ELSEVIER

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Hearing Research 202 (2005) 235–247

HEARING RESEARCH

www.elsevier.com/locate/heares

Attribute capture in the precedence effect for long-duration noise sounds

Liang Li [a,b,*], James G. Qi [b], Yu He [c], Claude Alain [c], Bruce A. Schneider [b]

[a] 100 1,

[b] 33 ., 1

[c] ., 2 1

Received 7 October 2004; accepted 13 October 2004
Available online 8 December 2004

Abstract

Listeners perceptually fuse the direct wave from a sound source with its reflections off nearby surfaces into a single sound image, located at or near the sound source (the precedence effect). This study investigated how a brief gap presented in the middle of either a direct wave or simulated reflection is incorporated into the fused image. For short (<9.5 ms) delays between the direct (leading) and reflected (lagging) waves, no sound was perceived from the direction of the lagging wave. For delays between 10 and 15 ms, both sounds were perceived, but the gap was heard only on the leading side. When the gap was only in the correlated lagging sound at short delays, it also was perceived as occurring on the leading side. Moreover, gap detection thresholds were the same for gaps in the leading and lagging sounds, suggesting that the perception of the gap was not suppressed, but rather incorporated into the leading sound. Finally, scalp event-related potentials were not associated with the precedence effect until the gap occurred. This suggests that cortical mechanisms are engaged to maintain fusion when attributes in direct or reflected waves change.

Precedence effect; Fusion; Reverberant environment; Correlation; Gap; Event-related potential

1. Introduction

In a reverberant environment, each sound source produces both a direct wavefront and numerous filtered and time-delayed reflections from the walls, ceilings and other surfaces. When the delay between the direct wave and a reflected wave is sufficiently long and the reflected wave is sufficiently intense, the reflected wave is perceived as a distinct auditory event (an echo), whose perceived location is usually different from that of the source. However, when the delays between the direct wavefront and its reflections are short (e.g., 1–10 ms or more, depending on the stimulus), the auditory system somehow gives "precedence" to the direct wavefront over its reflections so that the listener hears only a single fused sound whose point of origin is perceived to be at or near the location of the sound source. This phenomenon is called the "precedence effect" (Clifton and Freyman, 1989; Freyman et al., 1991; Shinn-Cunningham et al., 1993; Wallach et al., 1949; Zurek, 1980; for reviews see Blauert, 1997; Li and Yue, 2002; Litovsky et al., 1999; Zurek, 1987).

B&K, Brüel & kjær; ERP, event-related potential; HATS, head and torso simulator; IAC, Industrial Acoustic Company; RO, right loudspeaker was turned on only; L/U, left leading/uncorrelated; L/C, left leading/correlated; R/C, right leading/correlated; TDT, Tucker–Davis technologies

[*] Corresponding author. Tel.: +905 569 4628; fax: +905 569 4326/1 416 978 4811.

- liang@psych.utoronto.ca (L. Li).

0378-5955/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.heares.2004.10.007

The precedence effect reduces listeners' perception of multiple images by perceptually grouping correlated acoustic waveforms from different directions, thereby avoiding the perception of multiple sound images when only one source is present. Furthermore, because the fused image is perceived as originating at or near the location of the source, localization errors are reduced in reverberant environments. In experimental environments, the "direct" and "reflected" waves are usually produced by two spatially separated sound sources, and the shortest time delay between a direct and a reflected wave that produces a separate echo on certain percentage of experimental trials (usually between 50% and 80%) is called the echo threshold (Blauert, 1997, pp. 224–225).

Since a simulated reflection in an experimental environment is not heard as a separate auditory event when the lead/lag delay is below the echo threshold, it has been assumed that some inhibition or attenuation of information in reflected sounds, such as contralateral inhibition (Blauert, 1997, pp. 230–233), must take place in the precedence effect. For instance, a prevalent explanation is that the directional information associated with the reflected wave is suppressed (Blauert, 1997; Liebenthal and Pratt, 1999; Litovsky and Shinn-Cunningham, 2001; Rakerd et al., 2000; Yin, 1994; Zurek, 1980). This suppression hypothesis has dominated the search for neural correlates of the precedence effect. In most of the related physiological studies using either anesthetized or unanesthetized animals, suppressed neural responses to the lagging sound in the presence of the leading sound were treated as the neural correlates of the precedence effect (Fitzpatrick et al., 1995, 1999; Liebenthal and Pratt, 1999; Litovsky, 1998; Litovsky and Delgutte, 2002; Litovsky and Yin, 1998a,b; Litovsky et al., 1997; Yin, 1994).

However, suppression of the directional information in the reflection does not mean that the reflected wave is not heard because listeners are aware of the presence of reflections and even changes in them. For example, Freyman et al. (1998) have shown that listeners are as sensitive to intensity decreases in the lagging sound as to intensity increases in the leading sound, indicating that intensity information in the reflection is not suppressed. Also, hearing a reflection while presumably suppressing its directional information raises some puzzles as to how the perceptual system incorporates reflected waves into the percept of a single auditory event. For example, it is not clear how the intensities of a source and its reflections blend to determine the loudness of the "fused" sound image. Finally, Hartung and Trahiotis (2001) have developed a model for describing how monaural peripheral processing without an inhibitory mechanism may contribute to data obtained in binaural "precedence" experiments that use binaural pairs of transients as stimuli. Hence, it is evident that there is more to the precedence effect than simple inhibition.

Most studies on the precedence effect have used idealized brief acoustic stimuli, such as clicks or transient noise bursts, to avoid or reduce temporal overlap between the leading and lagging sounds (for a review see Litovsky et al., 1999). However, acoustic stimuli under normal circumstances are usually complex and last for more than a few hundred milliseconds. Therefore, it is important to study how the precedence effect works for long-duration stimuli, and determine how attributes that belong to reflections, and indeed may be unique to them, are incorporated into the fused image of the source.

In the present study, a transient gap, as a probe attribute, was inserted into an otherwise continuous steady-state broadband noise. Because this gap could be in the source (the leading sound) only, the reflection (the lagging sound) only, or both source and reflection, it should be easier to determine how this attribute of the direct wave and/or the reflection is detected and incorporated in the overall percept of the sound.

Introducing a single gap into either the leading or the lagging sound (but not both) is also interesting from the point of view of top-down control over the precedence effect. For example, a gap only in the lagging but not in leading stimulus is inconsistent with the lagging stimulus being an echo (a gap in a natural reflection should have its origin in the sound source), and could lead to a breakdown in the precedence effect. Moreover, if the gap is in the lagging stimulus only, and the leading and lagging stimuli remained fused into a single percept, will the listener perceive a break in the fused stimulus, or will the gap in the lagging stimulus be suppressed so that the listener hears a continuous fused stimulus? To investigate issues such as these, listeners were asked to describe their experience to the gap, which was introduced into the middle of either the leading or lagging sound.

As mentioned earlier, most neurophysiological studies on the precedence effect have mainly focused their efforts on determining the brainstem mechanisms involved in lag suppression in experimental animals (Fitzpatrick et al., 1995, 1999; Litovsky, 1998; Litovsky and Delgutte, 2002; Litovsky and Yin, 1998a,b; Litovsky et al., 1997; Yin, 1994). However, there is more to precedence than simple suppression of the location information of the lagging stimulus. For example, several studies have shown that listeners' knowledge and expectations about the room acoustics can strongly affect the precedence effect (Clifton, 1987; Clifton and Freyman, 1989; Clifton et al., 1994; Freyman et al., 1991). Repeated presentations of the leading and lagging clicks, which are not perceived to be fused at the beginning, can eventually cause fusion to occur, suggesting that following continued exposure to a reverberant environment, listeners can build up a new representation of the room acoustics consistent with the leading and lagging stimulus being produced by a single source. Moreover, once fusion is established, it is most readily broken when a change in the spatial relationship between the leading and lagging sounds is inconsistent with the knowledge of the room acoustics that has been acquired previously. Thus there is a strong higher-order cognitive component involved in the precedence effect. For this reason, human's cortical correlates of the precedence effect were investigated using the method of scalp event-related potential (ERP) recording. Since ERPs to a brief acoustic event can last a few hundred ms, in the present were introduced. The left loudspeaker led the right loudspeaker, and the time lag between them (called the lead/lag time) was reduced following responses indicating a perceived noise sound from the right loudspeaker, and increased following responses indicating that no noise sound was perceived from the right loudspeaker using a 3-down-1-up procedure (Levitt, 1971). All sessions were started with a 50 ms lead/lag time. Therefore, the longest lead/lag time, at which no sound image from the right loudspeaker was perceived (the "echo inaudible" criterion), was obtained. That an echo (not a reflection) is perceived or not is subjective, and listeners' responses cannot be categorized as either "correct" or "incorrect". Thus in this and the other two conditions of this experiment, no feedback was given to listeners.

In Condition Correlated/50-ms-gap, a 50-ms gap was introduced into the middle of each of the two correlated noises from the two loudspeakers. The delay between the onsets of the two gaps was equal to the delay between the leading and lagging sounds. The left loudspeaker was also the leading loudspeaker, and listeners, when presented with a stimulus, indicated by pressing one of two buttons whether they heard a gap in the sound coming from the right (lagging) source. Logically, of course, they could only hear a gap in the right-side noise if they heard a noise on the right. Hence, the question here is whether, when they heard a noise on the right (lead/lag delays > echo threshold), they also perceived a gap in the right-side noise, or whether the gap was only heard in the left-side (leading) noise. If they did not hear a gap in the noise coming from the right loudspeaker they were to press the other button. In other words, the lag time between the sounds from the two loudspeakers was reduced following responses indicating a perceived gap in the noise perceived on the right, and increased following responses indicating that they did not hear a gap on the right. The same 3-down-1-up procedure was employed (Levitt, 1971).

In Condition Uncorrelated/50-ms-gap, a 50-ms gap was introduced into the middle of each of the two uncorrelated noise sounds from the two loudspeakers, and the procedure was the same as that of Condition Correlated/50-ms-gap. There were four repetitions in each of the three conditions.

2.2.

In Condition Correlated/No-gap, when the lead/lag times were substantially longer than the individuals' echo thresholds, all 15 listeners perceived a distinct sound image originating from the right loudspeaker. Because the noise sound image originating from the left loudspeaker was always perceived, two spatially separate noise sounds were actually heard at the longer lead/lag times, one on the left and one on the right. When the lead/lag delays were substantially below the individuals' echo thresholds, only one noise sound image was heard as coming from the locus of the leading loudspeaker and no sound image as coming from the right loudspeaker was perceived. As shown in Fig. 2, the average echo threshold was approximately 9.5 ms.

When a gap was introduced into both the leading and lagging sounds in Condition Correlated/50-ms-gap, the average gap capture threshold was 15.6 ms (Fig. 2). The gap capture threshold in Condition Correlated/50-ms-gap was significantly longer than the echo threshold in the same condition ($_{1,14} = 5.769$, MSE = 47.617, $= 0.031$). At delays substantially longer than the gap capture threshold, listeners perceived a gap in the sound image associated with the right loudspeaker. At delays between the echo threshold, and the gap capture threshold, listeners perceived sounds from both the left (leading) and right (lagging) loudspeakers, but did not hear a gap in the lagging sound. Rather the gap was heard only in the leading sound. Finally, at delays shorter than the echo threshold listeners only heard a sound on the left with a gap in it. Hence, for intermediate delays (between 10 and 15 ms) in Condition Correlated/50-ms-gap, listeners heard two spatially separated continuous sound images (a direct wave and its echo), with a gap in the leading image, but not in the lagging image, even though both leading and lagging sounds contained a 50 ms gap.

In Condition Uncorrelated/50-ms-gap, listeners always perceived two spatially distinct sounds (one on the left and the other on the right), regardless of the

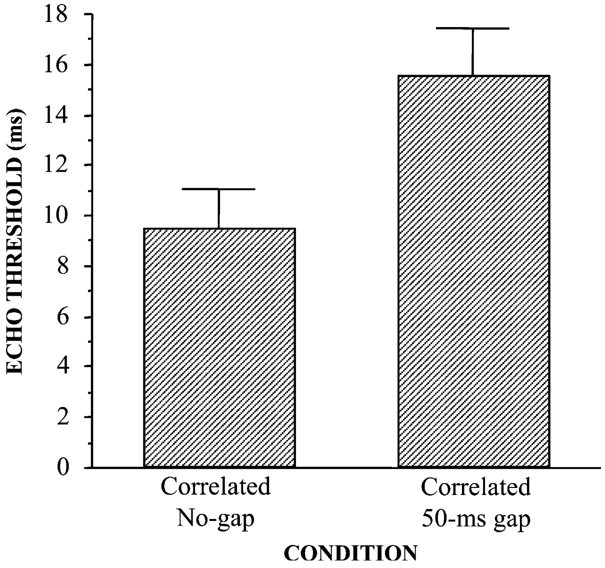

Fig. 2. Comparison of average attribute capture thresholds between the two conditions: (1) Condition Correlated/No-gap, the two noises from the two loudspeakers were correlated and no gaps were introduced. (2) Condition Correlated/50-ms-gap, a 50-ms gap was introduced into the middle of each of the two correlated noise sounds from the two loudspeakers. The error bars indicate the standard errors of the mean.

lead/lag time. Thirteen of the 15 listeners always heard gaps in both sounds at all delays, however, two of the listeners occasionally reported that they did not hear a gap in the lagging sound.

3. Experiment 2

In Experiment 1, when there were gaps in both leading and lagging correlated noises, and the lead/lag time was slightly longer than the echo threshold (10–15 ms), so that both the leading and lagging sounds were heard, listeners heard a gap in the leading but not in the lagging sound. A possible explanation of this phenomenon is that some attributes of the lagging sound (e.g., the presence of a gap) were being suppressed, even though the lagging sound was heard. If that were the case then it would be expected that attributes of the lagging sound would be even more suppressed when the lead/lag time was short enough that only a single fused sound was heard. To check whether attributes of the gap were suppressed when the lagging sound was clearly captured, in the second experiment gap detection thresholds (the shortest duration at which a gap was perceived), both when sounds were fused (echo capture) and when they were not, were determined.

To see whether a listener's sensitivity to a gap depended on whether or not fusion occurred, in Experiment 2 gap detection thresholds when fusion clearly happen (correlated noises, 2 ms delay) were compared to a condition when it did not (uncorrelated noises, 2 ms delay). If the gap appeared only on the lagging side and was suppressed when fusion occurred, then the gap detection threshold should be higher than when there was no fusion.

3.1.

3.1.1.

The fifteen people who participated in Experiment 1 also participated in this experiment.

3.1.2.

The apparatus and materials were same as in Experiment 1.

3.1.3.

Unlike Experiment 1, where there was a gap in the noises produced by both the left and right loudspeakers, in Experiment 2, the gap appeared only in the noise that was delivered from the right loudspeaker. The minimum size of the gap in the right-loudspeaker noise that could be detected using a single-interval staircase procedure was then determined, for both correlated and independent leading and lagging noises. Specifically, on each trial a stimulus with a gap in the sound emanating from the right loudspeaker was presented. If the listener responded that she/he heard the gap on three consecutive trials, the duration of the gap on the next trial was reduced. If, however, the listener indicated on a trial that they could not hear a gap, the duration of the gap on the next trial was increased, a 3-down (gap duration reduced)-1-up (gap duration increased) procedure (Levitt, 1971). [1]

In Condition RO, the right loudspeaker was turned on and the left loudspeaker was turned off. In Condition L/U, a right-side noise sound (with a gap) lagged 2 ms behind an uncorrelated left-side noise sound without a gap. [2] In Condition L/C, a right-side noise sound (with a gap) lagged 2 ms behind a correlated left-side noise sound without a gap. In Condition R/C, a right-side noise sound (with a gap)led, by 2 ms, a correlated left-side noise sound without a gap. There were four repetitions in each of the conditions. The maximum gap at the beginning of a session was 50 ms.

3.2.

As indicated in Fig. 3, the gap detection thresholds among Conditions L/U, L/C and R/C were similar, and the lowest gap detection threshold was obtained when only the right loudspeaker was operative (Condition RO). A one-way analysis of variance with repeated measures revealed that the differences in gap detection thresholds between these four conditions were significant ($_{3,42} = 5.146$, MSE = 6.030, $= 0.004$). Pairwise analyses indicated that Condition RO was significantly different from each of the other three conditions (< 0.005) but there were no significant differences among Conditions L/U, L/C and R/C (> 0.800). Hence

---

[1] We opted to use a single-interval staircase procedure rather than the more standard two-interval, forced-choice procedure for two reasons. First, the use of a two-interval technique would have more than doubled trial length from its current 3.05 s to more than 7 s (once an inter-stimulus interval was added), and we were concerned about tiring our volunteers. Second, we wanted to keep the testing situation as comparable as possible to that used in Experiment 1 (where we also used a single-interval staircase procedure) since we were using naïve listeners. Although thresholds determined using single-interval staircase procedures are subject to response biases, such biases are not a significant problem for comparisons of thresholds as long as these biases remain constant across comparisons. Because there is no reason to expect that a change from left leading to right leading, or from correlated to independent noises, or from the left loudspeaker "on" to the left loudspeaker "off" would affect the bias to report a gap, gap detection threshold differences among these conditions should accurately reflect relative (but perhaps not absolute) sensitivity to the presence of a gap.

[2] Because the two uncorrelated sounds did not fuse, it should not matter whether right or left was leading for detecting the gap in the middle of the right sound, especially when the gap occurred 1500 ms after sound onset. Thus for the gap detection test, Condition R/U should be equivalent to Condition L/U and was not included in the experimental protocol.

there was no indication that changing the left noise from lagging to leading affected the detection of a gap. Indeed, the gap detection threshold remained unchanged even when the two sounds were uncorrelated. These results are consistent with the notion that the detection of a gap in a stimulus depends only on the extent of the drop in acoustic energy present in the stimulus at the ears, since the degree of interaural correlation and the direction of the lag apparently had no effect on threshold.

To determine the nature of the local cues in the left and right ear that could signal the presence of a gap, a B&K head and torso simulator (HATS, 4128C) was placed at the position that would be occupied by the listeners' head. The signal at the location of the eardrum in the simulated head was then recorded for both left and right ears under two conditions using the B&K Pulse Platform. In the first condition, correlated noises were presented over both loudspeakers with the left loudspeaker leading the right loudspeaker by 2 ms. The lines with filled circles in Fig. 4 depict the long-term spectra of the left- (left panel) and right- (right panel) ear signals when both loudspeakers were playing. The lines with open squares depict the long-term spectra of the left (left panel) and right (right panel) ear signals when only the left loudspeaker was on (i.e., the condition that existed when there was a gap in the right loudspeaker). The differences between the two spectra in the left panel identify the left-ear spectral cues to the presence of a gap. The comparable differences in the right panel identify the right-ear spectral cues to the presence of a gap. Clearly, spectral differences in the right ear are much more pronounced than they are in the left ear, especially at the high frequencies (due to the head shadow effect). Because gap detections thresholds did not vary with the degree of interaural correlation, and because the spectral cues are much more pronounced at the right ear, it is reasonable to conclude that the detection of a gap was based on the processing of intensity information in the right ear. Hence, on the basis of spectral cues, it would

3 4 5

FREQUENCY IN kHz

be expected that the listener might hear the gap as occurring on the right. However, for gap durations just above threshold, listeners reported (after the session) that they perceived a gap in the sound source located to the left. In other words, the leading sound appeared to fully capture an attribute in the lagging sound. This capture effect was explored more systematically in the next experiment.

4. Experiment 3

Experiment 3 investigated how the precedence effect modified listeners' perceptions of a gap that appeared either in the lagging or leading sounds, but not both. Specifically, listeners were asked to report their impressions associated with gaps in Conditions L/U, L/C, and R/C (see Experiment 2 for the definitions of the three conditions).

.1.

.1.1.

Eleven listeners (four females and seven males) with normal and balanced pure-tone hearing participated in this experiment. Four young male listeners also participated in Experiments 1 and 2. The other 7 listeners included 4 young female listeners (19–31 years old), and 3 male listeners (34, 34 and 39 years old, respectively). The gap detection threshold for each of these 7 listeners, who did not participated in Experiments 1 and 2, was also measured under Condition L/U.

.1.2.

The apparatus and materials were the same as in Experiments 1 and 2.

.1.3.

Stimuli were presented in each of the three conditions (L/U, L/C, and R/C) at the following three different gap sizes: (1) 2 ms above each individual's gap-detection threshold (as determined in Experiment 2), (2) 20 ms, and (3) 50 ms. Thus there were 9 (3×3) condition/gap-size combinations. These combinations were presented in a random order for each listener. The lead/lag time was fixed at 2 ms, which was well below the echo threshold for each of the listeners.

After 5 stimulus presentations in each of the 9 condition/gap-size combinations, the listeners were asked to report their impressions about the gap that occurred in the middle of the noise by selecting an answer from the following 6 options: (1) a single gap, (2) a sudden burst of noise, (3) both a single gap and a noise burst, (4) two gaps, (5) two noise bursts, or (6) no change. They were then asked to report which loudspeaker(s) delivered the perceived gap(s) and/or which loudspeaker(s) delivered the perceived noise burst(s) (for the instructions to listeners, see Footnote [3]). Thus Options 1 and 2 were associated with perception of only one brief auditory event in the middle of the noise sound, and Options 3, 4, and 5 were associated with perception of 2 brief auditory events. Option 6 indicated that the participant did not perceive any event in the middle of the noise.

Noise burst options were incorporated into the response list because there were reasons to expect that listeners would hear a noise burst if there was any tendency for echo capture to break down during a gap. For example, if a gap were introduced into the leading stimulus only, there would be no leading stimulus during the gap to suppress the information as to the location of the lagging stimulus. Hence, one might expect to hear a brief noise burst from the location of the lagging stimulus.

.2.

All the 11 listeners reported that they perceived one or two sudden changes in the middle of the sound in all combined conditions. No participant used the "no change" response. However, one male participant appeared not to follow the instructions appropriately. [4] Thus this participant's data were not used. The results from the other 10 listeners appear in Fig. 5

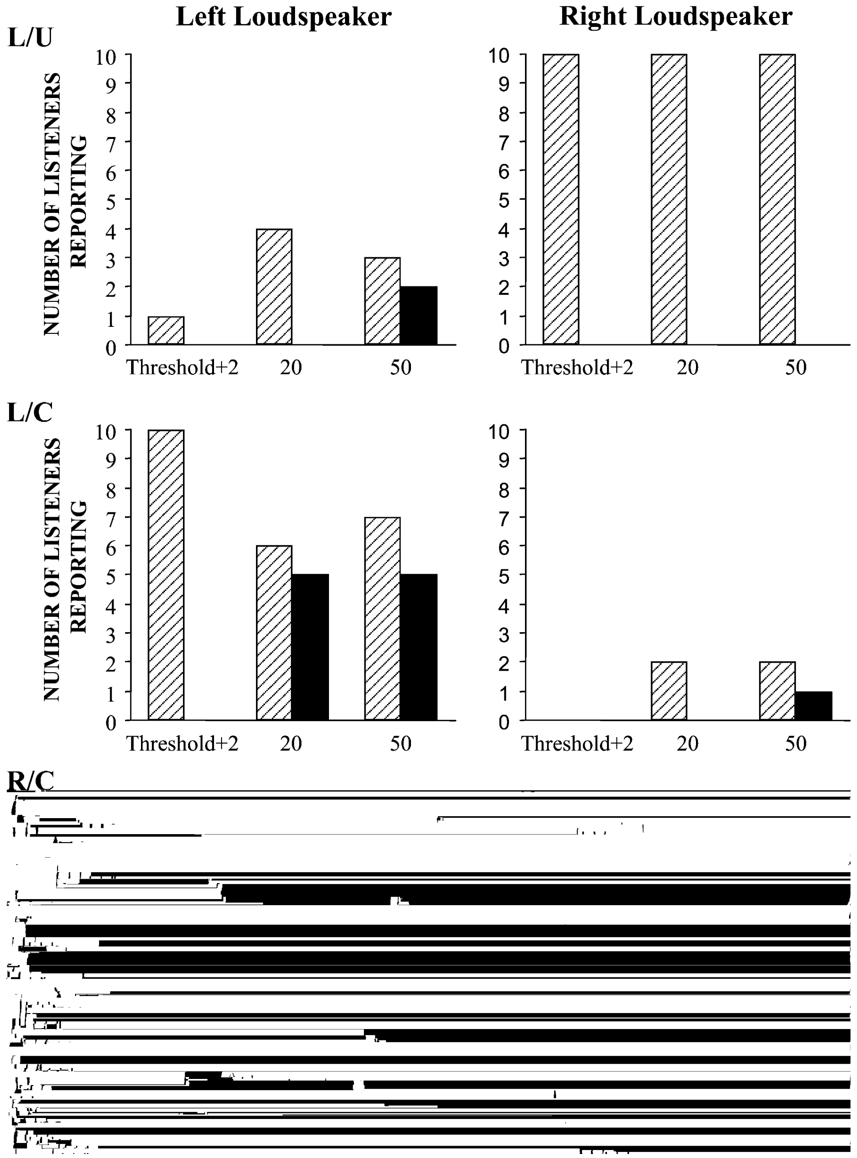

Fig. 5. Summary of listeners' perceptions of the gap in Conditions L/U (left leading/uncorrelated), L/C (left leading/correlated), and R/C (right leading/correlated). The gap was only in the sound from the right loudspeaker. The ordinates represent the numbers of listeners, who attributed a "gap" or a "noise burst" to a particular (left or right) loudspeaker at each of the three different gap sizes. Lighter bars indicate "Gap" responses, and darker bars indicate "Noise burst" responses.

durations, a gap was always perceived in the lagging sound.

In Condition L/C, the listeners predominately perceived a change in the sound coming from the left (leading) loudspeaker, even though the gap appeared only in the right (lagging) loudspeaker. When the gap size was near threshold, all the listeners reported that they perceived only a single gap image in the sound from the left loudspeaker. When the gap size was 20 or 50 ms, most listeners perceived either a gap or a noise-burst image as coming from the left loudspeaker. Only a small number of listeners reported that they perceived a gap or a burst image as coming from the right loudspeaker. Hence, when the gap is in the lagging sound and the sounds are correlated, listeners tend to incorporate any perceptual change occasioned by the gap into the fused image, which is perceived to be located on the leading side. In other words, perceptual changes evoked by a gap in the lagging sound are captured by the leading sound. It is interesting to note, that at the longer gap durations, listeners sometimes heard a noise burst, which they attributed (with one exception) to the leading stimulus. One possible explanation for this perception is that if the gap in the lagging stimulus is long enough, there is no location information coming from the lagging stimulus to suppress, and the circuitry responsible for the suppression of location information is disengaged. Consequently, when the gap is terminated, the perceptual system briefly treats the return of the lagging correlated stimulus as a new stimulus until it re-establishes the correlation between the leading and lagging stimulus and suppresses the perception of the lagging source. It is interesting to note, however, that this noise burst, rather than being attributed to the lagging stimulus is perceived as originating from the direction of the leading stimulus. In other words, it appears to be captured by the leading stimulus.

In Condition R/C, all the listeners perceived the gap as belonging to the right (leading) loudspeaker in the near-threshold condition. At the larger gap durations (20 and 50 ms), the listeners predominately perceived the gap (when it was heard as a gap) as belonging to the right (leading) loudspeaker, but they also reported hearing a noise-burst image as coming from the location of the left (lagging) loudspeaker. When there is a gap in the leading stimulus, there is no leading sound present to suppress the information as to the location of the lagging stimulus. Hence, one might expect to hear a brief noise burst during the gap from the location of the lagging stimulus until the perception of the lagging stimulus is suppressed. This is what appears to have happened here.

5. Experiment 4

To examine how the precedence effect modulates cortical responses to the probe gap, in Experiment 4, N1, P2, and long-latency sustained components of ERP responses to gaps were measured in Conditions L/U, L/C, and R/C, respectively.

.1.

.1.1.

All the 11 listeners from Experiment 3 and 1 new male young university student (21 years old) with normal and balanced pure-tone hearing participated in this physiological experiment. These listeners were instructed to remain awake and keep their eyes open, while they listened to the acoustic stimuli.

.1.2.

The apparatus and materials were same as in previous experiments. However, this ERP recording experiment was conducted in a different IAC sound-attenuated chamber that was equipped with 64-channel NeuroScan SynAmps (bandpass 0.05–50 Hz; 250 Hz sampling rate).

.1.3.

The size of the gap in the sound from the right loudspeaker was fixed at 50 ms and the delay between the sounds from the two loudspeakers was fixed at 2 ms. During the recording, all electrodes were referenced to the Cz site; for data analysis, they were re-referenced to an average reference. The analysis epoch included 200 ms of pre-stimulus activity and 3500 ms of post-stimulus activity following each of the 150 sound presentations for each of the three conditions: Conditions L/U, L/C, and R/C. Trials contaminated by excessive peak-to-peak deflection (±150 μV) at the electrodes not adjacent to the eyes were automatically rejected. ERP waveforms were then averaged separately for each site and conditions, and digitally low-pass filtered to attenuate the components with frequencies above 12 Hz. Although the number of stimulus-presentation trials was 150, the number of trials included in the average for each condition varied between listeners with the across-listener average being 116, 114, and 113 for Condition L/U, Condition L/C, and Condition R/C, respectively. For each individual average, ocular artifacts (e.g., blinks and lateral movements) were corrected by means of ocular source components using the Brain Electrical Source Analysis (BESA) software (Picton et al., 2000). ERP waveforms were quantified by computing mean values in selected latency regions, relative to the mean amplitude of the 200 ms pre-stimulus activity. All amplitude measurements were subjected to mixed ANOVA with condition and electrode as the two within-subject factors. Topographic voltage maps were examined using the 61 electrodes (the periocular electrodes were not included).

.2.

For the 9 central electrode sites (FC1, FCz, FC2, C1, Cz, C2, CP1, CPz, and CP2), there were no differences across these three conditions both for N1-P2 peak-to-peak amplitudes to sound onset ($_{2,22}$ = 0.238, MSE = 7.948, = 0.790) and for slow sustained potentials following sound onset ($_{2,22}$ = 1.308, MSE = 1.537, = 0.290) (Fig. 6). However, the N1-P2 responses to the gap did differ significantly across these three conditions ($_{2,22}$ = 9.129, MSE = 3.586, = 0.001) (Fig. 6). Pairwise comparisons indicate that the amplitude of N1-P2 response to the gap in Condition L/U was significantly smaller than that in Condition L/C (= 0.022) and that in Condition R/C (= 0.000), but the difference between Condition L/C and Condition R/C was not significant (= 0.125). Topographic voltage maps for the N1 component to the gap (Fig. 7) indicate that in Condition L/U, the highest negativity was widely distributed over the midline, but in both Condition L/C and Condition R/C, it became more concentrated over the right hemisphere. Hence, when a gap is introduced, the cortical response depends upon whether or not the two sounds were correlated or uncorrelated.

Moreover, there appeared to be ERP differences in the sustained responses following the gap (Fig. 6) between conditions. The average amplitude of the sustained responses 550–850 ms after the gap onset was analyzed. The results show that there were no significant differences in sustained responses for the two conditions (L/U and R/C) where the gap was correctly assigned to the right loudspeaker. However, the condition, in which the gap in the right sound was perceptually captured by the left sound (Condition L/C), differed significantly both from Condition L/U across all the 9 central sites (= 0.001) and from Condition R/C across the 3 fronto-central sites (FC1, FCz, FC2) (= 0.016). Hence, a long-latency and negatively shifted sustained response in the frontal cortical region following the gap appears to be associated with gap capture.

6. Discussion

Most previous studies of the precedence effect have used clicks or short noise bursts as acoustic stimuli to avoid or reduce the overlap between the leading and lagging stimuli. Here long-lasting sound segments were chosen for 3 major reasons: First, long-duration sound segments (e.g., speech or music) are more prevalent in everyday environments, therefore have greater ecological validity than idealized brief sounds for humans. Secwere uncorrelated and the delayed was below 8 ms, there were also a small proportion of trials on which fusion of the two bursts was perceived. In our experiments the two uncorrelated sounds did not fuse. On a few occasions, however, a gap that appeared only in the right (lagging) uncorrelated sound was also attributed to the leading sound, indicating that attributes of the lagging sound may occasionally be captured by the leading sound even when the two sounds are uncorrelated. Hence, although listeners never reported that the two independent sounds became fused, there is some indication of attribute capture by the leading sound. The disagreement concerning fusion between our data and those reported by Perrott et al. (1987) for uncorrelated noises may be due to the differences of stimulus parameters between the two studies, such as those in sound duration (50 ms vs. 3050 ms), onset/offset duration (0.2 ms vs. 30 ms), and loudspeaker separation (±20° vs. ±45°), etc.

In a reverberant environment, each sound reflection comes from a location that is usually different from that of the sound source, and not all attributes of reflections are suppressed by their sound sources (Clifton et al., 2002; Freyman et al., 1998; Perrott et al., 1987; Tollin and Henning, 1999). In the present study, if the gap attribute in the lagging sound had been suppressed by the correlated leading sound when the precedence effect occurred, the gap detection threshold in Condition L/C should have been higher than those in Condition L/U and Condition R/C, and the gap detection threshold in Condition R/C should have been lower than that in Condition L/U. However, our data show that gap detection thresholds were independent of whether the gap was in either the leading or lagging sound, and also independent of whether or not the leading and lagging sounds were correlated. These results are consistent with the hypothesis that gap detection depends primarily on the detection of an energy change in the ear on the side of the loudspeaker producing the gap. On the other hand, when the two sounds were correlated, a single compact sound image was perceived as coming from the leading side; when the two sounds were not correlated, more diffused sound images were perceived as coming from the both sides. Since there was no difference in gap detection between Conditions L/U, L/C, and R/C, there is no evidence in this experiment that sound-image compactness/diffuseness affects gap detection.

If information in these reflections is not being suppressed, then it has to be somehow perceptually incorporated into the fused image. The present study shows that when the two sounds are uncorrelated, the lagging sound is by and large not treated as the reflection of the leading sound by the auditory system, and two distinct noise images, coming from different directions are perceived, and the gap presented in the lagging sound is "correctly" perceived as coming from the lagging loudspeaker. The only exception to this statement is that sometimes, especially at the longer gap durations, the gap is also attributed to (captured) by the leading stimulus. In contrast, when the two sounds are correlated, the lagging sound is treated as a reflection of the leading sound, a single noise image is perceived, and attributes that appear only in the lagging sound are attributed to (captured by) the leading sound. This is not what we would expect on the basis of the physical cues to the location of the gap that are present when there is a gap only in the lagging sound. Fig. 4 shows that when there is a gap in the lagging (right side) source only, there is a corresponding drop in energy (especially in the high-frequency region) in the right ear, with little evidence of any change in the left ear. Hence, if the location of the gap were to be based on the ear with the most salient cues, one would expect the gap to be heard on the side of the lagging sound. Nevertheless, the gap is heard as occurring on the leading side. In other words, it is attributed to (captured by) the leading stimulus.

When the gap is only in the correlated lagging sound, the acoustic situation is ecologically anomalous, because the gap in the reflection should have its origin in the source. The ecological prediction is that a gap in the lagging sound would cause a temporary breakdown in the precedence effect. When the lagging loudspeaker becomes silent during the gap, there is no correlated signal coming from the lagging loudspeaker to be captured. Thus, when the gap terminates and the lagging loudspeaker is turned on again, the participant should initially perceive a new sound originating from the location of the lagging loudspeaker until the precedence of the leading sound is re-established. However, most of our listeners did not hear any sound change as coming from the location of the lagging loudspeaker. Rather they heard a gap or a burst-like image as coming from the leading loudspeaker. Since there is no physical gap in the sound from the leading loudspeaker, the gap in the sound from the lagging loudspeaker has no leading "partner" to "fuse" with. Moreover, hearing a gap or a burst-like image as coming from the leading loudspeaker cannot be caused by a peripheral effect, since there are no obvious differences in the sound spectra at the left ear (the ear on the side of the leading loudspeaker) between the condition when there is no gap in the lagging (right-side) stimulus versus when there is a gap in the lagging stimulus (see Fig. 4). Thus the shift of gap image from the lagging loudspeaker to the leading loudspeaker denotes the maintenance of the precedence effect during the period of the gap, and must involve a higher-order attribute capturing process.

On the other hand, when the gap is only in the leading sound that is correlated with the lagging sound, the acoustic situation is also ecologically anomalous, because a gap in a natural sound source will also appear in its reflections. Hearing a gap as coming from the leading loudspeaker and simultaneously a burst-like image as coming from the lagging loudspeaker indicates a transient disappearance of the precedence effect during the gap.

Our electrophysiological results suggest a tight link between subjective perception of the gap and neural responses to the gap. Surprisingly, there is no difference in ERP responses between the correlated and uncorrelated sound conditions until a gap occurs, even though the perceptual responses to the correlated and uncorrelated noise sounds are quite different. When the two long-duration sounds are correlated, the N1-P2 peak-to-peak response to the gap is enhanced and the N1-topographic-voltage map for the gap shifts laterally towards the right hemisphere, regardless of the gap being in the lagging or leading sound. Also, in the frontocentral region, a negatively shifted sustained ERP response following the gap embedded only in the lagging sound appears to be associated with the perceived capture of the gap. The present neurophysiological results suggest that there is a greater need for cortical involvement to maintain fusion of leading and lagging sounds when there is a break in one or the other, than to establish fusion at sound onset. This long-latency neural event following the occurrence of the gap also suggests that higher-order central processes are involved in attribute capture.

Clinical studies in humans suggest that both the cortex and the inferior colliculus are essential for the precedence effect. Cornelisse and Kelly (1987) reported that patients with lesions of the right temporo-parietal cortex were able to localize single clicks but could not localize the "fused" image of two spatially separated clicks, when the leading click was delivered from the left hemifield and the lagging click was delivered from the right hemifield. Litovsky et al. (2002) reported that a patient with lesions of the right inferior colliculus had substantially weaker echo suppression when the leading sound was delivered in the left hemifield. Hence it would be interesting to investigate attribute capture in patients with unilateral lesions of the central auditory system.

In summary, based on the data of the present study, three important features of attribute capture should be noted:

(1) Top-down higher-order processes are involved in attribute capture. A probe gap introduced in the leading stimulus can temporarily break the precedence effect whereas introducing a comparable gap in the lagging stimulus does not break the precedence effect in the majority of our listeners, even though both situations are ecologically anomalous. In addition, gap capture is associated with long-latency negatively-shifted slow potentials in the frontal area.

(2) Attribute capture is not an all-or-none process. For lead/lag delays between 9 and 15 ms, the location information concerning the lagging sound is not suppressed by the leading sound (a sound is still heard as coming from the direction of the lagging sound), but a gap in the lagging sound is, nevertheless, captured by the leading sound (a gap is heard in the leading sound but not in the lagging sound). This indicates that capture thresholds can differ for different attributes of the reflection (e.g., gaps in the lagging sound are more easily captured than other aspects of the sound). One may speculate that the degree to which the listener assigns spatially separate and distinct images to the leading and lagging sounds will depend on the extent to which different attributes of the lagging sound are incorporated into (captured) by the leading sound. According to this speculation, all of the attributes of the reflection would have to be captured in order for the listener to perceive only a single source.

(3) The introduction of a distinct feature such as a gap into a direct or reflected wave may be one way of probing cortical involvement in the precedence effect. In our study, identical ERP responses were elicited by both correlated and uncorrelated noises, even though listeners perceive correlated noises to be quite distinct from uncorrelated noises. One may speculate that the differences between the two are processed primarily by brain-stem mechanisms. However, the ERP to a gap differed substantially depending upon whether or not the noises were correlated. This suggests that while cortical involvement may not be necessary to distinguish between correlated and uncorrelated noises, it may be required to maintain and/or re-establish the perception of these two kinds of noise (especially, with respect to percepts related to precedence) once there is a break in either the leading or lagging noise. The use of gaps as probes may be a way of accessing the cortical mechanisms involved in the maintenance of percepts when there are sudden or unexpected changes in the sensory input. Thus, in order to more completely understand the neural mechanisms involved in the precedence effect, cortical neural correlates should be investigated in addition to the brainstem mechanisms.

Acknowledgements

We thank Jane W. Carey and Neda Chelehmalzadeh for their assistance during data acquisition. We also thank the following people who have reviewed previous versions of the manuscript and provided helpful comments and suggestions to improve its quality: Ann Clock

Eddins, William M. Hartmann, Jack B. Kelly and three anonymous reviewers. This work was supported by the Natural Sciences and Engineering Research Council of Canada, the Canadian Institutes of Health Research, the Canada Foundation for Innovation, and the Ontario Innovation Trust Fund.

References

Blauert, J., 1997. Spatial Hearing. MIT Press, Cambridge, MA.

Clifton, R.K., 1987. Breakdown of echo suppression in the precedence effect. J. Acoust. Soc. Am. 82, 1834–1835.

Clifton, R.K., Freyman, R.L., 1989. Effect of click rate and delay on breakdown of the precedence effect. Percep. Psychoph. 46, 139–145.

Clifton, R.K., Freyman, R.L., Meo, J., 2002. What the precedence effect tells us about room acoustics. Percep. Psychoph. 64, 180–188.

Clifton, R.K., Freyman, R.L., Litovsky, R.Y., McCall, D., 1994. Listeners' expectations about echoes can raise or lower echo threshold. J. Acoust. Soc. Am. 95, 1525–1533.

Cornelisse, L.E., Kelly, J.B., 1987. Neuropsychologia 25, 449–452.

Fitzpatrick, D.C., Kuwada, S., Batra, R., Trahiotis, C., 1995. Neural responses to simple, simulated echoes in the auditory brainstem of the unanesthetized rabbit. J. Neurophysiol. 74, 2469–2486.

Fitzpatrick, D.C., Kuwada, S., Kim, D.O., Parham, R., Batra, R., 1999. Responses of neurons to click-pairs as simulated echoes: auditory nerve to auditory cortex. J. Acoust. Soc. Am. 106, 3460–3472.

Freyman, R.L., Clifton, R.K., Litovsky, R.Y., 1991. Dynamic processes in the precedence effect. J. Acoust. Soc. Am. 90, 874–884.

Freyman, R.L., McCall, D.M., Clifton, R.K., 1998. Intensity discrimination for precedence effect stimuli. J. Acoust. Soc. Am. 103, 2031–2041.

Freyman, R.L., Helfer, K.S., McCall, D.D., Clifton, R.K., 1999. The role of perceived spatial separation in the unmasking of speech. J. Acoust. Soc. Am. 106, 3578–3588.

Hartung, K., Trahiotis, C., 2001. Peripheral auditory processing and investigations of the precedence effect which utilize successive transient stimuli. J. Acoust. Soc. Am. 110, 1505–1513.

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

Li, L., Yue, Q., 2002. Auditory gating processes and binaural inhibition in the inferior colliculus. Hear. Res. 168, 113–124.

Liebenthal, E., Pratt, H., 1999. Human auditory cortex electrophysiological correlates of the precedence effect: binaural echo lateralization suppression. J. Acoust. Soc. Am. 106, 291–303.

Litovsky, R.Y., 1998. Physiological studies on the precedence effect in the inferior colliculus of the kitten. J. Acoust. Soc. Am. 103, 3139–3152.

Litovsky, R.Y., Delgutte, B., 2002. Neural correlated of the precedence effect in the inferior colliculus: effect of localization cues. J. Neurophysiol. 87, 976–994.

Litovsky, R.Y., Shinn-Cunningham, B.G., 2001. Investigation of the relationship among three common measures of precedence: fusion, localization, and discrimination suppression. J. Acoust. Soc. Am. 109, 346–357.

Litovsky, R.Y., Yin, T.C.T., 1998a. Physiological studies of the precedence effect in the inferior colliculus of the cat: I. Correlates of psychophysics. J. Neurophysiol. 80, 1285–1301.

Litovsky, R.Y., Yin, T.C.T., 1998b. Physiological studies of the precedence effect in the inferior colliculus of the cat: II. Neural mechanisms. J. Neurophysiol. 80, 1302–1316.

Litovsky, R.Y., Rakerd, B., Yin, T.C.T., Hartmann, W.M., 1997. Psychophysical and physiological evidence for a precedence effect in the median sagittal plane. J. Neurophysiol. 77, 2223–2226.

Litovsky, R.Y., Colburn, H.S., Yost, W.A., Guzman, S.J., 1999. The precedence effect. J. Acoust. Soc. Am. 106, 1633–1654.

Litovsky, R.Y., Fligor, B.J., Traino, M.J., 2002. Functional role of the human inferior colliculus in binaural hearing. Hear. Res. 165, 177–188.

Perrott, D.R., Strybel, T.Z., Manligas, C.L., 1987. Conditions under which the Haas precedence effect may or may not occur. J. Audit. Res. 27, 59–72.

Picton, T.W., van Roon, P., Armilio, M.L., Berg, P., Ille, N., Scherg, M., 2000. The correction of ocular artifacts: a topographic perspective. Clin. Neurophysiol. 111, 53–65.

Rakerd, B., Hartmann, W.M., Hsu, J., 2000. Echo suppression in the horizontal and median sagittal planes. J. Acoust. Soc. Am. 107, 1061–1064.

Shinn-Cunningham, B.G., Zurek, P.M., Durlach, N.I., 1993. Adjustment and discrimination measurements of the precedence effect. J. Acoust. Soc. Am. 93, 2923–2932.

Tollin, D.J., Henning, G.B., 1999. Some aspects of the lateralization of echoed sound in man. II. The role of the stimulus spectrum. J. Acoust. Soc. Am. 105, 838–849.

Wallach, H., Newman, E.B., Rosenzweig, M.R., 1949. The precedence effect in sound localization. J. Acoust. Soc. Am. 62, 315–336.

Yin, T.C.T., 1994. Physiological correlates of the precedence effect and summing localization in the inferior colliculus of the cat. J. Neurosci. 14, 5170–5186.

Zurek, P.M., 1980. The precedence effect and its possible role in the avoidance of interaural ambiguities. J. Acoust. Soc. Am. 67, 952–964.

Zurek, P.M., 1987. The precedence effect. In: Yost, W.A., Gourevitch, G. (Eds.), Directional Hearing. Springer-Verlag, New York, pp. 85–105.