Perception of contrastive bi-syllabic lexical stress in unaccented and accented words by younger and older listeners

Sandra Gordon-Salant,[a] Grace H. Yeni-Komshian, Erin J. Pickett,[b] and Peter J. Fitzgibbons

Department of Hearing and Speech Sciences, University of Maryland, College Park, Maryland 20742, USA

(Received 14 May 2015; revised 3 November 2015; accepted 23 February 2016; published online 15 March 2016)

This study examined the ability of older and younger listeners to perceive contrastive syllable stress in unaccented and Spanish-accented cognate bi-syllabic English words. Younger listeners with normal hearing, older listeners with normal hearing, and older listeners with hearing impairment judged recordings of words that contrasted in stress that conveyed a noun or verb form (e.g., CONduct/conDUCT), using two paradigms differing in the amount of semantic support. The stimuli were spoken by four speakers: one native English speaker and three Spanish-accented speakers (one moderately and two mildly accented). The results indicate that all listeners showed the lowest accuracy scores in responding to the most heavily accented speaker and the highest accuracy in judging the productions of the native English speaker. The two older groups showed lower accuracy in judging contrastive lexical stress than the younger group, especially for verbs produced by the most accented speaker. This general pattern of performance was observed in the two experimental paradigms, although performance was generally lower in the paradigm without semantic support. The findings suggest that age-related difficulty in adjusting to deviations in contrastive bi-syllabic lexical stress produced with a Spanish accent may be an important factor limiting perception of accented English by older people. © *2016 Acoustical Society of America.*
[http://dx.doi.org/10.1121/1.4943557]

[MSS] Pages: 1132–1148

I. INTRODUCTION

Among the communication difficulties that older listeners experience is understanding English when spoken by non-native speakers (Burda *et al.*, 2003; Ferguson *et al.*, 2010). In the past, research on the effects of foreign accent on speech understanding by older listeners with or without hearing loss has focused on perception of single phonetic segments in monosyllabic English words spoken by native and Spanish-accented speakers (Gordon-Salant *et al.*, 2010a,b). This work has documented effects of age and hearing loss on perception of accented English at the segmental level. Additional work has shown that sentence understanding by older listeners, especially in noise, is compromised with the introduction of Spanish-accented English (Gordon-Salant *et al.*, 2013). It is reasonable to assume that older listeners' difficulty in processing Spanish-accented sentences is related not only to the difficulty in processing altered phonetic segments with foreign accent but also to the inability to accommodate deviations from English at the suprasegmental level. As a next step in this area of inquiry, this project examines the effect of speaker accent on perception of syllable stress in bi-syllabic words by younger and older listeners with and without hearing loss.

Stress in English is generally defined as prominence of syllables within words or words within sentences. The study of stress is a very broad field. This investigation is concerned with perception of emphasis of syllables within words as spoken with a foreign accent. Different languages vary in their pattern of stress placement within words and sentences. English is often referred to as a stress-timed language, with equal timing between stressed syllables in sentences (Pike, 1945), this results in variable stress patterns across syllables within a word. In contrast, Spanish is considered a syllable-timed language, in which equal timing is observed across syllables. This tends to produce equal stress across syllables within a word. Non-native speakers tend to introduce the phonology of their native language into the production of the second language, including segmental and suprasegmental stress patterns (Wenk, 1985; Peng and Ann, 2001). The implication of a syllable-timed language likely extends to the production of bi-syllabic words, in which it could be predicted that equal stress across the two syllables is more common in Spanish-accented English than in native English.

In English, there is a class of minimally contrastive bi-syllabic words in which the relative stress of the two syllables signals the meaning of the word as a noun or a verb. Lexical stress, in general, refers to the relative emphasis on one of many syllables in a word, in which the stressed syllable may signal the meaning of the word. The current study focuses on minimally contrastive bi-syllabic words, in which stress placed on the first vs the second syllable changes the meaning of the word, e.g., OBject and obJECT; stress placed on the first syllable conveys the noun form, and stress placed on the second syllable conveys the verb form (Fry, 1955,

[a]Electronic mail: sgsalant@umd.edu

[b]Present address: Rochester Institute of Technology, National Technical Institute of the Deaf, 52 Lomb Memorial Drive, Rochester, NY 14623, USA.

 © 2016 Acoustical Society of America

1958; Lieberman, 1960; Cutler, 1986, 2005; Zhang *et al.*, 2008).

A number of investigators have studied the acoustic correlates of minimally contrastive bi-syllabic lexical stress. Fry (1955) was the first to study the acoustic and perceptual correlates of five noun/verb pairs. His measures included vowel duration and intensity in the stressed and unstressed syllables as well as in the stressed and unstressed vowels, from which he created synthetic speech continua that varied either vowel duration ratio or vowel intensity ratio. Results of this investigation showed that both vowel intensity and vowel duration ratio convey stress in words, but that vowel duration ratio had a stronger impact than vowel intensity ratio on perception of lexical stress. In a subsequent study, Fry (1958) identified the importance of fundamental frequency (f0) as an additional cue to convey lexical stress, in which a syllable with a higher peak f0 or f0 movement was perceived as stressed. Taken together, Fry's work suggests a hierarchy of cues to lexical stress, where peak f0 and vowel duration ratio were judged to be more salient cues for stress than amplitude ratio. Since Fry's classic studies, numerous other investigators have examined the relative importance of these three cues to minimally contrastive bi-syllabic lexical stress in English (e.g., Lehiste and Peterson, 1959; Lieberman, 1960; Nakatani and Aston, 1978), and the general conclusion is that no single acoustic cue emerges as the dominant one for stress. Rather, all three cues (amplitude, duration, and f0) contribute to the perception of lexical stress. Moreover, Nakatani and Aston (1978) suggested that these multiple cues may be additive in expressing syllable stress. In addition to these three cues, investigators have observed that the vowel in the unstressed syllable is reduced relative to the vowel in the stressed syllable, thus producing a fourth cue to the perception of lexical stress (e.g., Cutler and Clifton, 1984; Cutler, 1986, 2005).

The expression of lexical stress may be different in the speech of non-native speakers. Chakraborty and Goffman (2011) reported that native speakers of Bengali who learn English as a second language transfer some, but not all, of their native language patterns to convey lexical stress (broadly defined) in English. Edmunds (2009) measured relative duration, intensity and f0 differences in stressed and unstressed vowels as produced by native English and Spanish-accented speakers. He reported that the accented speakers exhibited wider variation in all three measures compared to the native English speakers. The studies mentioned above did not focus on minimally contrastive bi-syllabic words. A study by Zhang *et al.* (2008) analyzed recordings of seven pairs of minimally contrastive English bi-syllabic words spoken by native speakers of English and of Mandarin. Acoustic measurements included syllable duration, average intensity, average f0, time of f0 peak, and the first and second formant frequencies. Although Mandarin speakers used four acoustic correlates of stress (duration, intensity, f0, and time of f0 peak) to express lexical stress, the relative use of these cues was different from that of native English speakers. Specifically, native speakers of Mandarin used a higher f0 on stressed syllables than native speakers of English, whereas peak f0 location, intensity, and duration did not show a language difference. Another study (Zuraiq and Sereno, 2007) similarly measured acoustic characteristics of eight minimally contrastive bi-syllabic words as spoken by native speakers of English and native speakers of Arabic who spoke English as a second language. The results showed that the native and non-native English speakers differed on the measure of f0, but not on the measures amplitude and duration. The non-native speakers exhibited a higher f0 ratio when the first syllable was stressed than when the second syllable was stressed, unlike the native speakers who showed no difference. Thus, the evidence suggests that non-native speakers of English exhibit different patterns of contrastive bi-syllabic lexical stress compared to native speakers of English. Henceforth, the term "lexical stress" will refer to minimally contrastive bi-syllabic words in which the stress pattern signals either a noun or a verb.

Currently, it is not known how cues used for lexical stress in English may be altered in speakers whose native language is Spanish. It may be expected that contrastive lexical stress is very different in native speakers of Spanish, because Spanish is considered a syllable-timed language, with equal stress on each syllable, whereas English is a stress-timed language, with stressed syllables occurring at regular intervals. Llisterri and his colleagues (Llisterri *et al.*, 2003; Alfano *et al.*, 2007) reported that in Spanish, stress is cued by f0 and duration, or f0 and amplitude, or by all three cues, but not by any one cue alone. Given differences in overall rhythm and timing of Spanish vs English, it could be predicted that differences in stressed and unstressed syllables would be less pronounced in native speakers of Spanish than in native speakers of English when considering the contributions of duration, f0, and intensity. Thus, it was expected that each of the three cues for lexical stress would be less pronounced in Spanish-accented English than in native English, or that the relative weight of the three cues to lexical stress would be different in Spanish-accented English compared to native English. Also, based on the work of Flege and Bohn (1989), Spanish-accented English is not characterized by vowel reduction in the unaccented syllable as it is in English, and therefore the absence of changes in vowel quality may serve as an additional cue.

Studies have examined listener sensitivity to deviant stress patterns as produced by non-native speakers. For example, Magen (1998) examined numerous phonetic and phonological factors that potentially affect perception of Spanish-accented English, and reported that native English listeners were sensitive to several aspects of accented speech, including lexical and phrasal stress. Additionally, Braun *et al.* (2011) showed that the appropriate production of unstressed vowels in English was an important factor in intelligibility of foreign-accented speech. Finally, Reinisch and Weber (2012) reported results suggesting that listeners can adapt to the lexical stress errors following a brief exposure to incorrectly stressed syllables in foreign-accented speech, implying that these listeners were sensitive to lexical stress errors.

One listener group that may have particular difficulty in perceiving cues for lexical stress by non-native speakers is older adults. First, it is now well-established that as people

age, they have difficulty perceiving auditory temporal cues (e.g., Fitzgibbons and Gordon-Salant, 1994, 1995, 2001; Grose *et al.*, 2006). Older listeners require longer changes in stimulus duration than younger listeners to discriminate isolated stimuli and even more dramatic changes for stimuli in sequences (Fitzgibbons and Gordon-Salant, 1995). Older listeners also require longer temporal cues that signal phonetic contrasts in nonsense syllables and monosyllabic words, compared to younger listeners (Strouse *et al.*, 1998; Tremblay *et al.*, 2002; Gordon-Salant *et al.*, 2008). Because duration is one cue to signal lexical stress, it is possible that older listeners' reduced sensitivity to timing information may limit perceived distinctions between stress and unstressed syllables. Second, older adults have shown more difficulty than younger adults in perceiving lexical stress that distinguishes compound words (e.g., "greenhouse") from adjective-word pairs (e.g., "green house") in English (Taler *et al.*, 2006). Third, older listeners exhibit greater difficulty compared to younger listeners in understanding Spanish-accented multisyllabic words (Gordon-Salant *et al.*, 2015) and sentences (Gordon-Salant *et al.*, 2013), which may be due, in part, to difficulty perceiving altered stress patterns associated with accent. Our studies have shown that the differences listed above in recognition of Spanish-accented English are even more pronounced in older listeners with hearing loss. As noted above, some of the difficulties older listeners experience in understanding accented speech may be related to underlying difficulties in accurately perceiving lexical stress in words, particularly because duration is one of the cues for lexical stress and discrimination of stimulus duration is often impaired among older people (Fitzgibbons and Gordon-Salant, 1995).

There is no standard paradigm for examining the perception of lexical stress using contrasting bi-syllabic words (verb vs noun) as stimuli. One paradigm presents a single bi-syllabic word and asks listeners to identify the stressed syllable from a choice of two alternative words that reflect the two possible stress patterns (Fry, 1958; Zhang *et al.*, 2008). This paradigm thus asks listeners to simply detect the location of stress in the stimulus in a single word without any linguistic or semantic support. A second paradigm presents a single bi-syllabic word and asks listeners to identify one of two sentences in which the word is correctly placed. In this case, the listener is provided with linguistic and semantic support. A final paradigm presents a phrase or sentence that includes a bi-syllabic word and asks listeners to identify whether or not the phrase or sentence is used appropriately (Cutler and Clifton, 1984). For example, the word OBject appearing in the phrase "to object" would not be appropriate, whereas the word obJECT appearing in the phrase "to object" would be appropriate. In this paradigm, the listener is given some linguistic support to aid their perception, but they also must recognize when the stimulus does not meet their expectation. It is possible that listeners' perception of lexical stress varies with the experimental paradigm; however, this has not been assessed previously. In this paper, we present results of two studies of lexical stress, both of which provided some linguistic support. In experiment 1, listeners heard a single word and selected one of two sentences that matched the stress pattern of the word (e.g., stimulus: COMpound; choice 1: "The army compound is controlled by the general"; choice 2: "Do not compound your mistakes by repeating them"). Option 1 is the correct response. In this paradigm, listeners are given linguistic and semantic support. In experiment 2, listeners heard a sentence that includes the bi-syllabic word and are asked to judge if the sentence is correct or incorrect (e.g., "The COMpound was easy to discuss" would be correct and "The comPOUND was easy to discuss" would be incorrect). This paradigm provides listeners with some linguistic context but no meaningful semantic support.

The principal objective of this study is to examine perception of lexical stress in bi-syllabic English words as produced by native and non-native speakers, and to assess the effects of age and hearing loss on perception of these words. Two hypotheses are examined in this study: (1) all listeners (younger and older) will show poorer perception of lexical stress for Spanish-accented English than for native English and (2) older listeners (with and without hearing loss) will show poorer perception of lexical stress for native and Spanish-accented English compared to younger listeners. These questions were examined in two experiments utilizing different paradigms, as described above.

II. GENERAL METHOD

A. Participants

The general criteria for participant recruitment were based on age, hearing, cognitive ability, and native language. Three groups of participants were recruited. The young normal-hearing listeners (YNH) were between 18 and 35 years of age and exhibited hearing sensitivity bilaterally within normal limits [$\leq$25 dB hearing level (HL), re: ANSI, 2010], from 250 to 4000 Hz. The older normal-hearing listeners (ONH) were between 65 and 90 years of age with the same requirements for auditory sensitivity. Finally, the older hearing-impaired listeners (OHI) were also 65–90 years, and had bilateral, symmetrical mild-to-moderate sloping sensorineural hearing losses typical of age-related hearing loss. Suprathreshold monosyllabic word recognition scores (Northwestern University Test No. 6) (Tillman and Carhart, 1966) were 80% or higher for all participants, as measured in quiet at 75 dB HL. All participants were required to exhibit normal middle ear function, as evidenced by tympanometric peak admittance, pressure peak, width, and equivalent volume within normal limits as established for adults (Roup *et al.*, 1998). Listeners in the two normal-hearing groups were also required to have acoustic reflexes present at 500, 1000, and 2000 Hz upon ipsilateral and contralateral stimulation, as well as present transient-evoked otoacoustic emissions. Listeners in the OHI group were required to have acoustic reflexes present at 500 and 1000 Hz only.

In addition to the auditory criteria described above, participants were required to pass a screening test of general cognitive awareness with a score of 26 or greater out of a maximum score of 30 (Mini Mental State Examination, MMSE) (Folstein *et al.*, 1975). Finally, only native speakers

of American English were selected to participate in the study.

There were additional criteria for participation in each experiment, based on reaching a criterion level of performance on practice blocks (described for each experiment). Listeners were evaluated on these practice blocks prior to enrolling in the study, and once selected for participation, they were assigned to the experiments in alternating order. Some listeners were permitted to enroll in the first experiment, even if they did not qualify for the second experiment.

B. Stimuli

The basic stimuli were pairs of contrastive two-syllable words with noun versus verb meaning denoted by the stress pattern. Initially, 52 two-syllable word pairs were selected. High-context sentences were created to incorporate these 104 words so that appropriate stress patterns would be achieved during recording. Eleven word pairs were eliminated due to grammatical issues and low accuracy rates from initial pilot testing. The remaining 41 pairs of words were used in the experiments and are shown in Appendix A. The high-context sentences in which they were recorded are shown in Appendix B. Analysis of word frequency for the 41 experimental word pairs using the SUBTLEX-US corpus (Brysbaert and New, 2009, 2015; Brysbaert *et al.*, 2012) showed that the average frequency of the noun forms was 292 (range = 3–3986), and the average frequency of the verb forms was 156 (1–2256). This reflects the distribution of nouns and verbs in English, in that nouns are more common than verbs (Johansson and Hofland, 1989).

Initially, five speakers recorded the high-context sentences: one native-English speaker and four native-Spanish speakers. All of the speakers were males, between the ages of 25 and 34 years, and all had completed at least two years of undergraduate coursework. The four non-native speakers of English were graduate students at the University of Maryland. Prior to recording, the speakers were instructed to speak at a conversational rate, avoid extraneous sounds, minimize vocal fry, use a natural pitch contour, maintain a fixed distance from the microphone, etc. Speakers then practiced reading the sentences, and then recorded each high-context sentence three times (total of 84 sentences). In addition, the speakers recorded four low-context sentences that were used as carrier sentences for experiment 2 (see Sec. IV A 2). Stimulus words were recorded within contextual sentences (e.g., "The food addict ate too much") in a sound-attenuated booth using a Shure SM63 dynamic microphone and a Marantz PMD661 solid state recorder. Using Cool Edit Pro Version 2 (Syntrillium Software Corporation), the target words were excised from each sentence and root-mean-square (rms) equalized. The low-context carrier sentences and the high context sentences were also edited into separate waveform files and equalized in rms level. Separate calibration tones of 1 kHz were created to be equivalent in rms level to that of the words and sentences.

A preliminary study was conducted to evaluate the degree of accentedness of the five speakers. The listeners for the preliminary study were 10 young monolingual English speakers with normal hearing who had no prior laboratory listening experience. The same four high-context sentences recorded by each of the five speakers were presented in randomized order to the listeners, for a total of 20 sentences. The listeners were asked to rate the degree of accent following each sentence presentation using a nine-point scale, with 1 representing "no foreign accent" and 9 representing a "very strong foreign accent" (Munro and Derwing, 2001; Riney *et al.*, 2000; Yeni-Komshian *et al.*, 2000). Average listener accent ratings were derived and revealed that two of the speakers had equivalent ratings. As a result, one of these speakers was dropped, and the remaining speakers had average ratings of 1.18 (native English), 4.68 (mild accent), 5.93 (mild accent), and 7.23 (moderate accent). In subsequent reporting of data for these speakers, their accent ratings are rounded to the nearest whole number and referenced to the maximum rating of 9, as in 1/9, 5/9, 6/9, and 7/9, respectively.

C. Acoustic analyses of stimuli

As noted in the Introduction, stressed syllables are manifested acoustically by longer duration, higher fundamental frequency (f0), and more intense amplitude (Lieberman, 1960; Pickett, 1999). In the minimally contrastive bi-syllabic words used in this experiment, stress is typically observed in the first syllable in nouns and in the second syllable in verbs. Several acoustic analyses were conducted on the speech samples produced by the native English speaker and the three accented speakers to determine the extent to which these speakers produced speech samples that were consistent with these expectations. To that end, Adobe Audition was used to measure duration, average f0, and intensity (average rms) in the first and second vowels of all of the two-syllable stimuli recorded by all four speakers. Results (means and standard deviations) are shown in Table I.

Statistical analyses were conducted to determine how the speakers differed on these measures. For the duration measures, a repeated measures analysis of variance (ANOVA) was conducted with three within-subjects variables [speaker (four levels), word form (two levels), vowel (two levels)]. Results showed significant main effects of speaker [$F(3,120) = 30.1$, $p < 0.001$], word form [$F(1, 40) = 16.4$, $p < 0.001$], and vowel [$F(1,40) = 40.6$, $p < 0.001$], and interactions between speaker × word form [$F(3, 120) = 5.5$, $p < 0.001$], speaker × vowel [$F(3, 120) = 4.2$, $p < 0.01$] and word form × vowel [$F(1, 40) = 500.9$, $p < 0.001$]. The three-way interaction was not significant ($p > 0.05$). *Post hoc* testing was conducted to examine the two-way interactions further, using a corrected alpha-level of 0.008. *Post hoc* t-tests of the vowels in nouns produced by each of the four speakers revealed that there were no significant differences in the duration of vowel 1 vs vowel 2 ($p > 0.008$, all comparisons). However, as expected, vowel 2 was significantly and consistently longer than vowel 1 in verbs ($p < 0.001$, all comparisons). To correct for the effect of the total duration of the bi-syllabic words, a relative difference limen (DL) measure was derived using the magnitude of the difference between the two vowels in proportion to the total duration of the two vowels combined [(vowel 1 – vowel 2)/

TABLE I. Means and standard deviations, in parentheses, of duration (in ms), mean f0 (in Hz), and intensity (rms) for the stressed and unstressed vowels of nouns and verbs spoken by the four talkers. The intended stressed vowel is indicated by italics.

	Nouns					
	Duration (ms)		Mean f0 (Hz)		Relative average intensity (dB)	
	Vowel 1	Vowel 2	*Vowel 1*	Vowel 2	*Vowel 1*	Vowel 2
Speaker 7/9 (moderate accent)	114.1 (31.8)	119.1 (29.2)	136.6 (10.6)	131.3 (16.3)	−23.14 (2.0)	−26.08 (2.2)
Speaker 6/9 (mild accent)	116.2 (27.2)	98.8 (42.0)	121.8 (9.2)	134.5 (15.1)	−22.61 (2.3)	−25.66 (3.0)
Speaker 5/9 (mild accent)	98.2 (28.3)	87.0 (30.4)	158.1 (15.4)	145.8 (21.7)	−22.71 (2.4)	−25.09 (2.4)
Speaker 1/9 (native English)	94.1 (26.3)	101.0 (34.7)	147.1 (20.4)	92.9 (9.9)	−20.86 (2.0)	−27.25 (3.0)
	Verbs					
	Duration (ms)		Mean f0 (Hz)		Relative intensity (rms)	
	Vowel 1	*Vowel 2*	Vowel 1	*Vowel 2*	Vowel 1	*Vowel 2*
Speaker 7/9 (moderate accent)	67.7 (25.3)	140.3 (34.3)	133.8 (15.7)	135.3 (10.4)	−26.15 (2.5)	−22.80 (1.8)
Speaker 6/9 (mild accent)	69.9 (31.2)	137.6 (42.3)	106.6 (11.4)	132.0 (9.2)	−27.21 (3.8)	−20.76 (2.4)
Speaker 5/9 (mild accent)	57.4 (24.6)	128.1 (32.7)	167.8 (23.3)	156.6 (13.9)	−25.64 (3.6)	−22.18 (2.5)
Speaker 1/9 (native English)	45.4 (25.7)	114.1 (46.2)	112.6 (16.3)	145.4 (22.6)	−27.90 (4.2)	−21.64 (2.2)

(vowel 1 + vowel 2)]. Results of an ANOVA for the relative duration DL measures in verbs showed a significant speaker effect [$F(3, 120) = 4.61$, $p < 0.01$), in which the native English speaker (speaker 1/9) was significantly different from two of the accented speakers (ratings of 7/9 and 6/9). None of the other differences between speakers were significant ($p > 0.05$).

A similar set of analyses was conducted for the f0 measures. The ANOVA of the average f0 values for vowel 1 vs vowel 2 in nouns and verbs for the four speakers revealed a significant main effect of speaker [$F(3, 120) = 145.3$, $p < 0.01$], $p < 0.01$] and the following significant interactions: speaker × word form [F[(3, 120) = 17.7, $p < 0.01$], speaker × vowel [$F(3, 120) = 64.2$, $p < 0.01$, word form × vowel [$F(1, 40) = 156.2$, $p < 0.01$], and speaker × word form × vowel [$F(3, 120) = 143.4$, $p < 0.01$]. To analyze the three-way interaction, paired samples t-tests were conducted to compare the f0 in vowel 1 and vowel 2 for each speaker and word form, using the corrected alpha level of 0.008. Results showed that the differences in f0 between vowels 1 and 2 were not significant for the most accented speaker, for either nouns or verbs. In addition, the significant differences in f0 observed for vowel 1 and vowel 2 were in the wrong direction for speaker 6/9 for nouns and for speaker 5/9 for verbs. The only speaker whose difference in f0 was consistent and significant, according to prediction, was the native English speaker (speaker 1/9: $p < 0.008$ for both nouns and verbs).

Finally, analyses of the intensity measures were conducted. The repeated measures ANOVA examining the effects of speaker, vowel, and word form revealed that none of the main effects were significant ($p > 0.05$, each main effect). However, significant interactions were observed between speaker × vowel [$F(3, 120) = 4.0$, $p < 0.01$], word form × vowel [$F(1, 40) = 327.0$, $p < 0.01$] and speaker × word form × vowel [$F(3, 120) = 22.2$, $p < 0.01$]. *Post hoc* analysis of the three-way interaction examined the differences in intensity in vowel 1 vs vowel 2 in nouns and verbs across the four speakers. All comparisons were significant ($p < 0.008$) and in the expected direction. However, this difference was twice as large in the productions of the native speaker than the accented speakers in all comparisons, with the exception of speaker 6/9 for verbs only.

In summary, these acoustic analyses revealed that the native English speaker exhibited consistent and strong differences in the expected direction, based on all the acoustic measures obtained for nouns and verbs, with the exception of differences in vowel duration in nouns. Furthermore, for this speaker, the relative duration measure was significantly stronger than for two of the accented speakers. In contrast, all of the accented speakers did not show the expected acoustic differences in half of the measures (i.e., duration in nouns, fundamental frequency in nouns and verbs). The relative duration measure showed a weaker difference limen for two of these speakers than for the unaccented speaker. Finally, the difference in intensity between the two vowels in nouns and verbs was about half as strong for the accented speakers than for the unaccented speaker. The acoustic analyses of the stimuli therefore showed that all accented speakers produced at least some of the expected differences conveying lexical stress. However, the native English speaker showed more pronounced distinctions in each of the three acoustic measures used to convey contrastive lexical stress than the accented speakers. These findings are generally in agreement with those reported by Zhang *et al.* (2008), for native speakers of Mandarin.

D. Stress patterns

A phonetically trained listener evaluated the stress patterns of the bisyllabic nouns and verbs produced by all of the speakers; correct stress patterns are Stressed (S)-Unstressed (U) for nouns and U-S for verbs. This analysis showed that all the stress patterns of the native speaker (1/9) were correct. The moderately accented speaker (7/9) exhibited an error rate of 20% in both his noun and verb productions. One mildly accented speaker (5/9) showed error rates of 5% for nouns and 10% for verbs, whereas the other mildly accented

speaker (6/9) showed an error rate of 5% for nouns and no errors for verbs. Thus, the majority of productions by the accented speakers followed the expected stress patterns.

E. General procedure

During the experiments, listeners were tested individually in a sound-attenuating chamber. Stimulus delivery and response collection were controlled using E-PRIME (Psychology Software Tools, Inc.) implemented on a PC. The stimuli were played from the PC through an external sound card (ASUS Xonar Essence One), and routed to a monaural insert earphone (EAR 3A) to the listener's better ear or preferred ear at a level of 85 dB sound pressure level (SPL). All stimuli for both experiments were presented in quiet. Listeners were tested over the course of multiple sessions; each session lasted 2 h. They were given frequent breaks during each session. Participants were reimbursed for their listening time at the rate of $12/h. The experimental protocol was approved by the Institutional Review Board at the University of Maryland.

III. EXPERIMENT 1

The aim of this experiment is to examine the effects of speaker accent and listener age and hearing sensitivity on the ability to identify nouns and verbs in contrasting stimuli that vary by lexical stress. In this experiment, the paradigm provided the listener with linguistic and semantic support.

A. Method

1. Participants

In addition to the general criteria for subject selection listed above in Sec. II, participants were required to pass a screening procedure to ensure that they could perform the task at a level that exceeds chance performance. To that end, participants listened to a training/practice block using 11 pairs of contrastive words. These stimuli were recorded by a native speaker of English who was not the same as the speaker used in the remainder of the experiment. The training block provided step-by-step instructions for participants during the first trial and gave correct-answer feedback. The practice block was the same as the training block, but did not include instructions or feedback. In each trial, a single word was presented and the listener was provided with the text of the two high-context sentences appearing on the screen, one corresponding to a high-context sentence for the verb form and one to a high-context sentence for the noun form. The position of the noun vs verb forms of the high context sentences on the screen varied from trial to trial. The listener's task was to identify the appropriate sentence for the target word. Listeners were required to achieve a criterion accuracy score of 15/22 items correct (68%). If the listeners did not meet the criterion on the first practice block, the training block was repeated, and then the listeners were re-tested on the practice block up to three more times until they achieved at least 68% correct.

A total of 64 participants listened to the training and practice blocks. Of the 15 young listeners with normal hearing, all passed the criterion. However, of the older listeners with normal hearing and with hearing loss, 21 of 23 and 23 of 26 passed the criterion, respectively. Some of these older participants did not enroll in the study. The final participant sample for this experiment included 15 young normal hearing listeners (ages 19 to 23 yr, mean = 21.07), 19 older normal-hearing listeners (ages 65 to 76 yr, mean = 68.68), and 18 older hearing-impaired listeners (ages 65 to 87 yr, mean = 74.0). Audiometric thresholds of the three listener groups are shown in Fig. 1.

Possible differences in task difficulty between participants recruited into the study and assigned to the three groups were examined. The first analysis compared the number of trials required to reach criterion performance across the three groups of listeners. Results of a one-way ANOVA showed that the effect of group was not significant [$F(2,51) = 2.5$, $p > 0.05$]. The second analysis compared listener accuracy scores when they reached criterion performance; this analysis showed that the group effect was not significant [$F(2,51) = 2.4$, $p > 0.05$]. Thus, at least for experiment 1, participants who met criterion were matched well for task difficulty on these two indices. A re-analysis of the data was conducted with the inclusion of data from participants who did not meet criterion performance. Values were nominally assigned to these participants to reflect the lowest level of performance observed among the group as a whole, with a maximum number of trials of 4 and an accuracy score of 40% assigned to these individuals. The one-way ANOVA for number of trials failed to reach significance [$F(2,63) = 2.0$, $p > 0.05$]. However, the ANOVA for accuracy scores indicated a significant effect of group [$F(2, 63) = 3.5$, $p < 0.05$]. *Post hoc* testing indicated that the young listeners with normal hearing achieved higher accuracy scores than the older listeners with hearing impairment ($p < 0.008$). The accuracy scores of the older normal hearing listeners did not differ from either of the other two groups ($p > 0.008$, both comparisons). These findings indicate that older hearing-impaired listeners retained for the experiment represent a somewhat high achieving group of individuals.

2. Stimuli

The acoustic stimuli for experiment 1 were isolated bi-syllabic words that varied in lexical stress. These words

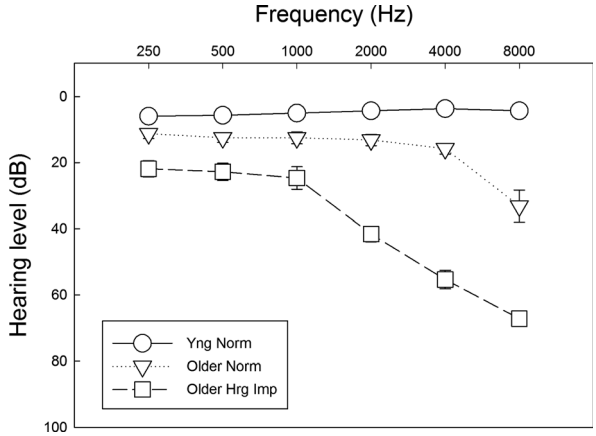

FIG. 1. Mean audiometric thresholds of the three listener groups who participated in experiment 1. Error bars indicate one standard error.

were excised from the recordings of the high-context sentences, described in Sec. II. There were 82 bi-syllabic words, consisting of contrasting 41 nouns and 41 verbs. The 82 stimulus words produced by each speaker were divided into two blocks, such that each block contained the same number of nouns and verbs. Similar blocks were created for each of the four speakers (one native English, and three non-native English).

3. Procedure

After listeners passed the screening criterion, described above, they were tested in the experiment. Listeners received instructions on the computer screen, and testing then commenced. A single bisyllabic word was presented to the listener, followed by a text presentation on the computer screen of the two high context sentences containing the word. The listener's task was to select the sentence that corresponded to the stimulus word. Listeners were tested in blocks of 41 words of one speaker. There were 8 blocks altogether (2 blocks/speaker $\times$ 4 speakers). The order of block presentation varied across listeners using a Latin Squares design. The timing of listener responses was self-paced.

B. Results

1. Accuracy

Overall accuracy in judging lexical stress for the three listener groups as a function of speaker is shown in Fig. 2. The figure generally indicates that the two older groups exhibited lower accuracy than the younger group, and that overall accuracy of all groups decreased with increasing accent. The three panels of Fig. 3 display recognition accuracy for nouns vs verbs as a function of speaker, for each of the three listener groups. These results suggest that accuracy in identifying lexical stress did not vary substantially for the noun vs verb forms, except when spoken by the moderately accented speaker. Listener accuracy in judging lexical stress

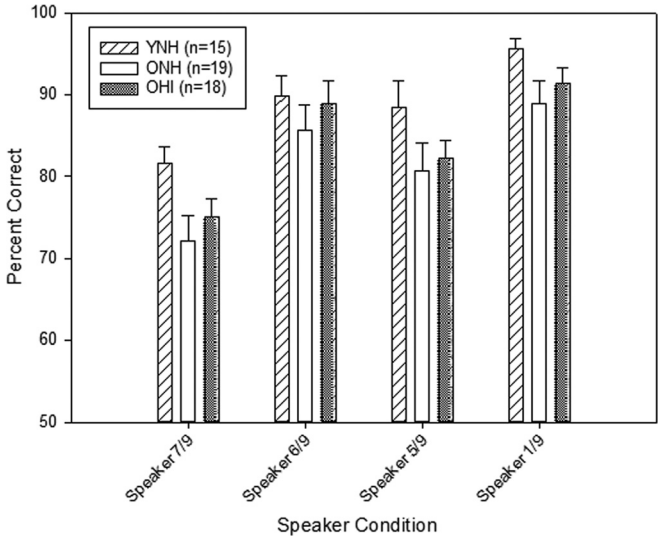

FIG. 2. Overall accuracy judgments of lexical stress by the three listener groups as a function of talker in experiment 1. Error bars indicate one standard error.

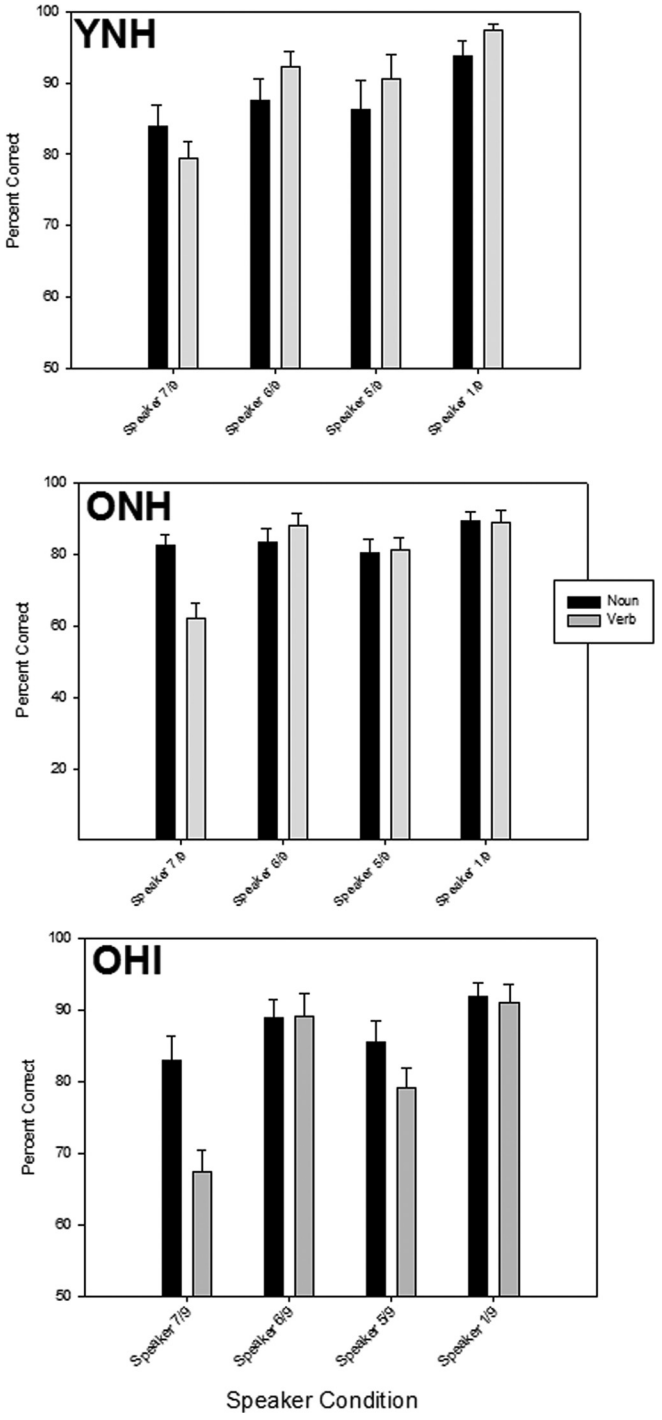

FIG. 3. Accuracy judgments of lexical stress in nouns vs verbs as spoken by the four talkers by the three listener groups in experiment 1. Error bars indicate one standard error.

was lower in verbs than in nouns as spoken by this speaker, especially by the two older groups.

The initial analysis examined the effects of speaker, word form (noun vs verb), and listener group. The percent correct accuracy scores derived for each listener in each condition were transformed using the arc-sine transformation and subjected to an ANOVA with two within-subjects variables (speaker and word form) and one between-subjects variable (listener group). Results are shown in Table II and reveal a significant main effect of speaker and significant

TABLE II. ANOVA results examining effects of speaker, word form, and listener group on listener accuracy in judging lexical stress (experiment 1).

Effect	df	F-value	*p*
Speaker	3, 147	149.77	<0.001[a]
Word form	1, 49	1.24	0.270
Group	2, 49	2.13	0.130
Speaker × word form	3, 147	29.33	< 0.001[a]
Speaker × group	6, 147	2.51	<0.05[a]
Word form × group	2, 49	4.0	<0.05[a]
Speaker × word form × group	6, 147	2.1	0.057

[a]Indicates significant effect.

two-way interactions between speaker × word form, speaker × group, and word form × group.

Subsequent analyses evaluated the significance of the speaker by word form interaction. All *post hoc* results for this and all other interactions are shown in Table III. *Post hoc* t-tests were conducted using Bonferroni corrections. The results showed that the word form effect was observed for two speakers: performance was significantly higher for nouns than verbs for the moderately accented speaker (rating of 7/9) whereas performance was significantly higher for verbs than nouns for one of the mildly accented speakers (rating of 6/9). The speaker effect varied slightly with word form. For verbs, performance for the moderately accented speaker (7/9) was significantly poorer than for all other speakers. In addition, performance for one mildly accented speaker (5/9) was poorer than performance for the other mildly accented speaker (6/9) and the native English speaker (1/9). For nouns, performance for the unaccented speaker (1/9) was better than for all other speakers. In addition, performance for the moderately accented speaker was poorer than for one of the mildly accented speakers (6/9). Thus, in general, accuracy in judging lexical stress was highest for the native English speaker and poorest for the moderately accented speaker for both word forms.

Post hoc analysis of the speaker x group interaction showed that there was a significant group effect for the moderately accented speaker (7/9); none of the other comparisons were significant. Multiple comparison testing showed that young listeners with normal hearing performed better than the older listeners with normal hearing when judging lexical stress for the moderately accented speaker. The effect of speaker was significant for each listener group. All groups judged the unaccented speaker (1/9) more accurately than all other speakers, with the exception of one mildly accented speaker (6/9) as judged by the older hearing-impaired group. Additionally, all groups judged lexical stress as spoken by the moderately accented speaker more poorly than all other speakers. Finally, the two older groups judged lexical stress more accurately for one mildly accented speaker (6/9) than the other mildly accented speaker (5/9). This difference was not significant for the young normal-hearing listeners.

Finally, *post hoc* testing of the word form x group interaction revealed that the group effect was significant for verbs only, in which the young normal-hearing listeners judged lexical stress significantly better than the two older groups

TABLE III. Results of *post-hoc* testing for accuracy in judging lexical stress (experiment 1).

Effect examined/comparison	F	t	df	*p*
Speaker × word form interaction				
Effect of word form (N-V), speaker 7/9		5.94	51	0.000[a]
Effect of word form (N-V), speaker 6/9		−2.92	51	0.005[a]
Effect of word form, speaker 5/9		0.370	51	0.713
Effect of word form, speaker. 1/9		−1.31	51	0.197
Effect of speaker, verbs				
Speaker 7/9 vs 6/9		−13.14	51	0.000[a]
Speaker 7/9 vs 5/9		−9.86	51	0.000[a]
Speaker 7/9 vs 1/9		−16.75	51	0.000[a]
Speaker 6/9 vs 5/9		5.30	51	0.000[a]
Speaker 6/9 vs 1/9		−2.7	51	0.009
Speaker 5/9 vs 1/9		−8.87	51	0.000[a]
Effect of speaker, nouns				
Speaker 7/9 vs 6/9		−2.9	51	0.006[a]
Speaker 7/9 vs 5/9		−0.60	51	0.554
Speaker 7/9 vs 1/9		−8.23	51	0.000[a]
Speaker 6/9 vs 5/9		2.51	51	0.015
Speaker 6/9 vs 1/9		−4.86	51	0.000[a]
Speaker 5/9 vs 1/9		−7.99	51	0.000[a]
Speaker × group interaction				
Effect of group, speaker 7/9	4.2		2, 51	0.02[a]
YN vs ON				0.02[a]
YN vs OHI				0.134
ON vs OHI				1.00
Effect of group, speaker 6/9	0.96		2, 51	0.39
Effect of group, speaker 5/9	2.84		2, 51	0.07
Effect of group, speaker 1/9	2.1		2, 51	0.13
Effect of speaker, YN				
Speaker 7/9 vs 6/9		−6.2	14	0.000[a]
Speaker 7/9 vs 5/9		−4.3	14	0.001[a]
Speaker 7/9 vs 1/9		−10.8	14	0.000[a]
Speaker 6/9 vs 5/9		1.32	14	0.207
Speaker 6/9 vs 1/9		−6.6	14	0.000[a]
Speaker 5/9 vs 1/9		−6.9	14	0.000[a]
Effect of speaker, ON				
Speaker 7/9 vs 6/9		−10.93	18	0.000[a]
Speaker 7/9 vs 5/9		−6.67	18	0.000[a]
Speaker 7/9 vs 1/9		−10.72	18	0.000[a]
Speaker 6/9 vs 5/9		4.43	18	0.000[a]
Effect examined/comparison	F	t	df	*P*
Speaker 6/9 vs 1/9		−3.85	18	0.000[a]
Speaker 5/9 vs 1/9		−6.18	18	0.000[a]
Effect of speaker, OHI				
Speaker 7/9 vs 6/9		−7.74	17	0.000[a]
Speaker 7/9 vs 5/9		−6.33	17	0.000[a]
Speaker 7/9 vs 1/9		−8.72	17	0.000[a]
Speaker 6/9 vs 5/9		5.293	17	0.000[a]
Speaker 6/9 vs 1/9		−1.40	17	0.178
Speaker 5/9 vs 1/9		−7.16	17	0.000[a]
Word form × group interaction				
Effect of group, verbs	8.45	2, 207		0.000[a]
YN vs ON				0.001[a]
YN vs OHI				0.003[a]
ON vs OHI				1.00
Effect of group, nouns	2.04	2, 207		0.133
Effect of word form (N-V), YN		−2.22	59	0.03[a]
Effect of word form (N-V), ON		1.50	0.75	0.14
Effect of word form (N-V), OHI		3.03	71	0.003[a]

[a]Indicates significant effect.

for the verbs. The two older groups did not differ from each other in judging lexical stress in verbs.

Taken together, the findings indicate that highest accuracy for lexical stress is seen for the native English speaker, followed by the mildly accented speakers and poorest accuracy is for the moderately accented speaker. Age-related differences in accuracy for judging lexical stress were largely seen for the moderately accented speaker. Additionally, the two older groups had more difficulty judging lexical stress in verbs than the younger group, although no age-related effects were observed for judging lexical stress in nouns.

2. Relationship of acoustic measures to accuracy

A multiple regression analysis was conducted in which the three acoustic measures obtained for each word were used as predictor variables, and the accuracy scores by all participants for each word were used as the outcome variables. Transformations of the acoustic measures were calculated to use in the analysis. These included a measure of relative duration [(vowel 1 – vowel 2)/(vowel 1 + vowel 2)], a measure of the difference in f0 (vowel 1 – vowel 2), and a measure of the difference in RMS (vowel 1 – vowel 2). A multiple regression was conducted separately for each speaker and each word form. The results indicated a complex pattern of findings, in which the primary acoustic measure contributing to the variance in judgments of lexical stress differed for nouns vs verbs and for speakers of different accents. No one measure emerged as the most important predictor of judgments of lexical stress. Of the eight analyses (4 speakers × 2 word forms), relative duration was retrieved as the most important predictor variable (3 times out of 8 comparisons) compared to the other two predictor variables (2 times out of 8 comparisons). In nouns, two significant predictor variables were retrieved for the unaccented speaker, one significant predictor variable was retrieved for each of the two mildly accented speakers, and no significant predictor variables were identified for the moderately accented speaker. Additionally, more variance in perceptual judgments was explained by the predictor variables for the unaccented speaker compared to the accented speakers. For verbs, there were two predictor variables retrieved for both the unaccented speaker and the moderately accented speaker, which were relative duration and difference in f0, whereas only one predictor variable was retrieved for each of the mildly accented speakers.

C. Discussion

The general conclusions of this first experiment are that listeners have difficulty judging contrastive lexical stress in English bi-syllabic words produced by native speakers of Spanish. Judgments of lexical stress tend to be less accurate as listener accent increases. Older listeners have more difficulty accurately judging lexical stress than younger listeners, particularly for verb stimuli.

These findings are observed when bi-syllabic words varying in lexical stress are presented with both linguistic and lexical support. The extent to which the results extend to situations where linguistic, but not semantic support is provided, is unknown. This is the main objective of experiment 2.

IV. EXPERIMENT 2

The aim of this experiment was to examine the effects of age and hearing loss on perception of lexical stress in bi-syllabic English words spoken by native English and native Spanish speakers when linguistic support is provided, but not semantic support. It was expected that older listeners would experience much greater difficulty than younger listeners in judging lexical stress when semantic cues were not available, because older listeners tend to rely on semantic cues to a large extent in order to resolve ambiguities in spoken language recognition in conditions with signal degradation (Schneider *et al.*, 2010).

A. Method

1. Participants

In order to participate in experiment 2, all participants were required to meet the general qualifications described earlier in Sec. II and additionally were required to pass a criterion level of performance on a practice block using stimuli and a task similar to those used in experiment 2 (described below). The intent of screening for the criterion level of performance was to avoid floor effects. The method for conducting the training and subsequent practice block was as follows. Noun-verb pairs (n = 11 pairs) that were not part of the experimental stimulus set were recorded by a native English-speaking male and presented in a paradigm similar to the actual experiment. For the training, the noun and verb forms of each word were presented one after the other in the same neutral sentence to enhance perception of the different stress patterns; feedback was provided. After completing the training block, the participants completed a practice block which consisted of the same 22 word stimuli presented in the training, but in random order and without feedback. Performance on the practice block was measured and compared against a criterion of 15/22 correct (68%). If the listeners did not meet the criterion on the first attempt, they were allowed to repeat the practice block up to three more times until they achieved at least 68% correct.

All of the young normal-hearing listeners who participated in the practice block met the criterion (n = 15). However, many of the older listeners who attempted to meet the criterion level of performance could not pass the criterion. Specifically, 6 of 21 older normal-hearing participants did not pass the criterion and 5 of 23 older hearing-impaired participants did not pass the criterion. Thus, overall, 25% of the older listeners could not pass the criterion. Three of the older hearing-impaired listeners dropped out of the experiment before completion. The ages of the listeners in the three groups who passed the practice block, enrolled in the experiment, and completed data collection were as follows: YNH group (n = 15; ages 19 to 23 yr, mean = 21.07), ONH group (n = 15; ages 65–76 yr, mean = 69.2), and OHI group (n = 15; ages 67–87 yr, mean = 74.64). All listeners who participated in experiment 2

also participated in experiment 1, but the administration order of the two experiments was alternated between participants.

As for experiment 1, the listeners recruited to the three groups in experiment 2 were compared on measures that could suggest relative difficulty of the basic task. Results of the one-way ANOVA on the number of trials to reach criterion performance indicated that the group effect was not significant [$F(2,44) = 0.81$, $p > 0.05$]. The second ANOVA examined accuracy scores achieved at the trial when the listeners reached criterion performance, and showed a significant group effect [$F(2,44) = 3.5$, $p < 0.05$]. However, *post hoc* multiple comparison testing with Bonferroni corrections failed to show a significant difference in accuracy between any of the groups tested. These data suggest that a highly selected group of older participants recruited into the study after meeting criterion performance did not differ in their performance on the basic lexical stress task used in experiment 2.

2. Stimuli

The stimuli for this experiment consisted of the bi-syllabic nouns and verbs that were described for experiment 1, which were subsequently inserted into four low context carrier sentences that were also recorded by each of the four speakers (one native English speaker and three native Spanish speakers). Two of these carrier sentences were appropriate for nouns ("The ___ was easy to discuss" and "Mr. Black discussed the ____") and two were appropriate for verbs ("Do not forget to ____" and "To ___ requires a lot of time"). The noun and verb forms of each stimulus word were inserted into each of the low-context sentences, creating a total of 328 sentences (41 stimulus words × 2 word forms × 4 sentences). Thus, half of the sentences had a bi-syllabic word stimulus that was congruent with the carrier sentence (e.g., "The CONsole was easy to discuss") and half of the sentences had a word stimulus that was incongruent with the carrier sentences (e.g., "To CONsole requires a lot of time").

3. Procedures

The 328 sentences were assigned to eight blocks, with each block consisting of one sentence per stimulus word. Each block contained a mix of the eight different types of sentences (4 low-context sentences × 2 word forms). The sentence types were divided as equally as possible among the eight blocks, such that approximately half of the stimulus words in each block appeared in congruent sentences and the other half appeared in incongruent sentences. Also, half of the sentences presented the target word near the beginning of the sentence and half presented the target word near the end of the sentence. The same eight blocks were used for each of the four speakers, for a total of 32 blocks. The 41 stimulus sentences in each block were presented randomly by E-Prime so that the stimulus order varied from block to block. Block order was randomized among participants using a Latin square design.

During the experiment, the listener was seated in the IAC booth in front of a computer monitor. Sentence stimuli were presented from a PC through an external sound card (ASUS Xonar Essence One) and routed to a monaural insert earphone (EAR 3A) at a level of 85 dB SPL. The monitor showed the low-context sentence with the target word in capital letters, while simultaneously the sentence containing the target word was presented through the earphones. The listener's task was to decide whether or not the capitalized word was pronounced correctly, by pressing one key for YES and another key for NO. Listeners were encouraged to respond as quickly as possible and to guess if they were unsure.

B. Results

Initially, overall accuracy in judging lexical stress by the three listener groups as a function of speaker and word form was assessed. Figure 4 displays the overall accuracy results of the three listener groups, collapsed across noun and verb word forms, and Fig. 5 shows accuracy of the three groups separately for the four speakers and both word forms. Overall, these figures show similar patterns to those observed for experiment 1. Specifically, accuracy was highest for the native speaker, somewhat lower for the two mildly accented speakers, and poorest for the moderately accented speaker. Effects of age group and word form are apparent, but they varied somewhat by speaker. The omnibus ANOVA confirmed these observations (see Table IV). There were significant main effects of speaker and word form, and significant interactions between speaker × word form, word form × group and speaker × word form × group. A summary of the results of *post hoc* analyses with Bonferroni corrections is shown in Table V. Essentially, these findings somewhat mirror those observed for experiment 1, indicating that judging lexical stress in accented speech generally decreases with increasing accent for all listener groups, judging lexical stress in verbs is poorer than in nouns for some of the accented speakers, and older listeners have more difficulty judging lexical stress than younger

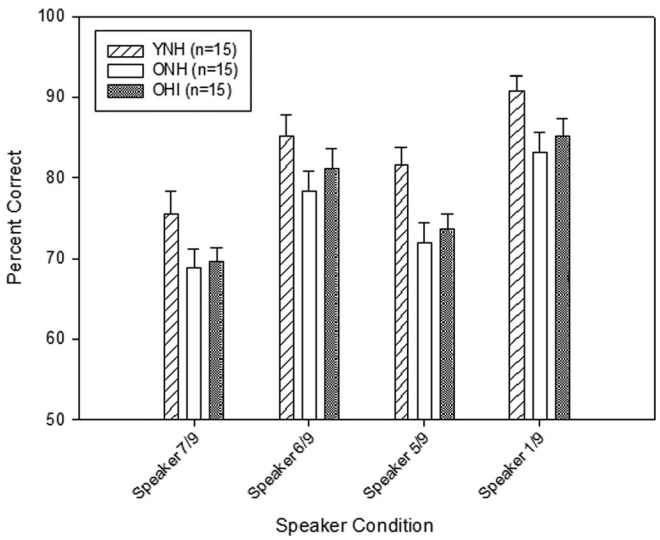

FIG. 4. Overall accuracy judgments of lexical stress by the three listener groups as a function of talker in experiment 2. Error bars indicate one standard error.

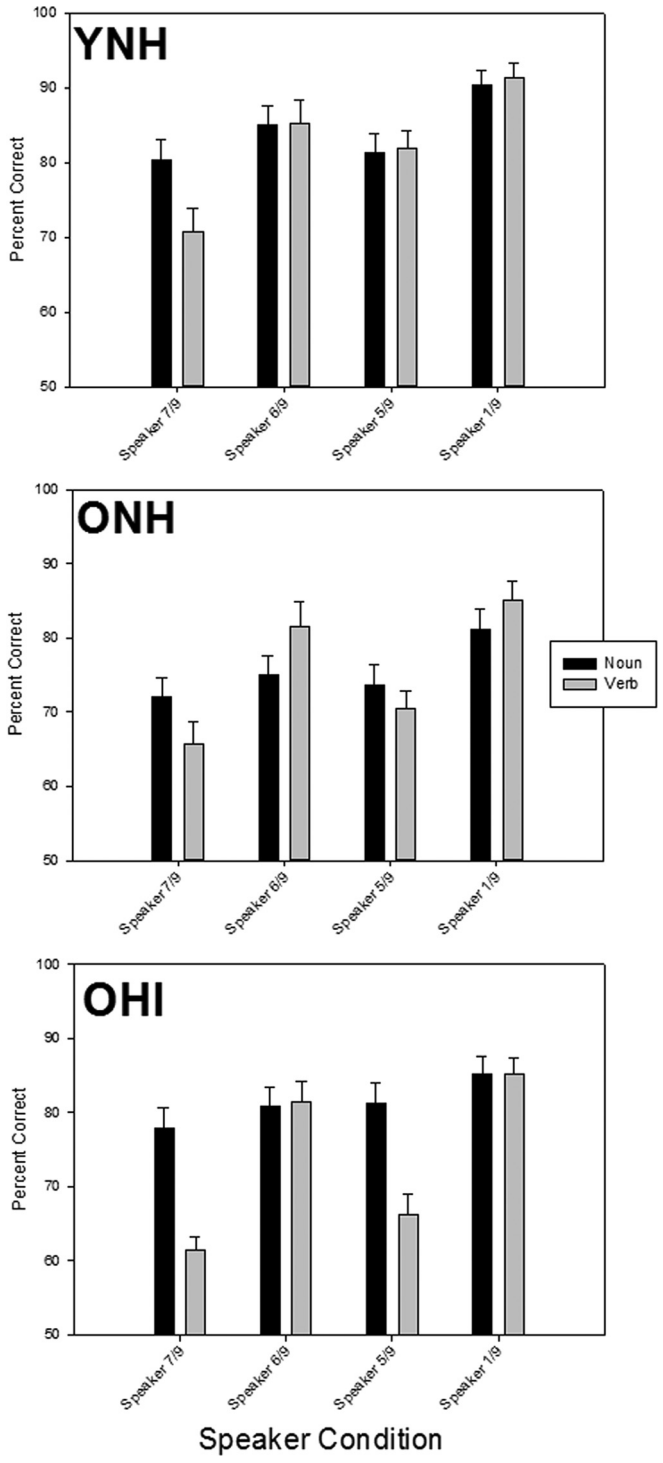

FIG. 5. Accuracy judgments of lexical stress in nouns vs verbs as spoken by the four talkers by the three listener groups in experiment 2. Error bars indicate one standard error.

listeners particularly for accented verbs. These findings are detailed in the next few paragraphs.

The effect of speaker for each listener group and word form was assessed with *post hoc* analyses with Bonferroni corrections. For young normal-hearing listeners judging verbs, accuracy was higher for the unaccented speaker than all accented speakers, and higher for the two mildly accented speakers than for the moderately accented speaker. In judging lexical stress in nouns, these listeners showed higher accuracy for the unaccented speaker than all accented speakers, and higher accuracy for one mildly accented speaker (rating = 6/9) than for the other mildly accented speaker (accent rating = 5/9) and the moderately accented speaker. For older listeners with normal hearing listening to verbs, accuracy was higher for the unaccented speaker and one mildly accented speaker (rating = 6/9) than for the other mildly accented speaker (rating = 5/9) and the moderately accented speaker (rating = 7/9). For nouns, these listeners showed higher accuracy for the unaccented speaker than for one mildly accented speaker (5/9) and the moderately accented speaker only. Older hearing-impaired listeners showed higher accuracy in judgments of perceptual stress in verbs for the native English speaker and one mildly accented speaker (6/9) than for the other mildly accented speaker (5/9) and the moderately accented speaker (7/9). They showed higher accuracy in nouns for the unaccented speaker than one mildly accented speaker (6/9) and the moderately accented speaker (7/9) as well as higher accuracy for one mildly accented speaker (5/9) than the moderately accented speaker (7/9). Thus, all listener groups showed a similar pattern with nouns and verbs, in which accuracy for lexical stress generally decreased with speaker accentedness, although the relative order of accuracy for the two mildly accented speakers was not consistent across group and word form.

TABLE IV. Results of ANOVA examining effects of speaker, word form, and listener group on listener accuracy in judging lexical stress (experiment 2).

Effect	df	F-value	p
Speaker	3, 126	215.69	<0.001[a]
Word form	1, 42	5.19	<0.05[a]
Group	2, 42	2.988	0.061
Speaker × word form	3, 126	36.13	<0.001[a]
Speaker × group	6, 126	1.011	0.421
Word form × group	2, 42	4.33	<0.05[a]
Speaker × word form × group	6, 126	3.66	<0.01[a]

[a]Indicates significant effect.

Post hoc analyses of the effect of word form showed that for both young normal-hearing listeners and older hearing-impaired listeners, lexical stress was judged more accurately for nouns than for verbs for the moderately accented speaker. Additionally, the older hearing-impaired listeners judged lexical stress in nouns more accurately than in verbs for one of the mildly accented speakers (rating = 5/9). In contrast, the older normal-hearing listeners judged lexical stress more accurately for verbs than for nouns for the native-English speaker.

The effect of listener group was assessed for each speaker and word form. One-way ANOVAs revealed that the effect of listener group was significant for verbs spoken by one of the mildly accented speakers (rating 5/9). Multiple comparison testing showed that the two older groups performed more poorly than the younger group, suggesting an age effect. Additional multiple comparison testing also showed the following age effects: young normal-hearing listeners judged lexical stress better than older hearing-impaired listeners for verbs spoken by the moderately accented speaker, and young normal-hearing listeners judged lexical stress better than the

TABLE V. Results of *post-hoc* testing for accuracy in judging lexical stress (Experiment 2).

Effect examined/comparison	F	t	df	*p*
Effect of speaker for each listener group and word form				
Young normal-hearing listeners[a]				
Verbs: speaker 7/9 vs 6/9		−11.982	14	0.000[a]
Verbs: speaker 7/9 vs 5/9		−7.900	14	0.000[a]
Verbs: speaker 7/9 vs 1/9		−14.973	14	0.000[a]
Verbs: speaker 6/9 vs 5/9		2.666	14	0.018
Verbs: speaker 6/9 vs 1/9		−3.810	14	0.002[a]
Verbs: speaker 5/9 vs 1/9		−6.736	14	0.000[a]
Nouns: speaker 7/9 vs 6/9		−4.633	14	0.000[a]
Nouns: speaker 7/9 vs 5/9		−0.754	14	0.463
Nouns: speaker 7/9 vs 1/9		−8.584	14	0.000[a]
Nouns: speaker 6/9 vs 5/9		3.880	14	0.002[a]
Nouns: speaker 6/9 vs 1/9		−7.276	14	0.000[a]
Nouns: speaker 5/9 vs 1/9		−9.757	14	0.000[a]
Older normal-hearing listeners[a]				
Verbs: speaker 7/9 vs 6/9		−7.149	14	0.000[a]
Verbs: speaker 7/9 vs 5/9		−2.801	14	0.014
Verbs: speaker 7/9 vs 1/9		−9.951	14	0.000[a]
Verbs: speaker 6/9 vs 5/9		5.750	14	0.000[a]
Verbs: speaker 6/9 vs 1/9		−16.51	14	0.121
Verbs: speaker 5/9 vs 1/9		−9.429	14	0.000[a]
Nouns: speaker 7/9 vs 6/9		−1.497	14	0.157
Nouns: speaker 7/9 vs 5/9		−1.425	14	0.176
Nouns: speaker 7/9 vs 1/9		−6.395	14	0.000[a]
Nouns: speaker 6/9 vs 5/9		0.779	14	0.449
Nouns: speaker 6/9 vs 1/9		−2.480	14	0.026
Nouns: speaker 5/9 vs 1/9		−5.508	14	0.000[a]
Older hearing-impaired listeners[a]				
Verbs: speaker 7/9 vs 6/9		−9.097	14	0.000[a]
Verbs: speaker 7/9 vs 5/9		−2.108	14	0.054
Verbs: speaker 7/9 vs 1/9		−12.196	14	0.000[a]
Verbs: speaker 6/9 vs 5/9		4.952	14	0.000[a]
Verbs: speaker 6/9 vs 1/9		−2.322	14	0.036
Verbs: speaker 5/9 vs 1/9		−5.608	14	0.000[a]
Nouns: speaker 7/9 vs 6/9		−2.697	14	0.017
Nouns: speaker 7/9 vs 5/9		−4.768	14	0.000[a]
Nouns: speaker 7/9 vs 1/9		−6.531	14	0.000[a]
Nouns: speaker 6/9 vs 5/9		0.688	14	0.503
Nouns: speaker 6/9 vs 1/9		−4.058	14	0.001[a]
Nouns: speaker 5/9 vs 1/9		−2.971	14	0.010
Effect of word form for each speaker within each group				
Young normal-hearing listeners[a]				
Speaker 7/9: nouns vs verbs		4.184	14	0.001[a]
Speaker 6/9: nouns vs verbs		−0.367	14	0.719
Speaker 5/9: nouns vs verbs		−0.190	14	0.852
Speaker 1/9: nouns vs verbs		−1.280	14	0.221
Older normal-hearing listeners[a]				
Speaker 7/9: nouns vs verbs		2.308	14	0.037
Speaker 6/9: nouns vs verbs		−2.389	14	0.032
Speaker 5/9: nouns vs verbs		1.627	14	0.126
Speaker 1/9: nouns vs verbs		−2.995	14	0.010[a]
Older hearing-impaired listeners[a]				
Speaker 7/9: nouns vs verbs		4.535	14	0.000[a]
Speaker 6/9: nouns vs verbs		−0.112	14	0.92
Speaker 5/9: nouns vs verbs		3.229	14	0.006[a]
Speaker 1/9: nouns vs verbs		0.488	14	0.633
Effect of group, for each speaker and word form				
Speaker 7/9, nouns	2.246		2, 44	0.118
Speaker 7/9, verbs	3.208		2, 44	0.050[a]
YN vs ON				0.528
YN vs OHI				0.046[a]
ON vs OHI				0.766
Speaker 6/9, nouns	3.608		2, 44	0.036[a]
YN vs ON				0.032[a]
YN vs OHI				0.806
ON vs OHI				0.383
Speaker 6/9, verbs	0.554		2, 44	0.579
Speaker 5/9, nouns	2.47		2, 44	0.097
Speaker 5/9, verbs	10.406		2, 44	0.000[a]
YN vs ON				0.006[a]
YN vs OHI				0.000[a]
ON vs OHI				0.863
Speaker 1/9, nouns	2.936		2, 44	0.064
Speaker 1/9, verbs	3.213		2, 44	0.050[a]
YN vs ON				0.116
YN vs OHI				0.089
ON vs OHI				1.000

[a]Indicates significant effect.

older normal-hearing listeners for nouns spoken by one mildly accented speaker (rating of 6/9).

During the course of the experiment, it became apparent that listeners had more difficulty judging lexical stress for incongruent carrier sentences and words than for congruent carrier sentences and words. [Congruent = correct answer is "yes"; Incongruent = correct answer is "no."] A second repeated measures ANOVA was therefore conducted to assess the effects of congruency, speaker, and group (see Fig. 6). Results showed significant main effects of congruency [$F(1,42) = 36.78$, $p < 0.001$], speaker [$F(3, 126) = 188.66$, $p < 0.001$], and group [$F(2, 42) = 3.33$, $p < 0.05$], and a significant interaction between congruency and speaker [$F(3, 126) = 34.61$, $p < 001$]. *Post hoc* analyses of the group effect showed that the young normal-hearing listeners achieved significantly higher scores than the older normal-hearing listeners ($p < 0.001$) and the older hearing-impaired listeners ($p < 0.01$), with no significant performance differences between the two older groups. The interaction between congruency and speaker was analyzed with t-tests to examine the effect of congruency for each speaker, collapsed across listener group. Results showed that accuracy was higher for congruent stimuli than for incongruent stimuli for all four speakers ($p < 0.001$, all speakers), but the congruency effect increased with increasing accent.

C. Discussion

1. Overall accuracy

The task of judging lexical stress in bi-syllabic words embedded in sentences that had no semantic support was generally difficult. Preliminary testing to determine if

participants could meet criterion performance in experiment 2 revealed that 30% of older participants could not meet criteria. However, all of the younger listeners met criterion performance. Therefore, the older listeners who participated in experiment 2 generally performed at a high level, and this is reflected in their overall accuracy judgments. Consistent with findings of experiment 1, all listener groups exhibited poorer scores with increasing speaker accent, with the difference being most pronounced between the native English speaker and the moderately accented speaker. Accuracy in judging lexical stress alternated for the two mildly accented speakers, depending on listener group and word form. Group effects were apparent for some of the speaker conditions, and when group effects were observed, the two older listener groups had lower accuracy in judging lexical stress than the young listener group. However, hearing loss effects were not observed, i.e., older hearing-impaired listeners did not perform differently than older normal-hearing listeners in any of the listening conditions. The source of the age effect in judging lexical stress is tentatively thought to be associated with age-related difficulty in discriminating differences in duration between stressed and unstressed vowels, as duration is one of the prominent cues that listeners use for judging lexical stress. If f0 or vowel amplitude were key indices of stress in these stimuli, then effects of hearing loss might have been observed, because hearing loss is accompanied by poorer spectral resolution, which would have affected perception of the f0 cue (Florentine *et al.*, 1980) and intensity discrimination, which would have affected perception of the vowel amplitude cue (Florentine *et al.*, 1993). However, effects of hearing loss were not observed in the present results, perhaps reflecting that listeners were not using these two cues to make judgments of lexical stress. Finally, effects of word form were observed for selected speakers and listeners, and were primarily associated with higher accuracy for nouns than verbs.

2. Congruency

The task used to judge lexical stress in this experiment required listeners to determine if a bi-syllabic noun or verb was correctly or incorrectly placed in a neutral sentence where the intended word was either a noun or verb. That is, the noun form "CONvert" would be correct in the sentence, "Mr. Black discussed the CONvert," but would be incorrect in the sentence, "To CONvert takes a lot of time." The former sentence is an example of a congruent sentence, and the latter is an example of an incongruent sentence. In general, listeners judged lexical stress much more accurately for congruent sentences than for incongruent sentences, although the magnitude of this advantage for congruent sentences varied somewhat with the speaker. Specifically, the congruency effect was significant but modest for the native English speaker, maximal for the moderately accented speaker, and between these two extremes for the two mildly accented speakers. This suggests that listeners were better able to discern the difference between a correct and incorrect sentence, based on lexical stress, for the unaccented speaker, and had progressively more difficulty as speaker accentedness increased. The implication of this finding is that if Spanish-accented speakers produce bi-syllabic words with incorrect stress patterns, as has been shown in experiment 1, then older listeners may not be able to perceive this deviation in a spoken sentence and fail to correct for the altered lexical stress pattern. Consequently, older listeners are likely to misinterpret the intended message.

The analysis of the congruency results also indicates that younger listeners judged lexical stress more accurately for both congruent and incongruent sentences than the two older listener groups. This finding may reflect better discrimination of the contrasting stress across the two syllables in bi-syllabic words by younger listeners than older listeners. As noted previously, multiple cues were employed by all speakers to convey lexical stress, although the contrast between stressed and unstressed syllables (and vowels) was more subtle for non-native speakers. It appears that younger listeners can distinguish even the subtle differences in syllable stress in both congruent and incongruent sentences better than older listeners. This may reflect either a difference in the ability to use multiple cues for stress between younger and older listeners or a difference in the ability to use a single, prominent cue, such as duration, by younger and older listeners.

V. GENERAL DISCUSSION AND SUMMARY

This investigation examined perception of contrastive lexical stress in bi-syllabic English words by younger and older listeners. One hypothesis stated that all listeners would show poorer perception of lexical stress for Spanish-accented English than for native English. The other hypothesis stated that older listeners would show poorer perception of lexical stress for native English and Spanish-accented English compared to younger listeners. Two different experimental paradigms were used to test these hypotheses. In general, the results of both experiments confirmed the hypotheses.

Both experiments showed that listeners' judgments of lexical stress declined with increasing speaker accent. Specifically, performance for the native speakers' productions was best, performance for the moderately accented speakers' productions was poorest, and performance for the two mildly accented speakers was between these two extremes. This general pattern was observed in both experiments. The acoustic analyses of the stimuli suggest that the accented speakers do not convey lexical stress in the same manner as the native speaker of English. Specifically, stressed and unstressed syllables in bi-syllabic words produced by the native speaker contrasted on duration, f0, and amplitude consistently, whereas those produced by the accented speakers were not as consistent and not as pronounced. These differences in the stimuli are the most probable explanation for the decline in judgment accuracy of lexical stress in the speech of accented speakers. These results generally support those of Zhang *et al.* (2008) who studied Mandarin-accented speakers of English. Additionally, the multiple regression analyses examining the variance accounted for by the three measures (duration, f0, intensity) in nouns and verbs across the four speakers generally

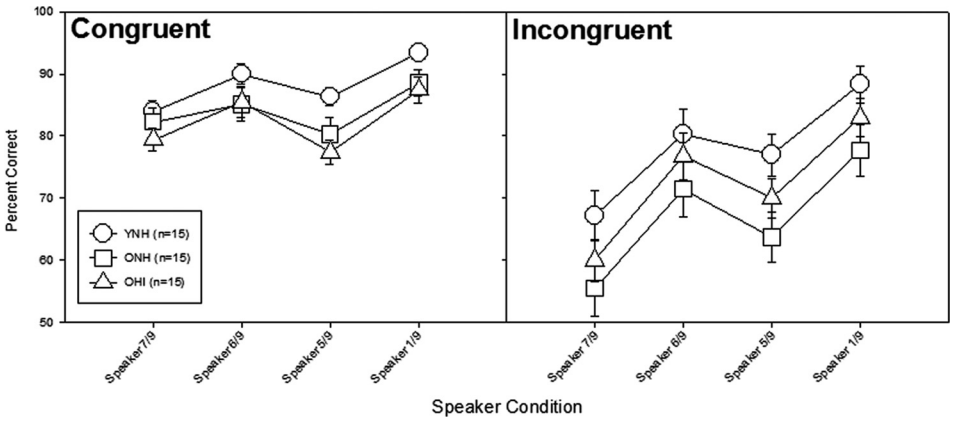

FIG. 6. Accuracy in judging lexical stress in congruent vs incongruent stimuli spoken by the four talkers. Error bars indicate one standard error.

supported the concept that multiple acoustic cues conveyed contrastive lexical stress in the stimuli produced by the native English speaker unlike most of the accented speakers. Of all of the word contrasts examined, relative duration of the stressed and unstressed syllables appeared to be the most frequent acoustic measure to account for significant variance in listener performance.

Experiments 1 and 2 also converged on the finding that older listeners, both with and without hearing loss, generally performed less accurately on judgments of contrastive bi-syllabic lexical stress than the young listeners for non-native speakers, although the three groups performed similarly for the native English speaker. This was observed in light of the difficulty in recruiting older participants who could pass a preliminary screening to demonstrate greater than chance performance on the two tasks. For both experiments, all young listeners who were screened reached criterion performance, but this was not the case for the older listeners. The older listeners who participated in the experiments could therefore be considered a biased sample of listeners who were exceptionally good at judging lexical stress for their age group; if an unbiased sample of older listeners could have been recruited (with a different screening procedure) and tested, it is likely that age differences in performance would have been even greater than those reported here. The source of the age effect in judging lexical stress accurately for the accented speakers may be either that the older listeners are less adept at using the multiple cues that convey lexical stress than younger listeners, or may have relied on one prominent cue for lexical stress, such as duration, and are less sensitive to this single cue than younger listeners. It appears that all three listener groups were able to make good use of the duration, f0, and amplitude cues that signal lexical stress that were available for the native English speaker, as evidenced by mean performance between 89% and 96% correct for all three groups in the first experiment. Although the younger listeners were also able to use the inconsistent and reduced cues produced by the accented speakers, as evidenced by scores of 82%–90% for these speakers, the older listeners appear to have been less proficient in successfully perceiving these multiple cues. This pattern of results suggests that when multiple, robust cues are available, older adults perceive contrastive lexical stress as well as younger adults, but when the acoustic cues for contrastive bi-syllabic lexical stress are minimized, as in the case of accented speakers, older listeners are less able to adjust to such deviations. Additionally, there were no differences in performance between the older normal-hearing listeners and the older hearing-impaired listeners in any of the conditions. It is assumed that the older hearing-impaired listeners are less sensitive to spectral and intensity cues for stress than the older normal-hearing listeners, as a consequence of poorer spectral and intensity resolution that accompanies sensorineural hearing loss, and thus these listeners must rely primarily on the duration cue to perceive lexical stress. That the two older groups showed no differences in perception of lexical stress tentatively suggests that the older normal-hearing listeners similarly attended primarily to the duration cue for lexical stress. The limited ability of older listeners to accurately distinguish the longer (stressed) from the shorter (unstressed) vowels in the bi-syllabic words may be related to age-related deficits in duration discrimination (Fitzgibbons and Gordon-Salant, 1995). A related problem is that the older listeners may have perceived the differences in the stress patterns, but were unable to assign the correct response alternative to a given stimulus (i.e., difficulty with labeling or confusion over noun vs verb forms).

Listeners tended to show different patterns of performance for the noun and verb forms in both experiments, and this varied with speaker. There were no differences in accuracy in judging lexical stress between nouns and verbs for the native English speaker. Multiple regression analyses showed that for this speaker, the variables retrieved were mean difference in f0 and relative duration, with relative duration accounting for more variance in verbs and difference in f0 accounting for more variance in nouns (data not shown). This finding reflects the substantial differences in mean f0 in nouns, and mean vowel duration in verbs, between the stressed and unstressed vowels by this speaker (see Table I). For the most heavily accented speaker, listeners (especially older) had more difficulty judging lexical stress in verbs than in nouns. The multiple regression analysis revealed that none of the acoustic measures contributed to accuracy in judging lexical stress in nouns by this speaker, perhaps reflecting the strikingly similar acoustic values in all measures in the stressed and unstressed vowels in nouns as

shown in Table I. It is unclear why listeners were able to judge lexical stress relatively well for nouns, given these results. For verbs, the multiple regression analysis retrieved the difference in f0 as the most important variable and relative duration as the second most important variable contributing to performance across all listeners. It appears, then, that for this speaker, listeners did not use the relative duration of stressed and unstressed vowels in verbs to judge lexical stress effectively, as they did for the native English speaker. The difference in duration between the vowels in the stressed and unstressed syllables, relative to the total vowel duration, was considerately reduced for the moderately accented speaker, and may be one reason for the poorer perception of lexical stress in verbs for this speaker. Moreover, the distinction in f0 in the stressed and unstressed vowels was minimal for verbs for this speaker, but was more distinct for the other speakers across the two word forms. These different stress patterns observed in the productions of the three native speakers of Spanish evaluated in this study underscore the notion that individual differences in allocating stress to bi-syllabic English words exist among non-native speakers, even those with the same native language.

Some differences were observed in listener performance across the two experiments. In general, performance was poorer in the second experiment than in the first experiment. One finding in support of this interpretation is that 25% of the older listeners who were screened for enrollment in the second experiment could not pass the screening, whereas 8% of the older listeners who were screened for enrollment in the first experiment could not reach criterion performance. A comparison of the overall accuracy scores in the first and second experiments (Figs. 2 and 4) also indicates that scores for each group in each speaker condition were lower for the second experiment than the first. Statistical analyses confirmed that performance was poorer in experiment 2 than experiment 1, for each listener group in each speaker condition, except for the younger group listening to the moderately Spanish-accented speaker ($p < 0.008$). The first experiment provided linguistic and semantic support by presenting listeners with two possible written sentences from which to select the correct placement of the target word. The two choices were meaningful sentences each of which included a bi-syllabic word, and these bisyllabic words contrasted in lexical stress. The listener's task was to decide which of the two sentences matched the lexical stress pattern heard in the target word. The second experiment provided linguistic support but not semantic support. Listeners heard a neutral sentence in which the bi-syllabic target word was embedded in the appropriate sentence position and were required to determine if the target word was the correct form of the word in that neutral sentence. Thus, listeners needed to determine if they heard the noun or verb form of the bi-syllabic word and if the structure of the neutral sentence required a noun or verb. The overall accuracy findings thus suggest that when listeners are provided with both linguistic and semantic support, they judge lexical stress better than when they are provided with linguistic support only. Most prior experiments evaluating perception of lexical stress have used only one method of assessment that provided linguistic support (e.g., Cutler and Clifton, 1984; Fry, 1958). The present results generally suggest that patterns of performance in judging lexical stress across different listening conditions (in this case, variations in degree of speaker accentedness) are similar when obtained through different experimental paradigms, but that the overall level of performance is poorer when fewer semantic-contextual cues are available. It was expected that older listeners would demonstrate substantially better performance when both lexical and semantic support were available compared to when lexical support only was available, based on prior findings indicating that older listeners take more advantage of contextual information (Pichora-Fuller *et al.*, 1995). Although all groups showed significantly better performance in the first experiment than the second, this difference was consistent for the older listeners, but not for the younger listeners. Additionally, the magnitude of the difference in performance between experiments 1 and 2 was greater for the two older groups compared to the younger group.

The current findings add new insight for understanding underlying sources of older listeners' difficulty in understanding Spanish-accented English sentences and multisyllabic words. These speech stimuli are characterized by deviations in the expected lexical stress patterns (i.e., stressed and unstressed syllables) relative to native English. The present results suggest that older listeners are challenged more than younger listeners in recognizing such incorrect stress patterns and adjusting their perception accordingly. These problems are exacerbated when semantic contextual cues are limited, especially for older listeners. Age-related deficits in temporal processing are tentatively thought to underlie the age-associated difficulties in accurately perceiving lexical stress produced by Spanish-accented speakers.

ACKNOWLEDGMENTS

This research was supported by grant #R01AG09191 from the National Institute on Aging of the National Institutes of Health. The authors are grateful to Hannah Willison, Kerrianne Costantino, and Rebecca Bieber for their assistance in data collection and analysis.

APPENDIX A: THE BI-SYLLABIC WORDS CONTRASTING IN LEXICAL STRESS USED IN THE EXPERIMENTS

accent	contract	pervert	refuse
addict	contrast	present	reject
address	convert	produce	subject
abstract	convict	progress	survey
combat	defect	project	suspect
compact	digest	protest	torment
compound	extract	rebel	transport
compress	impact	recall	upset
conduct	incense	recess	
conflict	object	record	
content	permit	refund	

APPENDIX B: HIGH CONTEXT SENTENCES USED FOR RECORDING THE STIMULUS WORDS

Nouns	Verbs
The short **abstract** was published in the journal.	You must **abstract** the meaning of this sentence.
His **accent** was hard to understand.	To **accent** her eyes she wore make-up.
The food **addict** ate too much.	Drugs might **addict** young people.
His **address** is on First Avenue.	Don't **address** the audience quietly.
The **combat** boots were very dirty.	They **combat** their illness every day.
Electric **compact** cars save on gas.	Big machines **compact** our trash weekly.
The army **compound** is controlled by the general.	Do not **compound** your mistakes by repeating them.
The cold **compress** will improve your headache.	You must **compress** air to inflate your tires.
His **conduct** in grade school was very good.	To **conduct** a train you must be awake.
One way to end a **conflict** is to shake hands.	The two news reports **conflict** with each other.
Study the **content** of this book.	The man is **content** with his job.
Our work **contract** is over.	Do not **contract** Asian flu.
A sharp **contrast** is seen with black and white.	You should **contrast** her book and the movie.
The Catholic **convert** prays to God every day.	Solar panels **convert** sunlight to energy.
The **convict** escaped from prison through a tunnel.	To **convict** the accused requires clear evidence.
He found a **defect** in his leather jacket.	The spy should **defect** to a friendly country.
She wrote a **digest** of the new book.	Make sure you **digest** your evening meal.
The star made a grand **entrance** onto the stage.	The beauty queen can **entrance** her admirers.
Flower **extract** is too expensive.	Surgeons **extract** tumors from patients.
The strong **impact** of the blow broke his jaw.	Harsh laws **impact** a person's way of life.
The smell of the burning **incense** was very strong.	His bad behavior will **incense** his step-father.
The French art **object** is at the museum.	The girl will **object** to cleaning her bedroom.
A learner's **permit** is obtained at age sixteen.	The parents **permit** their son to stay out late.
The police charged the **pervert** with assault.	Your ugly lies **pervert** the truth.
I bought a **present** for her first birthday.	I want to **present** my research results.
Fresh **produce** can be bought at the store.	Cars **produce** harmful gas emissions.
We will follow your **progress** on a daily basis.	The athlete wants to **progress** to a higher level.
The reconstruction **project** was delayed because of rain.	The director will **project** the film onto the screen.
The **protest** march was peaceful.	We **protest** against tax hikes.
The political **rebel** marched for freedom.	Some teenagers **rebel** against strict rules.
Honda had a **recall** of its hybrid cars.	The students must **recall** yesterday's lecture.
Boys play during **recess** at school.	The jury will **recess** for lunch.
The **record** shows that you have made a mistake.	To **record** your voice you need a microphone.
Your dad's tax **refund** is in the mail.	The store will **refund** your money now.
Put your plastic **refuse** in the blue trash bin.	His parents **refuse** to invite her to lunch.
The **reject** pile contains old textbooks.	Don't **reject** her offer of help.
The main **subject** of the story was love.	We will **subject** you to many trials.
The students will conduct a **survey** on cheating.	Helicopters are used to **survey** the damage.
The murder **suspect** is in prison.	The police **suspect** him of stealing.
It was a **torment** to sit through his long speech.	The brothers **torment** their sister every day.
The final score was an **upset** for the team.	The boy's behavior will **upset** his parents.

Alfano, I., Llisterri, J., and Savy, R. (**2007**). "The perception of Italian and Spanish lexical stress: A first cross-linguistic study," in *ICPhS 2007. Proceedings of the 16th International* Congress of Phonetic Sciences [CD-ROM], Saarbrücken, Germany, 6–10 August, pp. 1793–1796.

ANSI (**2010**). ANSI S3.6-2010, *American National Standard Specification for Audiometers (Revision of ANSI S3.6-1996, 2004)* (American National Standards Institute, New York).

Braun, B., Lemhofer, K., and Mani, N. (**2011**). "Perceiving unstressed vowels in foreign- accented English," J. Acoust. Soc. Am. **129**, 376–387.

Brysbaert, M., and New, B. (**2009**). "Moving beyond Kucera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English," Behav. Res. Meth. **41**, 977–990.

Brysbaert, M., and New, B. (**2015**). www.subtlexus.lexique.org (Last viewed 10/30/2015).

Brysbaert, M., New, B., and Keuleers, E. (**2012**). "Adding Part-of-Speech information to the SUBTLEX-US word frequencies," Behav. Res. Meth. **44**, 991–997.

Burda, A. N., Scherz, J. A., Hageman, C. F., and Edwards, H. T. (**2003**). "Age and understanding speakers with Spanish or Taiwanese accents," Percept. Motor Skills **97**, 11–20.

Chakraborty, R., and Goffman, L. (**2011**). "Production of lexical stress in non-native speakers of American English: Kinematic correlates of stress and transfer," J. Speech Lang. Hear. Res. **54**, 821–835.

Cutler, A. (**1986**). "*Forbear* is a homophone: Lexical prosody does not constrain lexical access," Lang. Speech **29**, 201–220.

Cutler, A. (**2005**). "Lexical stress," in *The Handbook of Speech Perception*, edited by D. Pisoni and R. Remez (Wiley-Blackwell, Hoboken, NJ), Chap. 11.

Cutler, A., and Clifton, C. E. (**1984**). "The use of prosodic information in word recognition," in *Attention and Performance X: Control of Language Processes*, edited by H. Bouma and D. G. Bouwhuis (Lawrence Erlbaum, Hillsdale, NJ), pp. 183–960.

Edmunds, P. (**2009**). "ESL speakers' production of English lexical stress: The effect of variation in acoustic correlates on perceived intelligibility and nativeness," Ph.D. dissertation, The University of New Mexico, Albuquerque, NM.

Ferguson, S. H., Jongman, A., Sereno, J. A., and Keum, K. A. (**2010**). "Intelligibility of foreign- accented speech for older adults with and without hearing loss," J. Am. Acad. Aud. **21**, 153–162.
Fitzgibbons, P., and Gordon-Salant, S. (**1994**). "Age effects on measures of auditory duration discrimination," J. Speech Hear. Res. **37**, 662–670.
Fitzgibbons, P. J., and Gordon-Salant, S. (**1995**). "Duration discrimination with simple and complex stimuli: Effects of age and hearing sensitivity," J. Acoust. Soc. Am. **98**, 3140–3145.
Fitzgibbons, P. J., and Gordon-Salant, S. (**2001**). "Aging and temporal discrimination in auditory sequences," J. Acoust. Soc. Am. **109**, 2955–2963.
Flege, J. E., and Bohn, O.-S. (**1989**). "An instrumental study of vowel reduction and stress placement in Spanish-accented English," Stud. Sec. Lang. Acquisit. **11**, 35–62.
Florentine, M., Buus, S., Scharf, B., and Zwicker, E. (**1980**). "Frequency selectivity in normally-hearing and hearing-impaired observers," J. Speech Hear. Res. **23**, 457–461.
Florentine, M., Reed, C. M., Rabinowitz, W. M., Braida, L. D., Durlach, N. I., and Buus, S. (**1993**). "Intensity discrimination in listeners with sensorineural hearing loss," J. Acoust. Soc. Am. **94**, 2575–2586.
Folstein, M. F., Folstein, S. E., and McHugh, P. R. (**1975**). "Mini-mental state: A practical method for grading the cognitive state of patients for the clinician," J. Psychiatr. Res. **12**, 189–198.
Fry, D. B. (**1955**). "Duration and intensity as physical correlates of linguistic stress," J. Acoust. Soc. Am. **27**, 765–768.
Fry, D. B. (**1958**). "Experiments in the perception of stress," Lang. Speech **1**, 126–152.
Gordon-Salant, S., Yeni-Komshian, G. H., and Fitzgibbons, P. (**2008**). "The role of temporal cues in word identification by younger and older adults: Effects of sentence context," J. Acoust. Soc. Am. **124**, 3249–3260.
Gordon-Salant, S., Yeni-Komshian, G. H., and Fitzgibbons, P. J. (**2010a**). "Recognition of accented English in quiet by younger normal-hearing listeners and older listeners with normal hearing and hearing loss," J. Acoust. Soc. Am. **128**, 444–455.
Gordon-Salant, S., Yeni-Komshian, G. H., and Fitzgibbons, P. J. (**2010b**). "Perception of accented English in quiet and noise by younger and older listeners," J. Acoust. Soc. Am. **128**, 3152–3160.
Gordon-Salant, S., Yeni-Komshian, G. H., Fitzgibbons, P. J., and Cohen, J. I. (**2015**). "Effects of talker accent and age on recognition of multisyllabic words," J. Acoust. Soc. Am. **137**, 884–897.
Gordon-Salant, S., Yeni-Komshian, G. H., Fitzgibbons, P. J., Cohen, J. I., and Waldroup. C. (**2013**). "Recognition of accented and unaccented speech in different noise backgrounds by younger and older listeners," J. Acoust. Soc. Am. **134**, 618–627.
Grose, J. H., Hall, J. W., and Buss, E. (**2006**). "Temporal processing deficits in the pre-senescent auditory system," J. Acoust. Soc. Am. **119**, 2305–2315.
Johansson, S., and Hofland, K. (**1989**). *Frequency Analysis of English Vocabulary and Grammar: Based on the LOB Corpus* (Clarendon Press, New York).
Lehiste, I., and Peterson, G. (**1959**). "Vowel amplitude and phonemic stress in American English," J. Acoust. Soc. Am. **31**, 428–435.
Lieberman, P. (**1960**). "Some acoustic correlates of word stress in American English," J. Acoust. Soc. Am. **32**, 451–454.
Llisterri, J., Machuca, M. J., de la Mota, C., Riera, M., and Rios, A. (**2003**). "The perception of lexical stress in Spanish," in *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona.
Magen, H. S. (**1998**). "The perception of foreign-accented speech," J. Phon. **26**, 381–400.
Munro, M. J., and Derwing, T. M. (**2001**). "Modelling perceptions of the comprehensibility and accentedness of L2 speech: The role of speaking rate," Stud. Sec. Lang. Acquisit. **23**, 451–468.
Nakatani, L. H., and Aston, C. H. (**1978**). "Perceiving the stress pattern of words in sentences," J. Acoust. Soc. Am. **63**, S55.
Peng, L., and Ann, J. (**2001**). "Stress and duration in three varieties of English," World Englishes **20**, 1–27.
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.
Pickett, J. M. (**1999**). *The Acoustics of Speech Communication* (Allyn and Bacon, Boston).
Pike, K. L. (**1945**). *The Intonation of American English* (University of Michigan Press, Ann Arbor, MI).
Reinish, E., and Weber, A. (**2012**). "Adapting to suprasegmental lexical stress errors in foreign-accented speech," J. Acoust. Soc. Am. **132**, 1165–1176.
Riney, T. J., Takada, M., and Ota, M. (**2000**). "Segmentals and global foreign accent: The Japanese flap in EFL," TESOL Quart. **34**, 711–737.
Roup, C. M., Wiley, T. L., Safady, S. H., and Stoppenbach, D. T. (**1998**). "Tympanometric screening norms for adults," Am. J. Audiol. **7**, 55–60.
Schneider, B. A., Pichora-Fuller, M. K., and Daneman, M. (**2010**). "The effects of senescent changes in audition and cognition on spoken language comprehension," in *Springer Handbook of Auditory Research: The Aging Auditory System: Perceptual Characterization and Neural Bases of Presbycusis*, edited by S. Gordon-Salant, R. D. Frisina, A. N. Popper, and R. R. Fay (Springer, New York), pp. 167–210.
Strouse, A., Ashmead, D. H., Ohde, R. N., and Grantham, D. W. (**1998**). "Temporal processing in the aging auditory system," J. Acoust. Soc. Am. **104**, 2385–2399.
Taler, V., Baum, S., and Saumier, D. (**2006**). "Perception of linguistic and affective prosody in younger and older adults," paper presented at the *28th Annual Conference of the Cognitive Science Society*, Vancouver, pp. 2216–2221.
Tillman, T., and Carhart, R. (**1966**). "An expanded test for speech discrimination utilizing CNC monosyllabic words, Northwestern University auditory test no. 6 technical report," USAF School of Aerospace Medicine, SAM-TR- 66-55, Brooks AFB, Texas.
Tremblay, K. L., Piskosz, M., and Souza, P. (**2002**). "Aging alters the neural representation of speech cues," NeuroReport **13**, 1865–1870.
Wenk, B. J. (**1985**). "Speech rhythms in second language acquisition," Lang. Speech **28**, 157–175.
Yeni-Komshian, G. H., Flege, J. E., and Liu, S. (**2000**). "Pronunciation efficiency in first and second languages of Korean-English bilinguals," Bilingualism: Lang. Cogn. **3**, 131–149.
Zhang, Y., Nissen, S. L., and Francis, A. L. (**2008**). "Acoustic characteristics of English lexical stress produced by native Mandarin speakers," J. Acoust. Soc. Am. **123**, 4498–4513.
Zuraiq, W., and Sereno, J. A. (**2007**). "English lexical stress cues I native English and non-native Arabic speakers," ICPhS I Saarbrücken, 6–10 August, pp. 829–832.