Effects of age and hearing loss on recognition of unaccented and accented multisyllabic words

Sandra Gordon-Salant,[a)] Grace H. Yeni-Komshian, Peter J. Fitzgibbons, and Julie I. Cohen[b)]
Department of Hearing and Speech Sciences, University of Maryland, College Park, Maryland 20742

(Received 5 August 2013; revised 7 January 2015; accepted 8 January 2015)

The effects of age and hearing loss on recognition of unaccented and accented words of varying syllable length were investigated. It was hypothesized that with increments in length of syllables, there would be atypical alterations in syllable stress in accented compared to native English, and that these altered stress patterns would be sensitive to auditory temporal processing deficits with aging. Sets of one-, two-, three-, and four-syllable words with the same initial syllable were recorded by one native English and two Spanish-accented talkers. Lists of these words were presented in isolation and in sentence contexts to younger and older normal-hearing listeners and to older hearing-impaired listeners. Hearing loss effects were apparent for unaccented and accented monosyllabic words, whereas age effects were observed for recognition of accented multisyllabic words, consistent with the notion that altered syllable stress patterns with accent are sensitive for revealing effects of age. Older listeners also exhibited lower recognition scores for moderately accented words in sentence contexts than in isolation, suggesting that the added demands on working memory for words in sentence contexts impact recognition of accented speech. The general pattern of results suggests that hearing loss, age, and cognitive factors limit the ability to recognize Spanish-accented speech. [http://dx.doi.org/10.1121/1.4906270]

[BRM] Pages: 884–897

I. INTRODUCTION

Older listeners exhibit difficulty understanding accented speech. Age-related performance decline for accented speech has been observed for recognition of Spanish-accented monosyllabic words and sentences presented in either quiet or noise (Burda *et al.*, 2003; Ferguson *et al.*, 2010; Gordon-Salant *et al.*, 2010a,b). Recent work has focused on the phonetic cues in Spanish-accented English that show age-related differences in perception, and findings suggest that phonetic contrasts cued by temporal distinctions are particularly difficult for older listeners (Gordon-Salant *et al.*, 2010a). Corresponding work examining how differences in syllable stress with accent affect perception of multisyllabic vs monosyllabic words by younger and older listeners is virtually non-existent. Given that stress and timing in a spoken message are powerful cues used by both younger and older listeners to support speech recognition (Wingfield *et al.*, 2000), it seems important to investigate the extent to which alterations in stress and timing with accent affect listener performance and whether or not age-related differences operate to influence performance.

Non-native speech can be viewed as a form of altered speech, in which there are deviations relative to native speech at the segmental as well as suprasegmental levels. The focus of the present paper is on Spanish-accented English, because the majority of non-native talkers in the U.S. are native talkers of Spanish (Shin and Kominski, 2010). Acoustic analyses of Spanish-accented English have shown reduced differences between stressed and unstressed vowels in multisyllabic words relative to native English talkers (Shah, 2004). These changes are consistent with the observation that Spanish and English differ in their general rhythmic patterns, in which equal stress across syllables is more predominant in Spanish than in English (Pons and Bosch, 2010). Traditionally, this has been denoted through reference to Spanish as a syllable-timed language with equal timing between successive syllables, and to English as a stress-timed language with equal timing between stressed syllables (Pike, 1945). The implication of a syllable-timed language likely extends to the production of multisyllabic words, in which it could be predicted that equal stress across syllables is more common in Spanish-accented English than in native English. There are many acoustic correlates of stress in English, including pitch, duration, vowel quality, and intensity (e.g., Lehiste, 1970; Beckman, 1986). Second language learners tend to transfer at least some aspects of the native-language stress-timing pattern to the new language being learned (Wenk, 1985; Peng and Ann, 2001), suggesting that native talkers of Spanish may produce English characterized by syllable timing rather than stress timing. Thus, the expected pattern of stress and unstressed syllables in native English may be altered in Spanish-accented English.

The challenge for listeners is to process this type of altered speech, adjust for the differences relative to native English, and derive accurate meaning. While young listeners with normal hearing are relatively good at this task (e.g., Gordon-Salant *et al.*, 2010a), older listeners, especially those with hearing loss, are less able to accurately recognize

[a)]Author to whom correspondence should be addressed. Electronic mail: sgsalant@umd.edu
[b)]Current address: Walter Reed National Military Medical Center, 8901 Wisconsin Avenue, Bldg. 102, Bethesda, MD 20889.

 0001-4966/2015/137(2)/884/14/$30.00

accented speech (Gordon-Salant *et al.,* 2010a, 2013). Performance deficits of older listeners may reflect age-related central auditory temporal processing limitations (Anderson *et al.,* 2012) coupled with hearing-loss-related spectral distortions that act to further alter the internal representation of the accented speech signal to be processed in the afferent auditory pathway. The older person then must resolve this altered signal using cognitive processing resources that also may be declining with age (e.g., working memory, speed of processing, selective attention, etc.), creating a speech processing task that substantially increases listening effort.

Recent work has examined the phonetic cues in Spanish-accented English that may be especially difficult for older listeners to perceive accurately (Gordon-Salant *et al.,* 2010a). Stimuli were monosyllabic words that varied in place, manner, and voicing in initial and final consonants. The findings showed that the perceptual errors of older listeners with normal hearing and with hearing loss are associated with temporal alterations in specific segmental cues with accent, such as vowel duration as a cue to post-vocalic voicing (e.g., *bit* vs *bid*), alignment of the onset of voicing and frication in initial voiced fricatives (e.g., *zoo* vs *Sue*), and silence duration as a cue to the affricate/fricative distinction (e.g., *chair* vs *share).* Those with hearing loss also exhibited place errors consistent with deficits in spectral resolution (e.g., *cope* vs *coat* and *weep* vs *wheat*).

The focus of the current investigation is the effect of accent on stress and timing within words, and how alterations with accent affect perception of words with an increasing length of syllables. Spanish-accented English talkers may produce appropriate stress-timing patterns for monosyllabic words in ongoing speech, but minimize expected stress-timing differences across stressed and unstressed syllables for multisyllabic words. This notion derives from the fact that Spanish is a syllable-timed language with equal stress on each syllable. Thus, monosyllabic words are appropriately stressed, but multisyllabic words produced with equal stress on each syllable would not reflect the expected variations in stressed and unstressed syllables that characterize native English. This accent-induced alteration in timing for multisyllabic words therefore may produce inappropriate stress on syllables within a word, disrupt prosodic phrasing in sentences, change speech rate, and compromise recognition. The first hypothesis to be examined in these experiments is that accent affects recognition of multisyllabic words more than monosyllabic words. It was assumed that monosyllabic words would be perceived more accurately than multisyllabic words, even with accent, because the duration of the vowel in a single syllable word is longer than the duration of the same vowel in multisyllabic words (Harris and Umeda, 1974). A potentially competing factor, however, is that multisyllabic words are typically recognized with higher accuracy than monosyllabic words (Fletcher and Steinberg, 1929), perhaps reflecting the reduced number of lexical competitors in multisyllabic words compared to monosyllabic words (Luce and Pisoni, 1998). One implication of the first hypothesis is that recognition of multisyllabic words is affected more by accent, thus reducing any lexical advantage associated with fewer lexical competitors in recognition of multisyllabic words compared to monosyllabic words.

A second hypothesis is that age-related differences in word recognition will be larger with an increasing number of syllables comprising the word. This hypothesis derives from previous findings indicating that speeding a portion of a message has a significant detrimental effect on performance of older but not younger listeners (Gordon-Salant and Fitzgibbons, 2004). It is expected that with an increase in the number of syllables, there will be a corresponding decrease in the duration of each syllable, which will be more difficult for older listeners to perceive because of auditory temporal processing deficits.

A final hypothesis is that recognition of accented multisyllabic words will be negatively affected by increasing the cognitive demand on the task. This was assessed in two ways: first, by introducing the target stimuli in the context of a sentence, and second, by increasing the variability in syllable components from trial to trial. Placing an accented word in the context of an accented sentence is expected to compound the alterations in timing and stress, disrupt processing, increase the memory demand, and distract older listeners' attention. Similarly, varying the number of syllables in the stimulus from trial-to-trial is expected to add to the cognitive load by requiring listeners to constantly adjust their attention to the number of syllables in the stimulus. In other words, listeners' expectation for the number of syllables in a word cannot be used to aid recognition with inter-trial stimulus variability. Age-related differences in monosyllabic word recognition have been shown in the ability to compensate for trial-to-trial variation in relevant acoustic-phonetic speech cues (e.g., talker, speaking rate, and overall amplitude), a phenomenon attributed to a combination of senescent changes in perceptual normalization and selective attention (Sommers and Barcroft, 2006). The current investigation will expand this concept by determining if older listeners are similarly affected by trial-by-trial variations in the acoustic-phonetic dimension of length of syllables, and the additional impact of talker accent on accessing relevant cues.

II. METHOD

A. Participants

Three groups of volunteers with 15 listeners per group participated in the experiments. They were all native talkers of American English who earned at least a high school diploma. One group consisted of young listeners (ages 18–26 years) with normal hearing sensitivity [pure-tone thresholds $\leq$20 dB hearing level (HL) re: ANSI, 2010, between 250 and 4000 Hz]. A second group included older listeners (65–80 years) with normal hearing sensitivity. The third group was composed of older adults with bilateral, symmetrical, mild-to-moderate, sloping sensorineural hearing losses typical of presbycusis. Average audiograms of the three listener groups are shown in Fig. 1. The participants were required to pass a screening test of general cognitive awareness (Short Portable Mental Status Questionnaire; Pfeiffer *et al.,* 1977) to qualify for inclusion in the study.

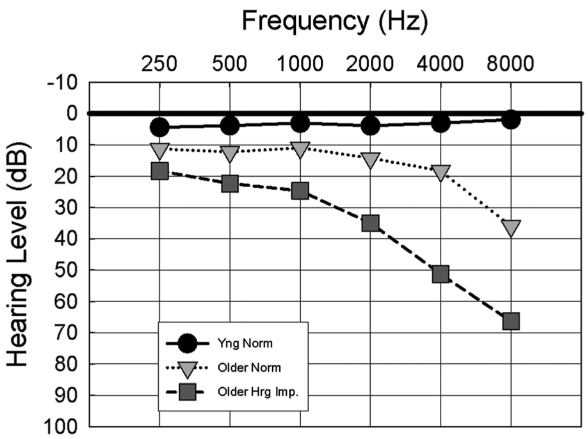

FIG. 1. Mean pure-tone thresholds (in dB HL, re: ANSI, 2010) of the three listener groups.

B. Speech stimuli

The speech stimuli were sets of one-, two-, three-, and four-syllable words that were constructed for this investigation. The words in each set were based on the same sequence of the first three or four phonemes (for example: art, artist, article, artificial). Extensive analyses of these word sets were conducted with the Kucera and Francis (1967) list of word frequencies, the Thorndike and Lorge (1944) frequency counts, and the English Lexicon Project (Balota *et al.*, 2007) word frequencies. Only sets of words with high word frequencies, reflecting common words, were selected. Every effort was made to select words that were comparable in frequency across the words in a set. The total number of sets was 31. The words in these sets were then assigned to lists, with lists 1, 2, 3, and 4 corresponding to all of the one-syllable words, two-syllable words, three-syllable words, and four-syllable words, respectively. List 5 was a mixed list, containing a sampling of words with varying number of syllables that appeared in random order on the list. These constituted the five lists of words presented in isolation. In addition, five lists were created with these same words appearing as the last word of a low-context sentence (e.g., *Mrs. White talked about the art.*).

The words and sentences were recorded in English by four male talkers between 21 and 29 years of age. One talker was a native talker of American English (unaccented), and the other three talkers were native talkers of Spanish (accented). The countries of origin of the three accented talkers were Ecuador, Costa Rica, and Nicaragua. All talkers were either undergraduate or graduate students at universities in the US, and all were enrolled in a program that was taught in English. Recordings were made in a quiet room using a Shure SM48 microphone, a Shure FP-42 pre-amplifier, an Audigy 2ZS 24-bit sound card and a PC. The single-word lists were recorded first, followed by the lists with the words embedded at the end of the sentences. Each word and sentence was recorded three times by each talker, and the second token was selected as the stimulus unless there was a frank pronunciation error, dysfluency, or other recording distortion (e.g., clearing the throat, vocal fry, etc.). In that case, the third token was selected as the stimulus. As part of a larger project involving these same talkers, a panel of five judges (naive native English-speaking young adults with normal hearing) rated the accentedness of other sentence recordings (IEEE sentences) of the four talkers on a scale of 1 to 5, with 1 = unaccented, 2 = slightly accented, 3 = mildly accented, 4 = moderately accented, and 5 = severely accented. The mean rating of the unaccented talker was verified as 1.04, and the ratings of the three accented talkers ranged from 2.75 to 4.41. From these ratings, one talker was selected with a mild accent (mean rating = 2.75; country of origin: Costa Rica) and one talker was selected with a moderate accent (mean rating = 3.69; country of origin: Nicaragua). The recordings from the talker with the heaviest accent (4.41) were not used further in the study because he produced many unintelligible words.

The final set of stimuli used in the experiment therefore consisted of 4 lists of 31 words/each and 4 lists of 31 sentences/each recorded by the 3 talkers (124 unique words and 124 unique sentences × 3 talkers). In addition, a fifth "mixed" list of 31 words and a corresponding "mixed" list of 31 sentences were created using a subset of words and sentences from the four fixed-syllable word lists (i.e., lists 1, 2, 3, and 4).[1] All recorded words and sentences were equated in root-mean-square (RMS) level and a 1 k-Hz calibration tone was created to be equivalent in RMS level to the sentences. To avoid learning effects, the order of the items (words or words in sentences) on each list was randomized separately for each talker.

C. Babble

The background noise used in the experiment was 12-talker babble created in our laboratory to emulate that used in the Speech Perception in Noise test (Kalikow *et al.*, 1977). The talkers were three males and three females between the ages of 20 and 26 years with normal hearing, none of whom served as talkers for the target speech stimuli. They were all native talkers of American English and university students. Each talker was recorded separately in a quiet room while reading different stories from *Grimm's Tales for Young and Old: The Complete Stories* (translated by Ralph Manheim). Two different recordings of each talker were then added to different tracks in Cool Edit Pro (Syntrillium Software). All 12 tracks were digitally mixed to create the 12-talker babble. The RMS level of the 12-talker babble was sampled, and a 1-kHz calibration tone was created to be equivalent in RMS level for purposes of calibration.

The words and sentences recorded by each talker were burned onto separate CDs. One channel of each CD consisted of the 5 lists of 31 words or sentences preceded by the associated calibration tone. The babble was burned onto the second channel of the CDs and was also preceded by the related calibration tone.

D. Cognitive measures

A set of cognitive measures was administered to all participants to determine if variation in cognitive abilities is related to the ability to recognize accented words of varying length of syllables presented in isolation vs sentence

TABLE I. Number and percent of deviations in stress patterns of the two accented talkers relative to the native English talker.

	2 syllables	3 syllables	4 syllables	Total
Mild accent	6/31 = 19.4%	6/31 = 19.4%	12/31 = 38.71%	24/93 = 25.8%
Moderate accent	21/31 = 67.4%	24/31 = 77.4%	22/31 = 73.3%	67/93 = 72.0%
Total	27/62 = 43.67%	30/62 = 48.4%	34/62 = 54.8%	91/186 = 48.9%

contexts. Four subtests of the Wechsler Adult Intelligence Scale (WAIS-III; Wechsler, 1997) were administered to all listeners. Two of the subtests (digit symbol and digit search) assess speed of processing, and two of the subtests (digit span and letter-number sequencing) assess working memory.

E. Procedure

Preliminary assessment of all participants entailed an audiometric evaluation (including acoustic immittance measures), the cognitive screen, and the four subtests of the WAIS-III. The experimental measures of speech recognition in noise were administered following this preliminary assessment. The CDs with the word lists or sentence lists recorded on one channel and the 12-talker babble on the other channel were played back on a CD player (Tascam CD-200) and routed to an Interacoustics AC40 audiometer. The speech level was fixed at 85 dB sound pressure level, while the babble was presented at a signal-to-noise ratio (SNR) of +10 dB. This SNR was established following extensive pilot testing in which accented and unaccented stimuli were presented to younger listeners with normal hearing at a range of SNRs to identify an SNR that approximated 80% correct, to avoid floor and ceiling effects across groups and experimental conditions. During the experiment, the stimuli and babble were mixed and presented to a single insert earphone (ER-3A).

The listener's task was to repeat the word or sentence presented, and these verbal responses were recorded for later verification. During the experiment, the tester graded the responses. A second judge listened to and graded the recorded responses, and in cases of discrepant scoring a third judge served as a tie-breaker. All judges were aware of the correct target word. Target words, presented in isolation or as the final word in sentence contexts, were scored and were judged correct if all phonemes were repeated correctly.

There were 30 conditions, which were derived from combinations of three talkers × five lists (one-syllable, two-syllable, three-syllable, four-syllable, mixed) × 2 contexts (isolated words, words in sentences). The order of the presentation of isolated word conditions and words-in-sentences conditions was counterbalanced across participants. Subsequently, a Latin Squares design was used to determine the order of talkers presented to listeners. Following the determination of talker order, a Latin Squares design was used to determine the order of the lists for each talker. Other listening experiments using non-speech signals were conducted with these same listeners (not reported in this paper). These other experiments were conducted following the first set of speech conditions (either words alone or words in sentences) and preceding the second set of speech conditions to minimize learning. On average, there was a two-week separation between collection of data for the first set of speech conditions and the second set.

Prior to testing, participants were administered a practice list consisting of five words or words in sentences spoken by each of the three talkers (native English, mildly accented English, moderately accented English). These practice sentences were not used in the experiment.

Listeners were tested over the course of two or three sessions of 2 h/each. They were provided frequent breaks and were reimbursed for their participation. The procedures used in this project were approved by the University of Maryland Institutional Review Board for Human Subjects Research.

III. RESULTS

A. Analyses of the stimuli

Prior to conducting the statistical analyses of listener recognition performance, several analyses of the stimuli were conducted. Initially, the stress patterns of the test stimuli (two, three, and four syllables) produced in isolation by the unaccented and accented talkers were evaluated by two trained raters following procedures similar to those described by Flege and Bohn (1989). The agreement in stress pattern analyses between the two raters was 98.33%, 95.96%, and 92.3% for the two-, three-, and four-syllable words, respectively. As a result, the analysis of stress patterns by just one of the raters is presented here. The number and percent of deviations in stress patterns, relative to the unaccented talker (i.e., the error patterns), observed for the mildly and moderately accented talkers are shown in Table I. The number of deviations appears to increase with length of syllables from 2 to 4 and with the degree of accent. A related-samples Wilcoxon Signed Rank test rejected the null hypothesis that the median of differences between the mild and moderate talkers equals 0 ($p < 0.05$). Additionally, a related samples Friedman two-way analysis of variance (ANOVA) by ranks test confirmed that the effect of length of syllables was significant ($p < 0.05$).

Specific stress pattern deviations of the two accented talkers, relative to the unaccented talker, are shown in Tables II, III, and IV for the two-, three-, and four-syllable

TABLE II. Stressed (S) and unstressed (U) syllable patterns for two-syllable words by the native English talker (top row), with deviations from native English timing patterns entered in the table. The frequency of occurrence of the native stress patterns among the 31 words is shown in parentheses.

Native talker pattern (frequency)	SU (22)	US (8)	SS (1)	SU (22)	US (8)	SS (1)
Accented talker pattern	Mild accent			Moderate accent		
SU			1			
US						
SS		5		15	6	

TABLE III. Stressed (S) and unstressed (S) syllable patterns for three-syllable words by the native English talker (top row), with deviations from native English timing patterns by the mildly and moderately accented talkers entered in the table. The frequency of occurrence of the native stress patterns among the 31 words is shown in parentheses.

Native talker pattern (frequency)	SUU (22)	USU (6)	UUS (2)	SSU (1)	SUU (22)	USU (6)	UUS (2)	SSU (1)
Accented talker pattern	Mild accent				Moderate accent			
SUU			2	1				
USU								
SSU	3							
UUS								
SSS					11	1	1	1
SUS					6		1	
USS						3		

words, respectively. For each table, the stress patterns shown in the top row are those used by the native English talker (numbers in parentheses indicate number of occurrences of each stress pattern among the test words), the stress patterns shown in the first column are those used by the accented talkers, and the numeric entries are the frequency of occurrences of each error pattern by each accented talker. Tables II, III, and IV show that the deviant stress patterns of the mildly accented talker for two-, three-, and four-syllable words are confined to other stress patterns used by the native talker of English observed within this corpus of words. However, the deviant stress patterns of the moderately accented talker for three- and four-syllable words (Tables III and IV) are primarily those that were not produced by the native talker of English, some of which are not typical of English. For both talkers, the deviant patterns consistently reflected a change from an unstressed syllable produced by the native talker of English to a stressed syllable. It is also noteworthy that for two- and three-syllable words (Tables II and III), the most common deviant stress pattern observed for the moderately accented talker was characterized by all stressed syllables, perhaps reflecting the syllable-timed nature of Spanish.

The second analysis involved transcriptions of the accented stimuli to identify deviations from the unaccented model. One highly trained listener transcribed all of the words produced in isolation by the two Spanish-accented talkers. The transcriptions of this listener were consistent with the transcriptions of a second listener who evaluated a subset of the words. A summary of the findings are presented here. For both talkers, mild and moderate, the most common transcription errors (relative to the unaccented speaker) were vowel raising and absence of vowel reduction, reflecting non-temporal changes that are observed in Spanish-accented English (Flege and Bohn, 1989). In addition, the moderately accented talker had frequent instances of absence of aspiration for the initial voiceless plosive and absence of final consonant release.

The duration of the initial syllable (root word) in each of the 31 word sets for the three talkers was analyzed acoustically using Adobe Audition. These duration results are presented in Table V. A repeated measures ANOVA was conducted to determine the main effects of length of syllables, talker, and the interaction between these two effects. There was a main effect of syllable [$F(3, 270) = 356.15$, $p < 0.001$], talker [$F(2, 90) = 18.239$, $p < 0.001$], and an interaction between syllable and talker [$F(6, 270) = 6.77$, $p < 0.001$]. *Post hoc* analyses (paired t-tests with Bonferroni corrections) revealed that the duration of the initial root word decreased with an increase in length of syllables, except for the two accented talkers for the three- and four-syllable words ($p < 0.001$). This finding indicates that initial syllable duration patterns changed more dramatically with increasing length of syllables for the unaccented talker than

TABLE IV. Stressed (S) and unstressed (U) syllable patterns for four-syllable words by the native English talker (top row), with deviations from native English timing patterns by the mildly and moderately accented talkers entered in the table. The frequency of occurrence of the native stress patterns among the 31 words is shown in parentheses.

Native talker pattern (frequency)	USUU (12)	UUSU (9)	SUUU (5)	SUSU (4)	SSUU (1)	USUU (12)	UUSU (9)	SUUU (5)	SUSU (4)	SSUU (1)
Accented talker pattern	Mild accent					Moderate accent				
USUU			1							
SUSU		6	1				3	1		
UUSU										
SUUU										
SSUU	4					5				
UUSS							1		1	
USUS						2				
SUSS							2		2	
SSSS						1				
USSS						2				
SSUS						2				

TABLE V. Mean initial syllable duration (in sec) of one-, two-, three-, and four-syllable words produced by the three talkers. Numbers in parentheses are the standard errors of the mean.

Talker	1-syllable words	2-syllable words	3-syllable words	4-syllable words
Unaccented	0.44 (0.02)	0.27 (0.01)	0.24 (0.01)	0.22 (0.01)
Mild accent	0.51 (0.01)	0.32 (0.01)	0.27 (0.01)	0.26 (0.01)
Moderate accent	0.36 (0.02)	0.26 (0.01)	0.21 (0.01)	0.19 (0.01)
Average	0.44 (0.01)	0.28 (0.01)	0.24 (0.01)	0.22 (0.01)

for the accented talkers. The findings may also reflect the expected change of the initial syllable from stressed for monosyllabic words to unstressed for multisyllabic words in native English that was not observed as consistently in the multisyllabic word productions of the accented talkers.

B. Percent correct word recognition performance

Percent correct recognition performance of the three listener groups for unaccented and accented words of varying length of syllables (lists 1, 2, 3, and 4) presented in isolation and in sentence contexts is shown in Figs. 2 and 3, respectively. The three panels of each figure display the results separately for each talker. Performance patterns generally show improvements with an increasing number of syllables for the unaccented and mildly accented talkers, but a somewhat different pattern is seen for the moderately accented talker. In addition, age and hearing loss effects appear to become more prominent with increasing talker accent.

In order to clearly examine the separate effects of listener group, talker accent, and length of syllables, and interactions between them, ANOVAs were conducted using repeated measures, split-plot factorial designs. In this design, the between-subjects variable was group (three levels) and the within-subjects variables were talker (three levels) and list (four levels). Findings of significant group effects were then evaluated with *post hoc* analyses to determine the effects of age (comparison of young normal-hearing listeners to older normal-hearing listeners) and the effects of hearing loss (comparison of older normal-hearing listeners to older hearing-impaired listeners). This has become a common design and analysis technique in research aiming to distinguish the effects of age and hearing impairment (e.g., Tremblay *et al.*, 2003; Hopkins and Moore, 2011; John *et al.*, 2012; Gordon-Salant *et al.*, 2013). Separate ANOVAs were conducted on arc-sine transformations of the word and sentence recognition scores. Tests of the assumption of sphericity were conducted (see the Appendix for detailed results), which did not alter the findings presented below.

For recognition of stimulus words presented in isolation, the ANOVA revealed significant main effects of talker [$F(2,84) = 321.03$, $p < 0.001$], list [$F(3, 126) = 342.22$, $p < 0.001$], and group [$F(2,42) = 18.48$, $p < 0.001$]. There were also significant two-way interactions between talker and list [$F(6, 252) = 12.23$, $p < 0.001$] and between list and group [$F(6,126) = 4.63$, $p < 0.01$]. All other two- and three-way interactions were not significant ($p > 0.05$). The talker × list interaction was analyzed further by collapsing data across listener groups. Paired samples t-tests revealed that for the native talker and mildly accented talker, performance on list 1 was poorer than on all other lists, performance on list 2 was poorer than on lists 3 and 4, and performance on list 3 was poorer than on list 4 ($p < 0.008$, using the Bonferroni correction). For the moderately accented talker, a

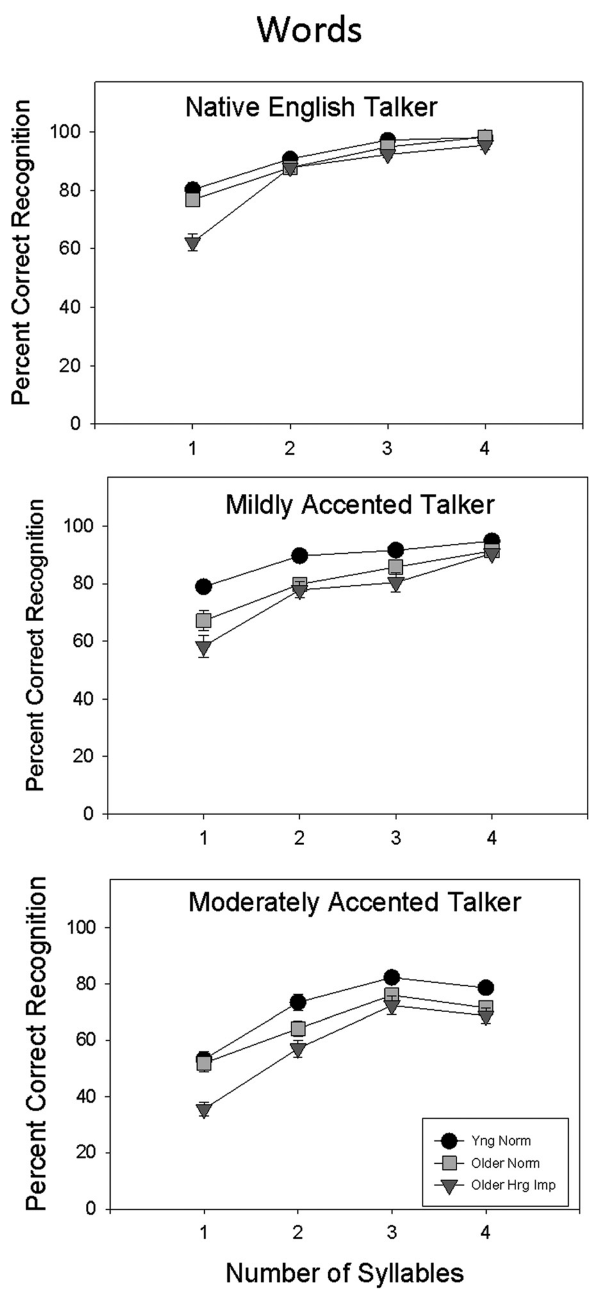

FIG. 2. Mean recognition of one-, two-, three-, and four-syllable words presented in isolation by three listeners groups. The three panels show data for stimuli recorded by three talkers varying in accent (panel a = no accent, panel b = mild accent, panel c = moderate accent).

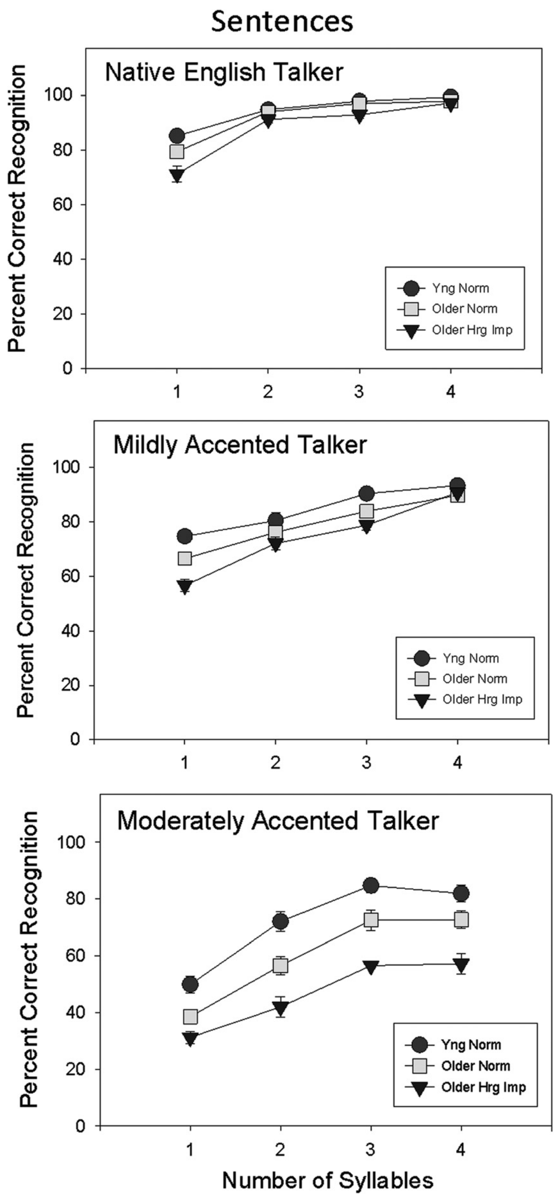

FIG. 3. Mean recognition of one-, two-, three-, and four-syllable words presented in sentence contexts by three listeners groups. The three panels show data for stimuli recorded by three talkers varying in accent (panel a = no accent, panel b = mild accent, panel c = moderate accent).

similar pattern was observed for lists 1 and 2 relative to the other lists (i.e., performance on list 1 < lists 2, 3, and 4, and list 2 < lists 3 and 4), but a somewhat different pattern was observed for lists 3 and 4 in which recognition of words on list 4 was poorer than recognition of words on list 3. Thus, listener performance improved progressively with increasing length of syllables for the native and mildly accented talkers, but listener performance declined with increasing syllables beyond three syllables for the moderately accented talker.

The list × group interaction was also analyzed further for each talker. For the native English talker, the effect of listener group was significant for list 1 only, with younger and older listeners with normal hearing showing better performance than the older hearing-impaired listeners ($p < 0.01$). For the mildly accented talker, group effects varied somewhat across lists: the young listeners with normal hearing showed better performance than older hearing-impaired listeners on all lists, but showed higher recognition scores than the older normal-hearing listeners for list 2 only ($p < 0.008$), with no significant difference in performance between the two older groups ($p > 0.05$). For the moderately accented talker, the group effect varied by list. For list 1 (one-syllable words), both the younger and older listeners with normal hearing performed better than the older hearing-impaired listeners ($p < 0.016$). For lists 2 and 4 (two- and four-syllable words, respectively), younger listeners performed better than older hearing-impaired listeners only ($p < 0.016$). However, there were no significant group differences for list 3 ($p < 0.016$).

The pattern of results was generally similar for recognition of words in sentences. The omnibus ANOVA revealed significant main effects of talker [$F(2,84) = 394.57$, $p < 0.001$]; list [$F(3, 126) = 399.41$, $p < 0.001$], and group [$F(2,42) = 22.05$, $p < 0.001$], significant two-way interactions between talker and group [$F(4,84) = 5.84$, $p < 0.001$] and talker and list [$F(6,252) = 7.71$, $p < 0.001$], and a significant three-way interaction between talker, list, and group [$F(12, 252) = 2.42$, $p < 0.01$]. Since performance generally declined for all groups with increasing accent, subsequent analyses of the three-way interaction focused on the effects of group and list separately for each talker. Analyses of results for the native English talker showed that for all listener groups, performance improved significantly and progressively with an increasing number of syllables (list $1 < 2 < 3 < 4$; $p < 0.001$, all comparisons). Multiple comparison testing with Bonferroni analysis showed that younger and older listeners with normal hearing exhibited significantly higher recognition scores than the older hearing-impaired group. For the mildly accented talker, effects of group were significant for lists 1 and 3, but not 2 and 4. For list 1, younger listeners performed better than the two older groups, and the older normal-hearing listeners performed better than the older hearing-impaired listeners. However, for list 3, the younger listeners performed better than the older hearing-impaired listeners only. The pattern of performance across the different lists varied with listener group when listening to the mildly accented talker. For all three groups, performance was poorer for list 1 than lists 2, 3, and 4. Additionally, both younger and older listeners with normal hearing showed poorer performance on list 2 than lists 3 and 4, but equivalent performance on lists 3 and 4. Older hearing-impaired listeners, however, exhibited poorer performance on lists 2 and 3 than on list 4. For the moderately accented talker, group effects varied with each list. Young normal-hearing listeners performed better than the older hearing-impaired listeners for all four lists. In addition, young normal-hearing listeners performed better than the older normal-hearing listeners for lists 1, 2, and 3. Finally, older normal-hearing listeners performed better than older hearing-impaired listeners for lists 2, 3, and 4. Thus, age and hearing loss effects emerged, but varied depending upon length of syllables. The list effect was consistent across listener groups: all groups showed lower accuracy on list 1 than on lists 2, 3, and 4, as well as lower accuracy in recognition of list 2 than list 4, with no differences between lists 3 and 4. In summary, the results of these analyses for recognition of words in isolation and in sentences suggest that recognition performance does not improve progressively with an increasing number of syllables (from 1 to 4) for words spoken with a moderate accent, and that age differences emerge with a more pronounced accent and for words with increasing length of syllables.

Visual comparison of Figs. 2 and 3 suggests that presenting words in isolation vs in sentences had no effect on recognition of stimuli produced by the native-English and mildly accented talkers, but appears to be quite different for stimuli produced by the moderately accented talker. The term "sentence context effect" is used here to refer to the effect of perceiving a word in a sentence vs perceiving the same word in isolation. The sentence context effect was analyzed for the moderately accented speaker only, because it appeared to be minimal for the other two talkers. These results are shown in Fig. 4. The young normal hearing listeners showed no significant effect of sentence context for any of the lists. The older listeners with normal hearing, however, showed significant effects of sentence context for lists 1 and 2, but not lists 3 and 4. Finally, the older hearing-impaired listeners showed significant effects of sentence context for lists 2, 3, and 4. In cases where the effect of sentence context was significant, recognition of words presented in sentences was consistently poorer than recognition of words presented in isolation.

One question of interest in this investigation was whether or not listeners experience more difficulty when presented with words on a list that varied in number of syllables compared to lists with words of fixed length of syllables. To evaluate this question, percent correct performance of the listeners was averaged across the four lists with a fixed number of syllables (average performance on lists 1, 2, 3, and 4) and compared to performance for the mixed list (list 5) for the words presented in isolation. These results are presented in Fig. 5. An ANOVA was conducted with one grouping variable and two within-subjects variables ("mix" and talker). Results revealed significant main effects of talker [$F(2,84) = 337.71$, $p < 0.001$] and group [$F(2,42) = 14.43$, $p < 0.001$], and a significant interaction between talker and mix [$F(2, 84) = 24.35$, $p < 0.001$]. Other main effects and interaction effects were not significant. Subsequent analyses (t-tests with Bonferroni corrections) examined the effect of mix separately for each listener group. Results revealed that the effect of fixed vs mixed length of syllables within a stimulus list (i.e., the effect of mix) was significant for the younger listeners when listening to the moderately accented talker and for the older listeners with normal hearing when listening to the two accented talkers. Recognition scores were higher for both groups when listening to lists of words that were fixed in length of syllables compared to the mixed lists for the moderately accented talker, but the opposite effect was observed for the older normal-hearing listeners when listening to the mildly accented talker. The effect of fixed vs mixed length of syllables for the older hearing-impaired listeners was not significant. In summary, the effects of increasing the cognitive demand of the task by comparing recognition for words presented in isolation and

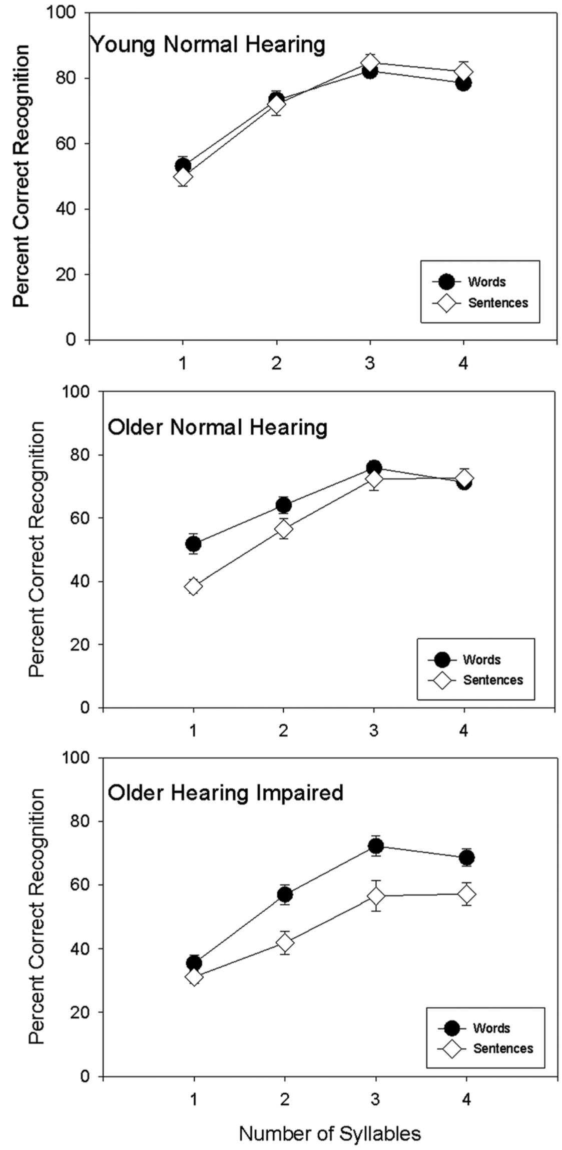

FIG. 4. Recognition of moderately accented words presented in isolation vs sentence contexts as a function of length of syllables by three listener groups.

in sentence contexts showed that older listeners had greater difficulty than younger listeners in sentence contexts. For the cognitive demand of trial-by-trial variation, listeners with normal hearing experienced greater difficulty with the mixed list than the fixed lists. This variable did not have an effect on performance of the older hearing-impaired listeners. Both of these cognitive effects were only seen for accented stimuli.

C. Multiple regression of participant predictor variables

Multiple regression analyses were conducted to determine which among a set of predictor variables accounted for the most variance in recognition of unaccented and accented words of varying length of syllables. The set of predictor variables included a high-frequency pure-tone average (HFPTA: mean of thresholds at 1 k, 2 k, 4 k Hz), age, and performance on two measures of speed of processing (digit

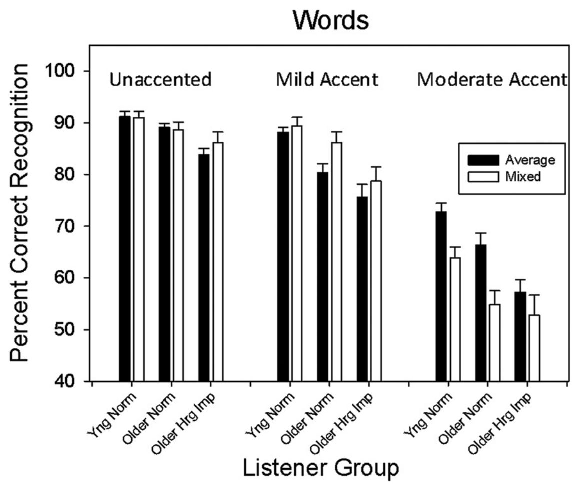

FIG. 5. Recognition of unaccented, mildly accented, and moderately accented words of varying length of syllables, averaged across conditions with fixed length of syllables ("average") vs a condition with trial-to-trial variation in length of syllables ("mixed") by three listener groups.

symbol score, symbol search) and two measures of working memory (digit span, letter-number-sequencing). The dependent variables, in separate analyses, were recognition scores for one-, two-, three-, and four-syllable words spoken by each of the three talkers in isolation and in sentence contexts. The results of the fully step-wise multiple regression analyses for words presented in isolation and in sentences are shown in Tables VI and VII, respectively. HFPTA was the single predictor variable retrieved in the analyses for one- and three-syllable words spoken in isolation, regardless of talker accent, as well as for two-syllable words spoken by the mildly and moderately accented talkers. However, a measure of processing speed (digit symbol) was identified as the single variable that accounted for significant variance for the four-syllable words spoken by the unaccented and moderately accented talkers, and age was the variable retrieved in the analysis for four-syllable words spoken by the mildly accented talker. The pattern of results is somewhat different for words spoken in sentence contexts (Table VII). While HFPTA was identified as the single or the most important variable accounting for performance scores in many conditions, a measure of working memory (either digit span score or letter-number-sequencing score) emerged as the single or

TABLE VI. Results of multiple regression analyses for unaccented, mildly accented, and moderately accented words of varying length of syllables presented in isolation. Significant predictor variables ($p < 0.01$) retrieved in the analysis with associated R^2 values are shown. (HFPTA = high-frequency pure-tone average; LNS = letter-number-sequence subtest of the WAIS-III, dig. symbol = digit symbol subtest of the WAIS-III.)

	Unaccented		Mild accent		Moderate accent	
	Variable	R^2	Variable	R^2	Variable	R^2
1-syllable	HFPTA	0.453	HFPTA	0.355	HFPTA	0.338
2-syllables	Dig. symbol	0.114	HFPTA	0.310	HFPTA	0.287
3-syllables	HFPTA	0.233	HFPTA	0.316	HFPTA	0.219
4-syllables	Dig. symbol	0.167	Age	0.260	Dig. symbol	0.208
Mix	HFPTA	0.119	HFPTA	0.327	Age	0.197
			LNS	0.424		

TABLE VII. Results of multiple regression analyses for unaccented, mildly accented, and moderately accented words of varying length of syllables presented in sentence contexts. Significant predictor variables ($p < 0.01$) retrieved in the analysis with associated R^2 values are shown. (HFPTA = high-frequency pure-tone average; LNS = letter-number-sequence subtest of the WAIS-III, dig. span = digit span subtest of the WAIS-III.)

	Unaccented		Mild accent		Moderate accent	
	variable	R^2	variable	R^2	variable	R^2
1-syllable	HFPTA	0.436	HFPTA	0.519	HFPTA	0.408
	Dig. span	0.435				
2-syllables	HFPTA	0.168	Dig. span	0.167	HFPTA	0.480
	Dig. span	0.258				
3-syllables	HFPTA	0.257	HFPTA	0.265	HFPTA	0.399
					LNS	0.457
4-syllables	Dig. span	0.172	Dig. span	0.136	HFPTA	0.447
Mix	HFPTA	0.316	HFPTA	0.315	HFPTA	0.511
			Dig. span	0.422	age	0.556

as one significant variable in a number of other conditions. To summarize, the step-wise multiple regression analyses suggest that, in addition to high-frequency hearing sensitivity accounting for a significant proportion of variance in word recognition, processing speed accounts for a significant proportion of variance for multisyllabic word recognition in isolation, but working memory accounts for a significant proportion of variance for multisyllabic word recognition in sentence contexts.

IV. DISCUSSION

A. Effect of talker accent on recognition of words with increasing length of syllables

The first objective of this experiment was to determine if accent affects recognition of multisyllabic words more than monosyllabic words. The expectation was that the non-native talkers' syllable-timing patterns deviate from English more as the number of syllables in the word increases. The findings of this study partially support this hypothesis. For the unaccented talker, the perceptual results showed that listeners' recognition performance improved incrementally and significantly with an increasing number of syllables. This pattern of performance was observed consistently for both words in isolation and words in sentence contexts. For the mildly accented talker, different performance patterns emerged depending on the presentation of the word in isolation or in a sentence context. For words in isolation, performance improved with increased numbers of syllables in the word stimuli. In sentence contexts, performance patterns were different across listeners groups, particularly for multisyllabic words. For the moderately accented talker, there was no improvement in performance for the four-syllable words compared to three-syllable words. Thus, while listeners' performance was better for accented multisyllabic words compared to monosyllabic words, their performance did not improve progressively and consistently with increasing length of syllables for accented words, unlike the performance pattern observed for unaccented words.

There are three possible explanations for the current findings. The first explanation derives from the lexical properties of the test words, including word frequency and lexical neighborhood density. Word frequency refers to the frequency of occurrence of the word in the language, and thus, recognition should be better for higher word frequency than lower word frequency. Lexical neighborhood density is derived from the number of lexical competitors, i.e., the number of words that differ from a given word by a single phoneme, either through addition, deletion, or substitution (Luce and Pisoni, 1998). As a result, speech recognition performance is expected to be better for words that come from sparse neighborhoods compared to words that come from dense neighborhoods. As noted in Sec. II, efforts were made to balance the word frequency of the one-, two-, three-, and four-syllable words to the extent possible. The average word frequencies for the four lists were 52.35, 53.16, 35.03, and 48.48 for the one-, two-, three-, and four-syllable words, respectively. Given the equivalence in word frequency between one- and two-syllable words, it appears that the large increments in performance from one- to two-syllable words could not be accounted for by word frequency. Analyses were also conducted for word densities of the monosyllabic and multisyllabic words using the eLexicon Project (Balota *et al.,* 2007). On average, the lexical density (number of lexical neighbors) of the one-, two-, three-, and four-syllable words was 23.09, 2.69, 0.85, and 0.16, respectively. Based on the large increments in performance between one- and two-syllable words, and more modest increments between two-, three-, and four-syllable words, it appears that neighborhood density was highly influential in predicting the pattern of results, at least for the unaccented talker. Listeners' performance, however, for the accented talkers is not predicted exclusively by lexical neighborhood density, because performance improvement was not observed for four-syllable words compared to two- and three-syllable words.

A second account of the pattern of performance is based on acoustic changes in word or syllable duration with increasing length of syllables and accent. The working hypothesis was that with an increase in the number of syllables, the duration of each syllable decreases. Thus, the duration of the initial syllable or common initial word in each stimulus set (e.g., *cap* in cap, captain, capital, and capacity) would be expected to decrease in duration with increments in the number of syllables in the target word. It was also expected that the initial common word duration would be shorter for accented talkers than the unaccented talker, based on previous reports comparing word duration of unaccented and Spanish-accented talkers (Shah, 2004). The acoustic analyses of the initial syllable duration demonstrated the expected decrease in initial syllable duration for the unaccented talker. For the two accented talkers, initial syllable duration decreased with increasing number of syllables in the word up to three syllables, with no further changes from three- to four-syllable words. Additionally, initial syllable duration for the moderately accented talker was shorter than that for the unaccented and mildly accented talkers for words comprised of one, three, or four syllables. In contrast, initial syllable duration was consistently longer for the mildly accented talker than for the unaccented talker, which may have contributed to the observation that listeners did not experience considerable difficulty understanding this talker.

A final explanation relates to the changes in stress patterns of the accented talkers compared to the unaccented talker, with increasing length of syllables comprising the words. Table I and the related analyses clearly show that the accented talkers' deviations in stress patterns increased with increasing length of syllables. Additionally, the moderately accented talker had considerably more deviant stress patterns compared to the mildly accented talker. Thus, the decline in listener performance with increasing length of syllables and accent, observed for all groups, may have been driven by the deviant stress patterns in three- and four-syllable words, especially as observed for the moderately accented talker. The changes in stress patterns observed for the two accented talkers most likely produced differences in vowel duration between the native English and accented English talkers. However, cues for stress are acoustically multidimensional. As the transcription analysis showed, the accented talkers' speech productions were also characterized by changes in vowel height and vowel reduction, relative to the native English talker. Thus, alterations in formant frequencies to cue vowel identity also may have contributed to the listeners' difficulty in accurately resolving accented speech.

Taking these factors together, word recognition performance as a function of length of syllables and accent appears to be most readily accounted for in part by lexical factors (neighborhood density) and in part by acoustic modifications to syllable duration that co-vary with length of syllables and accent. Changes in accented talkers' stress patterns with increasing length of syllables, relative to the native English talkers' stress patterns, also appear to play an important role in recognition of accented multisyllabic words. Further investigation is needed to elucidate which one of the multiple acoustic correlates of stress (i.e., frequency, intensity, and duration) is the predominant factor contributing to performance declines with accented multisyllabic words.

B. Effects of listener variables: Age, hearing sensitivity, and cognitive abilities

A second goal of this investigation was to determine if older listeners exhibit greater difficulty recognizing multisyllabic words spoken with an accent compared to younger listeners. The assumption was that multisyllabic words spoken by accented talkers would be characterized by reduced syllable duration relative to native English because of the combined effects of reduced syllable duration associated with multisyllabic words and altered syllable stress in Spanish-accented English (Shah, 2004). Because older listeners experience deficits in central auditory temporal processing for brief stimuli in sequences (Fitzgibbons and Gordon-Salant, 2004, 2011; Humes *et al.,* 2010), it was hypothesized that these temporal processing deficits by older listeners would emerge for recognition of accented multisyllabic words (comprised of relatively brief syllables) compared to either accented monosyllabic words or unaccented multisyllabic words. It was also predicted that both normal-hearing and hearing-impaired older listeners would show performance deficits for accented multisyllabic words, because age and not hearing loss effects are observed on measures of temporal discrimination in sequences (Fitzgibbons and Gordon-Salant, 2004, 2011). However, hearing loss effects were expected for unaccented and accented monosyllabic words, based on reduced spectral resolution of hearing-impaired listeners that limits accurate processing of consonants in stimuli presented at audible signal levels (Dubno *et al.,* 1982; Alexander and Kleunder, 2009).

The results are generally consistent with these predictions. Age-related performance differences were observed for multisyllabic words spoken by the accented talkers. Results for both words in isolation and words in sentence contexts revealed an absence of age effects for words of any length of syllables spoken by the unaccented talker, suggesting that the stress and timing patterns of native English could be used by the older listeners to support word understanding, consistent with previous findings by Wingfield *et al.* (2000). Differences in performance between younger and older normal-hearing listeners were also not observed for monosyllabic words spoken in isolation by the accented talkers. Indeed, the only significant age effects for words spoken in isolation were observed for two-syllable words spoken by the mildly accented talker. Age effects were quite prominent, however, for monosyllabic and multisyllabic words spoken in sentence contexts by the moderately accented talker. The general picture emerging is that the deviations with stress and timing associated with Spanish-accented English have a more deleterious effect on recognition of multisyllabic words by older listeners than younger listeners. This could be associated in part with observed age-related declines in neural synchrony (Anderson *et al.,* 2012), which limit the ability to precisely code rapid signal onsets in the central auditory pathways. Poor recognition of accented multisyllabic words may also be associated with difficulty controlling attention with disruptions and deviations in timing (Sommers, 1997; Hedden *et al.,* 2011) or adapting to modifications in expected (native English) stress patterns. Thus, it is possible that age-related declines in central processing of words with altered stress and increased number of syllables, as well as age-related decrease in attentional control for unexpected syllable stress and number, combine to limit the older listener's ability to accurately resolve accented multisyllabic words. Although listeners, both younger and older, benefit from reduced lexical competition in multisyllabic words, the ability to access this lexical information appears to be compromised when words are spoken with a moderate accent, especially for older listeners.

Older hearing-impaired listeners exhibited significantly poorer recognition of monosyllabic words than the two normal-hearing groups for monosyllabic words spoken in isolation, regardless of talker. Hearing loss effects were also observed for words spoken in sentence contexts, but this hearing loss effect was not confined to monosyllabic words. Rather, hearing-impaired listeners performed more poorly than the two normal-hearing groups for all lists spoken by the unaccented talker. For the mildly accented talker, older hearing-impaired listeners performed more poorly than older normal-hearing listeners for monosyllabic words only, but for the moderately accented talker, older hearing-impaired listeners performed more poorly than older normal hearing listeners for all multisyllabic words but not for monosyllabic words. It appears that producing a word in a sentence context compounds the effect of hearing loss among the older listeners, with words in sentence contexts contributing to greater performance deficits of older hearing-impaired listeners for multisyllabic words with accent. Cognitive decline is known to occur with the aging process, particularly in the domains of speed of information processing, selective attention, and working memory capacity (Salthouse, 1996; McDowd and Shaw, 2000; Zacks *et al.,* 2000). Recent evidence suggests that there are close links between cognitive decline and speech perception (Gates *et al.,* 1996; Humes *et al.,* 1994; Humes, 2007). To investigate the extent to which age-related factors (including cognitive decline) could explain recognition accuracy for unaccented and accented monosyllabic and multisyllabic words, multiple regression analyses were conducted. Results of the 30 separate analyses showed that high-frequency pure-tone average (HFPTA) was most often the only significant predictor of performance, consistent with prior findings obtained with unaccented speech (Humes, 2007). The variance accounted for by HFPTA, when identified as a significant predictor variable, ranged from 0.119 to 0.519. However, cognitive variables were also identified as significant predictor variables in some of the analyses, and the specific variable identified was strikingly different for words presented in isolation vs words presented in sentence contexts. Of the 15 analyses conducted for words presented in isolation, digit symbol was identified as the only significant predictor variable for three of the analyses and age was identified as the only significant variable for two analyses. The digit symbol subtest of the WAIS-III was identified as a significant variable for two of the three analyses conducted for four-syllable words, suggesting that speed of information processing is important for recognizing these relatively brief syllables in multisyllabic words. For words presented in sentence contexts, performance on the digital span subtest contributed significantly to performance in six of the 15 analyses, either as the single significant variable identified or as a second significant variable identified. Because the digit span subtest assesses working memory and was identified as a significant variable primarily for recognition of multisyllabic words, these findings suggest that recognition of multisyllabic words in the context of a sentence imposes a heavy demand on working memory. Moreover, performance on this working memory subtest was identified for recognition of unaccented and mildly accented speech, revealing that the added memory component for recognizing words in sentences is not limited to listening to accented talkers.

C. Effects of stimulus variability and context

The third and final goal of this investigation was to determine if recognition of accented words is influenced by increased cognitive demands, including trial-by-trial variability and sentence context. Listeners generally experience more difficulty perceiving stimuli that vary acoustically from trial to trial than stimuli that remain constant on a given acoustic dimension (Sommers and Barcroft, 2006). Moreover, older listeners exhibit poorer performance in conditions featuring trial-by-trial variation compared to younger listeners especially for variations in speech amplitude and presentation rate (Sommers, 1997). It was hypothesized that the effects of listener age on word recognition would be more prominent in conditions in which the number of syllables in the target word varied from trial-to-trial compared to an average of performance across fixed syllable conditions, especially for accented speech. Analysis of performance in the variable syllable ("mix") condition compared to the fixed syllable conditions revealed that both younger and older listeners with normal hearing performed more poorly in the mixed syllable condition than averaged performance in the fixed syllable conditions for the moderately accented talker only. This finding suggests that the predictability of the number of syllables in spoken word lists influences recognition accuracy for somewhat pronounced Spanish-accented English but not native English. Thus, speech stimuli spoken by a moderately accented talker that vary on the dimension of number of syllables apparently require listeners to adapt flexibly to this variability, thereby taxing attentional capabilities. The performance of the older hearing-impaired listeners did not show an effect of trial-by-trial variation for recognition of accented words, perhaps because these listeners' relatively poor recognition performance for the moderately accented talker obscured the effect of stimulus variability. Overall, the current findings may be another reflection of the impact of alterations in length of syllables and stress patterns that occur in accented English and which are difficult for listeners to resolve.

Another key objective was to compare recognition performance for isolated words vs words in sentence contexts to determine the impact of an accented or unaccented carrier phrase on a target word comprised of a varying number of syllables. At least one prior investigation showed no significant impact of sentence context on recognition of monosyllabic words recorded by mildly and moderately accented talkers whose native language was Spanish (Gordon-Salant *et al.,* 2010a). In the current experiment, however, it was assumed that accented talkers would revert to syllable timing for multisyllabic words when produced in the context of a sentence, possibly making these words more difficult to resolve in a context that is also characterized by deviant timing relative to native English. To examine this question, recognition scores were measured for words of varying length of syllables presented in isolation and in sentence contexts. Effects of sentence context were observed in the recognition scores of the two older groups, but not the younger group, for the moderately accented talker. Sentence context effects were observed for one- and two-syllable words for the older normal-hearing listeners and for all multisyllabic words for the older hearing-impaired listeners. Recognition of words in sentence contexts was poorer than recognition of words presented in isolation, whenever stimulus context effects were observed. It is possible that the sentence context for multisyllabic words created an added memory load, which had a greater impact on recognition by older listeners than younger listeners. This implication is supported by the results of the multiple regression analyses shown in Table VII. However, added memory demand is not the sole source of the observed sentence context effect, because the older listeners did not show a sentence context effect for unaccented speech. Rather, it appears that the combined effects of signal distortion associated with accent (i.e., altered stress and timing, as well as spectral alterations in vowels), coupled with placement of a stimulus word in a sentence context, which increases the memory and cognitive load, are particularly sensitive to age effects. As predicted, older hearing-impaired listeners showed this context effect for multisyllabic accented words, which are shown here to have greater alterations in stress and timing than monosyllabic words. However, older normal-hearing listeners showed the sentence context effect for one- and two-syllable words. The source of the different stimulus context effects for older normal and hearing-impaired individuals is not known.

V. SUMMARY AND CONCLUSIONS

This investigation of recognition for unaccented and Spanish-accented English words showed that expected increments in recognition scores with increasing number of syllables in a word are reduced by talker accent. Age-related differences in performance were observed for accented multisyllabic words, but not unaccented multisyllabic words, suggesting that deviations in stress patterns with accent are difficult for older listeners to resolve. These deviations in stress patterns include changes with timing that are likely difficult for older listeners to perceive, but also appear to include altered spectral characteristics. Finally, sentence context effects were observed for older, but not younger, listeners, in which recognition of accented words of varying length of syllables in sentence contexts was poorer than recognition of these same words presented in isolation. The interplay between stress and timing alterations in accented speech, age-related temporal processing deficits, and cognitive demand is offered as an interpretation of these results. The current findings suggest that altered word stress patterns in accented speech limit recognition performance, particularly for older listeners. These supra-segmental characteristics of accented English are likely to further diminish speech recognition beyond that attributed to deviations in isolated speech segments.

ACKNOWLEDGMENTS

This research project was supported by Grant No. R37 AG09191 from the National Institute on Aging. We are grateful to Rebecca Bieber, Alison Romano, and Niki Razeghi for their assistance in conducting the acoustic analyses of the stimuli.

APPENDIX

Mauchly's Test of Sphericity was conducted to evaluate the assumption of equality in variances of the differences between all pairs of groups (levels) of the repeated measures variables. Mauchly's Test of Sphericity was not significant for the main effects of talker and list for the omnibus analysis of word recognition in isolation [talker effect, Mauchly's $W(2) = 0.979$, $\chi^2 = 0.885$, $p > 0.05$; list effect, Mauchly's $W(5) = 0.763$, $\chi^2 = 10.99$, $p > 0.05$]. However, for the omnibus analysis of recognition of words in sentences, Mauchly's Test of Sphericity was significant for the talker effect [Mauchly's $W(2) = 0.751$, $\chi^2 = 11.74$, $p < 0.01$], but not for the list effect [Mauchly's $W(5) = 0.863$, $\chi^2 = 6.0$, $p > 0.05$]. In subsequent analyses, tests of within-subjects effects were evaluated with sphericity assumed, except in the assessment of the effect of talker for words presented in sentences. The Greenhouse–Geisser epsilon value was used to adjust the degrees of freedom in this one analysis.

[1]The stimuli are available from the authors upon request.

Alexander, J. M., and Kleunder, K. R. (**2009**). "Spectral tilt change in stop consonant perception by listeners with hearing impairment," J. Speech Lang. Hear. Res. **52**, 653–670.

Anderson, S., Parbery-Clark, A., White-Schwoch, W., and Kraus, N. (**2012**). "Aging affects neural precision of speech encoding," J. Neurosci. **32**, 14156–14164.

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), pp. 1–55.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchson, K. A., Kessler, B., Loftis, B., Neely, J. H., Delson, D. L., Smpson, G. B., and Treiman, R. (**2007**). "The English lexicon project," Behav. Res. Methods **39**, 445–459.

Beckman, M. E. (**1986**). *Stress and Non-Stress Accent (Netherlands Phonetic Archives No. 7)* (Foris, Dordecht, The Netherlands) [second printing, 1992 (Walter de Gruyter, Berlin)], pp. 1–239.

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

Dubno, J. R., Dirks, D. D., and Langhofer, L. R. (**1982**). "Evaluation of a hearing impaired listeners using a nonsense syllable test: II Syllable recognition and consonant confusion patterns," J. Speech Hear. Res. **25**, 141–148.

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. J., and Gordon-Salant, S. (**2004**). "Age effects on discrimination of timing in auditory sequences," J. Acoust. Soc. Am. **116**, 1126–1134.

Fitzgibbons, P. J., and Gordon-Salant, S. (**2011**). "Effects of interval repetition in tonal sequences on temporal discrimination by younger and older listeners," J. Acoust. Soc. Am. **129**, 1490–1500.

Flege, J. E., and Bohn, O.-S. (**1989**). "An instrumental study of vowel reduction and stress placement in Spanish-accented English," Stud. Second Lang. Acquis. **11**, 35–62.

Fletcher, H., and Steinberg, J. C. (**1929**). "Articulation testing methods," Bell Syst. Tech. J. **8**, 806–854.

Gates, G. A., Cobb, J. L., Linn, R. T., Rees, T., Wolf, P. A., and D'Agostino, R. B. (**1996**). "Central auditory dysfunction, cognitive dysfunction, and dementia in older people," Arch. Otolaryngol. Head Neck Surg. **122**, 161–167.

Gordon-Salant, S., and Fitzgibbons, P. J. (**2004**). "Effects of stimulus and noise rate variability on speech perception by younger and older adults," J. Acoust. Soc. Am. **115**, 1808–1817.

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., 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.

Harris, M. S., and Umeda, N. (**1974**). "Effect of speaking mode on temporal factors in speech: Vowel duration," J. Acoust. Soc. Am. **56**, 1016–1018.

Hedden, T., Van Dijk, K. R. A., Shire, E. H., Sperling, R. A., Johnson, K. A., and Buckner, R. L. (**2011**). "Failure to modulate attentional control in advanced aging linked to white matter pathology," Cerebral Cortex **22**, 1038–1051.

Hopkins, K., and Moore, B. C. (**2011**). "The effects of age and cochlear hearing loss on temporal fine structure sensitivity, frequency selectivity, and speech reception in noise," J. Acoust. Soc. Am. **130**, 334–349.

Humes, L. E. (**2007**). "The contributions of audibility and cognitive factors to the benefit provided by amplified speech to older adults," J. Am. Acad. Audiol. **18**, 590–603.

Humes, L. E., Kewley-Port, D., Fogarty, D., and Kinney, D. (**2010**). "Measures of hearing threshold and temporal processing across the adult lifespan," Hear. Res. **264**, 30–40.

Humes, L. E., Watson, B. U., Christensen, L. A., Cokely, C. G., Halling, D. C., and Lee, L. (**1994**). "Factors associated with individual differences in clinical measures of speech recognition among the elderly," J. Speech Hear. Res. **37**, 465–474.

John, A. B., Hall, J. W. III, and Kreisman, B. M. (**2012**). "Effects of advancing age and hearing loss on gaps-in-noise test performance," Am. J. Audiol. **21**, 242–250.

Kalikow, D. N., Stevens, K. N., and Elliott, L. L. (**1977**). "Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability," J. Acoust. Soc. Am. **61**, 1337–1351.

Kucera, H., and Francis, W. N. (**1967**). *Computational Analysis of Present-Day American English* (Brown University Press, Providence, RI), pp. 1–424.

Lehiste, I. (**1970**). *Suprasegmentals* (MIT Press, Cambridge, MA), pp. 1–194.

Luce, P. A., and Pisoni, D. B. (**1998**). "Recognizing spoken words: The Neighborhood Activation Model," Ear Hear. **19**, 1–36.

McDowd, J. M., and Shaw, R. J. (**2000**). "Attention and aging: A functional perspective," in *The Handbook of Aging and Cognition*, 2nd ed., edited by F. I. M. Craik and T. A. Salthouse (Lawrence Erlbaum and Associates, Mahwah, NJ), pp. 221–292.

Peng, L., and Ann, J. (**2001**). "Stress and duration in three varieties of English," World Englishes **20**, 1–27.

Pfeiffer, E. (**1977**). "A short portable mental status questionnaire for the assessment of organic brain deficit in elderly patients," J. Am. Geriatr. Soc. **23**, 443–441.

Pike, K. L. (**1945**). *The Intonation of American English* (University of Michigan Press, Ann Arbor, MI), pp. 1–203.

Pons, F., and Bosch, L. (**2010**). "Stress pattern preference in Spanish-learning infants: The role of syllable weight," Infancy **15**, 223–245.

Salthouse, T. A. (**1996**). "The processing-speed theory of adult age differences in cognition," Psychol. Rev. **103**, 403–428.

Shah, A. (**2004**). "Production and perceptual correlates of Spanish-accented English," in *Proceedings of the MIT Conference: From Sound to Sense: 50+ Years of Discoveries in Speech Communication*, MIT, Cambridge, MA, pp. C-79–C-84.

Shin, H. B., and Kominski, R. A. (**2010**). "Language use in the United States: 2007," American Community Survey Report No. ACS-127, U.S. Census Bureau, Washington, D.C.

Sommers, M. S. (**1997**). "Stimulus variability and spoken word recognition. II. The effects of age and hearing impairment," J. Acoust. Soc. Am. **101**, 2278–2288.

Sommers, M. S., and Barcroft, J. (**2006**). "Stimulus variability and the phonetic relevance hypothesis: Effects of variability in speaking style, fundamental frequency, and speaking rate on spoken word identification," J. Acoust. Soc. Am. **119**, 2406–2416.

Thorndike, E. L., and Lorge, I. (**1944**). *The Teacher's Wordbook of 30,000 Words* (Bureau of Publications, Teacher's College, Columbia University, NY), pp. 1–211.

Tremblay, K., Piskosz, M., and Souza, P. E. (**2003**). "Effects of age and age-related hearing loss on the neural representation of speech cues," Clin. Neurophysiol. **114**, 1332–1343.
Wechsler, D. (**1997**). *Wechsler Adult Intelligence Scale*, 3rd ed. (WAIS-III) (Pearson Assessment, San Antonio, TX).
Wenk, B. J. (**1985**). "Speech rhythms in second language acquisition," Lang. Speech **28**, 157–175.
Wingfield, A., Lindfield, K. C., and Goodglass, H. (**2000**). "Effects of age and hearing sensitivity on the use of prosodic information in spoken word recognition," J. Speech Lang. Hear. Res. **43**, 915–925.
Zacks, R. T., Hasher, L., and Li, K. Z. H. (**2000**). "Human memory," in *The Handbook of Aging and Cognition*, edited by F. I. M. Craik and T. A. Salthouse (Lawrence Erlbaum Associates, Mahwah, NJ), pp. 293–357.