Broadband Auditory Stream Segregation by Hearing-Impaired and Normal-Hearing Listeners

Susie Valentine
Jennifer J. Lentz
Indiana University, Bloomington

Purpose: To investigate the effects of hearing loss on auditory stream segregation of broadband inharmonic sounds.

Method: Auditory stream segregation by listeners with normal and impaired hearing was measured for 6-component inharmonic sounds ("A" and "B") using objective and subjective methods. Components in the A stimuli ranged between 1000 and 4000 Hz, whereas B stimuli were generated at the same frequency ratio but scaled upward in frequency relative to the A stimuli. In Experiment 1, streaming was measured by having listeners detect a delay inserted into a sequence of A and B stimuli (A_B_A_B_...) for B stimuli with different frequencies. In Experiment 2, streaming was measured using an ABA_ABA_... sequence, and the frequency of the B stimulus decreased until listeners reported that they could "no longer hear two separate streams."

Results: Experiment 1 indicated no significant differences between groups in the size of the just detectable delay and no significant interactions between group and the scaling factor between the B and A stimuli. Experiment 2 revealed no significant differences in streaming abilities between normal-hearing and hearing-impaired groups.

Conclusions: Overall, results indicate that listeners with normal and impaired hearing have similar auditory streaming abilities for broadband inharmonic complex stimuli.

KEY WORDS: stream segregation, hearing loss, psychoacoustics

It is widely known that sensorineural hearing loss detrimentally impacts speech perception in complex and noisy environments, but relatively few studies have focused on the influence of hearing loss on higher level processes that are likely involved in speech perception in noise (such as segregating a meaningful signal from an unwanted background). Furthermore, although loss of audibility is a primary contributor to perceptual deficits experienced by persons with hearing loss, it does not solely account for the communication deficits experienced by listeners with hearing loss. These difficulties might arise from poorer sound-segregation abilities in listeners with hearing loss than in listeners with normal hearing; such difficulties can be evaluated using a stream-segregation paradigm. The goal of the following study is to evaluate the effects of hearing loss on broadband auditory streaming with a particular emphasis on multitone inharmonic sounds.

Auditory stream segregation is the process of separating different sound sources from a complex sound environment into individual auditory objects (Bregman, 1990). A typical auditory streaming experiment uses two alternating sequences of stimuli (often tones) that differ in frequency. The two stimuli tend to be perceived in different auditory streams

1092-4388/08/5105-1341

when they have very different frequencies, whereas they tend to be perceived in a single stream when their frequencies are similar (Bregman & Campbell, 1971; Miller & Heise, 1950; Van Noorden, 1977). One explanation of this result is that two stimuli must excite separate neural populations to form two separate auditory streams (Beauvois & Meddis, 1996; Hartmann & Johnson, 1991; McCabe & Denham, 1997). A number of other stimulus manipulations also have been shown to produce a percept of two streams, such as differences in timbre, pitch, and intensity (for a review, see Bregman, 1990), with such manipulations possibly causing different neural populations to respond to the two stimuli.

However, more recent experiments have demonstrated that different neural populations need not be excited to produce two separate auditory streams. Vliegen, Moore, and Oxenham (1999) and Vliegen and Oxenham (1999) showed that two stimuli without detectable spectral differences could be streamed on the basis of their temporal structure (as generated by differences in fundamental frequency). Furthermore, stimuli with identical power spectra but different phase spectra can be separated into two auditory streams (Roberts, Glasberg, & Moore, 2002; Stainsby, Moore, & Glasberg, 2004), and timbre differences between stimuli can lead to the perception of two streams (Cusack & Roberts, 2000). Thus, it appears that two different neural populations need not be excited for the formation of different auditory streams. Given that spectral differences are not a prerequisite for two sounds to be perceived in separate streams, Moore and Gockel (2002) argued that as long as two sounds are perceptually different, they can be organized into different auditory streams. Such studies implicate a role for temporal processes in the formation of auditory streams and suggest that cochlear place cues are not a necessary requirement for streaming.

Research on listeners with sensorineural hearing loss also supports this idea. Rose and Moore (1997) measured stream segregation using a paradigm in which listeners heard a rapid sequence of sounds, in an ABA_ABA_ format, in which A and B represent tones of two different frequencies and the underscore (_) represents a quiet interval. When the frequency of the B sound is distant from the frequency of the A sound, listeners tend to perceive two separate streams (Van Noorden, 1977). When the frequency of the B sound is similar to the frequency of the A sound, listeners tend to hear the A and B sequence as one "galloping" auditory stream. Rose and Moore (1997) measured the frequency at which the B sequence cannot be heard as a separate stream from the A sequence (i.e., the "fission boundary") and found that most of their subjects with hearing loss had fission boundaries that were similar to those of the normal-hearing listeners, with only a few of the subjects with hearing loss having fission boundaries larger than those found in subjects with normal hearing. Because listeners with hearing loss generally have poorer frequency selectivity than those with normal hearing, Rose and Moore (1997) rejected the idea that the frequency selectivity of the auditory system dictates the fission boundary. Furthermore, Rose and Moore (2005) demonstrated that frequency discrimination thresholds are not related to the fission boundary, and Mackersie, Prida, and Stiles (2001) also showed no relationship between auditory filter bandwidths and the fission boundary. Both studies support the idea that frequency selectivity is not related to streaming ability for pure tones. However, broader auditory filters can lead to deficits in streaming based on the temporal structure of harmonic sounds (Stainsby et al., 2004). Given this, abnormal temporal processing that might be associated with hearing loss (Fitzgibbons & Gordon-Salant, 1987) might also contribute to deficits in streaming.

In contrast to the results of Rose and Moore (2005), Mackersie et al. (2001) measured stream-segregation abilities using tones in normal-hearing listeners and older individuals with hearing loss and showed that, on average, the older listeners with hearing loss had higher fission boundaries than did listeners with normal hearing. The mean ages of the two groups used in their experiment were vastly different, and although the effect of hearing loss was significant even after partialing out the effects of the ages of the listeners, it still cannot be verified that age differences between the two groups did not contribute to these results. Grimault, Micheyl, Carlyon, Arthaud, and Collet (2001) also suggested that aging could lead to a decreased ability to separate sounds into different auditory streams.

One aspect of all of the aforementioned studies is that the stimuli used to test streaming abilities in individuals with hearing loss were either pure tones or harmonic complexes. For most of the studies, frequencies for which a phase-locked response would have been present were tested, thus providing the ear a temporal code for stream segregation. The impaired auditory system does not have as robust a place code as the healthy auditory system and therefore might rely on a temporal code for the formation of auditory streams. In the experiments described here, we measured auditory streaming abilities in listeners with normal hearing and in listeners with hearing loss, but instead of using harmonic complexes or pure tones, we employed inharmonic complexes. These stimuli might elicit a local phase-locked response (as would occur for a single tone in the complex) but would not elicit a consistent phase-locked response across frequency. By eliminating the coherent temporal pattern across frequency, we can evaluate (a) whether inharmonic stimuli can be used to form separate perceptual streams and (b) whether hearing loss influences streaming abilities in the absence of an across-frequency temporal cue. No study has evaluated whether aperiodic multitonal complexes can be perceived in separate auditory streams and whether hearing loss influences streaming of aperiodic sounds. The two experiments presented here approach this question in two ways. To minimize bias effects within and across listeners, the first experiment adopts an objective (bias-free) method of measuring auditory streaming similar to that described by Cusack and Roberts (2000). In a second experiment, a traditional (but more subjective) method of measuring stream segregation is used.

Experiment 1

Using an objective stream segregation paradigm developed by Cusack and Roberts (2000), this experiment evaluated stream segregation abilities for broadband inharmonic stimuli that stimulate multiple frequency regions in the cochlea and have no consistent temporal pattern across frequency. In this experiment, listeners detected a temporal change in a sequence of two inharmonic sounds. Detecting this change was expected to be more difficult if the two sounds were perceived in different auditory streams. Performance on this task was compared between normal-hearing listeners and individuals with hearing loss.

Method

Observer characteristics. Participants were 8 normal-hearing listeners, ranging in age from 28 to 64 years, with a mean age of 48 years, and 7 individuals with hearing loss, ranging in age from 19 to 65 years, with a mean age of 47 years. Normal-hearing listeners had pure tone audiometric thresholds no greater than 20 dB HL (American National Standards Institute [ANSI], 1996) between 250 and 6000 Hz, inclusive. Individuals with hearing loss were selected so that the mean threshold at 2000, 3000, and 4000 Hz was between 40 and 65 dB HL in the test ear. This criterion ensured hearing loss at the stimulus frequencies (1000–4000 Hz) and prevented large losses from rendering any of the stimulus components inaudible (see the *Stimuli* section), although it should be noted that the sensation levels of the stimuli would vary across listeners and between groups. Hearing losses were moderate and bilateral. The site of lesion was presumed to be cochlear based on the agreement between air and bone-conduction thresholds as well as normal immittance audiometry. The audiometric configurations for all test ears (the better ear) together with the participant's age are reported in Table 1.

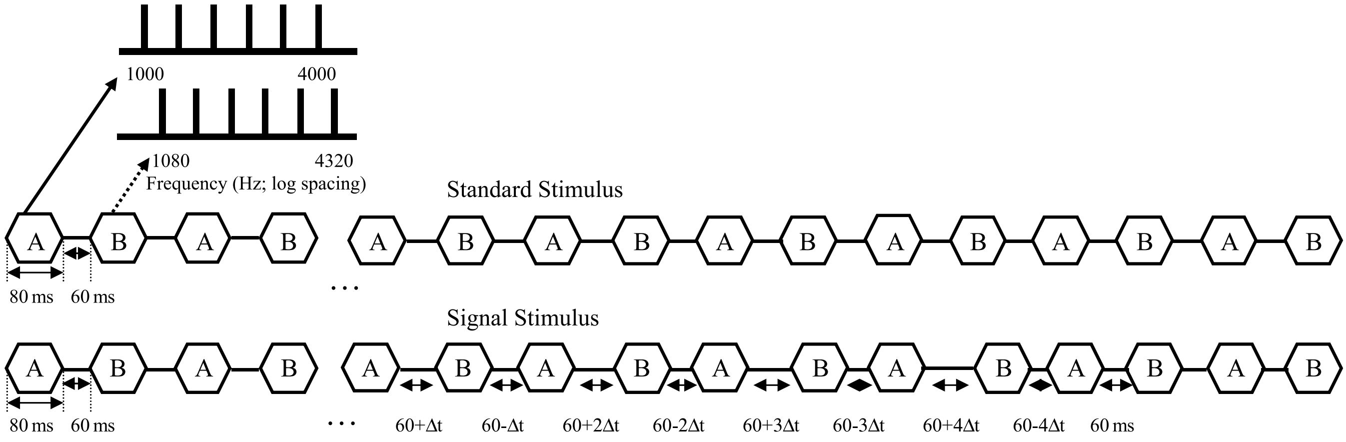
Table 1. Audiometric thresholds (db HL re: ANSI, 1996) of test ear for listeners with normal hearing (NH) and hearing impairment (HI) in Experiment 1.

			Frequency (Hz)						
Observer	Age	Test ear	250	500	1000	2000	3000	4000	8000
NH1	28	R	5	5	10	15	10	5	5
NH2	48	L	5	5	0	5	10	0	5
NH3	57	L	0	5	5	0	0	5	10
NH4	64	R	10	10	20	15	5	10	20
NH5	64	L	10	10	15	10	20	15	30
NH6	32	R	5	10	5	5	10	10	15
NH7	49	R	5	5	10	5	5	0	15
NH8	43	R	10	5	10	10	10	10	20
HI1	65	L	10	10	25	50	50	55	60
HI2	39	R	20	20	35	35	50	50	60
HI3	54	R	15	20	45	60	60	60	60
HI4	19	L	30	30	50	55	50	50	65
HI5	53	R	15	10	25	50	45	45	50
HI6	54	R	20	20	25	40	45	35	25
HI7	48	L	10	10	15	35	45	50	55

Note. R = right; L = left.

Stimuli. Two stimuli, A and B, were generated as the sum of six equal-amplitude sinusoids spaced equidistantly on a logarithmic scale spanning two octaves. The bandwidth of both stimuli was two octaves, with the A stimulus having frequencies of 1000, 1320, 1740, 2300, 3036, and 4000 Hz (frequency ratio = 1.32; see Figure 1). We tested high frequencies because these were where our participants had clinically significant hearing loss. Components in the B stimulus were generated using the same frequency ratio between successive harmonics as in the A stimulus but were scaled upward by a multiplicative factor of 1, 1.03, 1.06, 1.08, 1.1, 1.2, and 1.4 ($\boldsymbol{f}_B/\boldsymbol{f}_A$, where $\boldsymbol{f}_B$ and $\boldsymbol{f}_A$ denote the frequencies of the B and A stimuli). Thus, the lowest frequency in the B stimulus was 1000, 1030, 1060, 1080, 1100, 1200, or 1400 Hz. For example, when the B stimulus was generated at $\boldsymbol{f}_B/\boldsymbol{f}_A = 1.08$, its component frequencies were 1080, 1420, 1880, 2480, 3270, and 4320 Hz (see Figure 1). On each stimulus presentation, the starting phases of the component tones were selected randomly and independently from a uniform distribution ranging from 0 to 2π rad. The total duration of each stimulus (A or B) was 80 ms, including 16-ms cosine-squared rise/fall times. Each sequence consisted of 12 A stimuli and 12 B stimuli in the format A_B_A_B_A_B …, with A and B representing the different multitonal stimuli and the underscore (_) denoting a silent interval. In the standard sequence, the silent interval was always 60 ms (overall duration was 3.36 s). Signal sequences differed from the standard sequences in that on the 7th cycle of the sequence, a delay (Δt ms) was added to the silent interval between the A and B stimuli. Additionally, the silent interval between the B stimulus at the end of Cycle 7 and the A stimulus at the beginning of Cycle 8 was reduced

Figure 1. Schematic of the A_B_ paradigm used in this stream segregation task. The bottom illustration indicates the temporal sequence of A and B stimuli, with the underscore (_) denoting a silent interval. The top illustration shows the spectra of the A and B multitonal complexes in the condition where the frequencies in the B stimulus were generated by multiplying the frequencies of the A stimulus by 1.08.



by Δt ms. Over the next three A_B_ cycles, the delay between A and B was progressively increased by an additional Δt, leading to a final cumulative delay of 4Δt on the 10th A_B_ cycle (Figure 1). The reduction in the duration of the silent interval between the B stimulus of the 10th cycle and the A stimulus of the 11th cycle was 60–4Δt ms. On the 11th and 12th cycles, the silent delay reverted to the initial value of 60 ms (Figure 1). This procedure kept the duration of the signal and standard stimuli the same but is a modification of the procedure used by Cusack and Roberts (2000), in which the isochronous rhythm was maintained during the final 2 cycles.

Each component in the multitonal complexes was presented at a level of 80 dB SPL, thereby ensuring that all components were audible to the listeners with hearing loss by at least 10 dB when compared with audiometric thresholds (see Table 1). Signal and standard stimuli were generated and summed digitally and played through one channel of a 24-bit digital-to-analog converter (DAC; TDT System III RP2.1) at a sampling period of 4.096×10^{-5} s (sampling rate is about 24414 Hz). The output for the DAC was fed into a programmable attenuator that was adjusted to appropriately calibrate the stimuli, and then the output was fed into a single earphone of a Sennheiser HD 250 II Linear headset.

Procedure. A two-alternative forced-choice task was used to estimate the ability of listeners to discriminate between the standard sequence without the temporal delay and the signal sequence containing the temporal delay. Observers were seated in a sound-attenuating room and heard the standard and signal sequences separated by a 750-ms interstimulus interval. The first interval was as likely as the second interval to contain the signal sequence, with the remaining interval containing the standard sequence. Listeners indicated which interval contained the signal stimulus by responding on a button box. Correct-answer feedback was provided to the listener following each trial. Trial-by-trial Δt levels were chosen according to a two-down, one-up adaptive tracking procedure estimating the 70.1% correct point on the psychometric function (Levitt, 1971).

Threshold estimates for each condition were collected in blocks of two threshold estimates. For the first of the two threshold estimates, the starting signal strength (Δt) was 10 ms. Initially, Δt was adjusted by a multiplicative factor of 1.4 (up) and 1/1.4 (down), and after two reversals Δt was adjusted by a factor of 1.2. A track ended after five reversals, with the threshold estimate obtained as the mean of the delays at the final three reversal points. The second threshold estimate was measured using the same procedure as for the first threshold estimate; only the initial signal strength was set to be the previous estimate of threshold. Threshold estimates were typically based on 13 to 27 trials. This procedure was adopted in an attempt to shorten the duration of the second track, as the duration of an individual trial exceeded 6 s.

Thresholds were collected using a randomized block design, in which the frequency ratio of the B to the A stimuli (for all $f_A \neq f_B$) was selected at random. Two threshold estimates were obtained in this condition before a new frequency ratio was selected without replacement. Once all frequency ratios were tested, two new thresholds were run with the frequencies in the same random order. This process was repeated seven additional times for a total of 16 thresholds in each condition. Thresholds for the $f_B/f_A = 1.0$ condition were collected either in the beginning of the experiment or at the very end of the experiment. Reported thresholds are the mean of the last 12 thresholds collected. Thresholds were collected over four to five experimental sessions of 2 hr, for a total of 8–10 hr of listening. Of this time, the first 4 thresholds in each condition were treated as practice, for a total of approximately 2 hr of training. Subject HI7 was unavailable to complete the 1.0 frequency condition. Therefore, data for this subject consisted of frequency conditions in which the B stimulus did not have the same frequencies as the A stimulus ($f_A \neq f_B$).

Results

Figure 2 plots the just detectable delay (Δt) as a function of the frequency separation between the B and A stimuli (f_B/f_A) averaged across listeners in each group. Data obtained from listeners with hearing loss and with normal hearing are plotted as the unfilled squares and filled triangles, respectively. Error bars reflect standard errors of the mean across the 7 hearing-impaired listeners (six for $f_B/f_A = 1.0$) and the 8 normal-hearing listeners.

In general, Figure 2 shows that increasing frequency separation between the A and B stimuli leads to decreases in sensitivity to inharmonic complex stimuli for both groups of listeners. That is, as the B stimulus becomes more distinct from the A stimulus, performance degrades, suggesting that two streams are more likely to be perceived at the large frequency separations. A repeated-measures analysis of variance (ANOVA) with group membership as a between-subjects variable and the frequency separation between the B and A stimuli (f_B/f_A) as a within-subjects variable indicated a significant main effect of frequency, $F(1, 13) = 39.4$, $p < .001$. However, there was no significant main effect of group membership, $F(1, 13) = 2.46$, $p < .14$, nor a significant interaction between group membership and frequency, $F(1, 13) = 0.56$, $p < .82$. Note that the condition in which $f_B/f_A = 1.0$ was not included in the ANOVA was because Subject HI7 did not complete that condition. A t test conducted on the data obtained for $f_B/f_A = 1.0$ (8 normal-hearing listeners and 6 individuals with hearing loss) also indicated no significant effect of group membership, $t(12) = -0.64$, $p = .53$. The lack of a significant main effect or interaction involving group membership across the conditions could be due to large within-group variability, small group size, or no influence of hearing loss on auditory streaming abilities.

Figure 3 plots the individual data for better observation of variability across individuals. Data obtained from individuals with hearing loss are plotted with dotted lines and open symbols, and data obtained from normal-hearing listeners are plotted with solid lines and filled symbols. The different symbols indicate the average of the 12 individual threshold estimates obtained from different individuals. In Figure 2, it appeared that on average, listeners with hearing loss had poorer sensitivity than did listeners with normal hearing; however,

Figure 2. Average thresholds for 7 subjects with hearing impairment (HI; unfilled squares and dotted line) and 8 normal-hearing listeners (NH; filled triangles and solid line), plotted as a function of the frequency ratio between the A and B stimuli (f_B/f_A). Error bars represent the standard error of the mean.

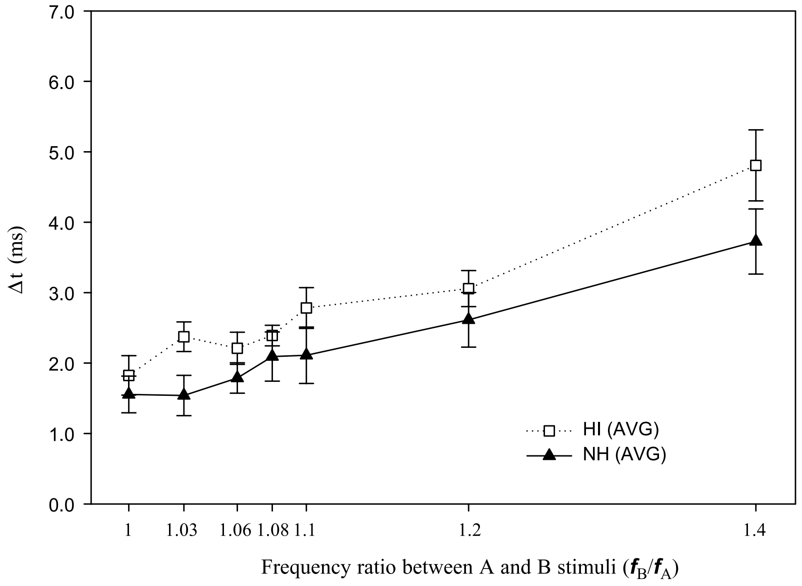
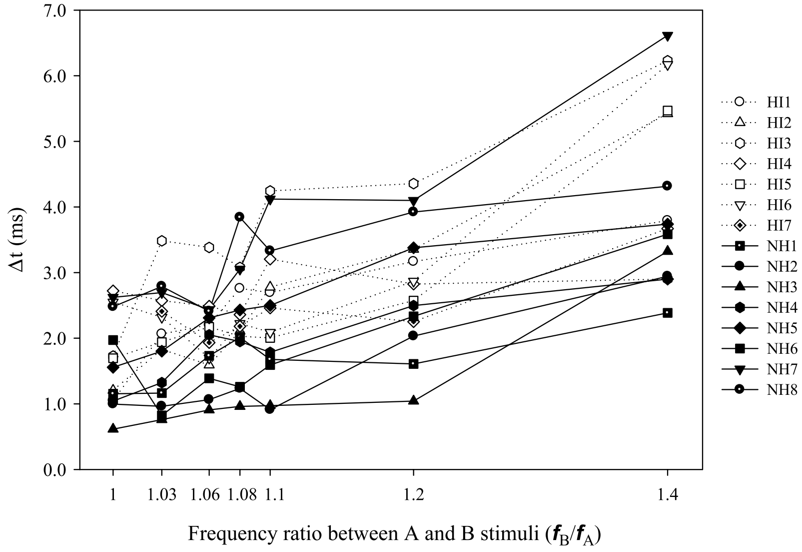


Figure 3. Individual thresholds for the subjects with hearing impairment (unfilled symbols, dotted lines) and normal-hearing listeners (filled symbols, solid lines), plotted as a function of the frequency ratio between the A and B stimuli (f_B/f_A).



that difference was not supported statistically. The lack of statistical significance may be due to the large variability evident in Figure 3, which shows the substantial overlap among data obtained from individual listeners. Of the normal-hearing listeners, 3 or 4 (NH1, NH2, NH3, and possibly NH4) have thresholds that are consistently better than any of the individuals with hearing loss. However, the other normal-hearing listeners have thresholds that are similar to the thresholds of individuals with hearing loss. Furthermore, 2 of the normal-hearing listeners (NH7 and NH8) have thresholds that are similar to the highest thresholds in the group of listeners with hearing loss. Nothing about these 2 observers stands out—they both had audiometric thresholds within normal limits, and they were not the oldest listeners within the normal-hearing group. One source of variability for the group of listeners with hearing loss may have been the differences in sensation levels between individuals; however, no consistent relationship was found between differences in sensation levels and the results. Additionally, had differences in sensation level contributed to the results obtained in this study, then it is likely that group differences would have been observed, which was not the case. The results of this experiment revealed large individual differences in streaming ability for listeners with normal and impaired hearing.

Discussion

This experiment used an objective streaming task to determine whether listeners with hearing loss have different streaming abilities than listeners with normal hearing for broadband, inharmonic sounds. This procedure is based on that of Cusack and Roberts (2000), who argued that detecting a temporal delay between two stimuli in an alternating sequence should be more difficult when two sounds are perceived in separate auditory streams. Applying this logic to Experiment 1, then, at large frequency separations, A and B stimuli are more likely to be in two separate auditory streams than when their frequencies are similar. When in separate streams, the differences in temporal patterns are more difficult to perceive. Because the experimental data indicate increases in the just detectable temporal delay between two stimuli with increasing frequency separations, streaming is demonstrated to occur for inharmonic stimuli as well as for the periodic stimuli used in the past (Grimault, Micheyl, Carlyon, Arthaud, & Collet, 2000; Grimault et al., 2001; Stainsby et al., 2004; Vliegen et al., 1999; Vliegen & Oxenham, 1999).

Further applying this logic to the data obtained from the group of listeners with hearing loss, then, we would expect one of two results. If individuals with hearing loss have poorer stream segregation abilities than do normal-hearing listeners, then the thresholds in the former group would be expected to be lower than the thresholds obtained in the normal-hearing group. That is, if two stimuli are less likely to be perceived in two separate streams, then this task should be easier for them. If individuals with hearing loss have similar stream segregation abilities to normal-hearing listeners, then one would anticipate no difference in their ability to accomplish this task. The results of this experiment support the latter idea to a greater extent than the former, as thresholds were not statistically different between the two groups of listeners. Furthermore, although some of the normal-hearing listeners have much better thresholds than do any of the individuals with hearing loss, this result is in the opposite direction to that anticipated.

In addition, it should be recognized that low thresholds on this task do not reflect only auditory streaming abilities, as the ability to detect a temporal difference between two sounds also impacts the magnitude of the just detectable delay. Listeners with poor temporal processing abilities may be less able to detect a temporal difference between two tasks, which would result in higher just-detectable delay thresholds. In this case, there is a possibility that listeners with good temporal processing are those who perform well, whereas listeners with poor temporal processing are those who perform poorly. To evaluate this possibility, we examined the condition in which the frequencies in the A and B stimuli were the same ($\boldsymbol{f}_B/\boldsymbol{f}_A = 1.0$). Thresholds in this condition would reflect temporal processing abilities, as the A and B stimuli would have been perceived in the same auditory stream. Figure 3 indicates that there was substantive variability in the thresholds obtained for this condition, with thresholds in the normal-hearing group ranging from 0.6 ms to 2.6 ms. Thus, our selection of subjects sampled a group with wide variation in their temporal-processing abilities. Data obtained from the individuals with hearing loss reflect similar variability in their performance, with thresholds ranging from 1.05 ms to 2.72 ms. Furthermore, as mentioned in the results, the *t* test on these data did not reveal a significant difference between the two groups. Large variability across subjects has been reported in studies investigating auditory stream segregation (Cusack & Roberts, 2000; Stainsby et al., 2004) and other temporal processes (Gordon-Salant & Fitzgibbons, 1999). Listeners typically show a broadly consistent pattern of data but show differences in the level of performance. This is essentially true for the results of this study, as all listeners showed a pattern of results in which their thresholds increased as the frequency of the B stimulus increased, but the overall performance level between subjects varied considerably.

To determine whether those with low thresholds for $\boldsymbol{f}_B/\boldsymbol{f}_A = 1.0$ also have low thresholds when $\boldsymbol{f}_A \neq \boldsymbol{f}_B$, we calculated correlations between subjects' threshold when the stimuli have the same frequencies ($\boldsymbol{f}_B/\boldsymbol{f}_A = 1.0$) and when the stimuli are separated by a frequency ratio of 1.4, the largest difference tested. A significant positive correlation was obtained for the normal-hearing subjects ($r = .76, p < .03$), but not for the 6 subjects with hearing loss who were tested on both conditions ($r = -.2, p = .71$). Thus, the normal-hearing listeners who have good temporal processing abilities (low thresholds for $\boldsymbol{f}_B/\boldsymbol{f}_A = 1.0$) also have low thresholds when the B stimulus has very different frequencies. In this case, the magnitude of the just detectable delay likely reflects both temporal processing abilities and auditory streaming abilities for all $\boldsymbol{f}_B/\boldsymbol{f}_A > 1.0$. As such, streaming abilities cannot be inferred from the magnitude of the just detectable delay alone. For the individuals with hearing loss, there was not a strong relationship between their performance when the stimuli had the same frequencies and when the stimuli were separated by a frequency ratio of 1.4. In this case, it is not clear whether absolute performance is indicative of only streaming abilities.

Given that the just detectable temporal delay on its own is not a clear indicator of streaming abilities, a better metric related to streaming ability would be the slope of the functions relating threshold (Δt) to the frequency ratio ($\boldsymbol{f}_B/\boldsymbol{f}_A$). Here, poorer streaming abilities would be indicated by a shallower slope, thereby reflecting an inability to separate two stimuli into two separate streams. Different streaming abilities, as reflected in the slope values, between individuals with normal hearing (average slope = 5.57 ms/octave separation) and individuals with hearing loss (average slope = 7 ms/octave separation) would appear in the statistical analysis as an interaction between group membership and frequency ratio. Because there was no significant interaction between group and frequency ratio ($p < .82$), we conclude that individuals with hearing loss have similar segregation abilities to normal-hearing listeners for inharmonic stimuli. Such a result is interesting because these stimuli do not evoke a clear pitch, nor would they elicit a coherent phase-locked response across frequency. Even though a listener might be required to rely more on a place code than in previous streaming experiments, individuals with hearing loss were still found to have similar streaming abilities to normal-hearing listeners.

These results obtained from inharmonic stimuli are along the same lines as those reported for pure tones. Specifically, Mackersie et al. (2001) and Rose and Moore (1997, 2005) showed that frequency selectivity measures do not relate to streaming abilities for pure tones. Of the three studies, only Mackersie et al. (2001) showed that individuals with hearing loss had poorer streaming abilities than did normal-hearing listeners. One large difference between the current study and theirs (in addition to the composition of the stimuli) is that we specifically selected listeners so that the mean ages of the normal-hearing group and the group of listeners with hearing loss were very similar, whereas the listeners with hearing impairment in Mackersie et al.'s (2001) study had a mean age of 73 years but the normal-hearing listeners had a mean age of 27 years. In the current study, a moderate correlation between participant age and streaming ability (the slope of the function relating threshold to the frequency difference, $\boldsymbol{f}_B/\boldsymbol{f}_A$) was found ($r = .47, p = .08$). Although this correlation is not significant, it does suggest a relationship between streaming abilities and age. Given that streaming abilities could degrade with increasing age, it is possible that the differences measured between the two groups in previous studies may not have been based solely on peripheral changes in the auditory system as a result of cochlear hearing loss.

In an informal follow-up study, we investigated the possibility that listeners were making judgments within a single auditory stream instead of across auditory streams. That is, listeners may have been making their decisions based only on the B stimulus and not using the A stimulus for comparison. The procedures for this experiment were identical to those used in the main experiment, but here, instead of the A stimulus being presented, silence was presented in its place. Again, listeners were asked to detect a small change in the duration of a temporal gap within the stream for three frequency conditions ($f_{\mathrm{B}}/f_{\mathrm{A}}$ =1.0, 1.2, and 1.4). Two normal-hearing listeners and two individuals with hearing loss participated in this follow-up; both of the individuals with hearing loss and 1 of the normal-hearing listeners participated in Experiment 1; the other normal-hearing listener was the first author of this article. The results showed that thresholds increased by 0.9 ms, on average, between $f_{\mathrm{B}}/f_{\mathrm{A}}$ =1.0 and 1.4. This increase in threshold is much smaller than observed in the data (a mean increase in threshold of 2.6 ms over the same range). Therefore, we interpret our experimental results as suggesting that listeners were comparing the A stimulus with the B stimulus in Experiment 1.

In summary, using an objective auditory streaming paradigm, individuals with hearing loss were found to have streaming abilities similar to normal-hearing listeners. However, this paradigm has one large confound, which is that temporal processing abilities might contribute to the measurement of streaming abilities. Therefore, the next experiment examines auditory stream segregation using a traditional measure of auditory streaming.

Experiment 2

Subjective streaming abilities are frequently measured using a paradigm in which two tones are alternating in a sequence (e.g., ABA_ABA_ … where A and B represent two tones with different frequencies and the underscore [_] denotes a silent interval; Rose & Moore, 1997; Van Noorden, 1977). In a traditional streaming experiment the frequency of the B tone changes while the frequency of A is held constant. The frequency at which the observer can no longer hear two separate streams is referred to as the fission boundary. The following experiment uses this procedure to determine whether individuals with hearing loss have similar streaming abilities to listeners with normal hearing for broadband inharmonic stimuli.

Method

Observer characteristics. Participants were 4 normal-hearing listeners, ranging in age from 25 to 50 years, with a mean of 36.5 years (SD = 9 years), and 4 individuals with hearing loss, ranging in age from 47 to 64 years, with a mean age of 49.5 years (SD = 10.7 years). No significant difference in the mean ages of the groups was present, $t(6) = -1.85$, $p = .114$. Selection criteria were the same as in Experiment 1. The audiometric configurations for all test ears together with the participant's age are reported in Table 2. Subjects HI1, HI2, and HI7 from Experiment 1 also participated in this experiment, but we were unable to test any of the same normal-hearing participants in this experiment.

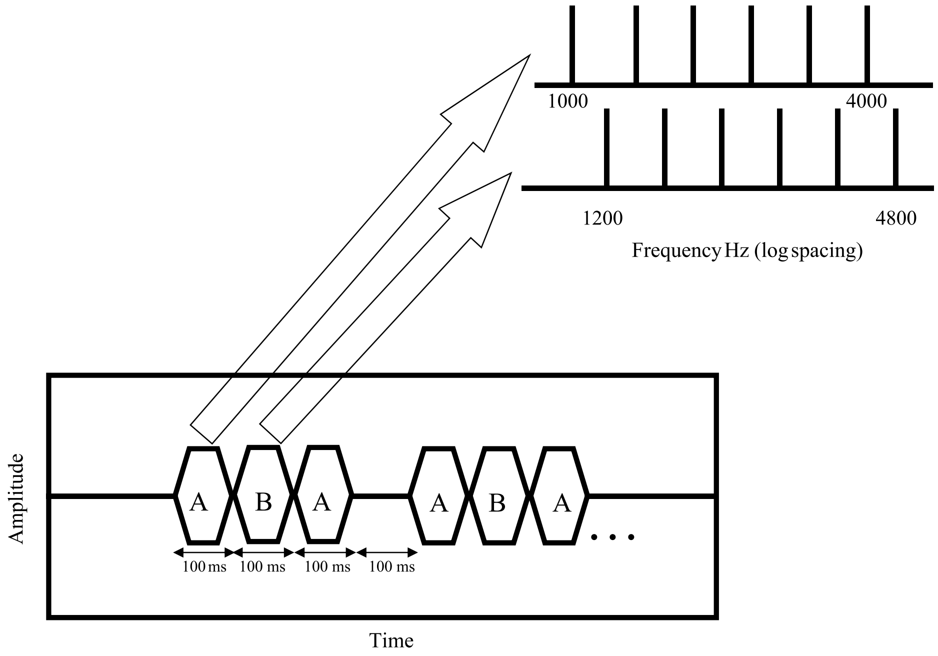
Stimuli. Stimuli consisted of a sequence of two alternating stimuli, A and B, in the following sequence: ABA_ABA_… ABA. Here, A and B represent inharmonic multitonal stimuli and the underscore (_) denotes a silent interval. As in Experiment 1, A and B stimuli had bandwidths of two octaves and were the sum of six equal-amplitude sinusoids spaced equidistantly on a logarithmic scale. Two frequency conditions were used in this experiment: HIGH and LOW. In the HIGH condition, frequencies of the components in the A stimulus ranged between 1000 and 4000 Hz (frequency ratio = 1.32; see Figure 4), whereas in the LOW condition, frequencies of the components in the A stimulus ranged between 250 and 1000 Hz (frequency ratio = 1.32). The B stimuli (which will be described in more detail in the *Procedure* section) were always higher in frequency than the corresponding A stimulus. The total duration of each stimulus (A or B) was 100 ms, including 15-ms cosine-squared rise/fall times. The silent interval (_) also had a duration of 100 ms.

Table 2. Audiometric thresholds (dB HL re: ANSI, 1996) of test ear for listeners with normal hearing and hearing loss in Experiment 2.

			Frequency (Hz)						
Observer	Age	Test ear	250	500	1000	2000	3000	4000	8000
NH9	47	R	5	10	0	5		5	
NH10	37	L	0	0	5	5	0	5	10
NH11	37	R	0	0	0	5	0	0	5
NH12	25	R	15	5	5	15	5	5	0
HI8	49	R	15	20	25	45	50	50	55

Note. Thresholds for listeners who participated in both experiments (HI1, HI2, and HI7) are listed in Table 1.

Figure 4. Schematic of the ABA_ paradigm with the corresponding power spectra of the A and B complexes for the HIGH condition.



Each multitonal complex was presented at a level of 85 dB SPL per component. A and B stimuli were generated and summed digitally and then played through one channel of a 24-bit DAC (TDT System III RP2.1) at a sampling period of 4.096×10^{-5} s (sampling rate is about 24414 Hz). The output for the DAC was fed into a programmable attenuator that was adjusted to appropriately calibrate the stimuli. The output was fed into the right headphone of a Sennheiser HD 250 II Linear headset.

Procedure. A frequency sweep procedure was used to measure the fission boundary (the frequency at which the two stimuli in the sequence could not be perceived as two streams). The frequencies of the A stimulus were kept constant, while the frequencies of the B stimulus were variable, but the frequency ratio between adjacent tones always remained the same. At the beginning of each sweep, the tones within the B stimulus started at frequencies higher than those in Stimulus A and progressed toward the frequencies in Stimulus A in an exponential manner, following the function:

$$\boldsymbol{f}_{\mathrm{B}}' = \boldsymbol{f}_{\mathrm{A}}(\boldsymbol{f}_{\mathrm{B}}/\boldsymbol{f}_{\mathrm{A}})^{1/1.08} \qquad (1)$$

In equation 1, $\boldsymbol{f}_{\mathrm{A}}$ represents the frequencies in the A stimulus, $\boldsymbol{f}_{\mathrm{B}}$ represents the frequencies in the B stimulus, and $\boldsymbol{f}_{\mathrm{B}}$ indicates the frequencies of the B stimulus in the subsequent ABA group (ABA). The experiment investigated two distinct frequency conditions: HIGH and LOW. At the beginning of the frequency sweep, the lowest frequency component of the B stimulus was selected at random to be either 1200, 1184, 1169, or 1157 Hz for the HIGH condition and to be either 375, 364, 354, or 345 Hz for the LOW condition. Observers were seated in a sound-attenuating room and were instructed to press a button when they "could no longer hear two separate streams." The fission boundary is described as the frequency ratio between the components in the B stimulus with respect to the components in the A stimulus ($\boldsymbol{f}_{\mathrm{B}}/\boldsymbol{f}_{\mathrm{A}}$) when the button is pressed.

Fission boundaries were collected using a randomized block design, in which the condition (either HIGH or LOW) was selected at random. Eight estimates of the fission boundary were obtained in the same condition before the other condition was selected. Once eight estimates of the fission boundaries were obtained in each condition, this process was repeated until a total of 128 fission boundary estimates were obtained for each condition. Reported fission boundaries are the average of the 128 fission boundary estimates collected.[1] Each participant received 2 hr of training, which equated to approximately 75 additional fission boundary estimates that were not used in data analysis.

---

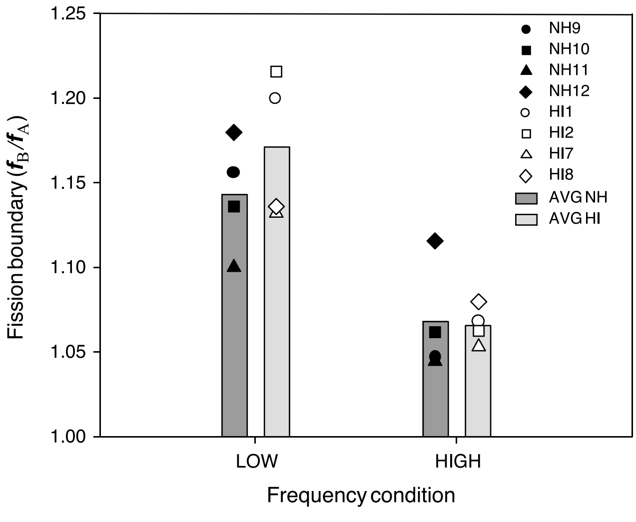
[1] Stimuli in Experiment 2 were created so that there were eight different stimuli with different phases. In order to obtain a sufficient number of thresholds from each different phase, we collected a total of 128 thresholds. Initially, we planned to look at any influence phase had on auditory stream segregation; however, preliminary evaluation showed no consistent difference in thresholds between phases.

Results and Discussion

Figure 5 plots the fission boundaries ($\boldsymbol{f}_B/\boldsymbol{f}_A$) for individual normal-hearing listeners and individuals with hearing loss as well as the average boundaries for each group in the LOW and HIGH conditions. Figure 5 indicates that the mean fission boundaries for individuals with hearing loss were very similar to those for normal-hearing listeners for these inharmonic stimuli. Additionally, between-subject variability can also be observed in Figure 5. Fission boundaries were also lower for the HIGH condition than for the LOW condition for both groups of listeners. A repeated-measures ANOVA with group membership as a between-subjects variable and the frequency range (HIGH or LOW) as a within-subjects variable indicated a significant main effect of frequency range, $F(1, 6) = 50.2, p < .001$, and no significant main effect of group or interaction involving group membership.

In general, the results indicate that normal-hearing listeners and individuals with hearing loss have similar fission boundaries for inharmonic stimuli. For the frequencies tested in the LOW condition, most of the individuals with hearing loss had audiometry thresholds within normal limits (25 dB HL or better) whereas for the HIGH condition, the individuals with hearing loss all had a sloping audiometric configuration. Because hearing thresholds for the listeners with hearing loss were better in the low frequency region, the LOW condition could be used as a "control" condition in case fission boundaries differed between groups in the HIGH condition. Regardless, there was no interaction between group status and the frequency range, providing further evidence that hearing loss does not influence streaming abilities. In general, these findings based on inharmonic stimuli support other experiments that have used pure tone and harmonic stimuli and have also shown that sensorineural hearing loss does not adversely affect a listener's ability to form auditory streams (Rose & Moore, 1997, 2005).

Figure 5. Average (AVG; bars) and individual (symbols) fission boundaries ($\boldsymbol{f}_B/\boldsymbol{f}_A$) for 4 subjects with hearing impairment (light gray bars and unfilled symbols) and 4 normal-hearing listeners (dark gray bars and filled symbols).



In addition to the result that the listeners with hearing loss and with normal hearing had similar fission boundaries, fission boundaries (measured in relative units) were found to be lower for the high-frequency stimulus than for the low-frequency stimulus. This finding is not consistent with results on the frequency dependence of fission boundaries using tonal stimuli, as fission boundaries in relative units have been measured to be fairly constant across frequency. Miller and Heise (1950) examined the perception of a tone sequence in which two frequencies were alternated at a rate of 10 tones per second. They determined the frequency separation at which the "trill seems to break" and found that the trill threshold was roughly equal in size to the auditory filter bandwidth associated with the lowest frequency in the trill. Rose and Moore (1997) also showed that the fission boundary for tones expressed in terms of the change in the equivalent rectangular bandwidth number between two tones (ΔE) tended to be similar from 250 to 2000 Hz. The average fission boundaries obtained here correspond to ΔE values of .79 for the low-frequency condition and .51 for the high-frequency condition. The fission boundaries for the high-frequency stimulus in this experiment are on the order of those reported by Rose and Moore (1997) for 1000-Hz pure tones, but the fission boundaries for the low-frequency stimulus are about 1.5 times higher. In this case, the higher fission boundaries might suggest that streaming for a low frequency, inharmonic stimulus is quite difficult for listeners with normal and impaired hearing. Furthermore, listeners were probably integrating information across the different frequency components of the complexes and not basing their streaming decisions on a single component.

These large fission boundaries are consistent with subject reports of having difficulty determining whether a "galloping" or "streaming" percept was present. Indeed, many subjects complained of not knowing when they could no longer hear two streams. In particular, these complaints suggest that a distinct fission boundary may not exist for inharmonic stimuli. One potential explanation for the weaker streaming percepts associated with inharmonic stimuli over harmonic stimuli may be related to the aperiodic nature of inharmonic stimuli. The harmonic stimuli used previously contain at least two cues indicating that the constituent frequency components originate from the same sound source: a consistent fundamental frequency across their bandwidth and shared onsets and offsets. Although psychophysical evidence indicates that simultaneous onsets provide a highly compelling cue that components originate from the same source (cf. Dai & Green, 1992; Hill & Bailey, 1997), it is possible that the multiple cues present for harmonic stimuli would lead to a more robust percept of a single sound. An inharmonic relationship between stimulus components has been demonstrated to influence the perception of multiplicity in double-vowel and harmonic mistuning studies (e.g., De Cheveigné, McAdams, & Marin, 1997; Peters, Moore, & Glasberg, 1983).

A large difference (besides the frequency range) between the low- and high-frequency stimulus is that the frequencies used for the high-frequency stimulus would be expected to be more equally spaced in the cochlea across the 1000–4000 Hz frequency range than the 250–1000 Hz in the low-frequency stimulus. Even though the high and low stimuli both span two octaves, the logarithmically spaced components in the low-frequency stimulus may not be equally spaced in the cochlea. This difference in spacing in the cochlea might lead to more difficulty streaming in the low frequency complex than in the high frequency complex. However, within-subject standard deviations calculated on the fission boundaries for the low and high conditions are similar, which suggests that the subjects experienced the same amount of variability or difficulty in determining the fission boundaries for each condition.

Experiment 1 demonstrated that temporal processing might play a role in measurement of streaming abilities. Temporal processing abilities could also influence performance on the subjective streaming task, as a "galloping" or "streaming" percept must also be dependent on temporal processes. As such, studies that have demonstrated that some listeners with hearing loss have poorer streaming abilities might be reflecting changes to temporal processing abilities. Whether listeners with hearing loss experience temporal deficits has yet to be definitively decided and could be related to reduced audibility (DeFilippo & Snell, 1986; Fitzgibbons & Gordon-Salant, 1987; Glasberg, Moore, & Bacon, 1987; Grose, Eddins, & Hall, 1989; Moore, Glasberg, Donaldson, McPherson, & Plack, 1989; Tyler, Summerfield, Wood, & Fernandes, 1982). Temporal processing deficits are more definitively associated with aging (Fitzgibbons & Gordon-Salant, 2004; Gordon-Salant & Fitzgibbons, 1999) and, as such, could explain, in part, the results of Mackersie et al. (2001) and Grimault et al. (2001), who showed that older individuals with hearing loss had larger fission boundaries than did young normal-hearing listeners. Grimault et al. (2001) did find differences in the streaming abilities of older listeners with hearing loss and those without. However, reduced temporal processing abilities cannot be ruled out as an explanative factor.

Taken together, these results provide additional support that stream segregation is not driven by peripheral factors. Furthermore, inharmonic complexes can be separated into different auditory streams, but such percepts might not be as robust as those evoked by pure tones or harmonic complexes.

Summary and Conclusion

The two experiments presented here investigated auditory stream segregation abilities in normal-hearing listeners and persons with hearing loss using broadband inharmonic stimuli. Inharmonic stimuli were selected so that robust phase-locked responses would not be present across frequency channels, and therefore, listeners would be more likely to rely on place cues for streaming. These stimuli were demonstrated to facilitate the formation of auditory streams. In addition, normal-hearing listeners and persons with hearing loss were demonstrated to have similar auditory streaming abilities. Furthermore, these experiments tested listeners with normal hearing and hearing loss with somewhat similar age ranges. Because on two separate metrics (an objective and a subjective streaming task) both groups of listeners had similar streaming abilities, results suggest that previously measured differences in fission boundaries may be due to differences in temporal processing that might be associated with aging or hearing loss.

Acknowledgments

This work was supported by Grant DC 005835 from the National Institute on Deafness and Communication Disorders. The authors thank Melissa Papesh for assistance with data collection and Larry Humes for valuable comments on this article and assistance in interpreting the data.

References

American National Standards Institute. (1996). *Specification for audiometers* (ANSI S3.6-1996). New York: Author.

Beauvois, M. W., & Meddis, R. (1996). Computer simulation of auditory stream segregation in alternating-tone sequences. *The Journal of the Acoustical Society of America, 99,* 2270–2280.

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound.* Cambridge, MA: MIT Press.

Bregman, A. S., & Campbell, J. (1971). Primary auditory stream segregation and perception of order in rapid sequences of tones. *Journal of Experimental Psychology, 89,* 244–249.

Cusack, R., & Roberts, B. (2000). Effects of differences in timbre on sequential grouping. *Perception and Psychophysics, 62,* 1112–1120.

Dai, H., & Green, D. M. (1992). Auditory intensity perception: Successive versus simultaneous, across-channel discriminations. *The Journal of the Acoustical Society of America, 91,* 2845–2854.

De Cheveigné, A., McAdams, S., & Marin, C. M. (1997). Concurrent vowel identification: II. Effects of phase, harmonicity, and task. *The Journal of the Acoustical Society of America, 101,* 2848–2856.

DeFilippo, C. L., & Snell, K. B. (1986). Detection of a temporal gap in low-frequency narrow-band signals by normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America, 80,* 1354–1358.

Fitzgibbons, P. J., & Gordon-Salant, S. (1987). Temporal gap resolution in listeners with high-frequency sensorineural hearing loss. *The Journal of the Acoustical Society of America, 81,* 133–137.

Fitzgibbons, P. J., & Gordon-Salant, S. (2004). Age effects on discrimination of timing in auditory sequences. *The Journal of the Acoustical Society of America, 116,* 1126–1134.

Glasberg, B. R., Moore, B. C. J., & Bacon, S. P. (1987). Gap detection and masking in hearing-impaired and normal-hearing subjects. *The Journal of the Acoustical Society of America, 81,* 1546–1556.

Gordon-Salant, S., & Fitzgibbons, P. J. (1999). Profile of auditory temporal processing in older listeners. *Journal of Speech, Language, and Hearing Research, 42,* 300–311.

Grimault, N., Micheyl, C., Carlyon, R. P., Arthaud, P., & Collet, L. (2000). Influence of peripheral resolvability on the perceptual segregation of harmonic complex tones differing in fundamental frequency. *The Journal of the Acoustical Society of America, 108,* 263–271.

Grimault, N., Micheyl, C., Carlyon, R. P., Arthaud, P., & Collet, L. (2001). Perceptual auditory stream segregation of sequences of complex sounds in subjects with normal and impaired hearing. *British Journal of Audiology, 35,* 173–182.

Grose, J. H., Eddins, D. A., & Hall, J. W., III (1989). Gap detection as a function of stimulus bandwidth with fixed high-frequency cutoff in normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America, 86,* 1747–1755.

Hartmann, W. M., & Johnson, D. (1991). Stream segregation and peripheral channeling. *Music Perception, 9,* 155–184.

Hill, N. I., & Bailey, P. J. (1997). Profile analysis with an asynchronous target: Evidence for auditory grouping. *The Journal of the Acoustical Society of America, 102,* 477–481.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America, 49,* 467–476.

Mackersie, C. L., Prida, T. L., & Stiles, D. (2001). The role of sequential stream segregation and frequency selectivity in the perception of simultaneous sentences by listeners with sensorineural hearing loss. *Journal of Speech, Language, and Hearing Research, 44,* 19–28.

McCabe, S. L., & Denham, M. J. (1997). A model of auditory streaming. *The Journal of the Acoustical Society of America, 101,* 1611–1621.

Miller, G. A., & Heise, G. A. (1950). The trill threshold. *The Journal of the Acoustical Society of America, 22,* 637–638.

Moore, B. C. J., Glasberg, B. R., Donaldson, E., McPherson, T., & Plack, C. J. (1989). Detection of temporal gaps in sinusoids by normally hearing and hearing-impaired subjects. *The Journal of the Acoustical Society of America, 85,* 1266–1275.

Moore, B. C. J., & Gockel, H. (2002). Factors influencing sequential stream segregation. *Acta Acustica United With Acustica, 88,* 320–333.

Peters, R. W., Moore, B. C. J., & Glasberg, B. R. (1983). Pitch of components of complex tones. *The Journal of the Acoustical Society of America, 73,* 924–929.

Roberts, B., Glasberg, B. R., & Moore, B. C. J. (2002). Primitive stream segregation of tone sequences without differences in fundamental frequency or passband. *The Journal of the Acoustical Society of America, 112,* 2074–2085.

Rose, M. M., & Moore, B. C. J. (1997). Perceptual grouping of tone sequences by normally hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America, 102,* 1768–1778.

Rose, M. M., & Moore, B. C. J. (2005). The relationship between stream segregation and frequency discrimination in normally hearing and hearing-impaired subjects. *Hearing Research, 204,* 16–28.

Stainsby, T. H., Moore, B. C. J., & Glasberg, B. R. (2004). Auditory streaming based on temporal structure in hearing-impaired listeners. *Hearing Research, 192,* 119–130.

Tyler, R. S., Summerfield, Q., Wood, E. J., & Fernandes, M. A. (1982). Psychoacoustic and phonetic temporal processing in normal and hearing-impaired listeners. *The Journal of the Acoustical Society of America, 72,* 740–752.

Van Noorden, L. P. (1977). Minimum differences of level and frequency for perceptual fission of tone sequences ABAB. *The Journal of the Acoustical Society of America, 61,* 1041–1045.

Vliegen, J., Moore, B. C. J., & Oxenham, A. J. (1999). The role of spectral and periodicity cues in auditory stream segregation, measured using a temporal discrimination task. *The Journal of the Acoustical Society of America, 106,* 938–945.

Vliegen, J., & Oxenham, A. J. (1999). Sequential stream segregation in the absence of spectral cues. *The Journal of the Acoustical Society of America, 105,* 339–346.

Received August 21, 2007

Revision received January 18, 2008

Accepted February 11, 2008

DOI: 10.1044/1092-4388(2008/07-0193)

Contact author: Susie Valentine, who is now at Starkey Hearing Research Center, 2150 Shattuck Ave., Suite 408, Berkeley, CA 94704. E-mail: susie_valentine@starkey.com.