, they did exhibit elevated hearing thresholds relative to the younger group. Furthermore, the older listeners' receptive vocabularies were superior. To more confidently assess the relative contribution of vocabulary knowledge and peripheral hearing loss to perceptual processing, follow-up studies are required that include younger and older listeners with varied hearing thresholds and a more distributed range of vocabulary knowledge. The older listeners in this study were also high performing—their working memory scores were equivalent to the younger listener groups, they had good hearing for their age, and they evidenced similar accuracy at comprehending dysarthric speech. This is not inconsistent with other studies of older listeners (e.g., Kidd and Humes, 2012; Taler *et al.*, 2010); however, including participants with wider ranges of general cognitive and memory abilities in addition to higher hearing thresholds will likely provide further insight into the role of hearing acuity, cognition and linguistic knowledge in speech perception. Given that vocabulary knowledge appears to be an influencing factor in listeners' ability to comprehend degraded speech signals, the linguistic content of experimental stimuli also requires further consideration. Future studies would benefit from stimuli that are explicitly developed to examine the contribution of linguistic knowledge to accuracy comprehending distorted speech (e.g., a variety of high and low frequency words). Finally, the study findings require replication.

In summary, we found that in younger listeners, vocabulary knowledge predicted the ability to accurately comprehend dysarthric speech. For older listeners, vocabulary also influenced performance, but this was moderated by hearing thresholds. Both groups used syllabic stress to inform their placement of word boundaries, but older listeners were less reliant on this strategy. We interpreted these findings to suggest that listeners with larger receptive vocabularies, with their presumed greater language experience, were better placed to leverage cue redundancies within the speech signal—but that this advantage was minimized in the presence of elevated hearing thresholds. The differing levels of use of stress cues to segmentation across the younger and older listener groups could not be attributed to specific individual differences; however, it seemed that a combination of larger receptive vocabularies and lessened hearing acuity may have led older listeners to prioritize lexical cues in their segmentation of dysarthric speech.

ACKNOWLEDGMENTS

We are grateful to the New Zealand Neurological Foundation (Grant 0827-PG) and Health Research Council of New Zealand (HRC09/251) for funding this work. We wish to thank the individuals with Parkinson's disease and their families for participating in this research, Professors Julie Liss and Jen Hay for advice and suggestions related to various aspects of this work, Jacqui Nokes for comments on an earlier version of this manuscript.

Alamsaputra, D. M., Kohnert, K. J., Munson, B., and Reichle, J. (**2006**). "Synthesized speech intelligibility among native speakers and non-native speakers of English," Augmentative Alternative Commun. **22**, 258–268.

Ángeles Serrano, M., Flammini, A., and Menczer, F. (**2009**). "Modeling statistical properties of written text," PLoS ONE **4**, 1–8.

Baayen, R. H., Davidson, D. J., and Bates, D. M. (**2008**). "Mixed-effects modeling with crossed random effects for subjects and items," J. Mem. Lang. **59**, 390–412.

Baayen, R. H., Piepenbrock, R., and Gulikers, L. (**1996**). *CELEX2* [cd-rom] (Linguistic Data Consortium, Philadelphia, PA).

Baum, S. R. (**2003**). "Age differences in the influence of metrical structure on phonetic identification," Speech Commun. **39**, 231–242.

Benichov, J., Cox, L. C., Tun, P. A., and Wingfield, A. (**2012**). "Word recognition within a linguistic context: Effects of age, hearing acuity, verbal ability, and cognitive function," Ear Hear. **33**, 250–256.

Boersma, P., and Weenink, D. (**2009**). "Praat: doing phonetics by computer," Ver. 5.1.05, http://www.praat.org/ (Last viewed 17 August 2012).

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

Choe, Y.-K., Liss, J. M., Azuma, T., and Mathy, P. (**2012**). "Evidence of cue use and performance differences in deciphering dysarthric speech," J. Acoust. Soc. Am. **131**, EL112–EL118.

Committee on Hearing and Bioacoustics and Biomechanics (CHABA) (**1988**). "Speech understanding and aging," J. Acoust. Soc. Am. **83**, 859–895.

Cooke, M. (**2006**). "A glimpsing model of speech perception in noise," J. Acoust. Soc. Am. **119**, 1562–1573.

Cooke, M., Garcia Lecumberri, M. L., and Barker, J. (**2008**). "The foreign language cocktail party problem: Energetic and informational masking effects in non-native speech perception," J. Acoust. Soc. Am. **123**, 414–427.

Cutler, A., and Butterfield, S. (**1992**). "Rhythmic cues to speech segmentation: Evidence from juncture misperception," J. Mem. Language **31**, 218–236.

Cutler, A., and Norris, D. (**1988**). "The role of strong syllables in segmentation for lexical access," J. Exp. Psychol. **14**, 113–121.

Darley, F. L., Aronson, A. E., and Brown, J. R. (**1969**). "Differential diagnostic patterns of dysarthria," J. Speech Hear. Res. **12**, 246–269.

de Lau, L. M. L., and Breteler, M. M. B. (**2006**). "Epidemiology of Parkinson's disease," Lancet Neurol. **5**, 525–535.

Dubno, J. R., Ahlstrom, J. B., and Horwitz, A. R. (**2000**). "Use of context by young and aged adults with normal hearing," J. Acoust. Soc. Am. **107**, 538–546.

Duffy, J. R. (**2005**). *Motor Speech Disorders: Substrates, Differential Diagnosis, and Management*, 2nd ed. (Elsevier Mosby, St. Louis, MS), 578 pp.

Dunn, L. M., and Dunn, D. M. (**2007**). *Peabody Picture Vocabulary Test*, 4th ed. (Pearson Education Incorporated, San Antonio, TX), 209 pp.

Fromont, R., and Hay, J. B. (**2008**). "ONZE Miner: The development of a browser-based research tool," Corpora **3**, 173–193.

Gordon-Salant, S., Yeni-Komshian, G. H., Fitzgibbons, P. J., and Barrett, J. (**2006**). "Age-related differences in identification and discrimination of temporal cues in speech segments," J. Acoust. Soc. Am. **119**, 2455–2466.

Grabe, E., and Low, E. L. (**2002**). *Durational variability in speech and the rhythm class hypothesis, in Laboratory phonology*, edited by C. Gussenhoven and N. Warner (Mouton de Gruyter, Ossining, NY), Vol. 7, pp. 515–546.

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

Heaps, H. S. (**1978**). *Information Retrieval: Computational and Theoretical Aspects* (Academic Press, Orlando, FL) 344 pp.

Janse, E., and Adank, P. (**2012**). "Predicting foreign-accent adaptation in older adults," Quart. J. Exp. Psychol. **65**, 1563–1585.

Jarvis, B. G. (**2010**). "DirectRT," Ver. 2010.3.109 [computer software] (Empirisoft Corporation, New York, NY).

Johnson, J. A., and Pring, T. R. (**1990**). "Speech therapy and Parkinson's disease: A review and further data," Br. J. Disord. Commun. **25**, 183–194.

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

Kidd, G. R., and Humes, L. E. (**2012**). "Effects of age and hearing loss on the recognition of interrupted words in isolation and in sentences," J. Acoust. Soc. Am. **131**, 1434–1448.

Lansford, K. L., Liss, J. M., Caviness, J. N., and Utianski, R. L. (**2011**). "A cognitive-perceptual approach to conceptualizing speech intelligibility deficits and remediation practice in hypokinetic dysarthria," Parkinson's Disease **2011**(150962).

Liss, J. M. (**2007**). "The role of speech perception in motor speech disorders," in *Motor Speech Disorders: Essays for Ray Kent*, edited by G. Weismer (Plural, San Diego, CA), pp. 195–231.

Liss, J. M., Spitzer, S., Caviness, J. N., Adler, C., and Edwards, B. (**1998**). "Syllabic strength and lexical boundary decisions in the perception of hypokinetic dysarthric speech," J. Acoust. Soc. Am. **104**, 2457–2466.

Mattys, S. L., Brooks, J., and Cooke, M. (**2009**). "Recognizing speech under a processing load: Dissociating energetic from informational factors," Cognit. Psychol. **59**, 203–243.

Mattys, S. L., and Liss, J. M. (**2008**). "On building models of spoken-word recognition: When there is as much to learn from natural "oddities" as artificial normality," Percept. Psychophys. **70**, 1235–1242.

Mattys, S. L., White, L., and Melhorn, J. F. (**2005**). "Integration of multiple speech segmentation cues: A hierarchical framework," J. Exp. Psychol. **134**, 477–500.

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

Munson, B. (**2001**). "Relationships between vocabulary size and spoken word recognition in children aged 3 to 7," Contemp. Issues Commun. Sci. Disord. **28**, 20–29.

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., and Chertkow, H. (**2005**). "The Montreal Cognitive Assessment (MoCA©): A brief screening tool for mild cognitive impairment," J. Am. Geriatric Soc. **53**, 695–699.

Pichora-Fuller, M. K. (**2008**). "Use of supportive context by younger and older adult listeners: Balancing bottom-up and top-down information processing," Int. J. Audiol. **47**, S72–S82.

Pichora-Fuller, M. K., Schneider, B. A., and Daneman, M. (**1995**). "How young and old adults listen to and remember speech in noise," J. Acoust. Soc. Am. **97**, 593–608.

Schvartz, K. C., Chatterjee, M., and Gordon-Salant, S. (**2008**). "Recognition of spectrally degraded phonemes by younger, middle-aged, and older normal-hearing listeners," J. Acoust. Soc. Am. **124**, 3972–3988.

Sheldon, S., Pichora-Fuller, M. K., and Schneider, B. A. (**2008**). "Priming and sentence context support listening to noise-vocoded speech by younger and older adults," J. Acoust. Soc. Am. **123**, 489–499.

Sommers, M. S., and Danielson, S. M. (**1999**). "Inhibitory processes and spoken word recognition in young and older adults: The interaction of lexical competition and semantic context," Psychol. Aging **14**, 458–472.

Souza, P. E. (**2000**). "Older listeners' use of temporal cues altered by compression amplification," J. Speech Lang. Hear. Res. **43**, 661–674.

Spitzer, S. M., Liss, J. M., and Mattys, S. L. (**2007**). "Acoustic cues to lexical segmentation: A study of resynthesised speech," J. Acoust. Soc. Am. **122**, 3678–3687.

Taler, V., Aaron, G. P., Steinmetz, L. G., and Pisoni, D. B. (**2010**). "Lexical neighborhood density effects on spoken word recognition and production in healthy aging," J. Gerontol. Ser. B **65**, 551–560.

Vaughan, N. E., and Letowski, T. (**1997**). "Effects of age, speech rate, and type of test on temporal auditory processing," J. Speech Lang. Hear. Res. **40**, 1192–1200.

Wechsler, D. (**1997**). *Wechsler Memory Scale*, 3rd ed. (Psychological Corporation, San Antonio, TX), 350 pp.

Weiss, D. J., Gerfen, C., and Mitchel, A. D. (**2010**). "Colliding cues in word segmentation: The role of cue strength and general cognitive processes," Lang. Cogn. Process. **25**, 402–422.

Wingfield, A., Peelle, J. E., and Grossman, M. (**2003**). "Speech rate and syntactic complexity as multiplicative factors in speech comprehension by young and older adults," Aging Neuropsychol. Cogn. **10**, 310–322.

Wingfield, A., Tun, P. A., and McCoy, S. L. (**2005**). "Hearing loss in older adulthood: What it is and how it interacts with cognitive performance," Curr. Directions Psychol. Sci. **14**, 144–148.

Wingfield, A., Wayland, S. C., and Stine, E. A. (**1992**). "Adult age differences in the use of prosody for syntactic parsing and recall of spoken sentences," J. Gerontol. **47**, P350–356.

Yorkston, K. M., Beukelman, D. R., and Hakel, M. (**1996**). "Sentence intelligibility test for Windows," [cd-rom] (Communication Disorders Software, Lincoln, NE).

Young, S., Evermann, G., Hain, T., Kershaw, D., Moore, G., Odell, J., Ollason, D., Povey, D., Valtchev, V., and Woodland, P. (**2002**). *The HTK Book: Revised for HTK*, Ver. 3.2 (Cambridge University Press, Cambridge, U.K.) 277 pp.