The Effects of Aging and Interaural Delay on the Detection of a Break in the Interaural Correlation Between Two Sounds

Liang Li,[1] Juan Huang,[1] Xihong Wu,[1] James G. Qi,[2] and Bruce A. Schneider[2]

Objectives: [illegible]

Design: [illegible]

Results: [illegible]

Conclusions: [illegible]

(Ear & Hearing 2009;30;273–286)

INTRODUCTION

Perhaps the most intriguing question in auditory scene analysis is how listeners are able to detect, identify, locate, and characterize individual sound sources in noisy, reverberant environments when they receive not only the sound waves that directly come from various sound sources, but also numerous filtered and time-delayed reflections from the walls, ceiling, and other surfaces (e.g., Bregman 1990; Koehnke & Besing 1996). In such environments, listeners, especially older adult listeners, often find it difficult to process acoustic signals (e.g., speech), even though they can function well in quiet situations (e.g., Cheesman et al. 1995; Dubno et al. 1984; Duquesnoy 1983; Gelfand et al. 1988; Gordon-Salant & Fitzgibbon 1995; Helfer & Wilber 1990; Nabelek & Robinson 1982; Nabelek 1988; Pichora-Fuller et al. 1995; Stuart & Phillips 1996). Here we investigated whether age-related decreases in some of the perceptual processes that support auditory scene analysis might be contributing to the difficulties that older adults experience in noisy, reverberant environments.

Auditory Scene Analysis

To perceptually separate a target from the background in reverberant situations, the auditory system of the listener has to be able to differentiate the group of correlated sound waves that belong to the target (the direct wave from the target source and its time-delayed and filtered reflections) from sound waves produced by other sound sources (which will not be as highly correlated with the direct wave emanating from the target). In other words, to efficiently process the signals coming from an attended sound source in a noisy, reverberant environment, the auditory system needs to conduct two major perceptual operations: (1) integrate the direct wave from the target sound with its correlated reflections; and (2) segregate the target sound waves from sound waves generated by other sources. If there are deficits in the first operation, the sound reflections themselves, rather than being perceptually integrated with the source, could split off (Blauert & Lindemann 1986) from the direct wave and be perceived as separate auditory events. If there are deficits in the second operation, information from other sources might be partially integrated with that of the target source, leading to confusion. Therefore, to be capable of determining whether or not two wavefronts, arriving at different times and from different directions, are from the same

---

[1] Department of Psychology, Speech and Hearing Research Center, Key Laboratory on Machine Perception, Peking University, Beijing, China; and [2] Department of Psychology, Centre for Research on Biological Communication Systems, University of Toronto at Mississauga, Mississauga, Ontario, Canada.

0196/0202/09/3002-0273/0 • Ear & Hearing • Copyright © 2009 by Lippincott Williams & Wilkins • Printed in the U.S.A.

sources or from different sources, the auditory system has to be able to recognize when a time-shifted version of one wave is highly correlated with another. If the auditory systems of older adults are less capable than those of younger adults at recognizing when a time-shifted version of one wave is correlated with another, the auditory scene of older adults will be more cluttered and confused than that of younger adults. This might explain why older adults are especially disadvantaged in highly reverberant environments.

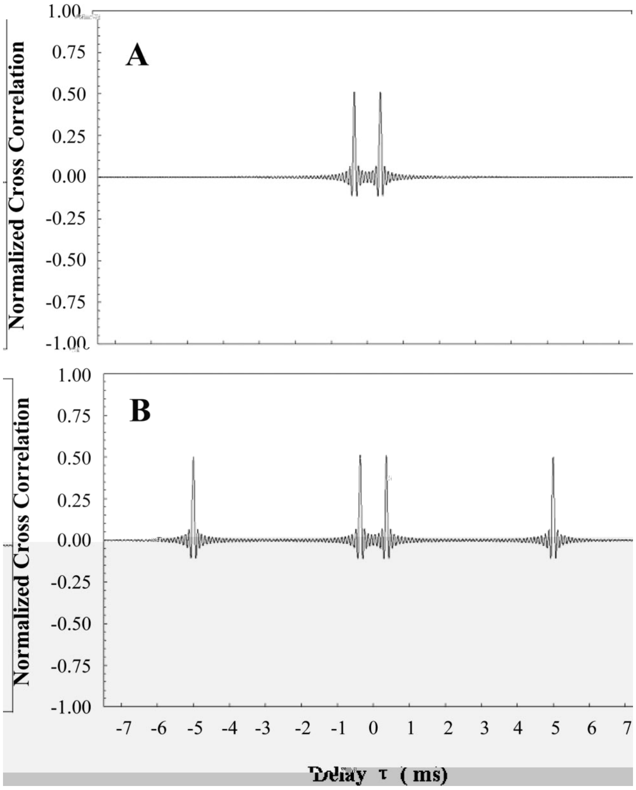
Integration of the Direct Wave and Its Reflections: The Precedence Effect

When the delay between the direct wave from the source and one of its reflections is sufficiently short (e.g., 5–10 ms or less, depending on the stimulus), all nonspatial attributes of the reflection are perceptually captured by the direct wavefront (e.g., Li et al. 2005), leading to a fused sound image whose point of origin is perceived to be at or near the location of the sound source. This phenomenon is called the precedence effect because the wavefront to arrive first takes precedence over the correlated wavefront (Blauert 1997; Li & Yue 2002; Litovsky et al. 1999; Wallach et al. 1949). The strength of this integration in a reverberant environment is largely determined by the delay between the direct and reflected waves. When this delay is sufficiently short (less than the echo threshold), the direct wave and the reflection are fused into a single image, in which the perceived location is at or near the location of the source. The spatial extent of the fused image usually exceeds that observed in the anechoic environment, an effect refe o-350.2(in299.9(the7ge)-395. 4(effec . 4(effe all en)-452.7(oa24mag tio0TD360(do0TD3G tive events, one at each ear. When the interaural correlation was 0.25, 0.50, or 0.75, listeners perceived one diffuse event in the median plane, and two additional ones lateralized symmetrically with respect to the median plane. In other words, the compactness, number, and placement of images depended on the degree of interaural correlation. It is not clear, however, whether there are age-related changes in the ability to detect or process interaural correlations. Nevertheless, we would expect that an age-related diminution in the ability to detect and process interaural correlations, especially when one of the sounds was delayed with respect to the other, could lead to a more fragmented auditory scene in older adults, which would increase the difficulty of attending to and processing information from the target talker.

Using Interaural Correlation to Detect Correlated Signals in the Sound Field

Detecting a correlation between two signals in the sound field is somewhat more complicated than detecting a cross-ear correlation under headphone conditions. Assume for the moment that we have two loudspeakers located 45 degrees to the left and right of the listener in an anechoic environment, playing independent band-limited white noises ($g(t)$ over the left loudspeaker and $h(t)$ over the right loudspeaker), both having bandwidth $W = 10$ kHz. To simplify the situation, we can measure, in the absence of the listener, the sound pressure at the positions that would be occupied by the listener's left and right ears. This is equivalent to assuming that the head does not cast a sound shadow so that only the delays between the sounds arriving at the near and far ears need to be considered (at 45 degrees, the delay, δ, is approximately 0.363 ms). In that case, the signal arriving at the position occupied by the left ear is $g(t) + h(t - 0.000363)$, whereas the signal arriving at the position occupied by right ear is $g(t - 0.000363) + h(t)$. The normalized cross-correlation function for this case is shown in Figure 1 (top panel). Note that the normalized cross-correlation function has two peaks at $\tau = -0.363$ ms and $\tau = 0.363$ ms. These two peaks represent the cross-correlation between the direct wave arriving at the near ear from an off midline source and the same wave arriving at the far ear. Note that these two peaks will always be present when the two loudspeakers are symmetrically displaced from the midline.

When the two noises are correlated and the left-loudspeaker noise leads the right-loudspeaker noise by γ seconds, the signal arriving at the left ear is $g(t) + g(t - \delta - \gamma)$, whereas the signal arriving at the right ear is $g(t - \delta) + g(t - \gamma)$, when measurements are taken in the absence of the head. Figure 1 (bottom panel) also plots the normalized cross-correlation function* for $\gamma = 5$ ms and $\delta = 0.363$ ms. Note that this cross-correlation function has two peaks on each side of $\tau = 0$, one corresponding to the interaural delay (0.0363 ms) and one corresponding to the delay between the correlated sounds played over the left- and right-loudspeakers (5 ms). As the loudspeaker delay is decreased, the peak in the cross-correlation function caused by this delay shifts accordingly (and becomes one when $\tau = 0$), whereas the two peaks caused by δ are unaffected by any delay between the loudspeakers. Hence, the listeners could discriminate between correlated and independent noises based on their ability to detect a peak in the cross-correlation function at a delay equal to that between the correlated sounds coming from the two loudspeakers.



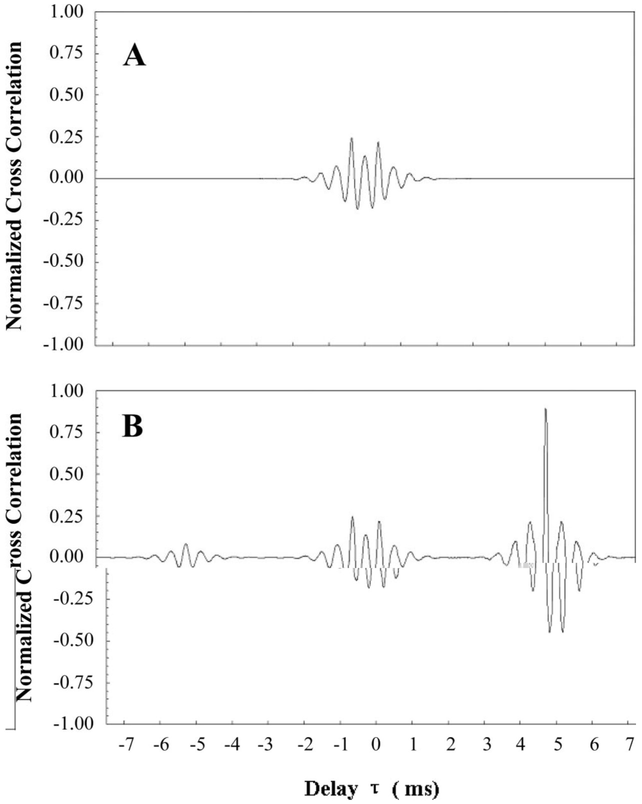
[illegible]

In Figure 1, it is assumed that there is no sound attenuation because of the shadow cast by the head. Figure 2 shows that when the head-related transfer functions are included in the computation of the normalized cross-correlation function, there is a decrease of the heights of the peaks because of the interaural delay, δ, an enhancement of the peak at $\tau = \gamma$ ms, and a substantial diminution of the peak at $\tau = -\gamma$ ms. However, the decreases in the peaks caused by the interaural delay are the same for both independent and correlated noises when the sound shadow is considered. As a result, these peaks convey no information as to whether or not the two sounds are correlated. Hence, the only way to determine whether or not the sounds are correlated from the cross-correlation function is to be able to sense the peak at $\tau = 5$ ms.

The situation will be further complicated if the loudspeakers are enclosed in a reverberant environment (e.g., a sound-attenuating chamber, as they were in these experiments), which will introduce other peaks caused by sound reflections. However, as a number of studies have indicated (e.g., Freyman et al. 1999; Kidd et al. 2005; Koehnke & Besing 1996; Zurek et al.

---

*To obtain a PDF file showing how the normalized cross-correlation functions in Figures 1 and 2 were computed, please contact Bruce Schneider.

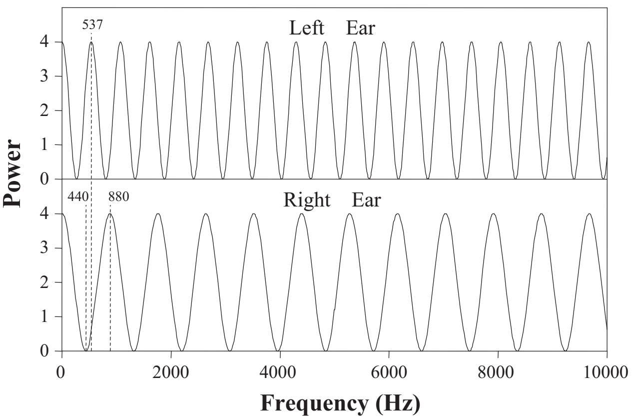


[illegible] $W =$ [illegible]

2004), the effect of adding these reflections is to increase the perceptual difficulties encountered by human observers and are unlikely to provide an additional cue that would aid them in discriminating between correlated and independent sounds. Finally, it should be noted that the cross-correlation functions shown in Figures 1 and 2 assume that the stimuli are infinite in duration. Cross-correlation functions computed over a shorter and more realistic time period would be, in general, broader than those depicted here.

[illegible]

In the sound field, the degree of correlation between the left and right noises is also revealed by the interference patterns that they create when the two waveforms add. If a band-limited white noise is added to itself after a delay of γ sec, the long-term power spectrum of their sum is no longer flat but rippled (comb filtering, Na in et al. 1979). If the spectrum level of the original noise is N_0, the spectrum level of the summed noise will be N_0 (2 + 2 cos [$2\pi f\gamma$]). However, if the two noises are independent, the long-term spectrum level is $2N_0$ for all frequencies within the bandwidth of the noise. Hence, when left and right correlated waveforms add, a ripple pattern will be observed in the spectrum, with the rate of modulation being determined by the delay.



[illegible] $W =$ [illegible] $N =$ [illegible] $=$ [illegible]

Figure 3 plots the long-term power spectra at the positions occupied by the left (top panel) and right (bottom panel) ears for a band-limited noise, $g(t)$, (10 kHz, $N_0 = 1$) played over a loudspeaker located 45 degrees to the left of the listener plus an identical version delayed by $\gamma = 1.5$ msec located 45 degrees to the right of the listener so that the interaural delay is again equal to 0.363 msec. If we ignore the sound shadow cast by the head, the signal arriving at the left ear is $g(t) + g(t - 0.0015 - 0.000363)$ and the signal arriving at the right ear is $g(t - 0.000363) + g(t - 0.0015)$. Hence, the power spectrum at the left ear is 2 + 2 cos ($2\pi f \times 0.001863$), and the power spectrum at the right ear is 2 + 2 cos ($2\pi f \times 0.001137$). By way of contrast, if the two noises are independent (again assuming no head shadow effects), the power spectrum has a uniform value of 2 across the entire spectrum. If the auditory system were to compare the outputs of a right ear monaural filter centered at 440 Hz to one centered at 880 Hz, the difference between the outputs of these two filters would be large when the noises were correlated and 0 when the noises were independent. Alternatively, if the auditory system were to compare the left- and right-ear monaural filters centered at 537 Hz, the interaural difference in the outputs of these two filters would be large when the left- and right-loudspeaker noises were correlated and negligible when they were independent.[1]

Hence, the auditory system could make use of both monaural and binaural spectral cues, as well as cross-ear correlation to determine whether or not a waveform arriving from one direction was a delayed version of another waveform that had arrived previously. Age-related changes in the ability to detect interaural spectral differences, a systematic ripple in the monaural spectrum, or age-related changes in the ability to detect an interaural correlation (especially when there was a

---

[1] This depiction assumes that the head casts no sound shadows. If the sound shadow is taken into consideration, the difference between peaks and troughs and the average power changes with frequency because of the HRTF. Hence, Figure 3 depicts an upper limit to the functional availability of these monaural and binaural spectral cues.

delay), could affect the ability of older adults to parse the auditory scene as effectively as younger adults.

The Aims of the Present Study

In experiment 1 of the present study, we assessed the age-related difference in the ability to detect a BIC when broadband noises are presented either over headphones or over loudspeakers. Note that when the BIC is presented over headphones, only binaural cues are available. However, when the same signals are presented in the sound field, the listeners could use comb-filtering effects to supplement the information obtained through interaural correlation. Hence, if listeners could use comb-filtering effects to detect a BIC, we would expect to find better performance in the sound field than under headphone presentation.

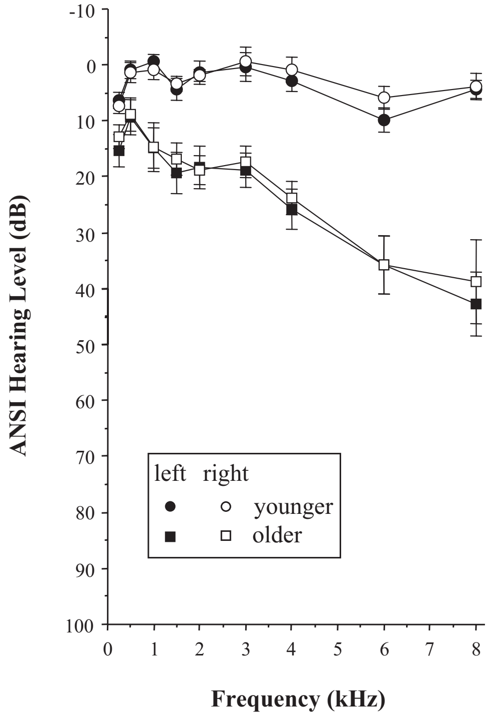
Based on the results of experiment 1, in experiment 2 we examined the longest interaural delay at which a BIC with a long duration (100 msec, which was well above the BIC-duration threshold at the zero interaural delay) was detectable, in both younger adults and older adults. We also examined the longest interloudspeaker delay where the change of intersound correlation could be detected to evaluate the degree to which monaural and binaural spectral cues would aid in the detection of a BIC.

MATERIALS AND METHODS

Experiment 1: BIC Duration Threshold at Zero Interaural Delay

Participants • Ten younger adults (6 females, 4 males, 19–21 yrs old, recruited from the University of Toronto at Mississauga) and 10 older adults (3 females, 7 males, 64–75 yrs old, recruited from the local community) participated in experiment 1. None of the participants had any history of hearing disorders, and none used hearing aids. All participants gave their written informed consent to participate in the experiments and were paid a modest stipend for their participation. The participants did not participate in experiment 2.

The younger adults and 6 of the 10 older adults had pure-tone, air-conduction thresholds less than 25 dB HL between 0.25 and 3 kHz. Four older adults had hearing levels at least at one of the test frequencies that were larger than 25 dB HL but less than 35 dB HL. Hearing thresholds for all participants were symmetrical (interaural difference less than 15 dB at each frequency). Figure 4 presents average hearing levels for both age groups as a function of frequency. Thresholds for all of the younger adults were well within the normal range. On average, the older adults' thresholds were 8 to 10 dB poorer than those of younger adults for frequencies less than 2 kHz. For frequencies higher than 2 kHz, threshold differences increased and differed by as much as 40 dB at the highest frequency tested. Although older adults with hearing in this range are usually referred to as having clinically normal hearing, they are best characterized as being in the early stages of presbycusis. Hence, they were likely experiencing subclinical declines in a number of auditory functions, including those related to temporal processing (e.g., Gordon-Salant & Fitzgibbons 1995, 1999; Schneider et al. 2002).



[illegible]

Sound chamber • During testing sessions, the participants were seated in a chair at the center of an Industrial Acoustic Company sound-attenuated chamber, whose internal dimensions were 283 cm in length, 274 cm in width, and 197 cm in height. The early decay times, which measured the time over the first 10 dB of the decay and are related to subjective judgments of reverberance (Bradley 1991), were 0.093, 0.135, 0.090, 0.079, 0.088, and 0.086 sec for frequencies of 125, 250, 500, 1000, 2000, and 4000 Hz, respectively.

Stimulus generation and delivery • Gaussian broadband noises (bandwidth = 0–10 kHz; sampling rate = 20 kHz), in which durations were 1000 msec, were digitally synthesized by generating 20,000 independent random normal deviates. Hence, the average spectrum of these digital noises was flat over the region from 0 to 10 kHz. Thirty millisecond, linear on- and off-ramps were applied to each noise burst. These digital signals were converted to analog forms using Tucker-Davis Technologies (TDT) DD1 digital-to-analog converters under the control of a Dell computer with a Pentium II processor. The analog outputs were low-pass filtered at 10 kHz with TDT FT5 filters, attenuated by two programmable attenuators (TDT PA4, for the left and right channels), and fed into a headphone buffer (TDT HB5). The outputs from the headphone buffer were either transduced by a pair of balanced headphones (Telephonic TDH-49P) or amplified via a Harman/Kardon power amplifier (HK3370) and then delivered from two balanced loudspeakers (Electro-Medical Instrument, 40 watts). The two loudspeakers were in the frontal azimuthal plane at the left and the right 45 positions symmetrical with respect to the median plane, respectively. The distance between each of the two loudspeakers to the center of the participants'

[redacted]

heads was 169 cm. Th

ear level for a seated

All the single-o

which was well abo

for both younger a

stimulation conditio

placed at the locatio

When the participant

no m mete e por

Procedure • Two 1

broadband noise

loudspeakers. The rig

of the interval was a

noise. The right-head

interval was also iden

noise except for the

middle of the 1000-m

pendent noise segmen

the BIC had an equal p

one of two interval

paradigm. The two inte

the offset of the first on … Fo

each interval, the noise co … e (to the

left loudspeaker) and the … right head-

phone (to the right loudspeak … same time. Fresh noise sounds were generated for each trial. The participant's task was to identify which of the two intervals contained the correlation break.

The participant initiated a trial by pressing a button on the response box. The starting BIC duration in a testing session was 100 ms. The BIC duration was decreased after three consecutive correct identifications of the interval containing the BIC and increased after one incorrect identification, using a three-down-one-up procedure (Levitt 1971). The initial step size of changing the BIC duration was 32 ms, and the step size was altered with each reversal in direction by a factor of 0.5 until the minimum size of 1 ms was reached. Feedback was provided at each trial. A test session was terminated after 12 reversals in direction, and the threshold for that session was defined as the averaged duration for the last eight reversals. Test sessions were repeated four times for each participant, and the

2: I

a • Ten younge

, recruited from the Unive

ga) and 11 older adults (7 females, 4 males, 63–75 years old, recruited from the local community) participated in experiment 2. None of the participants had any history of hearing disorders, and none used hearing aids. All participants gave their written informed consent to participate in the experiments and were paid a modest stipend for their participation. The criteria for participation in this experiment were the same as in experiment 1. The participants differed from those in experiment 1. Three of the female older participants could not reliably detect a long (100 ms) BIC, even though they had similar hearing levels with other older participants. Thus, data (including those of hearing levels) of these three older female participants are not reported here.

Figure 6 presents average hearing levels for both age groups as a function of frequency. Thresholds for all of the younger adults were well within the normal range. The older adults' thresholds were 8 to 10 dB poorer than those of younger adults for frequencies lower than 2 kHz. The threshold difference increased with frequency for frequencies higher than 2 kHz. The older participants are best characterized as being in the early stages of presbycusis.

C a b , a , a d d • The apparatus and materials used in experiment 2 were the same as those used in experiment 1, except that (1) tests were conducted in a different Industrial Acoustic Company sound-attenuated chambe (193 cm in length, 183 cm in width, and 198.5 cm in height), (2) the analog output from the headphone buffer were amplified via a different power amplifier (Technics, SA-DX950), and (3) the distance from each of the two loudspeakers to the center of the participant's head was 1.03 m. For the chamber used in experiment 2, the real decay times were 0.089, 0.035, 0.023, 0.044, 0.059, and 0.025 sec for frequencies of 125, 250, 500, 1000, 2000, and 4000 Hz, respectively.

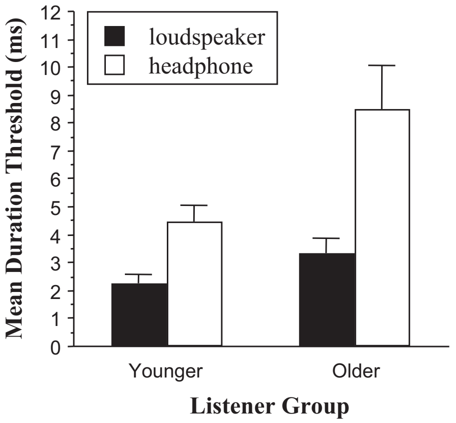
Procedure • Two 1000 ms intervals of correlated Gaussian broadband noises were presented either over headphones or loudspeakers. The right-headphone (loudspeaker) noise in one of the intervals was a copy of the left-headphone (loudspeaker) noise. The right-headphone (loudspeaker) noise in the other interval was also identical to the left-headphone (loudspeaker) noise except for the substitution of a long (100 ms) BIC introduced into the middle of the 1000 ms noise by simply substituting an independent noise segment in the left source. In each trial, the BIC had equal possibility to be randomly assigned to one of the two intervals of a 2IFC paradigm. The two intervals on a trial were separated by 1000 ms. For each interval, the 1000 ms noise coming from the left headphone (or the left loudspeaker) always led the 1000 ms noise coming from the right headphone (or the right loudspeaker) with the length of the intersound delay systematically manipulated (see below). That is, the intersound delay was applied to the whole waveform both onset and ongoing portions. Because the independent 100 ms noise segment associated with the BIC was always introduced in the center of the noise before the imposition of the signal delay, the uncorrelated segment itself was delayed in the right ear relative to the left by the same amount as the whole waveform delay. Fresh noise sounds were generated for each trial. The participant's task was to identify which of the two intervals contained the BIC.

The participant initiated a trial by pressing a button on the response box. The starting intersound delay in a testing session was 1 ms. The intersound delay was increased after three consecutive correct identifications of the interval containing the BIC and decreased after one incorrect identification using a three-up-one-down procedure (Levitt 1971). The initial step size of changing the intersound delay was 8 ms, and the step size was altered by a factor of 0.5 with each reversal of direction until the minimum size of 1 ms was reached. Feedback was provided at each trial. A test session was terminated after 12 reversals in direction, and the threshold for that session was defined as the average delay for the last eight reversals. Test sessions were repeated four times for each participant, and the best three thresholds were then averaged to obtain an estimate of the limit of each participant's ability to use waveform information available in the noise.

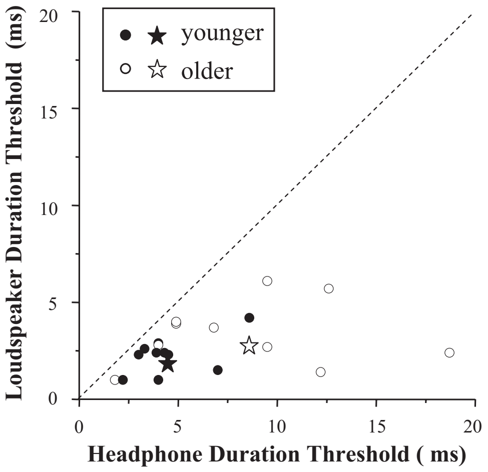
RESULTS

Experiment 1: BIC Duration Threshold at Zero Intersound Delay

Figure 7 shows the group averages of the shortest BIC durations at which the BIC could be detected under both the headphone-stimulation condition and the loudspeaker-stimulation condition for the two age groups. Under either the headphone- or the loudspeaker-stimulation condition, younger participants were able to detect shorter BIC than older participants, indicating a reduction in sensitivity to the BIC with age. Under the headphone-stimulation condition, on average, younger participants could detect a BIC approximately 4.5 ms long (median = 4 ms), whereas older participants could detect a BIC whose duration was approximately 8.5 ms (median = 8.1 ms). Under the loudspeaker-stimulation condition, the threshold for detecting the BIC was 2.3 ms (median = 2.4 ms) for the younger group and 3.4 ms (median = 3.2 ms) for the older group. The shortest BIC durations for individual participants under the two stimulation conditions are shown in Figure 8, Table 1 (for younger participants) and Table 2 (for older participants). Note that there is much more variability in threshold for older than for younger adults, with five of the older adults having duration thresholds within the range of those observed for younger adults. This increase in variability with age has been found in



[illegible]



[illegible]

TABLE 1. BIC 10 ( )

Participants	SM	SA	CL	CC	WL	IZ	NKN	MSD	VB	RP
Loudspeaker	4.2	2.3	2.4	2.6	1.0	2.9	1.0	2.4	1.5	2.3
Headphone	8.6	4.5	4.3	3.3	4.0	4.0	2.2	3.9	7.0	3.0

BIC, break in correlation.

other studies. For example, Schneider and Pichora-Fuller (2001) showed that whereas many older adults had gap detection thresholds that were within the range found for younger adults, a substantial number had thresholds in excess of this range.

A two between-subjects (younger, older) by two within-subjects (headphone, loudspeaker) mixed analysis of variance (ANOVA) did not reveal a significant interaction between age group (younger, older) and stimulus-presentation type (headphone, loudspeaker) ($F_{1,18} = 2.890$; MSE = 7.338; $p = 0.106$) but did verify that the main effects of stimulus-presentation type ($F_{1,18} = 18.385$; MSE = 7.338; $p < 0.001$) and age group ($F_{1,18} = 7.087$; MSE = 9.160; $p = 0.016$) were both significant. Hence, older adults have higher thresholds than younger adults, and there is insufficient evidence to reject the hypothesis that, in the sound field, comb filtering cues lower thresholds by the same amount in both younger and older adults when there is no delay between left and right noises.

An examination of Table 2 indicates the presence of a potential outlier in the headphone condition (participant AM). To check whether this outlier was responsible for the main effect of age, we repeated the ANOVA with this participant removed. The main effects of age and condition remained significant, and there was no interaction between age and condition. Hence, we have retained this possible outlier in the remaining analyses.

For younger participants, the correlation between the thresholds under loudspeaker presentation and that under headphone presentation was 0.521, which was not significant ($F_{1,8} = 2.987$; MSE = 0.734; $p = 0.122$). For older participants, the correlation between the thresholds under loudspeaker presentation and that under headphone presentation was 0.104, which was also not significant ($F_{1,8} = 0.088$; MSE = 3.056; $p = 0.774$).

To see whether the BIC thresholds were related to audiometric thresholds, we correlated BIC thresholds with pure-tone averages (PTAs, averaged across the two ears) for both low-frequencies (0.25–2 kHz, LF-PTA), and high-frequencies (3–8 kHz, HF-PTA) in both younger and older adults. None of these correlations were significant in either younger or older adults. For the younger adults, the correlations between BIC thresholds and LF-PTA were −0.1 ($p > 0.05$) and 0.156 ($p > 0.05$) for headphone and loudspeaker presentations, respectively; the correlations between BIC thresholds and HF-PTA were 0.541 ($p > 0.05$) and 0.262 ($p > 0.05$) for headphone and loudspeaker presentations, respectively. For older adults, the correlations between BIC thresholds and LF-PTA were 0.272 ($p > 0.05$) and −0.04 ($p > 0.05$) for headphone and loudspeaker presentations, respectively; the correlations between BIC thresholds and HF-PTA were 0.284 ($p > 0.05$) and 0.434 ($p > 0.05$) for headphone and loudspeaker presentations, respectively. Hence, there is very little evidence that BIC thresholds are correlated with either low- or high-frequency PTA in younger or older adults.

Experiment 2: The Maximum Intersound Delay

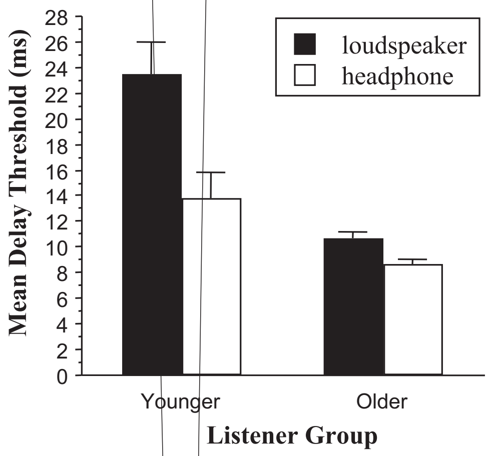
Figure 9 shows the group mean of the longest intersound delays at which younger or older participants were able to detect a 100 ms BIC. Under the headphone-stimulation condition, both the mean (13.8 ms) and median (11.9 ms) thresholds for younger participants were longer than those (mean = 8.6 ms; median = 8.7 ms) for older participants. Also, under the loudspeaker-stimulation condition, both the mean (23.5 ms) and median (26.1 ms) thresholds for younger participants were longer than those (mean = 10.6 ms; median = 11.2 ms) for older participants. Thus there was a substantial reduction in the ability to detect an intersound delay with age.

A two between-subjects (younger, older) by two within-subjects (headphone, loudspeaker presentation) ANOVA found that the interaction between age-group and stimulus-presentation type (headphone or loudspeaker) was significant ($F_{1,16} = 5.722$; MSE = 23.349; $p = 0.029$), as was the main effects of age group ($F_{1,16} = 19.959$; MSE = 36.299; $p < 0.001$), and stimulus-presentation type ($F_{1,16} = 13.149$; MSE = 23.349; $p = 0.002$). Separate ANOVAs for headphone and loudspeaker presentations showed that the age effect was significant for both loudspeaker ($F_{1,16} = 20.805$; MSE = 35.579; $p < 0.001$) and headphone-stimulation conditions ($F_{1,16} = 4.899$; MSE = 24.070; $p = 0.042$). Hence, the interaction effect indicates that the increment in performance going from headphone to loudspeaker conditions was larger for younger than for older adults.

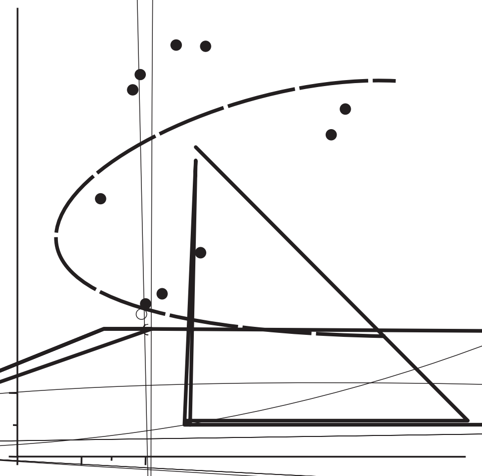
To further explore the nature of the interaction, we plotted the longest delay between left and right noises at which each individual could detect a 100 ms BIC in the sound field as a function of the longest delay they could detect a 100 ms BIC under headphone conditions (Fig. 10). The dotted line (slope = 1.0) represents what we would expect if there were no difference between headphone and sound field conditions. This figure shows that all participants but one performed better under sound-field conditions than under headphone conditions. Particularly, five of the younger adults performed markedly

TABLE 2. BIC 10 ( )

Participants	BR	AG	ES	BM	JZ	LW	GH	JSF	EW	AM
Loudspeaker	2.8	3.9	4.0	6.1	5.7	3.7	1.0	2.7	1.4	2.4
Headphone	4.0	4.9	4.9	9.5	12.6	6.8	1.8	9.5	12.2	18.7



better under sound-field conditions than under headphone conditions (those whose data points are farthest from the diagonal line). These results suggest that some younger participants (but no older ones) seem to derive a substantial benefit under sound field conditions (more than doubling the longest delay at which they could detect a BIC), even though they were not necessarily the best participants under either sound-field or headphone conditions. Hence, the greater improvement in the performance of younger adults when going from headphone to loudspeaker presentation can be attributed to the fact that half of the younger adults improved markedly, whereas the other half showed little improvement. The longest delays for individual participants under each of the two types of stimulation conditions are also shown in Table 3 (for younger participants) and Table 4 (for older participants). Unlike the case for duration thresholds, where there is more variability among the young than among the older listeners. Furthermore, there is no indication that older adults benefit from the loudspeaker presentation, whereas half of the younger adults exhibit a large benefit from the loudspeaker presentation.



For younger participants, the correlation between the thresholds under headphone-stimulation conditions and that under loudspeaker-stimulation conditions was 0.214, which was not significant ($F_{1,8} = 0.383$; MSE = 65.362; $p = 0.553$). For older participants, the correlation between the threshold under headphone-stimulation conditions and that under loudspeaker-stimulation conditions was 0.422, which was also not significant ($F_{1,6} = 1.299$; MSE = 2.919; $p = 0.298$).

To see whether the maximum interaural delays were related to audiometric thresholds, we correlated the interaural delays with PTA for both low (0.25–2 kHz, LF-PTA), and high (3–8 kHz, HF-PTA) frequencies. For the younger adults, the correlations between the longest delay at which a BIC was detectable and LF-PTA were 0.288 ($p > 0.05$) and 0.291 ($p > 0.05$) for headphone and loudspeaker presentations, respectively; the correlations between the longest delays and HF-PTA were 0.399 ($p > 0.05$) and 0.276 ($p > 0.05$) for headphone and loudspeaker presentations, respectively. For older adults, the correlations between the longest delays and LF-PTA were 0.282 ($p > 0.05$) and −0.15 ($p > 0.05$) for headphone and loudspeaker presentations, respectively; the correlations between the longest delays and HF-PTA were 0.338 ($p > 0.05$) and −0.27 ($p > 0.05$) for headphone and loudspeaker presentations, respectively. Hence, there is very little evidence that the longest interaural delay at which a 100 ms BIC can be detected is correlated with either low- or high-frequency PTA in younger or older adults.

DISCUSSION

[illegible] BIC [illegible]

In the present study, under headphone listening conditions with the 0 ms interaural delay, younger adult participants could detect a 4.5 ms BIC between Gaussian broadband noises (0–10,000 Hz), which is slightly larger than the mean threshold (2.34 ms) of the 1/0/1 interaural correlation change interval measured in eight participants (20–35 yr old) in the study by Boehnke et al. (2002) using a broader band noise (0–22,050 Hz), but smaller than the mean binaural gap threshold (5.3 ms) measured in six participants (whose ages were not provided) in the study by Akeroyd and Summerfield (1999) using bandpass noise (100–500 Hz). These results confirm that human listeners with normal hearing have a high sensitivity to a transient BIC when the interaural delay is zero. For older adults tested in the present study, their mean threshold of detecting the BIC under the headphone-stimulation condition was 8.5 ms, which was significantly larger than that for younger participants. Older adults were also much more variable than younger adults, a pattern that has been previously noted with relation to gap detection studies (Schneider & Pichora-Fuller 2001).

Older adults could be less sensitive to a BIC than younger adults because of age-related reductions in audiometric sensitivity. To investigate whether the age-related changes in the BIC thresholds were caused by age-related decreases in spec-

TABLE 3. T 10 ( )

Participants	DR	DV	CL	MR	ZN	TL	RC	FR	SM	CT
Loudspeaker	25.1	27.1	15.9	12.7	28.6	29.8	32.1	20.1	32.0	11.9
Headphone	24.5	25.6	14.3	11.3	9.0	9.6	12.4	6.5	14.7	10.0

tual sensitivity, we correlated the BIC threshold with audiometric threshold separately for younger and older adults at both high and low frequencies. These correlations, however, provided very little evidence for a relationship between audiometric hearing loss and sensitivity to BIC. Hence, it seems more likely that lower sensitivity to BIC are related to other age-related changes in the auditory system, such as a loss in neural synchrony. Previous studies have shown that older listeners with normal hearing have smaller masking level differences (MLDs) than younger-adult listeners (e.g., Grose et al. 1994; Olsen et al. 1976; Pichora-Fuller & Schneider 1991, 1992, 1998; Strouse et al. 1998). Pichora-Fuller and Schneider (1992) have suggested that smaller MLDs in older adults are caused by lower temporal synchrony between the two ears (i.e., an increase in temporal jitter; Durlach 1972). Hence, age-related lower temporal synchrony could account for both smaller MLDs and higher BIC threshold in older than in younger adults.

Previous functional magnetic resonance imaging and magnetoencephalography studies have suggested that in human the auditory cortex is involved in processing interaural correlation (e.g., Budd et al. 2003; Chait et al. 2005; Hall et al. 2005; Zimmer & Macaluso 2005). Thus, it is important in future studies to see if whether there are age-related alterations of the central representation of the change in interaural correlation at the cortical level.

Another possibility is that age-related changes in the ability to detect a BIC could reflect age-related changes in the size of the temporal windows over which interaural comparisons occur. Several investigators have proposed that binaural comparisons are performed within a temporal window applied to the inputs to the two ears (e.g., Bernstein et al. 2001; Moore et al. 1988). According to this notion, the auditory system effectively integrates binaural information falling within this temporal window. Hence, when there is a change in an interaural variable during this window, this integration process reduces the internal or effective value of this change.‡ For example, if observers were to center the temporal window at the midpoint of each of the two broadband noises presented on a 2IFC trial in experiment 1 (with the BIC occurring randomly in the center of one of these noises), they could compare the interaural information available in this window for each of the two noises to determine which one contained the BIC. Assuming that younger and older adults required the same amount of information to reach their threshold for detecting a BIC (e.g., the same difference in interaural correlation), age differences in the shape or width of the temporal window could lead to age differences in performance. For example, suppose the participants in experiment 1 applied a rectangular temporal window (a rectangular window is used here to simplify the description of how age differences in temporal windows size could account for age differences in detecting a BIC) to the time-varying interaural correlation. For the diotic noises without the BIC, the interaural correlation would be 1.0 for both age groups, independent of window size (assuming that the temporal windows are smaller than the length of the stimuli). However, the interaural correlation for a noise with a short BIC will depend on window size. Suppose the rectangular window sizes for younger and older adults were 4 and 8 ms, respectively. When a 6 ms BIC is presented, the interaural correlation of the windowed signal would be zero for younger adults but greater than zero for older adults because older adults would be computing interaural correlation over 8 ms of left- and right-ear signals where the correlation was 1.0 for the first and last ms of the 8 ms comparison and zero during the middle 6 ms. Hence the difference in interaural correlation between the noise segments with and without a BIC would be larger for younger than for older adults, leading to an age-difference in the ability to detect a BIC.

When the stimuli were presented over loudspeakers, the sound field provided certain additional cues, such as those induced by comb filtering effects (Na in et al. 1979). These cues could aid listeners to detect the transient break in intersound correlation. The data from experiment 1 suggest that both younger and older adults were able to use these cues to detect a shorter BIC when the cues were present (loudspeaker presentation) than they could when these cues were absent (headphone presentation). Moreover, even though older adults seemed to benefit more than younger adults from a switch from headphones to the sound field (Fig. 7, threshold decrease in older adults = 5.1 ms; threshold decrease in younger adults = 2.2 ms), the interaction of age group and stimulus-presentation type for the duration threshold was not statistically significant. Hence, when there is no delay between the left- and right-ear sounds, we cannot reject the hypothesis that younger and older adults benefit equally from the addition of sound-field cues.

Temporal Processing of Waveform Information (Headphone Presentation)

The present study also investigated how long waveform information is available to the listener by directly measuring the range of interaural delay in which a long-duration (100 ms) BIC is audible under headphone presentation (according to the

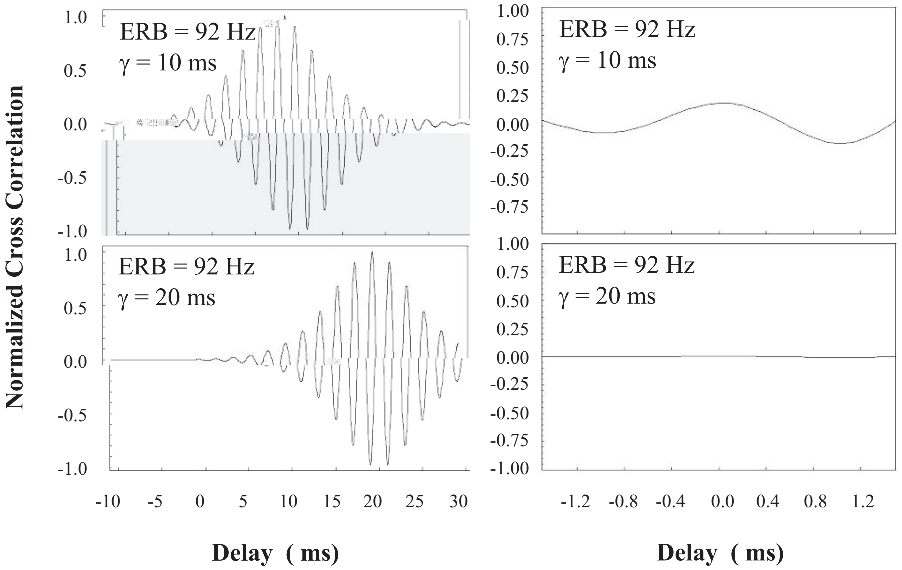
TABLE 4. T

( )

Participants	ARP	XL	IL	ML	JO	PL	BD	TL
Loudspeaker	11.1	9.9	12.3	7.8	12.0	8.4	11.3	12.3
Headphone	9.7	10.2	7.5	7.1	8.2	6.9	10.2	9.3

---

‡In the Bernstein et al. (2001) model, the smearing effect that the window has on binaural parameters is indexed by computing S, the area under the temporal window during the probe portion of the stimulus (e.g., a BIC), and dividing it by the total area under the temporal window during the entire stimulus. The internal or effective value of an interaural parameter is then assumed to be given by multiplying the external value by S.



$g\,t),\ (W =$ $g\,t - \gamma$ $\gamma =$ $\tau = -$ τ

results of experiment 1, at the zero interaural delay, the 100 ms duration was well above the BIC thresholds for all the younger and older participants). Two of the younger participants were able to detect the occurrence of the 100 ms BIC when the delay between the two ears was up to 25 ms in the headphone condition (Fig. 10). Note that delay thresholds are quite variable for younger adults, indicating a wide range of individual differences. Older adults, however, are much more uniform with respect to their ability to detect BIC at long delays. Recall, however, that long delay thresholds correspond to better performance. Hence age-related performance decrements would manifest themselves as lower thresholds. Because thresholds are bounded at the lower end by the value of 0, poorer performance in a group of older adults would tend to reduce the variance in this group, as is observed in Figure 10. Hence the pattern of results in experiment 2 suggests that as people age, their capacity to detect a change in correlation diminishes.

There seem to be two possible ways in which the auditory system of some young adults could bridge temporal delays greater than 15 ms between correlated left and right ear sounds. First, the cross-correlation function relating the output of matched, narrow band, left- and right-ear auditory filters could have substantial peaks within the range of delays that are physiologically realizable (-1.5 to 1.5 ms). If that were to occur, it would permit the auditory system to distinguish between correlated and independent noises, because the cross-correlation function for two independent noises would be zero for all delays.

To see how this could occur, let $y(t)$ be the output of a narrow-band, left-ear auditory filter to a broadband noise, $g(t)$. If the filter is linear and shift independent, then the output of the matching right-ear filter to $g(t - \gamma)$ is simply $y(t - \gamma)$. Therefore, we can compute a cross-correlation function on the outputs from these two filters. Figure 11 shows normalized cross-correlation functions, when the left- and right-ear noises are correlated, for delays $\gamma = 10$, and 20 ms, for the outputs of two matched gammatone auditory filters tuned to 500 Hz. The left panel plots the normalized cross-correlation function over a range of delays from -10 to 30 ms. The right panel plots the same function only over the range of delays that might be considered physiologically realizable. The parameters of this gammatone filter have been selected to provide the best fit to the spectral profile that characterizes a 500 Hz human auditory filter (Patterson 1976), and has an equivalent rectangular bandwidth of 92 Hz (454–546 Hz). Figure 11 indicates that if the observer could focus in on matched left- and right-ear filters at this bandwidth, the portion of the normalized cross-correlation function that is in the physiologically plausible range could possibly be used to discriminate left- and right-ear correlated noise from independent left and right-ear noise when the interaural delay is 10 ms but not when it is 20 ms. However, if the filter width is cut in half (Fig. 12), and the observer can focus in on this filter, then he or she could potentially perform this discrimination at interaural delays as long as 20 ms.

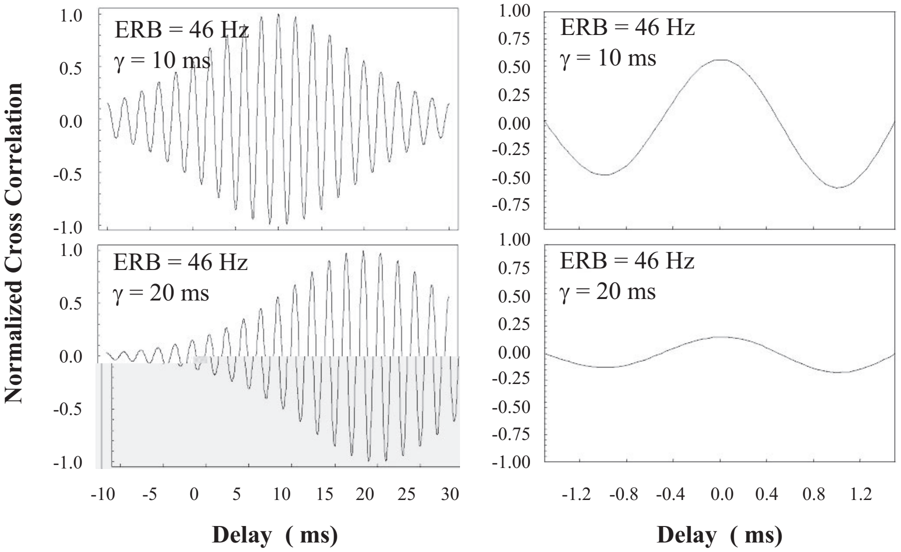
When stimuli are presented over headphones, it is interesting to note that narrow band filtering can account for delay thresholds $<$10 ms. Note that the delay thresholds for all of the older adults are less than 10 ms in the headphone condition, whereas the thresholds for six younger adults are greater than 10 ms in the same condition. Hence, it is possible that all of the older adults, and four of the younger adults use narrow band filtering to accomplish the task.

Hence, in order for the performance of some of the younger adults observed here to be based solely on cross-correlation of the outputs from matched auditory filters, it seems that these filters would have to be narrower than those previously observed. However, it might be possible to bridge longer interaural delays if narrow band filtering of the input at each ear is followed by propagation delays of several milliseconds (as in Durlach's 1972 EC model) before binaural comparisons are computed. Or it could be the case that nonlinearities of one sort or another in auditory processing could also help bridge these longer delays in some individuals. Another possibility is that higher-order central mechanisms could be involved in maintaining an auditory trace of the acoustic waveform.

The ability of some listeners to detect interaurally correlated sounds has also been found previously using indirect measures,

---

To obtain a PDF file showing how the normalized cross-correlation functions and average powers were computed for the outputs of these filters (Figs. 11–13), please contact Bruce Schneider.



g t), (W =

g(t − γ γ =

τ = − τ

such as those associated with judging sidedness of interaurally delayed noise (Blodgett et al. 1956; Cherry & Taylor 1954; Mossop & Culling 1998) or detecting signals in interaurally delayed noise (Langford & Jeffress 1964). Results of these earlier studies have suggested that a representation of the waveform may persist for up to 9 to 15 ms. However, to our knowledge, the present study is the first to use a BIC as the signal probe to directly measure the temporal extent of the representation of acoustic waveform information in both younger and older participants. The results of the present study show that older participants in headphone condition could detect the BIC only up to interaural delays of 10 ms or less, indicating age-related decline in the ability to detect interaural correlation over long delays.

Older listeners have smaller MLDs than younger listeners particularly when interaural delay is introduced. In the study by Pichora-Fuller and Schneider (1992), the threshold of detecting a 500 Hz pure tone against band-limited white noise (0.1–5 kHz) for older participants did not differ significantly from that for younger listeners when there was no interaural difference for the reference condition (N0). However, when MLDs were plotted as a function of the interaural delay of the noise maskers, the pattern of results differed significantly between younger and older listeners: There was no difference between the two age groups in the average MLD at the minimal interaural delay (0.25 ms), but the average MLDs of the younger group were larger than those of the older group at interaural delays equal to odd multiples of the half period of the signal frequency. Hence, older adults seem to be less able than younger adults to bridge interaural delays in at least two tasks: MLD and in the detection of a BIC.

It is also interesting to note that younger adults can detect a BIC at delays that exceed the maximum delay at which the lagging sound is fused with the leading sound (the precedence effect). The precedence effect reduces listeners' perception of multiple images in reverberant environments by perceptually grouping correlated acoustic waveforms from different directions. This perceptual grouping is based on capture of attributes of the reflections by the direct wave (Li et al. 2005). Thus, only a fused image is perceived as originating at or near the location of the source, and both localization errors and interference from the reflected waves are reduced (Litovsky et al. 1999). Because delays are always present between the direct and reflected waves coming from a sound source, the availability of aspects of the earlier-arriving waves would be essential if the reflected waves coming from different sites are to be perceptually fused with the appropriate sources. However, the present results indicate that younger adults are capable of accessing waveform information for durations that are longer than the fusion threshold for the precedence effect. For example, Li et al. (2005), using similar stimuli, have shown that for delays under 9.5 ms, the leading and lagging sounds were fused into a single sound whose origin was perceived to be at or near the location of the leading sound. For delays longer than 9.5 ms, younger listeners indicated that they heard two sounds, one coming from the location of the leading sound, the other from the location of the lagging sound. In the present study, BICs were observed for delays which exceed the fusion threshold, indicating that waveform information can be accessed for periods that are sometimes much longer than the fusion threshold.

The results of the present study also show that for both younger and older participants, the correlations between the longest delay under the headphone-stimulation condition and low- and high-frequency pure-tone average thresholds were not significant. Thus, the interlistener variation in performance can not be explained by the interlistener variation in hearing thresholds. Moreover, the study by Akeroyd and Summerfield (1999) has shown that when the center frequency of band-limited (100 Hz) noise was 2000 Hz, the mean BIC (binaural gap) detection threshold was larger than 100 ms. In other words, when the duration of a BIC is 100 ms, frequency components higher than 2000 Hz may not substantially contribute to the detection of the BIC between two correlated broadband noises. Thus, differences between the two age groups cannot be explained by the differences in hearing thresholds at high frequencies (≥3000 Hz).

REFERENCES