

Effects of listener age and native language on perception of accented and unaccented sentences

Rebecca E. Bieber, Grace H. Yeni-Komshian, Maya S. Freund, Peter J. Fitzgibbons, and Sandra Gordon-Salant^{a)}

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

(Received 4 June 2018; revised 6 November 2018; accepted 13 November 2018; published online 7 December 2018)

Degradations to auditory input have deleterious effects on speech recognition performance, especially by older listeners. Alterations to timing information in speech, such as occurs in rapid or foreign-accented speech, can be particularly difficult for older people to resolve. It is currently unclear how prior language experience modulates performance with temporally altered sentence-length speech utterances. The principal hypothesis is that prior experience with a foreign language affords an advantage for recognition of accented English when the talker and listener share the same native language, which may minimize age-related differences in performance with temporally altered speech. A secondary hypothesis is that native language experience with a syllable-timed language (Spanish) is advantageous for recognizing rapid English speech. Native speakers of English and Spanish completed speech recognition tasks with both accented and unaccented English sentences presented in various degrees of time compression (TC). Native English listeners showed higher or equivalent recognition of accented and unaccented English speech compared to native Spanish listeners in all TC conditions. Additionally, significant effects of aging were seen for native Spanish listeners on all tasks. Overall, the results did not support the hypotheses for a benefit of shared language experience for non-native speakers of English, particularly older native Spanish listeners. © 2018 Acoustical Society of America. <https://doi.org/10.1121/1.5081711>

[SHF]

Pages: 3191–3200

I. INTRODUCTION

One of the most common complaints of older listeners is difficulty understanding speech in challenging listening situations. These can take a variety of forms such as presence of background noise, a rapid rate of speech, or speech that is produced with a foreign accent. Older listeners are known to be particularly vulnerable to these challenges (Dubno *et al.*, 1984; Wingfield *et al.*, 1985; Ferguson *et al.*, 2010; Gordon-Salant *et al.*, 2010a,b), which may occur in isolation or in combination during daily listening.

Older listeners appear to be susceptible to difficulties in recognizing speech that has been temporally altered. For example, when listening to time-compressed speech, older listeners often show a disproportionate decrease in recognition of time-compressed sentences compared to younger listeners (Wingfield *et al.*, 1985; Gordon-Salant and Fitzgibbons, 2001). This age-related decrease in recognition of time-compressed speech may be exacerbated by additional acoustic degradations or reductions in supportive context (Wingfield *et al.*, 2006).

Another type of speech deviation that is becoming increasingly relevant for older listeners is recognition of foreign-accented speech. Accented English speech is characterized by numerous variations relative to native English, among which are alterations in stress and timing patterns at the supra-segmental level (Pike, 1945). Given that older people exhibit deficits in discriminating changes in timing

within stimulus sequences (Fitzgibbons and Gordon-Salant, 2001, 2010), it is expected that they would have difficulty recognizing foreign-accented speech that is characterized by alterations in stress and timing. Indeed, significant age effects have been documented for recognition of accented multi-syllabic words (Gordon-Salant *et al.*, 2015) and sentences (Gordon-Salant *et al.*, 2013).

It is also important to consider potential effects of listener and talker language background for these types of sentence recognition tasks. In this paper, *listeners* whose native language is English are referred to as L1 listeners, and listeners whose second language is English are referred to as L2 listeners. Similarly, *talkers* whose native language is English are referred to as unaccented talkers, and talkers whose native language is not English are referred to as accented talkers. There is evidence that listener language background may influence performance on speech recognition tasks, especially when the speech is presented in a noise background. Specifically, L2 listeners perform more poorly than L1 listeners when listening to unaccented speech in background noise, even when these differences are not seen when listening in quiet (Mayo *et al.*, 1997). Mayo and colleagues (1997) compared speech recognition performance in noise of monolingual English listeners with that of native Spanish listeners who acquired English at various stages in life (since infancy to post-puberty). All bilingual listeners required more favorable signal-to-noise ratios to achieve a criterion performance level compared to the monolingual listeners. Effects of native language for listening to other forms of degraded speech, such as time-compressed speech, have not been fully examined.

^{a)}Electronic mail: sgsalant@umd.edu

While L2 listeners may be at a disadvantage when listening to speech produced by unaccented talkers, some investigators have suggested that L2 listeners may receive a benefit when listening to English produced by an accented talker who shares a language background with the L2 listener. This idea has been termed the interlanguage speech intelligibility benefit (ISIB), and can be conceived of in various configurations (Bent and Bradlow, 2003; Hayes-Harb *et al.*, 2008; Xie and Fowler, 2013). Higher performance by L2 listeners compared to L1 listeners, when listening to accented speech, is known as the interlanguage speech intelligibility benefit for listeners (ISIB-L). On the other hand, when L2 listeners perform better for accented speech compared to unaccented speech, an interlanguage speech intelligibility benefit for talkers (ISIB-T) is observed. Partial support for the ISIB-T was reported by Bent and Bradlow (2003) in a study in which L2 listeners exhibited comparable recognition scores for accented and unaccented speech. These authors also proposed a “mismatched” ISIB, where a benefit for L2 listeners is observed even when the native language of the accented talker does not match that of the L2 listener. These ISIB effects have been documented for L2 listeners in quiet (Bent and Bradlow, 2003; Xie and Fowler, 2013) as well as for conditions with additional degradation (Brouwer *et al.*, 2012; Peng and Wang, 2016).

It is possible that this ISIB effect may minimize the difficulties that L2 listeners experience in recognizing accented English speech, including under conditions with additional acoustic distortions. For example, Peng and Wang (2016) found that on a composite speech comprehension task presented in reverberation and background noise, listeners who shared a language background with the accented talker were less affected by talker accent than those who did not share the native language. The effect of a possible ISIB on another form of speech degradation, rapid speech [as implemented by time compression (TC)], has not been examined to date.

Many of these prior investigations included participants who were younger and had normal peripheral hearing sensitivity. As noted above, there are well-documented deleterious effects of aging on speech perception tasks, particularly for speech that has been altered through the presence of background noise (Dubno *et al.*, 1984), foreign accent (Ferguson *et al.*, 2010; Gordon-Salant *et al.*, 2010a,b), or TC (Wingfield *et al.*, 1985; Gordon-Salant and Fitzgibbons, 2001). However, it is less clear how prior language experience modulates these aging effects because there have been no previous examinations of age effects for the ISIB-T and ISIB-L for sentence-length stimuli presented in acoustically degraded conditions. In the present study, therefore, both age effects and prior language experience effects are examined on one type of speech degradation: TC of speech.

A previous study investigating monosyllabic word recognition by younger and older L1 and L2 normal-hearing listeners (Gordon-Salant *et al.*, 2018) failed to find clear evidence of an ISIB effect. L1 listeners consistently showed higher word recognition scores for the accented talkers than did the L2 listeners. These findings were generally in contrast to the previous reports of ISIB-L and ISIB-T cited above. However, it is possible that the recent findings may

be unique to testing with monosyllabic stimuli. For sentence-level recognition tasks, listeners have access to different information, including semantic contextual cues and supra-segmental prosodic cues over the duration of the utterance. Access to these cues may be beneficial for listeners, particularly when the stimuli are produced with a foreign accent. Conversely, it is also possible that sentence-level stimuli may prove more challenging, as there is an inherent increase in working memory demands in a sentence recognition task compared to a word recognition task. Sentence-level stimuli also contain supra-segmental acoustic cues (i.e., prosody), which may be modified by the presence of a foreign-accent.

The principal objective of the present study is to examine the effects of listener native language and listener age on recognition of English sentence-length utterances as spoken by talkers of different native languages. Listeners were presented with sentences recorded by unaccented and accented talkers at the original speech rate and three speeded speech rates. The goal was to determine if native language experience minimizes expected age-related declines for recognition of accented speech presented at normal and fast rates, when the listener shares the same native language as the foreign-accented talker. According to Bent and Bradlow (2003), L2 listeners may have had greater exposure to accented speech than L1 listeners, affording them greater flexibility in listening to accented speech. Previous findings suggest that L2 listeners will perform as well as or better when listening to accented compared to unaccented speech, and will show similar or higher speech perception performance than L1 listeners for accented speech (i.e., Xie and Fowler, 2013). It was also expected that age effects will be reduced for the L2 listeners compared to the L1 listeners, assuming that the older L2 listeners could benefit from the ISIB-L. Additionally, the findings of Derwing and Munro (2001) might suggest that L2 listeners will not be as strongly affected as L1 listeners by modest degrees of TC of accented speech. These authors reported that native Mandarin Chinese listeners rated speech intelligibility of non-native English talkers higher when the accented speech was presented at faster-than-normal rates.

Previous work has shown that a number of cognitive abilities affect perception of degraded speech by older listeners. Principal among them is working memory capacity, which is related to perception of speech in noise (Akeroyd, 2008; Füllgrabe *et al.*, 2015), as well as to perception of time-compressed speech (Vaughan *et al.*, 2006) by older listeners. The cognitive domains of attention and processing speed are also related to speech recognition in noise (Zekveld *et al.*, 2011) and recognition of time-compressed speech by older listeners (Gordon-Salant and Fitzgibbons, 2001). Finally, younger and older listeners' knowledge of the language has been shown to relate to speech perception performance, particularly under degraded conditions (Zekveld *et al.*, 2011; Janse and Adank, 2012). In the present study, working memory, short-term memory, speed of processing, selective attention, and inhibitory control were assessed in addition to vocabulary knowledge. It was expected that speed of processing and, to a lesser extent,

working memory capacity would contribute to recognition of time-compressed speech by listeners of both native languages and ages (Vaughan *et al.*, 2006). Language proficiency, as measured by knowledge of English vocabulary, was also expected to contribute to speech perception performance (Tamati *et al.*, 2013).

II. METHOD

A. Participants

Data were collected from two groups of native Spanish-speaking participants: younger [L2 younger participants with normal hearing (YNH), ages 19–33 yr] and older [L2 older participants with normal hearing (ONH), ages 60–80 yr] and two groups of native English listeners: younger (L1 YNH, ages 20–27 yr) and older (L1 ONH, ages 65–76 yr). All participant groups comprised 15 listeners.

All listeners had normal hearing [pure-tone thresholds ≤ 20 dB hearing level (HL) re: ANSI, 2010, between 250 and 4000 Hz] as shown in Fig. 1. Participants also exhibited good-to-excellent monosyllabic word recognition scores and normal middle ear function as determined by tympanometric measures. Additional subject selection criteria included at least a high school education either in the U.S. or the country of origin and normal performance on a screening test of cognitive function [Mini-Mental Status Questionnaire (MMSE); Folstein *et al.*, 1975]. A Spanish version of this test (MMSE2) was administered to the L2 listeners. A passing score of 24 or higher was required for participation.

For L2 listeners, additional qualification criteria included Spanish as a native language, country of origin in a Spanish-speaking country, and at least 6 months of residency in the United States, with an age of arrival (AOA) of at least 12 years. All efforts were made to recruit listeners whose country of origin was in Latin America. In the younger L2 group, six listeners had a country of origin of Spain, while the remaining nine were from Latin American countries. In the older L2 group, all listeners had countries of origin

within Latin America. The mean years of residency for the L2 YNH and L2 ONH participants was 4.57 yr (standard deviation, s.d., 1.29) and 30.38 yr (s.d. 4.02), respectively. The mean AOA for the L2 YNH and L2 ONH participants was 20.27 yr (s.d. 1.2) and 37.94 yr (s.d. 3.72), respectively.

B. Stimuli

Stimuli included 200 low-predictability (LP) sentences from the revised speech perception in noise (R-SPIN) test (Bilger *et al.*, 1984). These stimuli are constructed to include one target word in the word-final position of a sentence containing no supportive semantic context (i.e., “Mr. White discussed the cruise”). There were 8 lists of 25 sentences/each. These low-predictability speech perception in noise (LP-SPIN) sentences were recorded by three male talkers: one unaccented talker and two accented talkers whose native language was Spanish. The unaccented materials were taken from the original speech perception in noise (SPIN) recordings, which were created in Cambridge, MA (Kalikow *et al.*, 1977; talker’s age not reported). The accented talkers were from Peru and El Salvador (ages 39 and 37 yr, respectively), and had achieved graduate to post-graduate levels of education at the time of recording.

The accented LP-SPIN sentences were rated for degree of accentedness and perceived speech intelligibility in two separate pilot studies. In the first pilot study, accentedness ratings were obtained from 19 young L1 listeners with normal hearing. Listeners were presented with five sentences per talker and asked to rate the degree of accentedness on a scale of 1–9, where 1 = no accent and 9 = extremely heavy accent, following the procedures described by Atagi and Bent (2011). Both talkers were rated in the “moderate” accent range (ratings of 5 and 6). In the second pilot study, a small cohort ($n = 5$) of young, normal hearing, L1 listeners provided ratings of perceived intelligibility of the two accented talkers. The listeners were presented with each of 200 LP-SPIN sentences for both accented talkers and asked to rate each sentence on a scale of 0%–100% intelligibility. Intelligibility ratings (by list) ranged from 84.80% to 89.20% for one accented talker, and 78.40% to 84.00% for the other. Following this pilot, it was decided that only the stimuli from the more intelligible talker would be used for this experiment. This talker had an average rating of accentedness of 5.74, and his country of origin was Peru.

The final LP-SPIN sentences recorded by the unaccented and accented talkers were time compressed using Praat (Boersma and Weenink, 2018) to the following TC ratios: 0%TC, 30%TC, 40%TC, and 50%TC. These TC ratios were selected based on pilot testing with young, normal hearing, L1 listeners to avoid floor effects. To summarize, each of the eight LP-SPIN lists was recorded by two talkers (L1,L2) and altered by four degrees of TC (0%,30%,40%,50%TC). These LP-SPIN sentences were equated in root-mean-square (rms) level and burned onto a compact disc (CD) together with a calibration tone, equal in rms (root-mean-squared) level to that of the sentences.

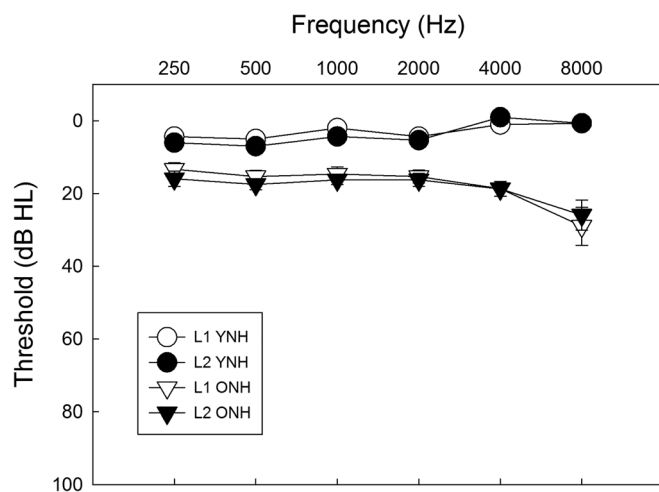


FIG. 1. Mean audiograms (with standard errors) of the four listener groups. Filled symbols represent L2 listeners and open symbols represent L1 listeners. YNH = younger participants with normal hearing; ONH = older participants with normal hearing.

C. Cognitive measures

All participants completed a number of cognitive measures to examine the relationships between cognitive ability and performance on the speech tasks. These measures included the forward and backward digit span from the Wechsler Adult Intelligence Scale (WAIS)-III (verbal short-term memory; [Wechsler, 1997](#)), listening span, an auditory version of the reading span test (working memory; [Daneman and Carpenter, 1980](#)), the symbol search and the digit symbol coding subtests of the WAIS-III (speed of processing; [Wechsler, 1997](#)), and the Flanker test (attention and inhibitory control; [Eriksen and Eriksen, 1974](#)). The Flanker test was downloaded from the National Institutes of Health (NIH) Cognitive Toolbox ([Weintraub et al., 2013](#)) for this study. All participants were also tested on a measure of knowledge of English vocabulary: the Receptive One-Word Picture Vocabulary Test (ROW-PVT; English version, [Brownell, 2012](#)).

D. Procedure

Prior to completing the speech experiments, all participants underwent an audiological evaluation and acoustic immittance testing. They also completed a language history questionnaire and a cognitive screening (MMSE). Those who met subject selection criteria completed the cognitive measures and receptive vocabulary test detailed above.

Listeners were seated in a sound-attenuating booth during testing. Stimuli were routed from a Tascam CD player (model CDRW-402, Montebello, CA) to a Crown D-75 amplifier (Elkhart, IL), Hewlett-Packard 350 D attenuator (Palo Alto, CA), Colbourn audio-mixer amplifier (model S82-24, Whitehall, PA), and delivered to a single EAR 3-A insert earphone (Etymotic Research, Elk Grove Village, IL). Stimulus presentation level was 85 dB sound pressure level (SPL) with speech presented in quiet. This presentation level was chosen in order to facilitate comparison with prior studies testing older listeners with TC speech (e.g., [Gordon-Salant and Fitzgibbons, 1993](#); [Gordon-Salant and Friedman, 2011](#)). Because all listeners had normal hearing, no perceptual differences in audibility were expected between the four listener groups.

There were eight listening conditions comprised of four TC ratios \times two talkers; each listener heard all eight listening conditions. The order of TC conditions was randomized with conditions additionally blocked by talker (i.e., half of the participants heard all four unaccented conditions first, and then all four accented conditions, and vice versa). There was random assignment of SPIN sentence list to listening condition with no repetition of a sentence list for a given listener. Following sentence presentation, listeners were asked to repeat the sentence out loud and write down the final word they heard. Participants' responses were recorded for verification, although written answers were considered the primary response. Following the experiment, L2 listeners were queried regarding their familiarity with the keywords used in the experiment. Each stimulus word was presented in written form, and participants were asked to indicate whether or not they had heard the word before (yes/no binary).

Results of the word familiarity ratings indicated that both L2 listener groups had heard over 80% of the words used in the experiment before.

E. Scoring and data analysis

Cognitive tests were scored as follows. Digit-symbol coding and symbol search (subtests of the WAIS), were scored based on total number of items correct. The digits forward and backward were scored based on number of trials (i.e., repetitions) completed correctly. The Flanker task was scored based on accuracy and reaction times (RTs); both the age-adjusted scale score and mean RT were recording for analysis. For the ROW-PVT, a raw score based on number of correct items was used for analysis.

Each LP-SPIN sentence was scored as correct or incorrect, based on recognition of the final keyword. The statistical model of speech recognition performance was derived with a generalized linear mixed effects regression analysis that included the following fixed effects and their interactions: listener native language (two levels: L1, L2; referenced to L1), talker native language (two levels: accented, unaccented; referenced to unaccented), age group (two levels: YNH, ONH; referenced to YNH), and TC ratio (continuous variable; reference is 0% TC). The model also contained the random effects of the participant and item intercepts.

Following the creation of the initial model, an additional model was built to evaluate the contribution of the various cognitive measures together with the fixed effects described above. The additional predictor variables evaluated were ROW-PVT English vocabulary score, the symbol search score, the digit symbol score, Flanker score, Flanker RT, digit-span forward, digit-span backward, and high-frequency pure-tone average (HFPTA; average of hearing thresholds at 1 kHz, 2 kHz, and 4 kHz). Summary statistics and group differences for these measures for each listener group can be found in Table II. Following a preliminary correlational analysis, one score from a representative measure from each cognitive domain was used as a predictor variable, in addition to high frequency pure-tone average. This reduction in predictor variables was completed to minimize collinearity among predictor variables representing similar domains. The final set of predictor variables evaluated included HFPTA, ROW-PVT, symbol search, Flanker RT, and digit-span forward.

III. RESULTS

A. Preliminary and cognitive measures

Performance of the four listener groups was compared on the preliminary measure of HFPTA, the language measure of ROW-PVT, and the cognitive measures of digit symbol coding, symbol search, digits-forward, digits-backward, Flanker (both raw score and RT), and L-SPAN. The purpose of these analyses was to examine the structure of the data to facilitate interpretation of the mixed effects regression modeling described below. Scores on these cognitive measures were compared across listener groups using two-way analyses of variance (ANOVAs). Results of the two-way

TABLE I. Age, HFPTA, and cognitive predictor summary statistics for all listener groups; mean (*s.d.*). All predictors were examined for group differences with two-way ANOVA; significance levels are reported; NS, nonsignificant.

Listener group	Age (yr)	HFPTA (dB HL)	Digit symbol coding (raw)	Symbol search (raw)	DigitsF (raw)	DigitsB (raw)	Flanker (age-adjusted scale score)	Flanker RT (ms)	ROW-PVT (raw)
L1 YNH	21.93 (1.79)	2.56 (2.08)	91.07 (18.17)	43.47 (9.63)	11.50 (2.61)	6.50 (2.32)	110.94 (9.27)	563.12 (117.29)	168.13 (10.11)
L1 ONH	69.13 (3.36)	15.11 (5.06)	65.40 (14.34)	30.33 (6.51)	10.82 (1.47)	6.64 (2.66)	103.94 (12.85)	916.73 (256.55)	180.27 (7.64)
L2 YNH	24.88 (3.96)	4.67 (2.90)	62.8 (21.58)	30.4 (12.52)	8.8 (2.35)	6.1 (2.18)	102.92 (12.56)	711.77 (230.42)	155.67 (10.17)
L2 ONH	68.53 (5.83)	17.44 (5.52)	39.60 (19.25)	18.67 (8.04)	7.92 (1.61)	4.00 (1.63)	87.91 (15.25)	1684 (976.89)	119.60 (28.64)
Effect of age group	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$	NS	NS	$p < 0.01$	$p < 0.001$	$p < 0.001$
Effect of language	NS	NS	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.05$	$p < 0.001$	$p < 0.001$	$p < 0.001$
Age \times language	NS	NS	NS	NS	NS	NS	NS	$p < 0.05$	$p < 0.001$

ANOVAs were examined using the Tukey honest significant difference (HSD) function with correction for multiple comparisons in *R*; results are summarized in Table I. As shown in Table I, there were significant effects of listener age group for age, HFPTA, digit symbol coding, symbol search, both Flanker measures, and ROW-PVT. There were significant effects of listener language group for digit symbol coding, symbol search, digits forward and backward, both Flanker measures, and the ROW-PVT. In addition, there were significant interactions for the measures of Flanker RT and ROW-PVT.

B. Linear mixed effects regression analysis

Recognition scores for the unaccented and accented sentences across the four TC ratio conditions by all listener groups are shown in Fig. 2. The final generalized linear mixed effects regression analysis included the between-subjects fixed effects of listener language (referenced to the native English listeners) and age group (referenced to the young listeners), and the within-subjects fixed effects of talker language (referenced to the native English talker) and TC ratio (referenced to 0% TC).

Random effects evaluated in the model were the by-participant intercept and the by-item intercept. Results are

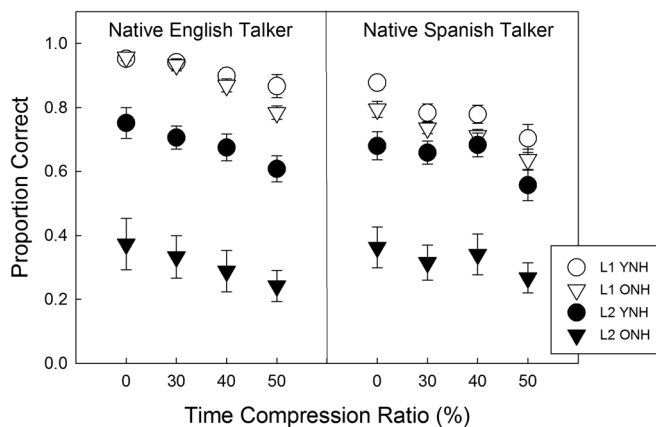


FIG. 2. Mean speech recognition scores for the sentence-final LP-SPIN words in four TC conditions, spoken by the unaccented (left) and Spanish-accented (right) talkers. Scores for all four listener groups are presented with error bars representing one standard error. Filled symbols represent L2 listeners and open symbols represent L1 listeners. YNH = younger participants with normal hearing; ONH = older participants with normal hearing.

shown in Table II. Significant fixed effects were listener language ($p < 0.001$), talker language ($p < 0.001$), and TC ratio ($p < 0.001$). Listener language was involved in significant two-way interactions with talker language ($p < 0.001$), TC ($p < 0.01$), and age group ($p < 0.05$). None of the higher level interactions of fixed-effects (i.e., three-way, four-way) were significant. Inclusion of both participant and item intercepts significantly increased the amount of variance explained by the model when compared to inclusion of either intercept alone ($p < 0.001$).

The contribution of cognitive, hearing, and vocabulary predictor variables was also examined. Model testing proceeded with comparison of models with and without each predictor variable; those variables that did not significantly improve model fit were removed. The predictors that significantly contributed to the fit of the model included ROW-PVT (vocabulary measure) scores and symbol search (speed of processing measure) scores. Inclusion of an interaction term for the ROW-PVT scores did not significantly improve model fit, but an interaction term for the symbol search scores did improve the model fit. Significant fixed effects involving the cognitive predictor variables included a main effect of ROW-PVT score ($p < 0.001$), and a main effect of symbol search score ($p < 0.01$). Symbol search was involved in significant two-way interactions with age group ($p < 0.001$) and TC ratio ($p < 0.01$).

Post hoc analyses of the two-way interactions of listener language \times talker language and listener language \times age

TABLE II. Results of generalized linear mixed effects model showing significant effects.

Effect	Estimate	Standard error	<i>z</i>	<i>p</i> -level
Listener language	-1.431	0.307	-4.664	<0.001
Talker language	-1.332	0.121	-11.021	<0.001
TC	-2.842	0.367	-7.734	<0.001
Listener language \times Talker language	1.073	0.15	7.174	<0.001
Listener language \times TC	1.384	0.442	3.133	<0.01
Listener language \times age group	-1.061	0.478	-2.222	<0.05
ROW-PVT	0.914	0.132	6.913	<0.001
Symbol search	0.0352	0.136	2.583	<0.01
Symbol search \times TC	0.555	0.178	3.113	<0.01
Symbol search \times age group	0.771	0.225	-3.428	<0.001

group were evaluated with the Tukey method using the function *lsmeans* for pairwise comparisons in *R*. The *post hoc* comparisons of the listener language by talker language interaction showed that the L1 listeners achieved higher performance for the unaccented talker compared to the accented talker ($z = 10.10, p < 0.001$), but there was no effect of talker accent for the L2 listeners ($p = 0.96$). The latter finding is somewhat supportive of the predictions of ISIB-T. Additionally, there was no overall performance difference between L1 and L2 listeners for the accented talker ($p = 0.84$), while L1 listeners performed higher than L2 listeners for the unaccented talker ($z = 5.904, p < 0.001$). The finding that L2 listeners did not achieve higher recognition scores for the accented talker than L1 listeners is not consistent with the predictions of ISIB-L.

The age group \times listener language interaction was driven by a significant age effect for the L2 listeners ($z = 3.193, p < 0.01$), in which younger L2 listeners performed better than the older L2 listeners. There was no age effect for L1 listeners ($p = 0.63$). Additionally, an effect of listener native language was significant for the older listeners ($z = 2.962, p < 0.05$), but not for the younger listeners ($p = 0.43$). Specifically, older L1 listeners recognized all of the speech materials better than older L2 listeners. The two-way interaction between listener language and TC arose from the differences in slope estimates for TC ratio for L1 listeners ($\beta = -2.842$) and L2 listeners ($\beta = -1.374$), reflecting a steeper decline in performance with increasing TC ratio for the L1 than the L2 listeners, regardless of talker language. Thus, L2 listeners were less sensitive to increments in TC ratio than L1 listeners.

The two-way interactions involving symbol search scores were also examined. Analysis of the symbol search by age group interaction indicated that there was a contribution of processing speed to speech recognition performance for younger listeners ($z = 0.196, p < 0.01$) but not for older listeners ($p = 0.44$). This finding suggests that younger listeners relied on processing speed ability to recognize time-compressed speech, but older listeners did not. Symbol search scores also interacted significantly with TC ratio ($z = 3.113, p < 0.01$), indicating that greater processing speed capacity benefitted listeners more as TC ratio increased.

C. Correlational analysis

As noted above, the regression model results indicated that English vocabulary score was a significant predictive variable for accented and unaccented speech recognition performance in various time-compressed conditions by all listener groups. In order to interpret the predictive nature of the English vocabulary score on speech recognition scores in the TC conditions, additional analyses examined the correlations between vocabulary score and speech recognition performance separately for the L1 and L2 listeners in the easiest and most difficult TC conditions (0% and 50% TC, respectively) for both the unaccented and accented talkers. Bonferroni corrections to adjust the significance level with multiple analyses were applied. The analyses showed that none of the correlations between English vocabulary score

and speech recognition score were significant for the L1 listeners. However, for the L2 listeners, all of the correlations were significant: unaccented talker at 0% TC ($r = 0.79, p < 0.001$), unaccented talker at 50% TC ($r = 0.67, p < 0.001$), accented talker at 0% TC ($r = 0.74, p < 0.001$), and accented talker at 50% TC ($r = 0.60, p < 0.001$). These correlations indicate that for L2 listeners, poorer knowledge of English vocabulary was highly related to poorer sentence recognition scores in conditions involving unaccented and accented talkers, with and without TC. It should be noted that the older L2 listeners had much higher variance in their ROW-PVT scores than the other listener groups (see Table I), suggesting that the importance of vocabulary as a predictor variable for these listeners may be related to the degree of variance. An additional question examined is whether or not L2 participants with a high English vocabulary score were any different from those with a low English vocabulary score on the magnitude of the ISIB-T. To that end, the L2 listeners were divided into high and low English vocabulary groups (those above the mean and those below the mean on the ROW-PVT) and their ISIB-T scores (accented score - unaccented score in each TC condition) were calculated. Independent samples *t*-tests were conducted to determine if there were significant differences in the ISIB-T for each of these two L2 listener groups. None of the analyses revealed a significant difference in the ISIB-T ($p > 0.05$, each analysis). That is, all difference scores were minimal, regardless of vocabulary knowledge, with average difference scores ranging from -3% to $+3\%$ across the four TC conditions (standard errors 2.6%), further confirming that the ISIB-T was minimal for this group of listeners.

IV. DISCUSSION

This experiment examined the effects of listener native language background and age on recognition of low-context sentences modified by talker accent, speech rate, or both. The principal hypothesis was that native speakers of Spanish (the L2 listeners) would have a benefit when listening to English produced by native Spanish talkers (the accented talkers), such that age-related differences often observed with native English (L1) listeners for accented speech and fast speech would be minimized. In general, the results do not support this hypothesis. A second hypothesis was that the cognitive domains of speed of processing and memory would contribute most to the variance in recognition scores for time-compressed speech regardless of talker native language, listener native language, and listener age. Only partial support of the predictions was found; younger listeners from both language groups appeared to benefit from greater processing speed capacity but not older listeners. For all participants, the findings indicate that the most important variable contributing to the variance in speech recognition scores for all participants combined was knowledge of English vocabulary. Further examination of these results is presented below.

A. Effect of listener and talker native language

The purpose of the experiment was to assess any possible advantageous effects of shared native language

background when L2 listeners are presented with accented speech that has been altered by increasing speech rate via TC. The significant interaction between listener language and talker language revealed by the linear mixed effects regression analysis indicated that L1 listeners outperformed L2 listeners for unaccented English, as expected, but that L1 and L2 listeners performed similarly for accented English sentences, regardless of the degree of TC. These findings suggest that the L2 listeners did not derive the expected ISIB-L (defined as L2 listeners recognizing accented speech better than L1 listeners), in contrast to findings described previously by other investigators (i.e., Imai *et al.*, 2005; Hayes-Harb *et al.*, 2008; Xie and Fowler, 2013).

One possible reason for the discrepant findings between the current study and those of the earlier studies is that the stimuli were quite different across studies: the previous studies presented word stimuli to listeners while the current study presented sentences. It is likely that for word stimuli, L2 listeners could easily resolve the phonological characteristics of their native language that carried over to foreign-accented English words, contributing to better recognition of accented speech by L2 listeners than was attained by L1 listeners. The use of sentences with minimal contextual cues in the current study was probably more challenging for the L2 listeners, who likely expended considerable processing resources to understand each word individually within the sentence. Any possible benefit of shared native language between these L2 listeners and the accented talkers may have been overwhelmed by the processing resources required of the sentence recognition task. A study by Major *et al.* (2002) showed that L1 listeners achieved higher listening comprehension scores than L2 listeners for passages spoken by accented talkers, providing partial support for the idea that listening to longer passages (sentences, running text) may limit a potential advantage afforded by a shared native foreign language for L2 listeners as evidenced by an ISIB-L.

A second possible reason for discrepant results across studies is that the English proficiency of the accented talker was different in the current study compared to those used in prior studies. One previous investigation suggested that the ISIB-L may be modulated by the L2 proficiency of the talker and listener (Hayes-Harb *et al.*, 2008). Hayes-Harb *et al.* (2008) found an ISIB-L for native Mandarin listeners (compared to native English listeners) in recognizing the English speech of native Mandarin talkers when both talkers and L2 listeners were rated as having low phonological proficiency in English. However, no such benefit was found with the speech of native Mandarin talkers who were rated as having high phonological English proficiency. Similarly, Xie and Fowler (2013) argue that the ISIB-L may be modulated by the language proficiency of both the talker and listener. In the current study, the accented talker was rated as having a moderate accent (rating of 5.74 on a scale of 1–9), and was highly intelligible to native English listeners, as noted previously. Listeners' proficiency in English was not strictly controlled for despite all listeners having spent at least six months in the U.S. at the time of testing. Thus, the discrepant findings between the current study and previous studies

could possibly be associated with the degree of accentedness or intelligibility of the accented talkers employed in the different studies, or with differences in English language proficiency of the L2 listeners. It should additionally be noted that, while the L2 listeners and the accented talker had a shared native language (Spanish), there may have been dialectal differences between the talker's and the listeners' native languages resulting from their differences in country of origin. It is possible that these differences in dialect precluded some benefit of shared native language.

The significant interaction between listener language and talker language was also evaluated for evidence of the ISIB-T. The ISIB-T predicts that L2 listeners recognize accented speech better than unaccented speech (Hayes-Harb *et al.*, 2008). In the current investigation, the L2 listeners did not show a significant difference in recognition of unaccented and accented sentences.

However, it is noted that L1 listeners did show a difference in performance between the accented and unaccented conditions: recognition of accented speech was significantly poorer than unaccented speech. These patterns thus contrast between the two listener language groups, suggesting that L2 listeners were not negatively impacted by the introduction of a foreign accent, unlike the L1 listeners. Thus, while the L2 listeners in this study did not show a significant advantage in sharing the native language of the accented talker, they showed the same recognition performance regardless of the native language of the talker. These findings may be construed as weak support for the ISIB-T predictions.

Taken together, the findings indicate that the predictions based on the interlanguage speech intelligibility benefit for either talkers or listeners were not strongly supported in this study. This phenomenon appears to be dependent on a variety of factors specific to the native language of the listener, native language of the talker, and the proficiency of both talkers and listeners, as demonstrated by the findings of Munro *et al.* (2006). The lack of a three-way interaction between listener language, talker language, and age implies that the absence of clear support for ISIB predictions may be extended to older listeners, who were not tested in any of the previous studies.

B. Effect of listener age

Both younger and older listener groups were included in the investigation in order to examine the effect of age on recognition of speech that was altered by talker accent and TC. The initial hypothesis proposed that prior experience with an accented talker's native language would afford an advantage for younger and older L2 listeners. This advantage was expected to minimize the typical age-related gap in recognition performance under acoustically degraded speech conditions. The findings did not support this hypothesis. Rather, the results indicated that there were significant and consistent effects of age for the L2 listeners, but not for the L1 listeners, an interaction which appears to be driven at least in part by English vocabulary knowledge. While all older L2 listeners were full-time permanent residents of the U.S. and

had lived in the U.S. for an average of 30.4 yr, differences in English vocabulary knowledge still showed a strong influence on the accuracy of their performance on the speech perception tasks described in this study. The overall poor pattern of performance observed for the older L2 listeners, regardless of talker native language and TC condition, further indicates that these listeners were severely and negatively impacted by the combined effects of aging and listening to speech in their L2.

C. Effect of TC

Sentences recorded by the unaccented and accented talkers were presented to listeners in various degrees of TC in order to examine the hypothesis that L2 listeners would not be as strongly affected as L1 listeners by temporal distortion of accented speech. This hypothesis was based on earlier findings by Peng and Wang (2016), showing that L2 listeners were less affected by noise and reverberation (another form of temporal alteration) on a speech comprehension task when there was a match between talker accent and listener accent. Specifically, a three-way interaction between listener language, talker language, and TC was predicted, in which the L2 listeners would show an advantage for time-compressed accented speech but not unaccented speech. This three-way interaction was not observed. Rather, a significant two-way interaction was observed between listener language and TC ratio. Thus, talker language did not affect the pattern of performance in different TC ratio conditions.

The two-way interaction between listener language and TC ratio revealed that the decline in recognition performance with increasing TC was more pronounced for L1 listeners than for L2 listeners. It is unclear why this pattern of performance was observed. Nevertheless, both groups exhibited an overall decline in speech recognition performance with increasing TC, even though the rate of decline differed between listener language groups. Additionally, the average level of recognition of time-compressed speech was much higher for the L1 listeners compared to the L2 listeners and was higher for the younger L2 listeners compared to the older L2 listeners. These patterns of performance did not vary with the talker's native language.

The present findings are somewhat in contrast to a related study by Munro and Derwing (2001), which reported that non-native English listeners (native language Mandarin Chinese) preferred slightly speeded accented speech (10%), but not highly speeded accented speech in judging speech comprehensibility. The implication of the current findings is that in clinical assessments of perception of speech presented at a rapid rate, L2 listeners are expected to exhibit systematic decrements in recognition of time-compressed speech, but may be slightly less affected than L1 listeners by increasing speech rate.

D. Linguistic and cognitive variables

The statistical models indicated that the ROW-PVT score was a significant contributing variable for all listener groups, across talker and TC conditions. This striking finding indicates that knowledge of the English lexicon is the

dominant variable contributing to the variance in recognition performance of both unaccented and accented talkers across all TC ratio conditions when all L1 and L2 listeners are included in the analysis. Vocabulary score was not found to interact significantly with the fixed effects examined (age group, listener language, TC ratio) confirming the importance of vocabulary knowledge for the measures assessed in this study. Nonetheless, this finding is influenced primarily by the performance of the L2 listeners, who demonstrated greater variability in vocabulary scores than the L1 listeners.

The importance of knowledge of English vocabulary on word recognition performance by L1 and L2 listeners has been well established (Owens, 1961; Bradlow and Pisoni, 1999; Banks *et al.*, 2015), and is a logical finding in this study given the low-context sentence stimuli used. The LP-SPIN sentences presented to listeners are brief, declarative sentences that do not include semantic contextual cues that could influence performance, and the final monosyllabic keywords in each sentence were chosen to be within a broad range of frequency of occurrence in English (Kalikow *et al.*, 1977). These factors likely necessitate a strong reliance on vocabulary knowledge. It should be noted, however, that the L2 listeners demonstrated a relatively high level of familiarity with the keyword stimuli (i.e., over 80%) in the post-test assessment. It is possible that the ROW-PVT score reflects a more general index of language knowledge. Some studies have suggested that vocabulary size may be related to listening comprehension proficiency in L2 listeners (Stæhr, 2009; Wang and Treffers-Daller, 2017), although the measure of vocabulary size is likely not a perfect proxy for the many dimensions of language knowledge and proficiency.

In contrast to the current findings, at least one prior investigation of sentence recognition by L1 and L2 listeners did not observe a correlation between English vocabulary familiarity and sentence recognition performance, even though the L1 and L2 listeners showed significant differences on measures of speech recognition and vocabulary knowledge (Tamati *et al.*, 2014). In this previous study, the sentences were characterized by a high degree of linguistic and indexical variability, and thus were quite different from the sentences used in the current study. It may be useful if future studies investigating speech recognition performance of L2 listeners made use of speech materials developed specifically for this population, such as the basic English lexicon (BEL) sentences (Calandruccio and Smiljanic, 2012).

Nonetheless, further analysis of the importance of vocabulary knowledge on speech recognition in the current study indicated that regardless of how well the L2 listeners performed on the vocabulary measure, their speech recognition performance did not show the advantages predicted by ISIB. Specifically, the comparison of L2 listeners who exhibited high vs low English vocabulary scores demonstrated that there were no differences in the magnitude of improvement in recognition of accented speech over unaccented speech (i.e., the ISIB-T).

The results also indicated that processing speed capacity relates to perception of unaccented and accented speech under conditions with increased speech rate, and contributes more strongly at higher speech rates. This finding may be

expected, considering the temporal demands inherent in perceiving fast speech. However, it appears that this cognitive component only contributed to speech recognition performance for the younger listeners. The lack of a significant effect of processing speed for the older listeners may be related to the well-documented decline in processing speed associated with aging (Salthouse, 2000); older listeners with a smaller processing speed capacity may not be able to rely on this resource for recognition of time-compressed speech.

Contrary to expectation, the current investigation did not find that measures of working memory, short-term memory, or attention contributed significantly to recognition of time-compressed unaccented and accented sentences by younger and older listeners. Many prior investigations have reported correlations between working memory and speech recognition in degraded listening conditions (Akeroyd, 2008; Füllgrabe *et al.*, 2015). One prior investigation showed a significant relationship between sequential working memory and perception of time-compressed unaccented speech for older L1 listeners with age-related hearing loss (Vaughan *et al.*, 2006). However, other investigations that assessed perception of accented speech by L1 listeners did not reveal associations between attention or working memory and recognition of accented sentences at natural speech rates (e.g., Gordon-Salant *et al.*, 2013). In the current study, the measures of vocabulary and processing speed had an effect on performance, whereas memory and attention did not. Thus, the relative importance of different cognitive domains appears to vary, depending upon the speech recognition task, form of degradation, and listener attributes.

V. SUMMARY AND CONCLUSIONS

This experiment examined perception of Spanish-accented and unaccented English sentences under various degrees of TC of the signal. Younger and older listeners who were non-native and native speakers of English were included as participants in order to examine the effects of age and prior language experience on perception. The findings presented here suggest that prior language experience generally does not afford an advantage when listening to accented speech, in that L2 listeners did not perform better than L1 listeners when recognizing accented speech. The findings also showed that age effects among L2 listeners are substantial for sentences presented with and without TC. Both L1 and L2 listeners were affected by increasing speech rate, although the L2 listeners exhibited more modest declines in recognition of time-compressed speech at fast rates compared to the L1 listeners. The results also demonstrated a strong influence of English vocabulary knowledge on speech recognition performance, although this variable did not contribute to better recognition of accented sentences over unaccented sentences for the L2 listeners. Speed of processing had an effect on recognition of natural-rate and time-compressed speech by younger listeners, but not older listeners; neither memory nor attention were important cognitive domains contributing to speech recognition performance. Collectively, the findings suggest that older L2 listeners are at a considerable disadvantage for listening to

English sentences, regardless of whether the sentences are spoken by an unaccented or an accented talker.

ACKNOWLEDGMENTS

This research was supported by Grant No. R01AG009191 from the National Institute on Aging, NIH. The authors are grateful to David Jara Ureta and Rosa Lemus for their assistance in data collection with the native Spanish listeners.

- Akeroyd, M. A. (2008). "Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults," *Int. J. Audiol.* **47**(Suppl 2), S53–S71.
- ANSI (2010). S3.6-2010, *American National Standard Specification for Audiometers* (revision of ANSI S3.6-1996, 2004) (American National Standards Institute, New York).
- Atagi, E., and Bent, T. (2011). "Perceptual dimensions of nonnative speech," in *Proceedings of the XVIIth International Congress of Phonetic Sciences*, Hong Kong, China, pp. 260–263.
- Banks, B., Gowen, E., Munro, K. J., and Adank, P. (2015). "Cognitive predictors of perceptual adaptation to accented speech," *J. Acoust. Soc. Am.* **137**(4), 2015–2024.
- Bent, T., and Bradlow, A. R. (2003). "The interlanguage speech intelligibility benefit," *J. Acoust. Soc. Am.* **114**, 1600–1610.
- Bilger, R. C., Nuetzel, J. M., Rabinowitz, W. M., and Rzeckowski, C. (1984). "Standardization of a test of speech perception in noise," *J. Speech Lang. Hear. Res.* **27**, 32–48.
- Boersma, P., and Weenink, D. (2018). "Praat: Doing phonetics by computer (version 5.4.09) [computer program]," <http://www.praat.org> (Last viewed 5/10/2018).
- Bradlow, A. R., and Pisoni, D. B. (1999). "Recognition of spoken words by native and non-native listeners: Talker-, listener- and item-related factors," *J. Acoust. Soc. Am.* **106**, 2074–2085.
- Brouwer, S., Van Engen, K., Calandruccio, L., and Bradlow, A. R. (2012). "Linguistic contributions to speech-on-speech masking for native and non-native listeners: Language familiarity and semantic content," *J. Acoust. Soc. Am.* **131**, 1449–1464.
- Brownell, R. (Ed.). (2012). *Receptive and Expressive One-Word Vocabulary Test*, 4th ed. (Pearson, London).
- Calandruccio, L., and Smiljanic, R. (2012). "New sentence recognition materials developed using a basic non-native English lexicon," *J. Speech Lang. Hear. Res.* **55**, 1342–1355.
- Daneman, M., and Carpenter, P. A. (1980). "Individual differences in working memory and reading," *J. Verb Learn. Verb Behav.* **19**, 450–466.
- Derwing, T., and Munro, M. J. (2001). "What speaking rates do non-native listeners prefer?," *Appl. Linguist.* **22**(3), 324–337.
- Dubno, J. R., Dirks, D. D., and Morgan, D. E. (1984). "Effects of age and mild hearing loss on speech recognition," *J. Acoust. Soc. Am.* **76**, 87–96.
- Eriksen, B. A., and Eriksen, C. A. (1974). "Effects of noise letters upon the identification of a target letter in a nonsearch task," *Percept. Psychophys.* **16**, 143–149.
- Ferguson, S. H., Jongman, A., Sereno, J. A., and Keum, K. (2010). "Intelligibility of foreign accented speech for older adults with and without hearing loss," *J. Am. Acad. Audiol.* **21**(3), 153–162.
- Fitzgibbons, P. J., and Gordon-Salant, S. (2001). "Aging and temporal discrimination in auditory sequences," *J. Acoust. Soc. Am.* **109**, 2955–2963.
- Fitzgibbons, P., and Gordon-Salant, S. (2010). "Age-related differences in discrimination of temporal intervals in accented tone sequences," *Hear. Res.* **264**, 41–47.
- 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.
- Füllgrabe, C., Moore, B. C. J., and Stone, M. A. (2015). "Age-group differences in speech identification despite matched audiometrically normal hearing: Contributions from auditory temporal processing and cognition," *Front. Aging Neurosci.* **6**, 1–25.
- Gordon-Salant, S., and Fitzgibbons, P. (1993). "Temporal factors and speech recognition performance in young and elderly listeners," *J. Speech Hear. Res.* **36**, 1276–1285.

- Gordon-Salant, S., and Fitzgibbons, P. (2001). "Sources of age-related difficulty for time compressed speech," *J. Speech Lang. Hear. Res.* **44**, 709–719.
- Gordon-Salant, S., and Friedman, S. H. (2011). "Recognition of rapid speech by blind and sighted older adults," *J. Speech Lang. Hear. Res.* **54**, 622–631.
- Gordon-Salant, S., Yeni-Komshian, G. H., Bieber, R., Jara Ureta, D., Freund, M. S., and Fitzgibbons, P. J. (2018). "Effects of listener age and linguistic experience on recognition of accented and unaccented English words," *J. Speech. Lang. Hear. Res.* in press.
- 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 age and hearing loss on recognition of unaccented and accented 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 maskers by younger and older listeners," *J. Acoust. Soc. Am.* **134**, 618–627.
- Hayes-Harb, R., Smith, B. L., Bent, T., and Bradlow, A. R. (2008). "The interlanguage speech intelligibility benefit for native speakers of Mandarin: Production and perception of English word-final voicing consonants," *J. Phonetics* **36**, 664–679.
- Imai, S., Walley, A. C., and Flege, J. E. (2005). "Lexical frequency and neighborhood density effects on the recognition of native and Spanish-accented words by native English and Spanish listeners," *J. Acoust. Soc. Am.* **117**, 896–907.
- Janse, E., and Adank, P. (2012). "Predicting foreign-accent adaptation in older adults," *Quart. J. Exp. Psych.* **65**(8), 1563–1585.
- 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.
- Major, R. C., Fitzmaurice, S. M., Bunta, F., and Balasubramanian, C. (2002). "The effects of nonnative accents on listening comprehension: Implications for ESL assessment," *TESOL Q.* **36**, 173–190.
- Mayo, L. H., Florentine, M., and Buus, S. (1997). "Age of second-language acquisition and perception of speech in noise," *J. Speech Lang. Hear. Res.* **40**, 686–693.
- Munro, M. J., and Derwing, T. M. (2001). "Modeling perceptions of the accentedness and comprehensibility of L2 speech the role of speaking rate," *Studies Sec. Lang. Acq.* **23**(4), 451–468.
- Munro, M., Derwing, T., and Morton, S. (2006). "The mutual intelligibility of L2 speech," *Studies Sec. Lang. Acq.* **28**(1), 111–131.
- Owens, E. (1961). "Intelligibility of words varying in familiarity," *J. Speech Hear. Res.* **4**, 113–129.
- Peng, Z. E., and Wang, L. M. (2016). "Effects of noise, reverberation and foreign accent on native and non-native listeners' performance of English speech comprehension," *J. Acoust. Soc. Am.* **139**(5), 2772–2783.
- Pike, K. L. (1945). *The Intonation of American English* (University of Michigan Press, Ann Arbor, MI).
- Salthouse, T. A. (2000). "Aging and measures of processing speed," *Bio. Psychol.* **54**(1-3), 35–54.
- Stæhr, L. S. (2009). "Vocabulary knowledge and advanced listening comprehension in English as a foreign language," *Studies Second. Lang. Acquisit.* **31**(4), 577–607.
- Tamati, T. N., Gilbert, J. L., and Pisoni, D. B. (2013). "Some factors underlying individual differences in speech recognition on PRESTO: A first report," *J. Am. Acad. Audiol.* **24**, 616–634.
- Tamati, T. N., and Pisoni, D. B. (2014). "Non-native listeners' recognition of high-variability speech using PRESTO," *J. Am. Acad. Audiol.* **25**, 869–892.
- Vaughan, N., Storzbach, D., and Furukawa, I. (2006). "Sequencing vs. non-sequencing working memory in understanding of rapid speech by older listeners," *J. Am. Acad. Audiol.* **17**, 506–514.
- Wang, Y., and Treffers-Daller, J. (2017). "Explaining listening comprehension among L2 learners of English: The contribution of general language proficiency, vocabulary knowledge and metacognitive awareness," *System* **65**, 139–150.
- Wechsler, D. (1997). *Wechsler Adult Intelligence Scale Third Edition (WAIS III)* (The Psychological Corporation, San Antonio, TX).
- Weintraub, S., Dikmen, S. S., Heaton, R. K., Tulsky, D. S., Zelazo, P. D., Bauer, P. J., Carlozzi, N. E., Slotkin, J., Blitz, D., Wallner-Allen, K., Fox, N. A., Beaumont, J. L., Mungas, D., Nowinski, C. J., Richler, J., Deocampo, J. A., Anderson, J. E., Manly, J. J., Borosh, B., Havlik, R., Conway, K., Edwards, E., Freund, L., King, J. W., Moy, C., Witt, E., and Gershon, R. C. (2013). "Cognition assessment using the NIH Toolbox," *Neurology* **80**, S54–S64.
- Wingfield, A., McCoy, S. L., Peelle, J. E., Tun, P. A., and Cox, C. L. (2006). "Effects of adult aging and hearing loss on comprehension of rapid speech varying in syntactic complexity," *J. Am. Acad. Audiol.* **17**(7), 487–497.
- Wingfield, A., Poon, L. W., Lombardi, L., and Lowe, D. (1985). "Speed of processing in normal aging: Effects of speech rate, linguistic structure, and processing time," *J. Gerontol.* **40**, 579–585.
- Xie, X., and Fowler, C. A. (2013). "Listening with an accent: The interlanguage speech intelligibility benefit in Mandarin speakers of English," *J. Phon.* **41**, 369–378.
- Zekveld, A. A., Kramer, S. E., and Festen, J. M. (2011). "Cognitive load during speech perception in noise: The influence of age, hearing loss, and cognition on the pupil response," *Ear Hear.* **32**(4), 498–510.