

Intelligibility of interrupted sentences at subsegmental levels in young normal-hearing and elderly hearing-impaired listeners^{a)}

Jae Hee Lee^{b)} and Diane Kewley-Port^{c)}

Department of Speech and Hearing Sciences, Indiana University, Bloomington, Indiana 47405

(Received 16 December 2007; revised 8 October 2008; accepted 13 October 2008)

Although listeners can partially understand sentences interrupted by silence or noise, and their performance depends on the characteristics of the glimpses, few studies have examined effects of the types of segmental and subsegmental information on sentence intelligibility. Given the finding of twice better intelligibility from vowel-only glimpses than from consonants [Kewley-Port *et al.* (2007). "Contribution of consonant versus vowel information to sentence intelligibility for young normal-hearing and elderly hearing-impaired listeners," *J. Acoust. Soc. Am.* **122**, 2365–2375], this study examined young normal-hearing and elderly hearing-impaired (EHI) listeners' intelligibility of interrupted sentences that preserved four different types of subsegmental cues (steady-states at centers or transitions at margins; vowel onset or offset transitions). Forty-two interrupted sentences from TIMIT were presented twice at 95 dB SPL, first with 50% and second with 70% of sentence duration. Compared to high sentence intelligibility for uninterrupted sentences, interrupted sentences had significant decreases in performance for all listeners, with a larger decrease for EHI listeners. Scores for both groups were significantly better for 70% duration than for 50% but were not significantly different for the type of subsegmental information. Performance by EHI listeners was associated with their high-frequency hearing thresholds rather than with age. Together with previous results using segmental interruption, preservation of vowels in interrupted sentences provides greater benefit to sentence intelligibility compared to consonants or subsegmental cues.

© 2009 Acoustical Society of America. [DOI: 10.1121/1.3021304]

PACS number(s): 43.71.Ky, 43.71.Es, 43.66.Sr [MSS]

Pages: 1153–1163

I. INTRODUCTION

In everyday listening situations, both the target speech and background noise continuously fluctuate, especially when the noise consists of competing speech or dynamic environmental noise. Because these fluctuations are usually independent, the speech signal may be only partially audible or even completely inaudible depending on the relative levels of the target and the noise. Listeners must then integrate individual glimpses of target information within the dips or valleys of fluctuating noise in order to understand the target's entire meaning. Previous studies of temporal interruption demonstrated that when young normal-hearing (YNH) listeners glimpse the full spectrum of speech, their performance is relatively high for 8–10 Hz of silence or noise interruption at a 50% duty cycle, regardless of the type of stimulus materials such as monosyllabic words (Miller and Licklider, 1950; Kirikae *et al.*, 1964), sentences (Bergman *et al.*, 1976; Bergman, 1980; Nelson and Jin, 2004; Iyer *et al.*, 2007; Li and Loizou, 2007), and connected discourse passage (Powers and Speaks, 1973).

Two studies have examined the ability of elderly hearing-impaired (EHI) listeners to integrate glimpses in target sentences that were either periodically interrupted (Gordon-Salant and Fitzgibbons, 1993) or segmentally (consonants versus vowels) interrupted (Kewley-Port *et al.*, 2007). As expected, both studies have found that EHI listeners showed poorer recognition of interrupted sentences than did YNH listeners. Although the overall presentation level of the speech portions glimpsed was above the hearing thresholds of EHI listeners up to 4 kHz (i.e., 85–90 dB SPL), the results of both studies also indicated that variability in intelligibility scores was better accounted for by high pure-tone threshold average (PTA) (averaged hearing thresholds at 1, 2, and 4 kHz) than by age.

Kewley-Port *et al.* (2007) reported that intelligibility for both YNH and EHI listeners was two times better when only vowels remained in sentences compared to when only consonants remained (this study is referred to as KBL07 in the rest of this article). This finding replicated previous results for YNH listeners by Cole *et al.* (1996) of a ratio of 2:1 for intelligibility in vowel-only sentences versus consonant-only sentences for various types of interruption noises (no noise, harmonic complexes, and white noise). Together, these results suggest that as long as sentences are reasonably audible, intelligibility of segmentally interrupted speech for both YNH and EHI listeners depends strongly on the type of segmental information glimpsed. Because the importance of high-frequency information for consonants has been long

^{a)} Portions of the data were presented at the 151th Meeting of the Acoustical Society of America and at the fourth Joint Meeting of the Acoustical Society of America and the Acoustical Society of Japan.

^{b)} Author to whom correspondence should be addressed. Present address: Department of Audiology, Institute of Audiology, Hallym Institute of Advanced International Studies, Korea. Electronic mail: leejaehee@hallym.ac.kr

^{c)} Electronic mail: kewley@indiana.edu

emphasized for the clinical amplification in the hearing aids, these findings of the greater contribution of vowels compared to consonants for sentence intelligibility are noteworthy.

The motivation for the present study was to examine contributions of glimpsing vowels versus consonants from somewhat different but theoretically based definitions of vowel and consonant information. Traditionally steady-state acoustic cues have relatively less spectral change over time and specify vowel information (Peterson and Barney, 1952), whereas dynamic transition cues have more spectral changes over time and specify consonant information (Lieberman *et al.*, 1957). However, another theoretical point of view (dynamic specification theory) (Strange and Bohn, 1998) documented in a series of studies by Strange and colleagues (Strange *et al.*, 1976; Strange, 1989; Jenkins *et al.*, 1994) stresses the importance of dynamic transition cues in vowel perception. Their studies concentrated primarily on the contribution of dynamic transition information at CVC margins to vowel identification. Thus different theories of what specifies vowel versus consonant information suggest that there exist regions of subsegmental information in speech that contribute differentially to intelligibility. Four subsegmental regions, steady-state centers versus dynamic margins, and onset versus offset transitions were the focus of this study. The outcome should demonstrate whether there are more important information-rich subsegmental regions of speech that result in better sentence intelligibility. If this was found for EHI listeners, then those regions should be considered to be preserved or enhanced in the future design of speech processors for hearing assistive devices for EHI listeners.

Few studies have been conducted on the ability of older listeners to use dynamic cues glimpsed from nonsense syllables, and the results were not in agreement with each other. Fox *et al.* (1992) reported an age-related decrement in the ability to use dynamic cues in CVC margins for vowel and consonant perception among various age groups with relatively normal hearing, supporting age-related deficit hypothesis. Ohde and Abou-Khalil (2001), however, found the similar abilities among young, middle-aged, and older adults who had near-normal hearing for their age (i.e., less than 40 dB HL at 4000 Hz in older adults) to use these dynamic formant transition cues for vowel and consonant perception, thereby not supporting age-related deficits in using dynamic cues. Dorman *et al.* (1985) reported that differences in performance among YNH, elderly normal-hearing (ENH), and EHI listener groups were not consistent across various phonetic identification tasks, but rather were varied depending on the type of vowels and consonants contrasted. We note, however, that it is not clear how these previous, somewhat conflicting, results for older listeners' use of dynamic versus static cues in CVC syllables for vowel identification can be generalized to sentence intelligibility for EHI listeners.

The present study employs the same TIMIT sentences (Garofolo *et al.*, 1990) used by KBL07 and examines how four regions of subsegmental glimpsing cues might differentially contribute to sentence intelligibility for both YNH and EHI listeners. Centers or margins within each segment were selected as the first and the second subsegmental target re-

gions to focus on the contributions of quasi-steady-state versus transition information to sentence intelligibility and were primarily motivated by Strange's theory. Full vowel onset or offset transitions were selected as the third and the fourth target regions based on long-standing results that acoustic cues for consonants in CVC syllables have been found to be more salient at the onsets of syllables compared to the offsets (Redford and Diehl, 1999). Two different durations of glimpsing (50% and 70% durations of each segment) were employed.

The objectives of the present study were to examine the following questions: (1) would EHI listeners show significantly reduced ability to integrate subsegmental cues glimpsed from interrupted but audible (95 dB SPL) sentences than YNH listeners; (2) would dynamic transition cues result in equivalent intelligibility in sentence recognition in listeners compared to quasi-steady-state cues, similar to the role of dynamic cues shown for vowel identification; (3) what impact would different subsegmental cues (i.e., phoneme steady-state at centers versus transitions at margins; vowel onset versus offset transition regions) have on the ability of listeners to recognize interrupted sentences; and (4) how would performance improve as the duration of speech portions glimpsed increases from 50% to 70%. In addition, correlational analyses examined the relation between individual differences in the performance of EHI listeners with the variables of hearing loss or age.

II. GENERAL METHODS

A. Overview of experimental design

To investigate the effect of four regions of subsegmental cues on sentence intelligibility for YNH and EHI listeners, a mixed design was developed with three variables: two between-subject variables (two listener groups and four stimulus conditions) and one within-subject variable (two durations of glimpsing, 50% and 70%). The 42 TIMIT test sentences (Texas Instruments/Massachusetts Institute of Technology) (Garofolo *et al.*, 1990) that were used in KBL07 were employed in this study as test materials. These 42 sentences were interrupted with the same speech-shaped noise (SSN) in KBL07, but in four different ways to preserve four different subsegmental cues depending on the regions within each segment. The four conditions (see Fig. 2) preserved four different subsegmental regions focusing on either phoneme steady-state or three transition cues as follows: (1) the center region of vowel and consonant segments that are generally quasi-steady-states in each segment (CENTER), (2) the two margin regions of each vowel and consonant segments where generally the formant transitions in each segment are found (MARGIN), (3) the final portion of a consonant and the initial portion of following vowel incorporating the vowel onset transitions (ONSET), and (4) the final portion of a vowel and the initial portion of the following consonant incorporating the vowel offset transitions (OFFSET). Each test sentence was presented twice with two durations allowing glimpses of the target speech, 50% (50% speech-on and 50% noise-on, alternately) and 70% (70% speech-on and 30% noise-on, alternately) of the duration of

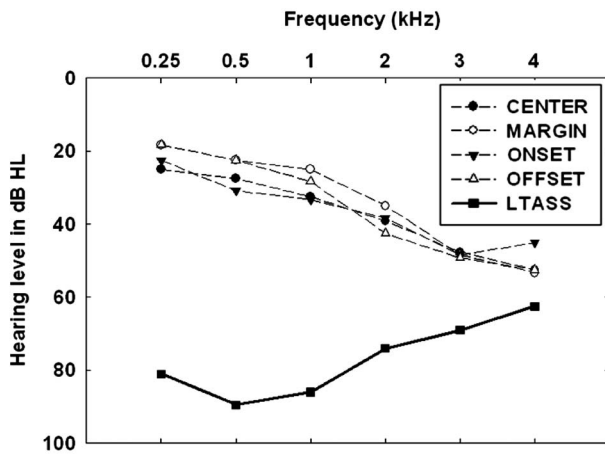


FIG. 1. Average pure-tone thresholds for the test ear of EHI listeners across CENTER, MARGIN, ONSET, and OFFSET conditions ($N=6$ per condition) displayed by broken lines in dB HL (ANSI, 1996). Solid line displays the LTASS in dB HL when calibrated at the 95 dB SPL level.

each segment. In a pilot study, the 50% duration yielded near zero performance for some EHI listeners. Therefore, the 70% duration of the target sentence was added as a second presentation to both listener groups. Details of the methods follow.

B. Participants

Twenty-four YNH and 24 EHI listeners were paid to participate. All listeners were native American-English speakers. Each listener group was gender balanced (12 males and 12 females). YNH and EHI listeners were recruited from Indiana University and from the Indiana University Hearing Clinic, respectively. To participate, all listeners were required to have a passing score ($>27/30$) on the Mini-Mental Status Examination (MMSE) (Folstein *et al.*, 1975) for cognitive status, a score of 5 or greater on the auditory forward digit span test, and a score of 4 or greater on the auditory backward digit span test for normal short-term memory. YNH listeners ranged in age from 20 to 35 years ($M=27$ years), and had pure-tone thresholds no greater than 20 dB HL at octave intervals from 250 to 8000 Hz (ANSI, 1996). EHI listeners ranged from 65 to 80 years of age ($M=74$ years). The hearing criteria for EHI listeners were normal middle ear status, and postlingual, bilateral, high-frequency, mild-to-moderate hearing loss of cochlear origin with pure-tone thresholds less than 60 dB HL from 2000 to 4000 Hz. With a quasi random assignment, participants listened to the 42 test sentences in one of the four conditions. To reduce individual variability in performance, age and hearing loss were matched carefully across the four conditions. In Fig. 1, each of the dashed lines shows the average air-conduction pure-tone thresholds for the tested ear in EHI listeners for each of the four conditions (CENTER, MARGIN, ONSET, and OFFSET). The solid line displays the level of the long-term-averaged speech spectrum (LTASS) of target speech in dB HL. As shown, target speech was presented to EHI listeners above pure-tone thresholds from 250 to 4000 Hz (e.g., at least 9 dB above pure-tone threshold at 4000 Hz). Results of Levene's test for equality of variances and a one-way analy-

sis of variance (ANOVA) revealed that there were homogeneous variances as well as no significant differences in both age and pure-tone thresholds for six EHI listeners in each of the four conditions.

C. Stimuli

KBL07 used 42 sentences (21 male and 21 female speakers) from the TIMIT corpus as test material. Speakers were from the North Midland dialect region that matched the catchment area of the participants in this study, namely, Indianapolis, IN, and further north.

In the TIMIT database, segmental boundaries and phonetic transcriptions were established by expert phoneticians. KBL07 verified the segmental boundaries provided by the TIMIT corpus. KBL07 added three minor rules appropriate for identifying the vowels and consonants in sentences: (1) stop closure symbols were combined with the following stop and treated as a single consonant; (2) syllable $V+[r]$ was considered as a single rhotocized vowel; and (3) the glottal stop $[q]$ occurred between two vowels such that; $[VqV]$ was treated as a vowel. These three rules were also used in the present study. As in the previous study, in order for all the TIMIT sentences needed to be sufficiently audible for both listener groups, 95 dB SPL was used as the signal level to present the sentences after digitally scaling them to a constant rms (root-mean-square) value (for more details, see Sec. II E).

D. Processing for interrupted sentences

1. Speech information in interrupted sentences

The test sentences were interrupted at specified subsegmental intervals with low-level SSN in four different conditions. Each condition presented one of four different regions of subsegmental glimpsing (CENTER, MARGIN, ONSET, and OFFSET) with either a 50% or 70% duration. Figure 2 shows temporal waveforms of an example CVC word "mean" extracted from a test sentence. The top waveform of "mean" has no interruption. The remaining waveforms of "mean" show four regions of subsegmental glimpsing in which 50% of glimpsing duration was applied.

As shown in Fig. 2, the 50% duration yields two pairs of conditions, CENTER/MARGIN (second and third waveforms) and ONSET/OFFSET (fourth and fifth waveforms) that present complementary acoustic pieces of the subsegmental intervals. The CENTER regions preserved 50% of the center portions of vowel and consonant segments, which represented subsegmental information containing the quasi-steady-state parts of vowels and consonants. As a complementary condition, the MARGIN regions preserved 25% of each of the two margin portions of the vowel and consonant segments, which contained mostly the spectral transitions of the vowels and consonants, similar to stimuli used by Strange *et al.* (1976). The ONSET and OFFSET conditions focused on different types of vowel transitional information. The ONSET condition preserved vowel onset information by capturing transitional information from the last 50% of a consonant preceding a vowel and the initial 50% of the following vowel. The complementary OFFSET preserved

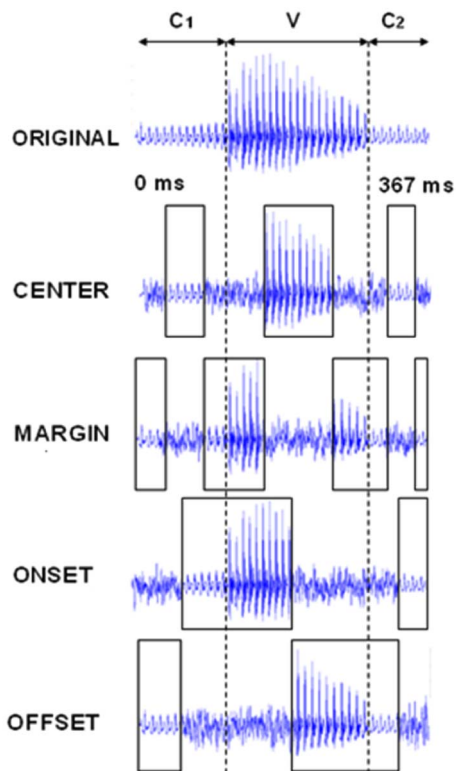


FIG. 2. (Color online) Temporal waveforms of the C_1VC_2 syllable “mean” (/min/) extracted from a test sentence, “What did you mean by that rattlesnake gag?”. Top waveform of “mean” is from the original sentence with no interruption. Portions inside the boxes from second to fifth waveforms display four different subsegmental cues preserved in each condition (CENTER, MARGIN, ONSET, and OFFSET from second to fifth waveforms) for the 50% proportion of duration, while the regions outside the boxes are replaced with SSN.

vowel offsets by capturing the last 50% of a vowel and the initial 50% of the following consonant. As expected, the TIMIT sentences include not only CVCs shown in Fig. 2 but also many consonant and vowel clusters such as CCVC, CVCC, and CVVC. The clusters were considered as a single syllable (for example, $CC \rightarrow C$, $VV \rightarrow V$).

The purpose of the 70% of glimpsing duration was to reduce the amount of interruption in sentences such that this duration would elevate the near floor performance of some EHI listeners in the 50% duration observed in pilot testing. 10% more information on either side of the 50% duration was added to each preserved subsegmental unit to comprise the 70% glimpsing duration of target speech. MATLAB scripts were used in conjunction with the TIMIT boundaries to calculate the glimpsing duration and insert the noise for all four conditions.

2. Noise in interrupted sentences

The SSN was generated by MATLAB and was used to replace parts of the sentences. The SSN shape was based on a standard LTASS (ANSI, 1969) that had a flat shape of 0–500 Hz and a -9 dB/octave roll-off above 500 Hz. The present study attempted to set the level of SSN to be low relative to the vowel, yet be audible to EHI listeners with mild-to-moderate hearing loss. Presumably the low-level

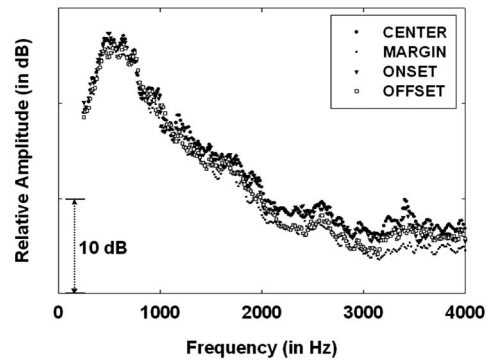


FIG. 3. Root-mean-square amplitude spectra of the concatenated 42 sentences processed in the four conditions (CENTER, MARGIN, ONSET, and OFFSET) as a function of frequency. Note that all sentences used in the current study were low-pass filtered at 4400 Hz.

noise would smooth out the somewhat choppy sentences, reduce boundary transients, and encourage phoneme restoration (Warren, 1970). After informal listening in a pilot test, 75 dB SPL (i.e., 20 dB lower than the average level of 95 dB SPL) was chosen to be the level of the SSN.

E. Calibration

Similar calibration procedures from KBL07 were administered to verify signal levels. First, scripts were written in MATLAB to verify that all the 42 test sentences had similar average rms levels (i.e., within ± 2 dB). Second, a MATLAB script was used to find the most intense vowel across all sentences and then iterated that vowel to produce a calibration vowel of 4 s. To avoid effects of hearing loss beyond 4000 Hz in EHI listeners, all sentences were filtered by a low-pass finite impulse response filter that was flat to 4000 Hz with a 3 dB cutoff at 4400 Hz and a 200 dB/octave steep slope in a Tucker Davis Technologies (TDT) PF1. The sound level for the calibration vowel was also low-pass filtered, and its sound level was set to 100 dB SPL through ER-3A insert earphones in a HA-2.2 cm³ coupler using a Larson Davis model 2800 sound level meter with linear weighting. Relative to the loudest calibration vowel, the mean of the distribution of the loudest vowels in the other sentences was 95 dB SPL, being the nominal level referenced to this study. An additional low-level background noise was continuously presented during testing. The purpose of this noise was to reduce transients between speech and noise. This noise was generated by the TDT WG2 and was also low-pass filtered at 4400 Hz. The level of this noise was reduced by more than 50 dB compared to the calibration vowel (100 dB SPL) measured using the same equipment described above.

Spectral analyses were used to verify that the long-term spectra for each of the conditions were similar. All 42 test sentences were concatenated together after eliminating pauses. Long-term spectra were calculated with a Hanning window and the Baum–Welch algorithm in MATLAB. We confirmed very similar spectral envelopes (i.e., within ± 3 dB) across the frequency range of 0–4000 Hz for each of the four conditions (CENTER, MARGIN, ONSET, and OFFSET), as shown in Fig. 3.

F. Test procedures

1. Stimulus presentation

All screening tests for hearing and cognitive functions were administered before testing. Each listener was instructed about the tasks using written and verbal instructions. Test stimuli were controlled by TDT system II hardware connected to a personal computer and were presented through ER-3A insert earphones to listeners in a sound-treated booth. Test sentences were presented to the better ear of the hearing-impaired listeners (generally the right ear) and to the right ear for the normal-hearing listeners. Six familiarization sentences consisted of two unprocessed sentences (i.e., no interruption) and four processed sentences. The four processed sentences were presented two times, with 50% and 70% durations of glimpsing corresponding to each experimental condition, and then the unprocessed sentence was given as feedback.

The 42 test sentences were randomized and then presented in one fixed order. Listeners heard the 42 sentences twice, first with the 50% duration and second with the 70% duration. After each presentation listeners were asked to respond by repeating verbally any words they thought they heard from the test sentence. All responses were recorded by a digital recorder. Listeners were encouraged to guess, regardless of whether the responded words or partial words made sensible sentences. No feedback was provided. The correctly identified words were scored by the experimenter during testing and then rechecked by a linguist from the recorded responses later. Experimental testing lasted 1 h for YNH listeners and 1.5 h for EHI listeners.

2. Scoring and data analysis

The number of correctly identified words was counted and scored as percentage of correct words relative to the total number of words in the sentences. All the words were scored as correct only when they exactly matched with the target words (i.e., incorrect for morphological variants). The purpose of this word scoring was to compare overall ability to understand interrupted sentences between YNH and EHI listeners. All the correct words were scored from the recorded responses by a second scorer who was a native American-English listener and a linguistics doctoral student. All discrepancies for the scoring of correct words were resolved by consensus between experimenter and the second scorer using the recorded responses. Word scores in percentage were transformed into rationalized arcsine units (RAUs) (Studebaker, 1985) for all statistical analyses. Statistical tests were based on a general linear model ANOVA with repeated-measures in SPSS statistical software (version 14.1; SPSS Inc., Chicago, IL)

III. RESULTS

A. Intelligibility for interrupted sentences

The scores of all listeners were obtained for the two uninterrupted (100% duration) sentences from the familiarization task. The averaged score of these two familiarization sentences presented at 95 dB SPL was 100% for YNH and

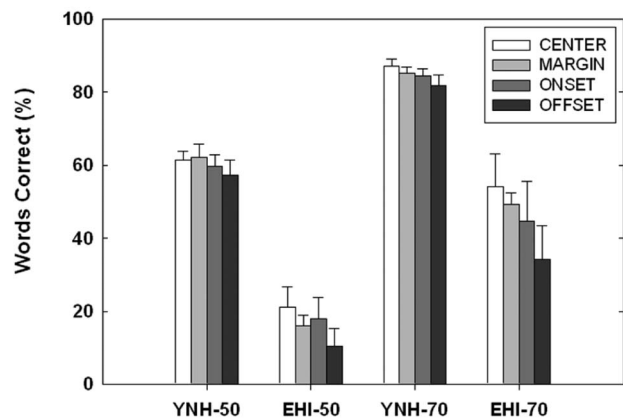


FIG. 4. Mean number of words correct in percentage in each of four conditions for YNH and EHI listeners with 50% duration (YNH-50 and EHI-50) and for YNH and EHI listeners with 70% duration (YNH-70 and EHI-70). Error bars indicate standard errors.

92% for EHI listeners. This accuracy was comparable to the results of KBL07 that reported a range of 97–100% for YNH and 88–99% for EHI listeners for the 14 uninterrupted sentences at 95 dB SPL. Thus full sentences at 95 dB SPL were reasonably audible up to 4000 Hz for both YNH and EHI listeners. Figure 4 shows the mean percentage of words correct in sentences across four conditions. The bars from left to right display scores obtained from YNH and EHI listeners with the 50% duration (YNH-50 and EHI-50) and then with the 70% duration (YNH-70 and EHI-70).

In general, EHI listeners performed poorer than YNH listeners, regardless of duration. For the 70% duration across all conditions, YNH listeners identified words in sentences about 85% correct, whereas EHI listeners identified words only about 46% correct. Compared to scores for uninterrupted sentences, 30% noise interruption resulted in a 15% decrease in scores (from 100% to 85% correct) for YNH but a 46% decrease (from 92% to 46% correct) for EHI listeners. For the 50% duration across all conditions, YNH listeners identified words in sentences about 60% correct, whereas EHI listeners identified words only about 16% correct. Compared to uninterrupted sentences, a 50% interruption yielded a 40% decrease (from 100% to 60% correct) for YNH but a 76% decrease (from 92% to 16% correct) for EHI listeners.

An ANOVA with repeated-measures was calculated for two between-subject variables (two groups \times four conditions) and one within-subject repeated-variable (two durations) with the dependent variable of words correct in RAU. Results showed a significant ($p < 0.05$) main effect of listener group [$F(1, 40) = 112.5$] and a significant main effect of duration [$F(1, 40) = 1321.9$] but no significant main effect of condition [$F(3, 40) = 1.5, p = 0.23$]. As expected, results indicate that YNH outperformed EHI overall, and performance with the 70% duration was better than that with 50% duration. Unexpectedly performance was similar across four different subsegmental conditions. Although scores showed that EHI performed the best in CENTER and the worst in OFFSET regardless of the duration presented, a significant effect of condition was not obtained due to large individual differences and the small sample size ($N = 6$) of EHI listener as-

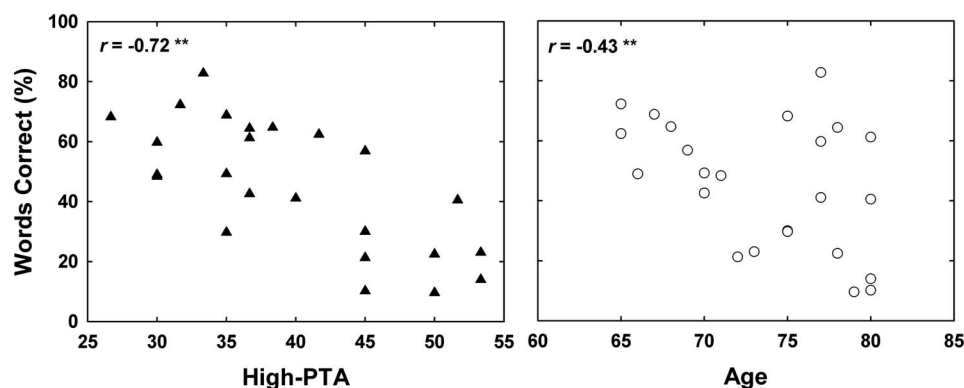


FIG. 5. Scatter plots of data for word scores by EHI-70 as a function of high-PTA (from 1, 2, and 4 kHz) in the left panel and as a function of age in the right panel.

signed to each of four conditions. There was one significant two-way interaction between group and duration [$F(1,40) = 13.4$], indicating that the amount of benefit for the increased information in the 70% duration over the 50% duration was greater for EHI compared to YNH listeners.

B. Individual differences for EHI listeners

Not surprisingly, large individual variability was observed for temporally interrupted sentences for EHI listeners even though sentences were audible. Two previous studies in which interrupted sentences were presented to EHI listeners at high levels, 85 dB SPL (Gordon-Salant and Fitzgibbons, 1993) and 95 dB SPL (KBL07), reported stronger correlations between hearing loss and sentence intelligibility scores compared to that for age. The present study also examined the relation between age and hearing thresholds, either averaged or individual. First, correlational analysis results showed that age was not significantly correlated with high-frequency pure-tone thresholds (high-PTA) averaged at 1, 2, and 4 kHz ($p=0.06$, $r=0.39$). There was also no significant correlation between age and individual hearing thresholds from 0.25 to 8 kHz (with r ranging from 0.26 to 0.37 and p values from 0.07 to 0.67). With this in mind, the main correlational analyses investigated whether the large EHI individual differences in intelligibility are better predicted by high-PTA or age. Only scores for the 70% duration were analyzed because scores for the 50% duration were at floor for some listeners. Percentages of correct word scores of all EHI listeners ($N=24$) were averaged across the four subsegmental conditions (given no significant condition effect) and transformed to RAU.

High-PTA had a high negative correlation ($p < 0.001$, $r = -0.72$), and age had a weaker, but significant, negative correlation ($p < 0.004$, $r = -0.43$) with word scores. This relationship between hearing loss and word scores for EHI-70 is displayed in the left panel of Fig. 5, while the right panel plots relationship between age and performance. Forward stepwise regression analyses showed that high-PTA accounted for 51% of the variance in word scores [$F(1,22) = 22.97$, $p < 0.001$], while age was not a significant predictor ($p = 0.14$). Apparently the large spread of word scores from the EHI listeners ranging 75–80 years, as displayed in Fig. 5, underlies weaker correlations obtained for age compared to high-frequency hearing loss. This dominant contribution of high-PTA rather than age suggests that factors associated

with hearing loss contributed to decreased sentence intelligibility by EHI listeners. Thus, the higher-than-normal speech levels did not eliminate the negative effects of hearing loss on EHI performance.

IV. GENERAL DISCUSSION

A. Intelligibility with subsegmental versus segmental interruption

Forty-two TIMIT sentences were temporally interrupted allowing two types of speech portions to be glimpsed, either four subsegmental cues in this study or two segmental cues in a previous study [Kewley-Port *et al.*, 2007 (denoted as KBL07)]. In the present study, sentences were subsegmentally interrupted by SSN, preserving either quasi-steady-state versus three types of transition cues, with either a 50% or 70% of duration. In KBL07, segmental interruption replaced either consonants or vowels in the target sentences by SSN, resulting in vowel-in (VIN) and consonant-in (CIN) conditions. As expected, both studies reported a high performance for uninterrupted sentences for both YNH and EHI listener groups, but for interrupted sentences a significant reduction in sentence intelligibility was obtained for all listeners, with a larger drop for EHI. However, the contribution of subsegmental versus segmental cues to intelligibility of interrupted sentences was not the same. The contribution of four different subsegmental cues to sentence intelligibility was significantly affected by the duration of glimpsing but not by the regions of subsegmental information (steady-state or dynamic transitions). Unlike subsegmental cues, the types of segmental cues contributed differentially to sentence intelligibility (i.e., 2:1 better performance in VIN than in CIN conditions in KBL07). Below a direct comparison of the two studies was attempted because the stimuli and methods are similar (i.e., the same 42 sentences, the same overall presentation level, and similar criteria of hearing status and age for EHI listeners).

In order to compare the two studies, the approximate duration of vowels versus consonants preserved in sentences was calculated relative to total sentence duration using the segment boundaries in the TIMIT database. Specifically, the proportion of the sum of the duration of all vowels in the VIN condition relative to sentence duration was approximately 45%, while the proportion of consonant duration in the CIN condition was 55%. Note that this 55% of glimpsing

TABLE I. The first row shows the duration (%) in each condition across current and KBL07 studies. The score of words correct (%) with standard errors (SEs) is displayed in the second row for both YNH and EHI listeners.

Duration (%) (condition)	YNH			EHI		
	45 (VIN)	50 (Sub-50)	55 (CIN)	45 (CIN)	50 (VIN)	55 (Sub-50)
Words correct (%) (SE)	65.06 (1.36)	60.13 (1.60)	51.59 (2.44)	40.13 (3.79)	16.29 (2.43)	19.96 (4.14)

duration for CIN compared to 45% of VIN was due to frequently occurring consonant clusters and while the average duration of individual consonants was actually less than that of the average vowel.

Table I shows the scores of words correct (%) for segmental conditions (VIN with 45% and CIN with 55% duration) and scores averaged across the four subsegmental conditions with 50% duration that were labeled as Sub-50. Although durations of 45%, 50%, and 55% are not particularly different in relation to the glimpsing opportunities of the target sentence, the highest word score was for VIN with the shortest, 45%, duration compared to the other conditions with longer durations. This pattern of greater performance in VIN than in others was more obvious in EHI listeners. This supports the previous finding (KBL07) that vowels contribute more to intelligibility of interrupted sentences than consonants, and relative to the current study more than any of four subsegmental cues.

To examine the differences shown in Table I in more detail, an additional one-way ANOVA measure was administered separately for YNH and EHI listener groups, with one between-subject variable (three conditions, VIN, Sub-50, and CIN) and the dependent variable of words correct in RAU. As expected from Table I, a significant ($p < 0.05$) main effect of condition was found for YNH [$F(2,53)=12.1$] and for EHI [$F(2,53)=12.8$] listener groups. Results of a Bonferroni post-test indicated that differences in scores were significant between VIN and CIN and between Sub-50 and CIN but not significant between VIN and Sub-50 for YNH listeners. For EHI listeners, significant differences in performance were found between VIN and Sub-50 and between VIN and CIN, but not between Sub-50 and CIN conditions. These results confirm a strong benefit of vowels compared to other cues as the most important glimpsing source for EHI listeners. Note that a 10% difference in scores between Sub-50 ($N=24$) and CIN ($N=16$) reached significance in YNH listeners in Table I, although a 20% difference in scores between CENTER ($N=6$) and OFFSET ($N=6$) subsegmental conditions in EHI listeners (see Fig. 4) did not reach a statistical significance in the current study.

Power analysis was used to determine if the sample size ($N=6$) was too small between conditions to correctly accept the null hypothesis. A simple power analysis between the two most extreme conditions, CENTER ($M=54\%$) and OFFSET ($M=34\%$), revealed that although Cohen's d showed that effect size was large (1.05), power was only 0.446. To raise power to 0.80 (or 0.90) a sample size of $N=11$ (or $N=15$) would be required, approximately double the

current sample size. Moreover, partial eta squared values observed from the SPSS ANOVA showed that the condition factor accounted for a very low, 13.1%, overall variance for EHI listeners, with even lower variance, 9.2%, for YNH listeners. These analyses suggest that our small sample size is not the reason underlying the negligible effect of the condition factor but rather that for both EHI and YNH listeners the manipulation of subsegmental information had no significant effect in these sentences.

Root-mean-square long-term amplitude spectra of the concatenated sentences in VIN, CIN, and Sub-50 were compared to investigate level differences across these conditions. Figure 6 shows that VIN sentences (displayed by unfilled circles) had overall 10 dB of level advantage than CIN sentences (displayed by unfilled squares), as reported in KBL07. The level of the concatenated sentences across the four subsegmental conditions (Sub-50, displayed by filled triangles) was very similar to the level of VIN sentences (i.e., within ± 3 dB). Considering this 6–7 dB of level advantage but the lower scores of EHI found in Sub-50 than in CIN, we speculate that the two times more frequent rate of interruption occurring in subsegmental condition than in segmental condition may be more challenging for EHI listeners to integrate the glimpses of the target speech. This hypothesis may be related to previous reports of auditory temporal processing deficits in older listeners [see reviews in Gordon-Salant (2005)].

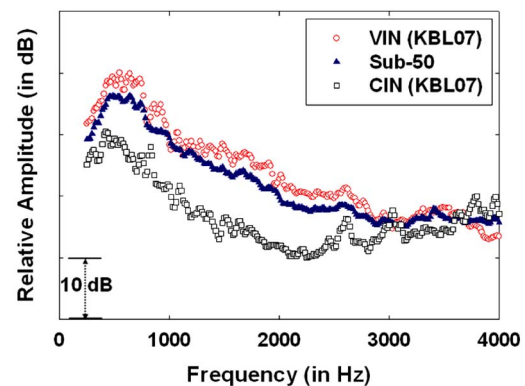


FIG. 6. (Color online) Spectra of the concatenated sentences used in VIN and CIN conditions (from KBL07) displayed as unfilled circles and unfilled squares, respectively. The filled triangles show the amplitude spectra of the concatenated sentences across the four subsegmental conditions with 50% duration (Sub-50).

B. Intelligibility of interrupted speech of young and older listeners

This section compares the present results with previous literature concerning intelligibility of interrupted sentences by younger and older listeners when 50% of duration was interrupted by either periodic or nonperiodic sound. For comparison purposes, we calculated from six randomly selected sentences that the approximate periodic rate of subsegmental interruption in our study for the 50% duration was about a 10 Hz rate, i.e., ten times of interruption per second. Because earlier reports using periodic interruption applied various interruption rates (0.1–10 000 Hz of interruption), glimpsing duration (6%–75% duration), and monaural versus binaural presentation, we selected only results that used a range of 8–10 Hz rate of periodic interruption, a 50% duration, and monaural presentation for comparison with our data.

60% to almost 100% correct intelligibility was found in young listeners with normal hearing when periodic interruption was applied to monosyllabic words (Miller and Licklider, 1950; Kirikae *et al.*, 1964), sentences (current data; Korsan-Bengtzen, 1973; Bergman *et al.*, 1976; Bergman, 1980; Nelson and Jin, 2004; Iyer *et al.*, 2007), and connected passages (Powers and Speaks, 1973). Findings from various studies suggest an apparent benefit for an approximately 10 Hz interruption rate compared to very low interruption rate (e.g., 1–2 Hz interruption). This 10 Hz periodic interruption rate might allow YNH listeners to have several glimpses at some essential segmental or subsegmental information regardless of the type of stimulus materials, whereas an entire syllable or word might be lost with a very low, 1–2 Hz, rate of interruption.

Only a few studies examined how well older listeners could integrate the glimpses of target sentences interrupted with the 8–10 Hz interruption rate using elderly listeners either with sensorineural hearing loss (current; Korsan-Bengtzen, 1973) or with near-normal hearing (Bergman *et al.*, 1976, Bergman, 1980). Similar to our study, elderly listeners in the studies above showed very low but consistent accuracy, 15%–16% correct, for identifying words in interrupted sentences. Given the high accuracy in performance by younger listeners across studies, we conclude that almost floor scores for interrupted sentences in older listeners indicate their general inability to successfully integrate information glimpsed from the target speech, regardless of types or difficulty of test materials.

An earlier study of Bergman *et al.* (1976) had 185 adult listeners ranging from 20 to 80 years of ages and showed a consistent drop in performance as a function of age (80%, 73%, 45%, 32%, 22% to 15% correct from 20 to 80 years of age). Based on this systematic drop as a function of age, they concluded that older listeners were less successful in integrating interrupted sentences due to age-related changes either in the auditory central nervous system or at a more cognitive level. Although this systematic drop in performance has significant implications, we note that Bergman *et al.* (1976) used a relatively lax hearing criteria for normal hearing (i.e., “35 dB at 0.5, 1, and 2 kHz and 40 dB at 4 kHz”) and did not investigate hearing status as a factor.

It should be noted that our study did not include either young hearing-impaired listeners or ENH listeners to control for possible confounding contributions of aging and hearing loss. The correlational results in the current study, however, revealed that high-frequency hearing thresholds better predicted the intelligibility of interrupted sentences than did age, consistent with correlational results reported in previous findings (Gordon-Salant and Fitzgibbons, 1993; KBL07). This stronger role of peripheral high-frequency hearing loss rather than age as a predictor has been well documented in various measures of speech understanding. Humes (2007) compared results of various speech measures, showing strong correlation between high-PTA and simple level-raised speech performance. Therefore, it seems inconclusive whether the poorer performance of elderly listeners occurred mainly from age-related changes reported in earlier studies by Bergman. We note that while age-related cognitive deficits have been found with time-compressed speech (Wingfield *et al.*, 1985; Gordon-Salant and Fitzgibbons, 2004; Wingfield *et al.*, 2006), those results do not appear to make clear predictions for our temporal interruption studies at normal speech rates. While these issues are complex, performance in this study was primarily predicted by hearing loss. Thus our approach of using a high presentation level was worthwhile for demonstrating that signal level was not sufficient to eliminate the negative effects of high-frequency audibility when EHI participants listened to interrupted speech.

Based on the previous findings, we expected that YNH listeners’ ability to integrate target sentences would be equivalent or even better from transition cues than from steady-state cues based on the importance of dynamic cues to vowel identification described by Strange [e.g., see Strange and Bohn (1998)]. On the other hand, EHI listeners were expected to benefit more from steady-state information compared to transition cues, based on age-related deficits in using dynamic cues discussed by Fox *et al.* (1992). No significant disadvantages to use four subsegmental cues was shown for our EHI participants, although we note that the high variability in EHI listener performance may have obscured possible subsegmental condition effects in this study. Our finding of no significant differences among steady-state and three transition regions of subsegmental cues to sentence intelligibility for both YNH and EHI listeners has some theoretical and clinical implications. Most of the previous studies on the role of dynamic transition cues primarily used monosyllabic words or nonsense words. Although this approach yielded well-controlled laboratory data that focused on the specific aims of each project, it was not known whether findings from word or syllable scores would generalize to sentence recognition because of various redundant cues preserved in sentences, as well as complex top-down processing used to comprehend incomplete sentences. Currently findings with syllables versus sentences do not correspond, suggesting that the contributions of transition cues found at the syllable or word level are different from contributions of transition cues to overall sentence recognition. Thus, future research on the importance of specific speech cues for EHI listeners should be established not only with words or syllables but also with

TABLE II. Examples of the four error types for incorrect words (phonetically matched word, MW; phonetically unmatched word, UW; phonetically matched pseudoword, MP; and phonetically unmatched pseudoword, UP).

Target sentence	Incorrect word response	Error type
<i>Her study</i> of history was persistently pursued.	<i>Or studies...</i> persistently pursued.	MW
But it was a hopeful <i>sign</i> , he <i>told</i> himself.	But it was a hopeful <i>time</i> , he <i>called</i> himself.	MW
<i>No one</i> will even suspect that it is your work.	<i>Tell them</i> we suspect that is your work.	UW
<i>In a</i> way, he couldn't blame her.	<i>There's no</i> way he couldn't blame her.	UW
<i>Instead</i> of that he was engulfed by bedlam.	<i>Insap...</i> that.	MP
What elements of our behavior are <i>decisive</i> ?	Behavior... <i>side-sive</i> ?	MP
But the problems cling to pools, as any pool owner <i>knows</i> .	Any pool owner <i>newvins</i> .	UP
But that <i>explanation</i> is only partly true.	That <i>lup-nik</i> is.....true.	UP

sentence materials, thereby avoiding overgeneralization of the role of speech cues at the word level to overall speech understanding by EHI listeners. A potential implication of our findings is that algorithms to enhance subsegmental cues focusing on transitions in hearing assistive devices for EHI listeners may be less beneficial than expected in clinical practice if evaluation materials exclude sentences.

C. Perceptual strategies in processing subsegmentally interrupted sentences

Two additional analyses were conducted, which detailed differences between the performance strategies of YNH and EHI listeners. Clearly, high-frequency hearing loss of EHI listeners causes a quantitative reduction in intelligibility of interrupted sentences. However, EHI listeners may also have used qualitatively different strategies to integrate information in interrupted sentences that might be related to age-related cognitive deficits. To determine if qualitatively different strategies were used, first the error distributions of incorrect words were compared. Second, the rankings of easier to harder sentences among the 42 test sentences across groups were examined for qualitative differences.

For the analysis of the distributions for incorrect words, we categorized incorrectly identified words into one of four error types. This analysis was based on a native American-English linguist's phonetic transcriptions of incorrect responses. The four types of errors for incorrect words were phonetically matched words (MWs), phonetically unmatched words (UWs), phonetically matched pseudowords (MPs), and phonetically unmatched pseudowords (UPs). The MW errors occurred when listeners' incorrect responses were apparently activated from meaningful words that sound similar to test words. For UW errors, listeners incorrectly responded with meaningful words; however their responses were pho-

netically dissimilar with the test words. MP and UP errors occurred when listeners' responses were meaningless pseudowords, as shown in Table II. If pseudoword errors were phonetically similar to the target, they were "matched" i.e., MP responses, while phonetically dissimilar responses, were "unmatched," i.e., UP. Table II displays examples of the incorrect word responses collected from listeners in the experiment corresponding to each of the four error types. Categorizing incorrect responses into meaningful words versus meaningless pseudowords was used to examine whether EHI listeners frequently guessed sounds they might have heard because the task was so hard for them. As shown in Table III, the order of the error distribution was similar across groups regardless of duration (i.e., MW > UW, MP > UP). Incorrect word responses for both listener groups occurred mostly from MWs, indicating that the strategies to recognize interrupted speech were very similar between YNH and EHI listeners as long as interrupted sentences were reasonably audible to listeners.

For the comparison of sentence rankings between groups, the 42 sentences were rank ordered based on correct words averaged, across conditions for each of the durations. Spearman rank correlation coefficients were computed within and across groups, as shown in Table IV. All sentence rankings were significantly and positively correlated between groups both across durations and within the same duration. Lower coefficients were found, as expected, when scores were either near floor for EHI-50 or near ceiling for YNH-70. However, a very high strength of correlation was found between 50% and 70% durations within each group ($r=0.80$, $p<0.01$ for YNH-50 and YNH-70; $r=0.94$, $p<0.01$ for EHI-50 and EHI-70). That is, the most understandable/difficult sentences with 50% duration were

TABLE III. Distribution (%) of four types of errors for the incorrect word responses pooled across conditions.

Group-duration (%)	Error type			
	MW (%)	UW (%)	MP (%)	UP (%)
YNH-50	71.2	25.5	2.8	0.5
EHI-50	67.3	31.6	0.9	0.2
YNH-70	75.0	19.8	4.8	0.4
EHI-70	70.2	27.1	2.0	0.7

TABLE IV. Spearman rank correlations (r) for the sentence ranking ($N=42$) between groups. The first column and row show the mean group performance [Mean=mean of words correct (%) across conditions and SE = standard errors]. Significant correlations are marked with $** (p < 0.01)$.

Group-Duration (%)	YNH-50	YNH0-70
	(Mean=60%; SE=3.3)	(Mean=85%; SE=2.1)
EHI-50	0.65**	0.43**
(Mean=16%; SE=4.8)		
EHI-70	0.71**	0.54**
(Mean=46%; SE=8.2)		

also the most understandable/difficult sentences with 70% duration for both YNH and EHI listeners. In addition, the correlations in Table IV reveal a strong relation ($r=0.71$) between sentence rankings for YNH-50 and EHI-70, indicating consistency of sentence rankings across two groups. Note that the significant but more moderate correlations in Table IV occurred where ceiling or floor performance was evident. Overall the Spearman correlations revealed that strategies for processing the interrupted sentences were quite similar across groups, regardless of hearing status.

To summarize, additional analyses revealed that YNH and EHI listener groups had similar error-type distributions across four error categories, as well as the positive relation between sentence rankings. This suggests that the negative effect of high-frequency hearing loss on the understanding of interrupted speech at a high presentation level resulted in quantitative, rather than qualitative, differences in performance between groups. These results further support that poorer performance of EHI listeners in processing interrupted sentences is unlikely to occur because of qualitatively different perceptual strategies that could be accompanied by age-related cognitive deficits but, rather, is largely caused by factors related to their peripheral hearing loss.

V. CONCLUSIONS

Following our previous study on sentence intelligibility with segmental interruption (Kewley-Port *et al.*, 2007), this study examined the ability of YNH and EHI listeners to comprehend interrupted sentences when different regions of subsegmental information (i.e., centers, margins, vowel onset and offset transition regions) preserved 50% or 70% of sentence duration. The major findings were as follows:

- (i) Despite high intelligibility of sentences without interruption, EHI listeners had less successful auditory integration of interrupted speech signals than YNH listeners, regardless of the region and the duration of subsegmental cues glimpsed.
- (ii) As expected, both groups performed better with a longer duration (70% versus 50% duration), resulting in improvement of 25% for YNH and of 30% for EHI listeners.
- (iii) Different types of subsegmental information had similar effects on intelligibility of interrupted sentences for both YNH and EHI listeners. Thus, dynamic transition cues do not benefit YNH listeners more than quasi-steady-state cues do, at least when the task involves the use of these cues to identify sentences from partial information. Apparently EHI listeners are not substantially impaired in using dynamic transition information compared to steady-state information preserved in interrupted sentences.
- (iv) Individual differences in word scores for EHI listeners were better predicted by their high-frequency hearing loss than by their age, despite the high presentation level of 95 dB SPL. This cautions us that a high presentation level does not ameliorate the negative effects of audibility in processing of interrupted sentences.

Additional analyses revealed similarity in error distributions of incorrect words and sentence rankings between groups. This indicates that reduced hearing at high frequencies may cause EHI listeners to have quantitatively worse performance than YNH but not qualitatively different perceptual strategies in processing interrupted sentences. Specifically, changes in perceptual strategies that might be attributed to age-related cognitive decline were not observed in EHI participants with our particularly challenging interrupted sentence task. Moreover, results from our two studies (current and KBL07) demonstrated that vowels contribute more to sentence intelligibility than did other cues for both YNH and EHI listener groups. Although reduced audibility of EHI listeners negatively affects their ability to integrate sentences interrupted with noise, preservation of vowel-only information compared to consonant-only or other subsegmental information has significant and substantial benefit for sentence understanding by EHI listeners. This motivates new ideas for the design of algorithms of speech processors for hearing aids, specifically to maximize the intelligibility of vowels. That is, algorithms for compensating hearing loss should preserve vowel information as much as possible in order to maximize possible resources that EHI listeners need when processing temporally interrupted speech information, a situation found in everyday listening environments.

ACKNOWLEDGMENTS

This research was supported by the National Institutes of Health Grant No. DC-02229 awarded to D.K.P. and the National Institute on Aging Grant No. AG022334 awarded to Dr. Larry E. Humes. The authors thank Brian W. Riordan for his assistance for the additional analysis of word scoring. They are also very appreciative of contributions made by Dr. Larry E. Humes to this project.

- ANSI (1969). "American National Standard Methods for Calculation of the Articulation Index," ANSI S3.5-1969, American National Standards Institute, New York, NY.
- ANSI (1996). "Specifications for audiometers," ANSI S3.6-1996, American National Standards Institute, New York, NY.
- Bergman, M. (1980). "Effects of physical aspects of the message," in *Aging and the Perception of Speech* (University Park, Baltimore), pp. 69–78.
- Bergman, M., Blumenfeld, V. G., Cascardo, D., Dash, B., Levitt, H., and Margulies, M. K. (1976). "Age-related decrement in hearing for speech: Sampling and longitudinal studies," *J. Gerontol.* **31**, 533–538.
- Cole, R. A., Yan, Y. H., Mak, B., Fany, M., and Bailey, T. (1996). "The contribution of consonants versus vowels to word recognition in fluent speech," in *Proceedings of the ICASSP'96*, pp. 853–856.
- Dorman, M. F., Marton, K., Hannley, M. T., and Lindholm, J. M. (1985). "Phonetic identification by elderly normal and hearing-impaired listeners," *J. Acoust. Soc. Am.* **77**, 664–670.
- 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.
- Fox, R. A., Wall, L. G., and Gokcen, J. (1992). "Age-related differences in processing dynamic information to identify vowel quality," *J. Speech Hear. Res.* **35**, 892–902.
- Garofolo, J. S., Lamel, L. F., Fisher, W. M., Fiscus, J. G., Pallett, D. S., and Dahlgren, N. L. (1990). "DARPA TIMIT acoustic-phonetic continuous speech corpus CDROM," National Institute of Standards and Technology, NTIS Order No. PB91-505065.
- Gordon-Salant, S. (2005). "Hearing loss and aging: New research findings and clinical implications," *J. Rehabil. Res. Dev. Clin. Suppl.* **42**, 9–24.
- Gordon-Salant, S., and Fitzgibbons, P. J. (1993). "Temporal factors and speech recognition performance in young and elderly listeners," *J. Speech*

- Hear. Res. **6**, 1276–1285.
- 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.
- 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.
- Iyer, N., Brungart, D. S., and Simpson, B. D. (2007). “Effects of periodic masker interruption on the intelligibility of interrupted speech,” *J. Acoust. Soc. Am.* **122**, 1693–1701.
- Jenkins, J. J., Strange, W., and Miranda, S. (1994). “Vowel identification in mixed-speaker silent-center syllables,” *J. Acoust. Soc. Am.* **95**, 1030–1043.
- Kewley-Port, D., Burkle, T. Z., and Lee, J. H. (2007). “Contribution of consonant versus vowel information to sentence intelligibility for young normal-hearing and elderly hearing-impaired listeners,” *J. Acoust. Soc. Am.* **122**, 2365–2375.
- Kirikae, I., Sato, T., and Shitara, T. (1964). “A study of hearing in advanced age,” *Laryngoscope* **74**, 205–220.
- Korsan-Bengtson, M. (1973). “Distorted speech audiometry: A methodological and clinical study,” *Acta Oto-Laryngol., Suppl.* **310**, 1–75.
- Li, N., and Loizou, P. C. (2007). “Factors influencing glimpsing of speech in noise,” *J. Acoust. Soc. Am.* **122**, 1165–1172.
- Liberman, A. M., Harris, K. S., Hoffman, H. S., and Griffith, B. C. (1957). “The discrimination of speech sounds within and across phoneme boundaries,” *J. Exp. Psychol.* **54**, 358–368.
- Miller, G. A., and Licklider, J. C. R. (1950). “The intelligibility of interrupted speech,” *J. Acoust. Soc. Am.* **22**, 167–173.
- Nelson, P. B., and Jin, S. H. (2004). “Factors affecting speech understanding in gated interference: Cochlear implant users and normal-hearing listeners,” *J. Acoust. Soc. Am.* **115**, 2286–2294.
- Ohde, R. N., and Abou-Khalil, R. (2001). “Age differences for stop-consonant and vowel perception in adults,” *J. Acoust. Soc. Am.* **110**, 2156–2166.
- Peterson, G. E., and Barney, H. L. (1952). “Control methods used in a study of the vowels,” *J. Acoust. Soc. Am.* **24**, 175–184.
- Powers, G. L., and Speaks, C. (1973). “Intelligibility of temporally interrupted speech,” *J. Acoust. Soc. Am.* **54**, 661–667.
- Redford, M. A., and Diehl, R. L. (1999). “The relative perceptual distinctiveness of initial and final consonants in CVC syllables,” *J. Acoust. Soc. Am.* **106**, 1555–1565.
- Strange, W. (1989). “Dynamic specification of coarticulated vowels spoken in sentence context,” *J. Acoust. Soc. Am.* **85**, 2135–2153.
- Strange, W., and Bohn, O. S. (1998). “Dynamic specification of coarticulated German vowels: Perceptual and acoustical studies,” *J. Acoust. Soc. Am.* **104**, 488–504.
- Strange, W., Verbrugge, R. R., Shankweiler, D. P., and Edman, T. R. (1976). “Consonant environment specifies vowel identity,” *J. Acoust. Soc. Am.* **60**, 213–224.
- Studebaker, G. A. (1985). “A ‘rationalized’ arcsine transform,” *J. Speech Hear. Res.* **28**, 455–462.
- Warren, R. M. (1970). “Perceptual restoration of missing speech sounds,” *Science* **23**, 392–393.
- Wingfield, A., McCoy, S. L., Peelle, J. E., Tun, P. A., and Cox, L. C. (2006). “Effects of adult aging and hearing loss on comprehension of rapid speech varying in syntactic complexity,” *J. Am. Acad. Audiol* **17**, 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.