# Age-related differences in identification and discrimination of temporal cues in speech segments

Sandra Gordon-Salant and Grace H. Yeni-Komshian
*Department of Hearing and Speech Sciences, University of Maryland, College Park, Maryland 20742*

Peter J. Fitzgibbons
*Department of Hearing, Speech, and Language Sciences, Gallaudet University, Washington, D.C. 20002*

Jessica Barrett
*Department of Hearing and Speech Sciences, University of Maryland, College Park, Maryland 20742*

(Received 30 August 2005; revised 10 January 2006; accepted 11 January 2006)

This study investigated age-related differences in sensitivity to temporal cues in modified natural speech sounds. Listeners included young noise-masked subjects, elderly normal-hearing subjects, and elderly hearing-impaired subjects. Four speech continua were presented to listeners, with stimuli from each continuum varying in a single temporal dimension. The acoustic cues varied in separate continua were voice-onset time, vowel duration, silence duration, and transition duration. In separate conditions, the listeners identified the word stimuli, discriminated two stimuli in a same-different paradigm, and discriminated two stimuli in a 3-interval, 2-alternative forced-choice procedure. Results showed age-related differences in the identification function crossover points for the continua that varied in silence duration and transition duration. All listeners demonstrated shorter difference limens (DLs) for the three-interval paradigm than the two-interval paradigm, with older hearing-impaired listeners showing larger DLs than the other listener groups for the silence duration cue. The findings support the general hypothesis that aging can influence the processing of specific temporal cues that are related to consonant manner distinctions. *© 2006 Acoustical Society of America.* [DOI: 10.1121/1.2171527]

PACS number(s): 43.71.Lz, 43.71.Es, 43.71.Gv, 43.66.Mk [JHG] Pages: 2455–2466

## I. INTRODUCTION

In recent years, compelling evidence has accumulated to show that elderly listeners experience difficulties in understanding speech with altered timing characteristics. For example, older listeners exhibit poorer recognition performance compared to younger listeners for sentences that are presented at an increased rate relative to natural-rate speech (Wingfield *et al.*, 1985; Gordon-Salant and Fitzgibbons, 1993; Vaughan and Letowski, 1997). Additionally, older listeners have more difficulty than younger listeners in accurately recognizing sentences with artificial disruptions in overall prosody (Gordon-Salant and Fitzgibbons, 1997).

One form of natural disruption in timing and prosody occurs in speech produced by non-native speakers of English. Accented English is characterized by numerous temporal alterations that modify individual phonemes as well as the overall timing of the sequence of syllables and words in a spoken message. Some specific temporal changes in accented English at the segmental level include variations in voice-onset time (VOT) (Flege and Eefting, 1988; MacKay *et al.*, 2000), syllable stress (Adams and Munro, 1978), and vowel duration (Fox *et al.*, 1995). Another alteration identified in accented English is overall sentence duration (Guion *et al.*, 1997). Disruptions at both the segmental and suprasegmental levels potentially act to vary the global accent of speech. At least one investigation has shown a significant age-related difference between younger (20–39 years) and older (60+years) adults in recognizing English spoken with different degrees of accentedness (Burda *et al.*, 2003). However, this prior study did not elucidate the source of the age-related deficit as deriving primarily from altered segmental characteristics or suprasegmental features.

Converging support for the notion that older listeners have a reduced capacity to process temporal changes in speech may come from psychoacoustic investigations with nonspeech stimuli. Older listeners demonstrate larger gap detection thresholds than younger listeners, particularly for gap intervals placed near the onset or offset of a noise marker (e.g., He *et al.*, 1999; Snell and Hu, 1999). They also show longer duration discrimination thresholds than younger listeners for tonal stimuli and silent intervals between tonal markers (Fitzgibbons and Gordon-Salant, 1994), as well as for silent intervals between noise-band markers of disparate frequencies (Lister *et al.*, 2002). One study has extended these results to show an effect of age on discrimination of silent gaps inserted in synthetic speech signals (Lister and Tarver, 2004). These findings suggest that identification and discrimination of discrete speech signals that are dependent on duration cues should reveal substantial age-related differences.

The overall objective of this preliminary investigation is to identify some age-related perceptual deficits for isolated temporal acoustic cues in consonants and vowels that are necessary for accurate word identification, and that may be characterized by considerable variation and/or deviation in non-native speech. The temporal acoustic cues selected for this study include vowel duration as a cue for final consonant

 0001-4966/2006/119(4)/2455/12/$22.50  

voicing (Denes, 1955; Peterson and Lehiste, 1960; Luce and Charles-Luce, 1985), voice-onset time (VOT) as a cue for initial stop consonant voicing (Lisker and Ambramson, 1964), glide duration as a cue for the stop-consonant vs glide distinction (Diehl, 1976), and duration of a silent interval to cue the fricative vs affricate distinction (Dorman *et al.* 1979). Relatively little is known about the ability of older listeners to perceive many of these temporal cues in natural speech. For synthetic speech continua with varying VOT, older listeners have been shown to perform differently than younger listeners on identification and discrimination tasks (Strouse *et al.*, 1998). Age-related differences in identification and discrimination of natural speech signals with variation in VOT and other temporal cues have not been reported previously.

A second purpose of the current study is to distinguish age-related difficulties from those attributed to hearing impairment. The primary research questions of interest relate to whether or not elderly listeners have difficulties with temporal cues in speech as a consequence of reduced audibility of high frequency sounds or a diminished capacity to discriminate the duration of speech cues. Additionally, prior studies with nonspeech signals suggest that specific stimulus parameters may be necessary to reveal effects of age and hearing impairment in temporal judgments. For example, older listeners show larger discrimination thresholds for unfilled signals (e.g., silent gaps between tones) than for filled signals (e.g., tones) (Fitzgibbons and Gordon-Salant, 1994). Additionally, hearing loss effects have been observed for stimuli with shorter reference durations but not for stimuli with longer reference durations (Fitzgibbons and Gordon-Salant, 1994). Previous studies have suggested also that identification performance may be predictable from discrimination performance, at least for judgments about tonal temporal order in comparable sequences (Fitzgibbons and Gordon-Salant, 1998). The extent to which these principles apply to older listener's discrimination and recognition of natural speech is unclear at present.

To investigate these issues, duration judgments for selected temporal cues for vowels and consonants were measured. Stimulus parameters include reference duration and type of segmental cue (filled, as in voiced speech segments, vs unfilled, as in affricate silence duration). Discrimination and identification of these temporal cues for single word stimuli were evaluated in separate conditions. The acoustic cues selected for the present investigation include VOT as a cue for initial stop consonant voicing, vowel duration as a cue for final stop consonant voicing, silence duration as a cue to distinguish sibilant from affricate, and transition duration as a cue for the stop vs glide distinction. Listener groups include elderly listeners with normal hearing, elderly listeners with hearing impairment, and noise-masked young listeners in which the noise masking is sufficient to equate thresholds with those of the elderly listeners with normal hearing. Comparisons of performance between the noise-masked young listeners and elderly listeners with normal hearing provide an estimate of age-related effects, and comparisons of performance between the elderly normal-hearing listeners and elderly hearing-impaired listeners provide an estimate of the effects of hearing impairment. Thus, if hearing loss is the primary source of the problem, then the elderly hearing-impaired listeners should perform more poorly than the young and elderly normal-hearing listeners, and if age is the primary source of the problem, then the elderly hearing-impaired and elderly normal-hearing listeners should perform more poorly than the young normal-hearing listeners with equivalent masked thresholds to the elderly normal-hearing group. Based on previous findings of substantial age-related deficits in duration discrimination for non-speech signals, and the importance of processing temporal cues for speech understanding, we hypothesize that older listeners will exhibit a deficit compared to younger listeners, on identification and discrimination of temporal cues in individual speech segments. Moreover, it is expected that for all listeners, perception of single speech segment durations will be poorer for speech segments of shorter reference durations than longer reference durations.

## II. METHOD

### A. Subjects

One objective of this study was to separate the effects of hearing loss from the effects of age on perception of temporal differences in speech signals. To that end, three listener groups participated in the experiments. The first group was elderly listeners (64–77 years, mean age=71 years; $n=16$) with normal hearing, defined as pure tone thresholds $\leq 25$ dB HL, from 250 to 4000 Hz (see ANSI, 1996). The second group consisted of elderly listeners (64–80 years; mean age=72.5 years; $n=15$) with mild-to-moderate or mild-to-moderately severe, sloping sensorineural hearing loss and good or excellent monosyllabic word recognition scores ($>80\%$). The third group of listeners was comprised of young listeners (18–33 years, mean age=21.4 years; $n=15$) with normal hearing sensitivity. Because the thresholds of the young listeners were considerably better than those of the older listeners with normal hearing, shaped noise masking was presented to the young listeners to shift their thresholds to be equivalent (within 5 dB), on average, to those of the elderly listeners with normal hearing. The average thresholds that served as targets were based on mean thresholds of elderly normal-hearing listeners who participated in several previous investigations. The software program COOL EDIT PRO (v. 2.0, Syntrillium Software; Graphic Equalizer Filter Option) was used to create the shaped noise, which was then burned onto a compact disc for presentation during the preliminary audiometric tests for verification of the shifted thresholds. This shaped noise also was presented to the young normal-hearing listeners throughout the experiments. Table I presents the average pure tone thresholds (250–4000 Hz) of the three listener groups.

Other preliminary audiometric criteria included normal middle ear function, as assessed by tympanometry, and acoustic reflex thresholds that were within the 90th percentile for a given pure tone threshold (Gelfand *et al.*, 1990). All listeners were tested using transient-evoked otoacoustic emissions (TEOAEs) to document their presence in listeners with normal hearing and their absence in listeners with hear-

TABLE I. Pure tone thresholds in dB HL (see ANSI 1996) in the test ear of the three subject groups. Data shown are group means, with standard deviations in parentheses.

| Group | Frequency (Hz) | | | | |
|---|---|---|---|---|---|
| | 250 | 500 | 1000 | 2000 | 4000 |
| Young masked | 14.7 (5.2) | 13.3 (3.6) | 12.0 (3.2) | 14.7 (3.5) | 20.0 (4.6) |
| Elderly normal | 12.8 (7.5) | 10.0 (6.3) | 10.9 (6.9) | 14.4 (7.3) | 23.1 (6.0) |
| Elderly hearing impaired | 16.7 (7.9) | 19.7 (10.1) | 24.3 (11.8) | 38.0 (10.5) | 52.3 (7.0) |

ing impairment. The absence of TEOAEs and the presence of acoustic reflex thresholds at expected levels confirmed a cochlear site of lesion among the listeners with hearing loss.

There were several additional criteria for subject selection. All listeners were required to be native speakers of English. They were in general good health, with no history of stroke, Parkinson's disease, or neurologic impairment, and possessed sufficient motor skills to provide a response using a computer mouse click. Additionally, all listeners passed a screening test for general cognitive awareness (Pfeiffer, 1977).

### B. Stimuli

Four sets of contrasting word pairs were selected from the Revised Speech Perception in Noise Test (R-SPIN; Bilger *et al.*, 1984). Each word pair contrast relied on a single acoustic duration cue. A continuum was constructed for each pair that varied along the relevant single acoustic duration cue. An adult American male speaker with a General American dialect recorded the target word pairs in isolation as well as the low probability R-SPIN sentences that contained the eight target words. Recordings were made in a quiet room directly onto a PC, using a microphone and sound card with 16-bit resolution (Audigy Sound Blaster). The stimuli were analyzed and edited from the original recordings using waveform editing software (COOL EDIT PRO 2.0 at 44.1 kHz sampling rate). Additionally, the sound spectrogram display of a second waveform editing program, WEDW (available online at www.asel.udel.edu), was used to verify the temporal characteristics of the modified stimuli.

In this study, four pairs of contrasting words produced in isolation were used to generate the target speech continua. The four sets of contrasting words and the associated single acoustic duration cue included: (a) BUY/PIE, varying in VOT; (b) WHEAT/WEED, varying in vowel duration; (c) DISH/DITCH, varying in closure duration between the initial CV and the final consonant; and (d) BEAT/WHEAT, varying in transition duration of the initial consonant. Continua of each of these contrasting word pairs were created from one of the endpoint natural productions, by varying the relevant duration cue in equal intervals. Pilot data using these stimuli (described below) were collected from 11 young normal-hearing listeners, and verified that the endpoint signals from each continuum were clearly identified as the intended targets.

For the BUY/PIE continuum, the unmodified natural endpoint of the continuum was BUY (0 ms VOT, Stimulus #1). To create Stimulus #2, 10 ms of aspiration was excised from the natural PIE sample and inserted at a zero-crossing between the burst release and the onset of voicing of the original BUY stimulus. Aspiration was used, rather than a silent gap, because natural productions of voiceless plosives in English are aspirated (Klatt, 1975). Stimulus #3 was created by inserting an additional 10 ms of aspiration to stimulus #2 immediately after the burst and contiguous with the onset of aspiration from stimulus #2. This process continued until the creation of Stimulus #7, which had a VOT of 60 ms. Thus, stimuli on this continuum varied in VOT in 7 steps separated by 10 ms intervals, from 0 ms (BUY) to 60 ms (PIE). The duration of the endpoint stimulus, BUY, was 310 ms. The duration of Stimulus #7 endpoint, PIE, was 370 ms.

The WHEAT/WEED continuum is based on vowel duration. Several steps were taken in creating this continuum. The endpoint stimulus was a hybrid, consisting of a natural token of WEED, in which the /d/ release was excised and replaced by a high amplitude release taken from the final burst of a natural token of WHEAT. This stimulus was perceived as WEED, and represents the token with the longest vowel duration of the continuum. The next stimulus in the continuum had a shorter vowel duration and was created by cutting a single pulse between two zero-crossings from the steady-state portion of the preceding stimulus. Additional stimuli were created by excising single pulses from the steady-state portion of the preceding stimulus on the continuum. Each excised pulse was 7–8 ms in duration. Thus, the WHEAT/WEED continuum varied vowel duration in discrete steps corresponding to one glottal pulse ($\Delta t = 7-8$ ms). A 9-step continuum, rather than a 7-step continuum, was created to establish reliable endpoints in the pilot study. The endpoint signal with the shortest vowel (perceived as WHEAT) had a total duration of 249 ms (93 ms steady-state vowel duration), whereas the endpoint signal with the longest vowel (perceived as WEED) had a total duration of 311 ms (steady-state vowel duration of 155 ms).

The DISH/DITCH continuum is based on the duration of the silent interval between the vowel and the fricative /ʃ/. The endpoint stimulus was a hybrid in which the initial stop, vowel, and closure duration, were taken from the natural token, DITCH. The burst and fricative of /tʃ/ were excised and replaced by the fricative /ʃ/ of a natural token of DISH. Thus, this stimulus, which was perceived as DITCH, consisted of an initial stop, a vowel, closure, and a final fricative, and represents the longest stimulus on the continuum. Subsequent stimuli were created by excising 10 ms of silence from the closure period. There were seven steps on this continuum: the shortest stimulus, DISH, had a 0 ms closure duration, whereas the longest stimulus, DITCH, had a closure duration of 60 ms. The total duration of the shortest stimulus, DISH, was 483 ms, and that of DITCH was 543 ms.

The final continuum, BEAT/WHEAT, varied the transition duration of the initial consonant. The natural production of WHEAT, minus the first 50 ms following stimulus onset, constituted one endpoint of the continuum. This stimulus was clearly perceived as "Wheat" by a group of pilot listeners. The duration of the transition, measured from the onset of the stimulus to the steady-state location of the vowel, was 51 ms. The next stimulus was created by excising one glottal pulse of 7–8 ms, taken from the onset of the preceding stimulus. Subsequent stimuli had additional pulses of 7–8 ms removed from the onset of the preceding stimulus, to shorten the duration of the initial transition. There were 7 steps in this continuum. The transition durations of BEAT and WHEAT were 7 and 51 ms, respectively. The endpoint signal, BEAT, had a total duration of 247 ms, and the endpoint signal, WHEAT, had a total duration of 289 ms.

After all of the stimuli were created for each continuum, the rms levels of each stimulus were analyzed. The stimuli were then scaled in level to be equivalent while avoiding any peak-clipping of any single stimulus. A calibration tone was also created that was equivalent in level to the rms levels of the stimuli. Each stimulus from each continuum was stored on the computer for later retrieval during online stimulus presentation and data collection.

### C. Procedure

The four stimulus continua were presented to the listeners using three different experimental paradigms: identification, two-interval discrimination, and three-interval discrimination. The identification and two-interval (same/different) discrimination tasks were designed to be somewhat comparable to standard identification and discrimination experiments using speech continua (e.g., Pisoni and Lazarus, 1974; Strouse *et al.*, 1998). The three-interval discrimination task was intended to be comparable to the paradigm typically used in psychoacoustic experiments conducted in our laboratory.

Each of the experimental paradigms was implemented using experiment generator and controller software (ECos2a, Avaaz Innovations). The identification paradigm followed a standard procedure in which the stimuli within each continuum were presented in 10 blocks, with a random order presented in each block, for a total of 10 presentations of each stimulus. Following each stimulus presentation, the computer monitor displayed the written words of the two endpoint stimuli, with each written word occupying one half of the screen in a vertical display. Listener responses were self-paced, and consisted of placing a mouse arrow on either the left or right half of the monitor to identify the stimulus word (using the left mouse click). Listeners were encouraged to guess if they were unsure of the stimulus perceived. The interval between the listener's response and the subsequent stimulus presentation was 4 s. Pilot testing with several older listeners indicated that the computer response was accomplished easily. There were four identification conditions, corresponding to the four stimulus continua.

The two-interval discrimination paradigm (A-X) presented two stimuli on each trial with a 500-ms interobservation interval, and required listeners to judge whether the two stimuli were the same or different. The first stimulus was the standard endpoint stimulus and the second stimulus was the comparison stimulus. The standard stimulus was chosen to be the shortest endpoint stimulus for each continuum. The comparison stimulus presented in the initial trial was the longest possible comparison stimulus relative to the standard (i.e., the opposite endpoint stimulus on the continuum). The listener's task was to identify whether the two stimuli were the same or different, by clicking on one of these two terms displayed on a computer monitor using the computer mouse. Feedback was not provided. Following two correct responses, the comparison stimulus selected was one step closer to the standard stimulus in the continuum. The procedure continued following the adaptive rule, which stipulated a decrease in stimulus duration following two consecutive correct responses and an increase in stimulus duration following each incorrect response (Levitt, 1971). Each change in duration was always one step, with the step size fixed for each continuum, as described previously. This adaptive procedure continued until there were eight reversals in stimulus selection. The mean of the midpoint of the final four reversals was calculated to determine the discrimination threshold of 70.7% correct discrimination. The stimulus pairs included two identical stimuli (i.e., "catch" trials) in order to facilitate threshold determination for listeners who displayed perfect discrimination performance. If the listener displayed perfect performance (i.e., they identified Stimuli 1 and 2 as different, and correctly identified the two identical endpoint stimuli on the catch trials as the same), then a nominal discrimination threshold of 1 (as in step #1, JND=0 steps) was assigned to that run. Three of these trial-run threshold estimates were collected from each subject for each of the four, two-interval discrimination conditions. An analysis was made of the threshold results to determine if stable performance was achieved between trial runs 2 and 3. Stable performance was defined as two threshold estimates that varied by less than 1 step. If stable performance was not observed, additional trial runs were conducted in an effort to observe stable performance, with a maximum of five trial runs presented. Thresholds values, collected in relative step sizes, were later converted to ms, based on the duration corresponding to the step size for each continuum.

The adaptive three-interval paradigm is a three-interval cued two-alternative forced-choice discrimination procedure. Each listening trial consisted of three observation intervals with an interobservation interval of 500 ms. The first interval on each trial contained the reference signal for a given condition, and the second and third intervals contained the reference and comparison signals in either order with equal probability. The listening intervals of each trial were marked with a visual display. Subjects responded with a mouse click to identify the interval (either 2 or 3) corresponding to the stimulus that was different from the standard stimulus presented in the first interval. The adaptive rule described for the two-interval discrimination paradigm was followed for the three-interval discrimination paradigm to derive the 70.7% correct discrimination threshold. A minimum of three and a maximum of five trial-run threshold estimates were presented to enable individual listeners to reach stable performance.

The order of the three paradigms (identification, two-interval discrimination, three-interval discrimination) was randomized across subjects. A strategy was followed to facilitate the listener's retention of the task associated with each paradigm, and to avoid confusion. To that end, all of the speech continua were presented in a given paradigm before introducing the next paradigm. Additionally, the order of presentation of the four speech continua was randomized for each paradigm, and with a different randomization for each subject. Prior to commencing each experimental paradigm, listeners were provided with some listening practice with the same paradigm, using a novel continuum of speech stimuli. During the experiments, the stimuli were routed from the laboratory computer and sound card to an amplifier (Crown D75A), an audio mixer-amplifier (Colbourn S82-24), and delivered to the listener's better ear through a single insert earphone (Etymotic ER3A) at 85 dB SPL. For the noise-masked young listeners, the CD containing the shaped noise was played back on a compact disk player (Tascam CD-160), amplified (Crown D-75A), attenuated (Hewlett-Packard 350 D), and mixed through the audio-mixer amplifier with the speech signal. The masking noise was presented at an overall level of 50 dB SPL, which was determined to be sufficient to shift the listener's thresholds to the desired levels on the basis of pilot testing. The stimuli and masking noise were calibrated daily.

Listeners were seated in a sound-attenuating chamber for all of the experiments. They were tested in 2 h sessions, with frequent breaks. The entire procedure, including preliminary audiometric assessment and practice sessions, was completed in approximately 6 h. Listeners were reimbursed for their participation in the experiment.

## III. RESULTS

Each listener's identification functions were reviewed to ensure that both target endpoint stimuli for each continuum were identified at a probability higher than chance level. For three of the continua, a few listeners failed to identify both endpoint stimuli on the continuum. The data for these listeners in these conditions were removed from subsequent data analysis (2 listeners for BUY/PIE, 3 listeners for WHEAT/WEED, and 3 listeners for BEAT/WHEAT; these listeners

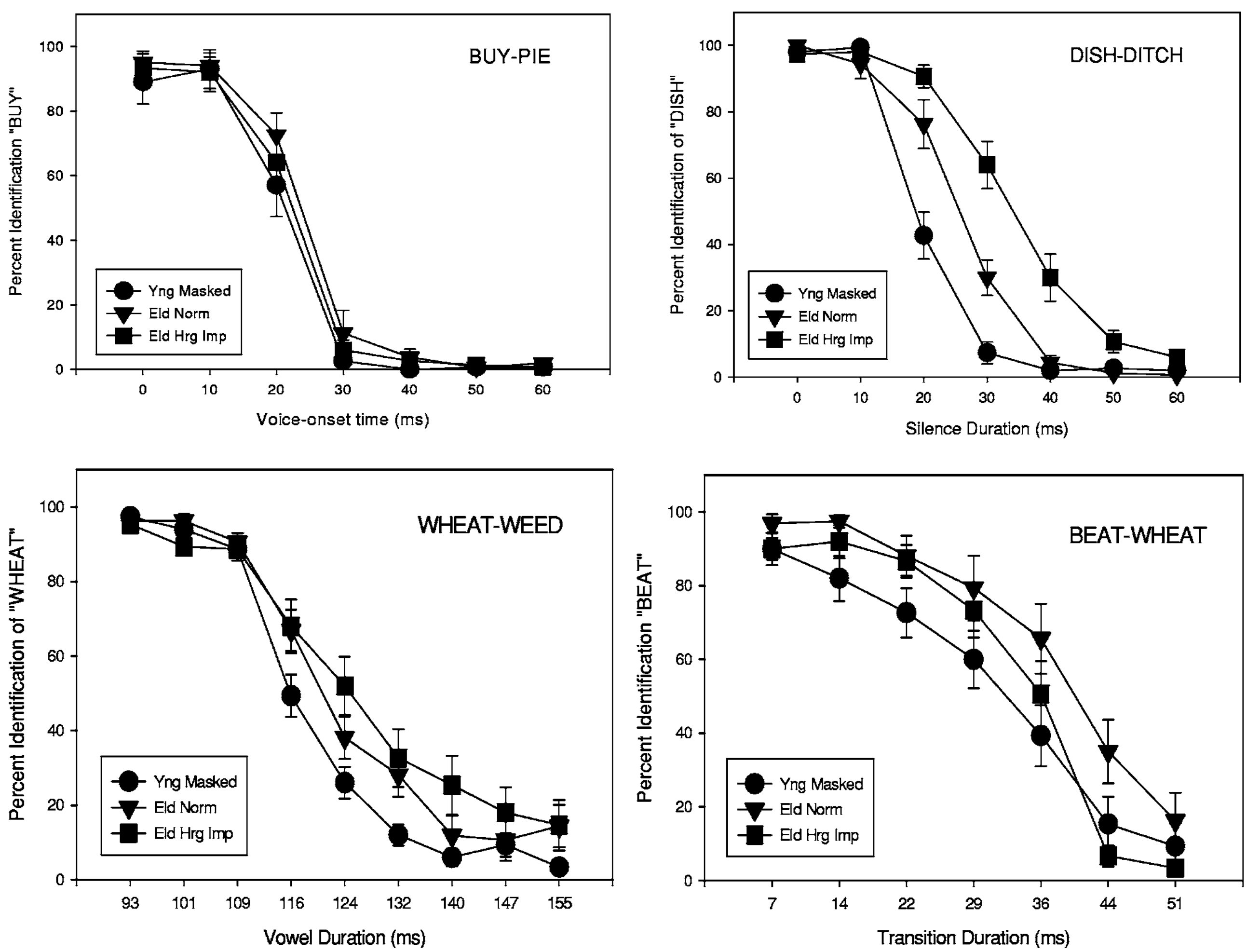


FIG. 1. Identification functions for the three listener groups for the BUY-PIE continuum (top left panel), DISH-DITCH continuum (top right panel), WHEAT-WEED continuum (bottom left panel), and BEAT-WHEAT continuum (bottom right panel). Error bars show one standard error of the mean.

TABLE II. Crossover points for the relevant acoustic cue, in ms, of the three subject groups for the four speech continua. Data shown are group means, with standard deviations in parentheses.

| | Speech continuum | | | |
|---|---|---|---|---|
| Group | BUY/ PIE | DISH/ DITCH | WHEAT/ WEED | BEAT/ WHEAT |
| Young masked | 21.39 (4.37) | 20.00 (4.15) | 118.78 (3.34) | 28.56 (10.92) |
| Elderly normal | 22.91 (4.43) | 25.31 (6.24) | 121.57 (5.56) | 37.33 (7.69) |
| Elderly hearing impaired | 22.36 (3.43) | 34.77 (6.00) | 122.97 (7.83) | 33.87 (6.69) |

were distributed across the three groups). The average identification functions of the remaining listeners from the three groups for each of the four speech continua are shown in Fig. 1. Individual listener identification functions for each continuum were analyzed separately for the slope (a) and $y$-intercept (b) values by performing a linear regression using data points on the linear portion of the function (between 80% and 20% correct, approximately). The number of points used to perform the linear regression varied between subjects both within and across continua. The 50% crossover point, $x$, for each function was subsequently calculated using the formula

$$x=(50-b)/a.$$

The average crossover points for the three subject groups across the four speech continua are shown in Table II. Separate one-way analyses of variance (ANOVA) were conducted for each speech continuum to determine the effect of listener group on the crossover points. Significant group effects were observed for the DISH/DITCH ($p<0.01$) and BEAT/WHEAT ($p<0.05$) continua, but not for the BUY/PIE or the WHEAT/WEED continua ($p>0.05$). Follow-up analyses of the significant group effects were conducted with the Bonnferoni method. The crossover values were significantly different between each pair of groups for the DISH/DITCH continuum. For the BEAT/WHEAT continuum, the crossover values were significantly different for the young noise-masked listeners and the elderly normal-hearing listeners.

The slope data for the three subject groups, derived from the identification functions, are shown in Table III. A higher slope value suggests a clearer distinction between one speech sound and the contrasting speech sound in a continuum. A review of Table III suggests that large standard deviations are observed in the slope data, particularly for the BUY/PIE and BEAT/WHEAT continua. ANOVAs were conducted on the slope values, using a one-way design, and the results failed to reveal a significant main effect of listener group for the BUY/PIE, WHEAT/WEED, and BEAT/WHEAT continua. However, the group effect was significant for the DISH/DITCH continuum, with the young noise-masked listeners showing a larger slope value than the elderly hearing-impaired listener group ($p<0.05$). In addition, the overall rank order across the BUY/PIE and WHEAT/WEED continua follows a similar pattern in which the young noise-masked group had steeper slopes than the two elderly groups.

Initial analysis of the discrimination results was directed toward evaluation of the reliability of the data. To that end, thresholds of all subjects were compared across the trial blocks for each condition. For most subjects, no significant changes in discrimination performance were observed after three trial blocks. A few subjects required up to five practice blocks before stable performance was observed, but there were no observable differences in the training required of young and elderly subjects. The discrimination threshold for each subject was taken as the average threshold from the final three trial blocks, which is a comparable procedure to that used in our psychoacoustic experiments. Discrimination thresholds were converted to difference limens (DLs), relative to the duration of the reference signal.

As data were collected, it became apparent that the discrimination thresholds were quite different for the two- vs the three-interval procedures, for all listener groups. Figure 2 presents the DLs obtained in the two procedures for the three listener groups, across the four different speech continua. Repeated measures ANOVAs were conducted separately for each continuum using the DL data, with one within-subjects factor (discrimination procedure) and one between-subjects

TABLE III. Average absolute slope values (% change in recognition performance/step; step size varies for each continuum) for the three subject groups for the four speech continua. Standard deviations are shown in parentheses.

| | Speech continuum | | | |
|---|---|---|---|---|
| Group | BUY/ PIE | DISH/ DITCH | WHEAT/ WEED | BEAT/ WHEAT |
| Young masked | 69.1 (24.1) | 59.0 (18.1) | 39.2 (17.6) | 36.5 (21.7) |
| Elderly normal | 67.7 (26.8) | 48.0 (12.4) | 29.6 (14.8) | 47.2 (15.7) |
| Elderly hearing impaired | 57.6 (25.2) | 40.8 (19.6) | 26.8 (10.8) | 48.7 (25.7) |

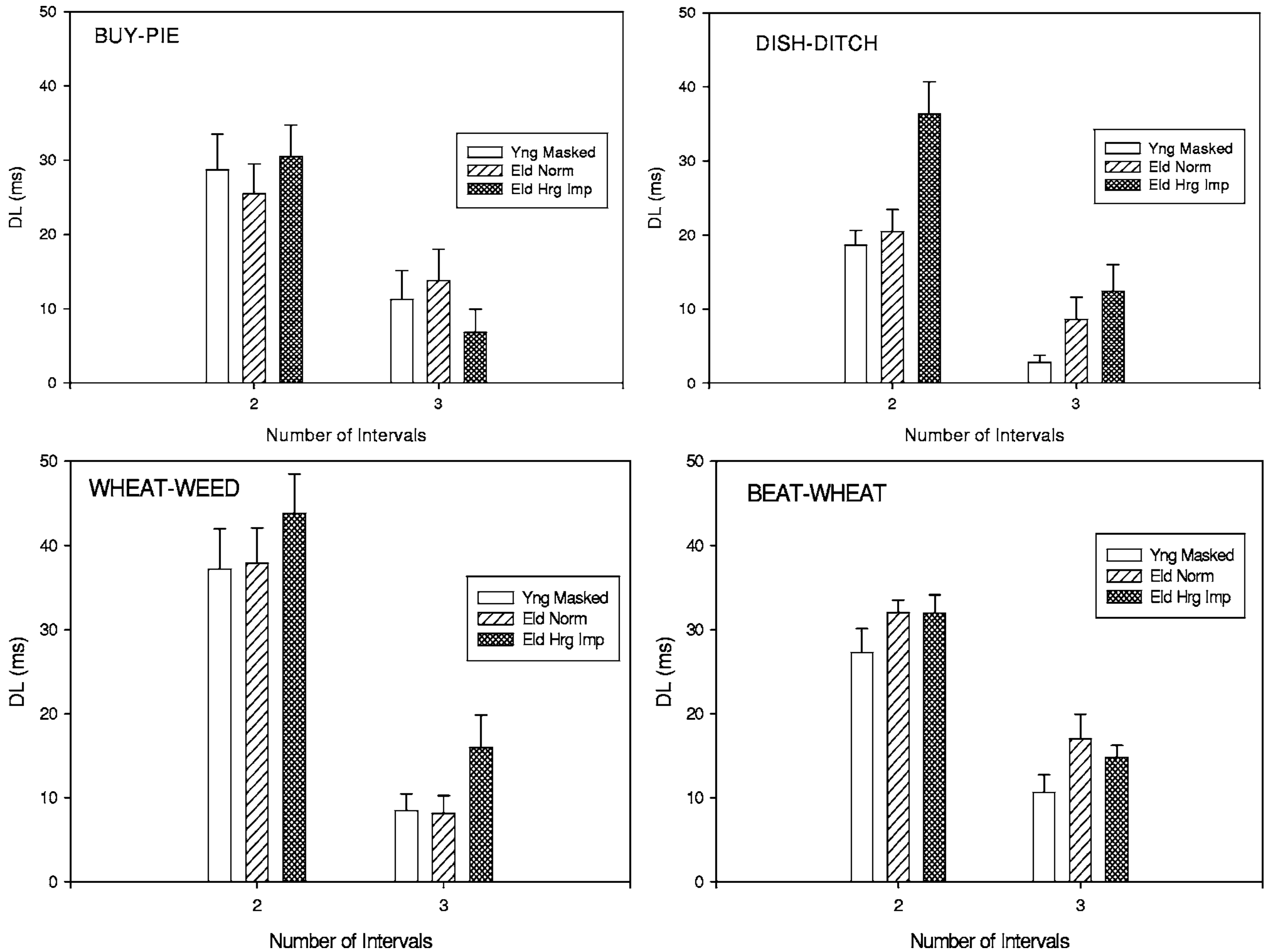


FIG. 2. Mean discrimination thresholds (DLs), in ms, for the three listener groups in the two-interval same-different paradigm and the three-interval, cued, two-alternative forced choice paradigm, for the BUY-PIE continuum (top left panel), DISH-DITCH continuum (top right panel), WHEAT-WEED continuum (bottom left panel), and BEAT-WHEAT continuum (bottom right panel). Error bars show one standard error of the mean.

factor (listener group). Results revealed a significant main effect of procedure (number of intervals) for all four speech continua ($p < 0.01$), with smaller DLs measured consistently for the three-interval procedure compared to the two-interval procedure. The main effect of listener group was significant for the DISH/DITCH continuum only ($p < 0.01$). There were no interactions between listener group and procedure for this continuum. Multiple comparison testing indicated that the elderly hearing-impaired group had higher average DLs than the young noise-masked group and the elderly normal-hearing group, on this continuum.

## IV. DISCUSSION

### A. Identification performance

The principal experimental question concerned whether or not age-related differences would be observed on identification judgments of modified natural speech tokens that varied along a single temporal parameter to represent two phonetic categories. The tokens of four different contrasting speech continua were presented in an effort to identify specific temporal cues that are most sensitive to auditory aging. Results from the varying experiments generally support the notion that age does influence perception of specific temporal acoustic cues in natural speech tokens, but these findings were not uniform across the different speech continua.

The DISH/DITCH continuum revealed the strongest age-related findings. The variable duration cue in this continuum was the silent interval that preceded the final sibiliant /ʃ/, as a cue for the sibilant/affricate distinction, ranging from 0 ms (/ʃ/) to 60 ms (/tʃ/). The total duration of the endpoint stimulus DISH was 483 ms, and was the longest reference duration of all of the continua presented in this experiment. Crossover values for the young masked, elderly normal-hearing, and elderly hearing-impaired listeners were 20 ms, 25.31 ms, and 34.77 ms, respectively. These findings provide additional information regarding phonetic boundaries, and are consistent with the 30 ms boundary reported previously for young normal-hearing listeners, in judging stimuli from a SHOP/CHOP continuum that varied the duration of a silent interval inserted between the offset of the carrier phrase "please say" and the onset of the sibilant /ʃ/ in "shop" (Dorman *et al.*, 1979). Statistical analyses of the current data showed that the crossover values for the young listeners were significantly different from those of the two elderly groups, confirming the presence of a significant age effect. Thus, elderly listeners required a longer silent interval than younger listeners to identify the stimulus as DITCH. This finding supports the hypothesis that elderly listeners require longer duration acoustic cues than younger listeners for accurate recognition of target speech signals. One possible reason for the identification performance differences between the younger and older listeners was the use of noise masking for the younger listeners, which might have affected their perception of the silent interval cue distinguishing DISH from DITCH. To address this issue, an analysis was conducted to compare the identification performance of the 15 young listeners with noise masking to that of 11 young listeners with normal hearing (without masking) who participated in pilot testing with the same stimuli and task. The results revealed no significant differences in either crossover points ($t=0.803, p>0.10$) or slopes ($t=0.14, p>0.10$). Thus, it appears that the low level of noise masking presented to the young listeners did not affect their perception of the stop closure that distinguished stimuli along the DISH/DITCH continuum. Additionally, the crossover values for the elderly hearing-impaired group were significantly larger than those for the elderly normal-hearing group, indicating that the presence of hearing loss among elderly listeners also affected performance. A temporal processing deficit was not predicted for the hearing-impaired listeners, particularly for relatively long baseline duration stimuli. However, the target acoustic cues in the DISH/DITCH distinction involve relatively high frequency, weak energy signals that often are not perceived accurately by hearing-impaired listeners (Owens *et al.*, 1972). It is possible that the distorted perception of the final sibilant by the elderly hearing-impaired listeners obscured their ability to utilize the silent interval cue effectively for the sibilant/affricate distinction. This suggestion is supported somewhat by the finding of significant correlations between the DISH/DITCH crossover points and the listeners' high frequency thresholds ($r=0.558, p<0.01$ at 2000 Hz; $r=0.718, p<0.01$ at 4000 Hz). The slope data for the DISH/DITCH continuum suggest that the young listeners demonstrated steeper identification function slopes (59.0) than those observed for the elderly normal-hearing listeners (48.0) and elderly hearing-impaired listeners (40.8). Statistical analyses partially supported this observation, and revealed a significant group effect in which the slopes of the young noise-masked listeners were significantly steeper than those observed for the elderly hearing-impaired listeners. The group effects shown for both the crossover points and the slopes for this continuum strongly suggest that older listeners with and without hearing impairment have difficulty perceiving the silence duration cue.

The present finding of an age effect for phoneme category boundaries on a continuum that varies silence duration is somewhat different from results reported by Dorman and co-workers (1985). In that investigation, listeners identified stimuli in a "slit-split" continuum in which the duration of a silent interval inserted between the sibilant noise and the vocalic portion was varied from 20 to 120 ms. Young normal-hearing listeners showed a shorter category boundary than elderly hearing-impaired listeners; however, the variability in performance of the elderly normal-hearing and elderly hearing-impaired listener groups obscured possible age effects in that investigation. In another investigation, which made use of a "rabid" to "rapid" 4-step continuum, older listeners required longer silence durations than younger listeners to perceive the "rapid" tokens (Price and Simon, 1984). The range of silence closure employed in this 4-step continuum was 35–125 ms. Thus, there appears to be converging evidence from several different investigations and different speech samples that elderly listeners require longer silent intervals to perceptually classify stimuli that are cued by a relatively longer silent interval.

Age-related differences were observed for the BEAT/WHEAT continuum. The varying acoustic parameter in this continuum was the duration of the formant transitions, which serves as a cue for the voiced stop/glide distinction. Liberman *et al.* (1956) showed that listeners perceived a voiced stop for formant transitions <40 ms, and perceived a glide for formant transitions >40 ms, in a continuum of synthetic tokens of /*bε*/ to /*wε*/. In the current experiment, the formant transition duration for the natural token BEAT was 7 ms, and that of WHEAT was 51 ms. Crossover points were 28.56 ms (young noise-masked listeners), 37.33 ms (elderly normal-hearing listeners), and 33.87 ms (elderly hearing-impaired listeners), suggesting that the change in percept from the voiced stop to the glide occurred at relatively brief formant transition durations for the young listeners, and that the elderly listeners required a longer formant transition duration to perceive a glide. Statistical analyses indicated a significant age effect, with the young noise-masked listeners exhibiting smaller crossover values than the elderly normal-hearing listeners. This finding also tends to support the hypothesis of an age-related deficit in processing brief temporal cues for speech. The slope data for the BEAT/WHEAT continuum did not show significant group effects, suggesting that the change in perceived category boundaries occurred at a similar rate for the three groups, despite a difference in the crossover value required by younger and older listeners to alter their stimulus identification. Dorman *et al.* (1985) did not observe an age effect for identification judgments of synthetic stimuli that constituted a /bæ/ to /wæ/ continuum, created by varying the duration of the initial formant transitions from 40 to 90 ms. The stimuli in the Dorman *et al.* synthetic continuum varied in several important ways from the natural stimuli used in the current experiment, including the type of stimuli (synthetic vs natural), number of formants (two formants in the Dorman *et al.* study vs three or more formants in the present study) and the range in duration of the formant transitions (40–90 ms in the Dorman *et al.* study vs 7–51 ms in the present study).

Group effects were not observed for the BUY/PIE and the WHEAT/WEED continua. The temporal cue that varied in the BUY/PIE continuum was VOT, or the interval from the burst release to the onset of voicing, which ranged from 0 to 60 ms. Identification functions of the three listener groups revealed crossover points of approximately 21 ms, 23 ms, and 22 ms for the young noise-masked, elderly normal, and elderly hearing-impaired listeners, respectively, which are generally consistent with those observed in other studies with young normal-hearing listeners where the VOT continuum consisted of bilabial stops (Lisker and Abramson, 1967; Pisoni and Lazarus, 1974). Statistical analyses showed no group differences for the identification function crossover points and the slopes. Strouse *et al.* (1998) also did not find an age effect for phonetic boundaries in identification functions for a synthetic /ba/ to /pa/ continuum varying in VOT, although slope values were significantly different between younger and older groups. The divergent findings regarding the slope values between the Strouse *et al.* investigation and the present study could be attributed, at least in part, to the use of synthetic vs natural speech in creating the VOT continua. Additionally, the large variance in slope values for the BUY/PIE continuum in the present investigation may have been great enough to obscure possible differences in the means of these groups.

The WHEAT/WEED continuum was comprised of stimuli with varying vowel duration to cue final stop consonant voicing. Peterson and Lehiste (1960) reported that vowel duration in naturally produced English words averaged 197 ms preceding a voiceless final stop consonant and 297 ms preceding a voiced stop, with a ratio of vowel before voiceless consonant to vowel before voiced consonant of 2:3. The perceptual relevance of vowel duration as a cue to word final voicing has been shown for word-final stops, fricatives, and consonant clusters (Raphael, 1972). The absolute vowel durations required for a change in the percept of final consonant voicing depend on the specific vowel used; hence, the vowel durations in the current continuum are not readily comparable to those reported in other studies. For example, the vowel in "kit" vs "kid" is shorter than the vowel in "bat" vs "bad." In the present experiment, crossover values corresponded to vowel durations of 118.78, 121.57, and 122.97 for the young masked, elderly normal, and elderly hearing-impaired listener groups, respectively. Although the average data suggest a tendency for young listeners to show a shorter crossover point and steeper slope for this continuum than the two older groups, statistical analyses failed to reveal any significant differences in the crossover points or the slope values between groups for the vowel duration cue. The slope values for the WHEAT/WEED continuum ranged from 26.8 to 39.2, suggesting that listeners did not perceive a strong contrast between the two tokens in this continuum. Indeed, our pilot testing during stimulus development dictated that we employ a 9-step continuum, rather than a 7-step continuum, to ensure that listeners identified the endpoint stimuli with reasonable accuracy. The identification function for WHEAT/WEED, shown in Fig. 1, suggests that the young noise-masked listeners achieved nearly perfect accuracy in identifying the endpoint stimuli; however, the two older listener groups did not show perfect identification of the endpoint stimulus "WEED." This lack of a clear percept of the exemplar word stimulus by older listeners may have contributed to more variable results and a resulting inability to observe significant group effects for this continuum.

A comparison of the identification data across the different speech continua reveals some interesting observations. First, significant age effects, when observed, are associated with smaller crossover values for younger listeners than for older listeners. This finding tentatively indicates that young listeners require relatively brief alterations in target stimulus duration to alter their phonetic percept. Second, the strongest age-related differences are observed for the continuum with the longest baseline duration (483 ms). All other continua were characterized by baseline reference durations of 310 ms or less. Third, the type of duration cue, filled or unfilled, does not appear to be linked consistently with the presence or absence of age effects in the identification data. For example, age effects were observed for identification of the BEAT/WHEAT continuum which requires processing of a filled cue (transition duration), but age effects were not observed for the WHEAT/WEED continuum which requires processing of another filled cue (vowel duration). Moreover, age effects were observed for the DISH/DITCH continuum which requires processing of the duration of an unfilled cue (silence). Fourth, the location of the target cue within the word is not clearly associated with age-related effects. The target cues for DISH/DITCH and WHEAT/WEED are located in the medial position of the words, but age-related effects were observed for the former continuum only. The target cues for BUY/PIE and BEAT/WHEAT are located in the initial position of the words, but age-related effects were observed for the latter continuum only. Taken together, the significant group findings in the present experiment were related to temporal cues for consonant manner and not to temporal cues for consonant voicing. Other factors, such as cue position, reference duration, and filled vs unfilled cue interval, are not consistently related to the observation of age effects in identification of the limited set of speech contrasts examined in the present experiments.

### B. Discrimination performance

Discrimination performance was measured with two different paradigms to facilitate comparison of the current results with those obtained by other investigators. The two-interval paradigm was intended to be relatively comparable to a standard same-different speech discrimination paradigm, whereas the three-interval paradigm was designed to be analogous to psychoacoustic measurement with nonspeech signals. Discrimination thresholds obtained with the two paradigms were quite different; statistical analyses revealed that the thresholds measured with the three-interval paradigm were significantly smaller than those measured with the two-interval paradigm for all listener groups across all four speech continua (Fig. 2). Comparable findings were reported by Pisoni and Lazarus (1974) in a study comparing performance on two different discrimination tasks intended to elicit categorical vs noncategorical perception. Listeners in the current experiments provided anecdotal reports suggesting that same-different procedures are performed on the basis of stimulus identification whereas three-interval forced choice procedures are based on acoustic differences between speech stimuli. Additionally, the discrimination thresholds measured for the two-interval paradigm are generally similar in value to the crossover points measured in the identification paradigm, lending further support to the notion that listeners judged the identity of the two stimulus words while performing this discrimination task.

The three-interval results tend to approximate some of the relevant findings from psychoacoustic experiments. A possible psychoacoustic analog for the DISH/DITCH continuum is the measurement of the threshold of a silent gap inserted between two markers. Detection thresholds for a silent gap inserted in the midpoint of a broadband marker are approximately 2–3 ms for young listeners with normal hearing (Plomp, 1964). The average threshold value measured for young listeners for the DISH/DITCH continuum in the current experiment is 2.77 ms, which has the same approximate value as that observed for nonspeech broadband stimuli.

The three-interval discrimination task for the WHEAT/WEED and BEAT/WHEAT continua is similar to a duration discrimination measure for filled (tonal or noise) stimuli. Duration discrimination thresholds for signals of approximately 250 ms are about 25 ms, or 10% of the reference duration on average, for young listeners (Small and Campbell, 1962; Abel, 1972). The DLs measured for the WHEAT/WEED continuum range from 8.15 ms to 15.95 ms, and those measured for the BEAT/WHEAT continuum range from 10.63 ms to 14.78 ms. Given that the reference duration of the endpoint speech stimulus for each of these continua is approximately 250 ms, the listeners in these experiments demonstrate better DLs than those that have been observed for nonspeech signals of equivalent reference duration. For the WHEAT/WEED continuum, the relevant reference duration may be the vowel duration cue (93 ms) rather than the entire word stimulus, in which case the results generally conform to the expected DL of 10%. The reference transition duration cue for BEAT/WHEAT is only 7 ms. Duration discrimination for such brief reference signals are not commonly examined in psychoacoustic studies but can be expected to be considerably larger than the 10% value associated with longer references (Abel, 1972). Generally, these findings suggest that in the three-interval paradigm, listeners based their judgments on discrete acoustic differences in the stimuli, rather than on labels assigned to the three stimuli presented on a trial. Moreover, their performance appears to follow trends observed in psychoacoustic data for analogous stimuli.

Performance patterns for the different listener groups derived from the two different discrimination paradigms were somewhat consistent with those observed in the identification paradigm. An examination of Fig. 2 suggests that the elderly hearing-impaired group showed larger discrimination thresholds than the other groups for the DISH/DITCH continuum. There also appears to be an age-related trend for larger discrimination thresholds in the BEAT/WHEAT continuum. Statistical analyses confirmed a significant main effect of listener group for the discrimination thresholds measured for the DISH/DITCH continuum, but not for any other continuum. Multiple comparison testing of the group effect revealed that the DLs for the elderly hearing-impaired listeners were greater than those for the two normal-hearing groups for both the two- and three-interval discrimination procedures. This finding agrees with the hearing loss effect observed in the analysis of the identification functions obtained for the DISH/DITCH continuum. The absence of pure age effects in the discrimination data differed from results observed in gap detection studies conducted with nonspeech stimuli (e.g., Schneider *et al.*, 1994; Snell, 1997). However, the age effects observed previously for gap detection are relatively small ($<3$ ms difference) and require small stimulus changes in the measurement procedures to be observed. It is possible that the larger step size employed in the present discrimination procedures obscured observation of possible age effects in these conditions. Other stimulus variables differentially influence gap detection performance by younger and older listeners, such as location of the gap in the stimulus (He *et al.*, 1999) or the location of the stimulus itself within a sequential context (Fitzgibbons and Gordon-Salant, 1995). In these more complex conditions, older listeners generally exhibit substantially larger temporal discrimination thresholds than those measured for stimuli presented in isolation. Further evaluation of younger and older listener's discrimination performance for temporal acoustic cues in speech may also reveal substantial age-related differences when the speech targets are embedded within extended sentence-length stimuli.

## V. SUMMARY AND CONCLUSIONS

This investigation of identification and discrimination for natural speech stimuli varying in temporal acoustic cues by young and elderly listeners revealed the following:

(1) Older listeners show different identification functions than younger listeners for some, but not all, temporally based speech continua. Age effects were observed for the speech continua that varied acoustic temporal cues underlying the perception of consonant manner of articulation. These continua included DISH/DITCH, in which silence duration as a cue for the sibilant/fricative distinction was varied, and BEAT/WHEAT in which the transition duration for the voiced stop/glide distinction was varied. The present experiment did not reveal age effects for speech continua that varied acoustic cues for stop consonant voicing in either the initial or final position of a word. These continua included BUY/PIE, in which VOT served as the primary cue for initial consonant voicing, and WHEAT/WEED, in which vowel duration was the principal cue for final consonant voicing.

(2) For speech continua in which an age-related difference was observed, younger listeners demonstrated shorter duration crossover points than older listeners. These findings suggest that older listeners require longer target temporal acoustic cues to form judgments about stimulus identity.

(3) Discrimination thresholds for the three-interval task were dramatically lower than for the two-interval task. These results indicate that listeners may be able to distinguish fine acoustic differences in speech that are not perceptible to them in a same-different paradigm. This increased acuity may be due, in part, to the listener's knowledge that one of the two comparison stimuli must be different from the initial reference stimulus.

(4) Discrimination thresholds for the temporal acoustic cues in speech agree reasonably well with discrimination thresholds obtained for nonspeech signals, when similar

paradigms are used. The results suggest that all listener groups are quite sensitive to subtle temporal changes in speech when a labeling judgment is not required and the target speech signals are presented in isolation. Although age-related effects were not observed in these preliminary discrimination data, a more perceptually relevant test would involve discrimination judgments for target speech signals presented in the context of ongoing sequential speech.

(5) Effects of hearing loss were observed consistently for the DISH/DITCH speech continuum, when presented in the identification paradigm and both discrimination paradigms. The source of the hearing loss effect for the silence duration cue may be associated with the hearing-impaired listeners' diminished capacity to detect the onset of the final sibilant in these speech tokens, because of the loss of high frequency audibility.

In conclusion, the findings support the hypothesis that older listeners experience a subtle deficit in processing changes in some acoustic temporal cues for discrete speech signals. For these signals, the data presented in this report provide further evidence that older listeners exhibit deficits in auditory temporal processing, and that these deficits contribute to alterations in speech identification.

## ACKNOWLEDGMENTS

This research was supported by an individual research grant (R37AG09191) from the National Institute on Aging. The authors are grateful to Michele Spencer for her assistance in data collection and analysis, and to two anonymous reviewers for their valuable comments on an earlier version of this manuscript.

Abel, S. (**1972**). "Duration discrimination of noise and tone bursts," J. Acoust. Soc. Am. **51**, 1219–1223.

Adams, C., and Munro, R. R. (**1978**). "In search of the acoustic correlates of stress: Fundamental frequency, amplitude, and duration in the connected utterance of some native and non-native speakers of English," Phonetica **35**, 125–156.

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

Bilger, R. C., Nuetzel, J. M., Rabinowitz, W. M., and Rzeczkowski, C. (**1984**). "Standardization of a test of speech perception in noise," J. Speech Hear. Res. **27**, 32–48.

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

Denes, P. (**1955**). "Effect of duration on the perception of voicing," J. Acoust. Soc. Am. **27**, 761–764.

Diehl, R. (**1976**). "Feature analyzers for the phonetic dimensions 'stop vs continuant,' " Percept. Psychophys. **19**, 267–272.

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.

Dorman, M. F., Raphael, L. J., and Liberman, A. M. (**1979**). "Some experiments on the sound of silence in phonetic perception," J. Acoust. Soc. Am. **65**, 1518–1532.

Fitzgibbons, P., and Gordon-Salant, S. (**1994**). "Age effects on measures of auditory duration discrimination," J. Speech Hear. Res. **37**, 662–670.

Fitzgibbons, P. J., and Gordon-Salant, S. (**1995**). "Duration discrimination with simple and complex stimuli: Effects of age and hearing sensitivity," J. Acoust. Soc. Am. **98**, 3140–3145.

Fitzgibbons, P., and Gordon-Salant (**1998**). "Auditory temporal order perception in younger and older adults," J. Speech Lang. Hear. Res. **41**, 1052–1060.

Flege, J. E., and Eefting, W. (**1988**). "Imitation of a VOT continuum by native speakers of English and Spanish: Evidence for phonetic category formation," J. Acoust. Soc. Am. **83**, 729–740.

Fox, R. A., Flege, J. E., and Munro, J. (**1995**). "The perception of English and Spanish vowels by native English and Spanish listeners: A multidimensional scaling analysis," J. Acoust. Soc. Am. **97**, 2540–2551.

Gelfand, S., Schwander, T., and Silman, S. (**1990**). "Acoustic reflex thresholds in normal and cochlear-impaired ears: Effect of no-response rates on 90th percentiles in a large sample," J. Speech Hear Disord. **55**, 198–205.

Gordon-Salant, S., and Fitzgibbons, P. (**1997**). "Selected cognitive factors and speech recognition performance among young and elderly listeners," J. Speech Lang. Hear. Res. **40**, 423–431.

Gordon-Salant, S., and Fitzgibbons, P. J. (**1993**). "Temporal factors and speech recognition performance in young and elderly listeners," J. Speech Hear. Res. **36**, 1276–1285.

Guion, S. G., Flege, J. E., Liu, H., and Yeni-Komshian, G. (**1997**). "Sentence duration as an index of overall proficiency in an *L2*," J. Acoust. Soc. Am. **101**, 3128.

He, N.-J., Horwitz, A. R., Dubno, J. R., and Mills, J. (**1999**). "Psychometric functions for gap detection in noise measured from young and aged subjects," J. Acoust. Soc. Am. **106**, 966–978.

Klatt, D. H. (**1975**). "Voice onset time, frication, and aspiration in word initial consonant clusters," J. Speech Hear. Res. **18**, 686–706.

Levitt, H. (**1971**). "Transformed up-down methods in psychoacoustics," J. Acoust. Soc. Am. **49**, 467–477.

Liberman, A. M., Delattre, P. C., Gerstman, L. J., and Cooper, F. S. (**1956**). "Tempo of frequency change as a cue for distinguishing classes of speech sounds," J. Exp. Psychol. **52**, 127–137.

Lisker, L., and Abramson, A. (**1967**). "Some effects of context on voice onset time in English stops," Lang Speech **10**, 1–28.

Lisker, L., and Abramson, A. S. (**1964**). "A cross-language study of voicing in initial stops: Acoustical measurements," Word **20**, 384–422.

Lister, Besing, and Koehnke (**2002**). "Effects of age and frequency disparity on gap discrimination," J. Acoust. Soc. Am. **111**, 2793–2800.

Lister, J., and Tarver, K. (**2004**). "Effect of age on silent gap discrimination in synthetic speech stimuli," J. Speech Lang. Hear. Res. **47**, 257–268.

Luce, P. A., and Charles-Luce, J. (**1985**). "Contextual effects on vowel duration, closure duration, and the consonant/vowel ratio in speech production," J. Acoust. Soc. Am. **78**, 1949–1957.

MacKay, I. R. A., Flege, J. E., and Piske, T. (**2000**). "Persistent errors in the perception and production of word-initial English stop consonants by native speakers of Italian," J. Acoust. Soc. Am. **107**, 2802.

Owens, E., Benedict, M., and Schubert, E. D. (**1972**). "Consonant phonemic errors associated with pure tone configurations and certain kinds of hearing impairment," J. Speech Hear. Res. **15**, 308–322.

Peterson, G. E., and Lehiste, I. (**1960**). "Duration of syllable nuclei in English," J. Acoust. Soc. Am. **32**, 268–277.

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

Pisoni, D. B., and Lazarus, J. H. (**1974**). "Categorical and noncategorical modes of speech perception along the voicing continuum," J. Acoust. Soc. Am. **55**, 328–333.

Plomp, R. (**1964**). "Rate of decay of auditory sensation," J. Acoust. Soc. Am. **36**, 277–282.

Price, P. J., and Simon, H. J. (**1984**). "Perception of temporal differences in speech by "normal-hearing" adults: Effects of age and intensity," J. Acoust. Soc. Am. **76**, 405–410.

Raphael, L. J. (**1972**). "Preceding vowel duration as a cue to the perception of the voicing characteristic of word-final consonants in American English," J. Acoust. Soc. Am. **51**, 1296–1303.

Schneider, B. A., Pichora-Fuller, M. K., Kowalchuk, D., and Lamb, M. (**1994**). "Gap detection and the precedence effect in young and old adults," J. Acoust. Soc. Am. **95**, 980–991.

Small, A. M., and Campbell, R. A. (**1962**). "Temporal differential sensitivity for auditory stimuli," Am. J. Psychol. **75**, 401–410.

Snell, K. B. (**1997**). "Age-related changes in temporal gap detection," J. Acoust. Soc. Am. **101**, 2214–2220.

Snell, K. B., and Hu, H.-L. (**1999**). "The effect of temporal placement on gap detectability," J. Acoust. Soc. Am. **106**, 3571–3577.

Strouse, A., Ashmead, D. H., Ohde, R. N., and Grantham, D. W. (**1998**). "Temporal processing in the aging auditory system," J. Acoust. Soc. Am. **104**, 2385–2399.

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

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.