Perception & Psychophysics
2007, 69 (5), 641-653

Subjectively divided tone components in the gap transfer illusion

Kyoko Kanafuka
Nippon Telegraph and Telephone West Corporation, Osaka, Japan

Yoshitaka Nakajima and Gerard B. Remijn
Kyushu University, Fukuoka, Japan

Takayuki Sasaki
Miyagi Gakuin Women's University, Sendai, Japan

and

Shunsuke Tanaka
Matsushita Electric Industrial Co., Osaka, Japan

When a long glide with a short temporal gap in its middle crosses with a continuous short glide at the temporal midpoint of both glides, the gap is perceived in the short glide instead of in the long glide. In the present article, we tested possible explanations for this "gap transfer illusion" by obtaining points of subjective equality of the pitches and durations of the two short tones that are subjectively divided by the gap. The results of two experiments showed that neither an explanation in terms of envelope patterns nor explanations in terms of combination tones or acoustic beats could account for the perception of the short tones in the gap transfer illusion. Rather, the results were compatible with the idea that the illusory tones were formed by the perceptual integration of onsets and offsets of acoustically different sounds. Implications for the perceptual construction of auditory events are discussed.

In order to derive information about the environment, the auditory system must be able to process a mixture of sounds. The auditory system has to identify each individual sound of a mixture by segregating it into several auditory streams that correspond to sound sources in our surroundings (Bregman, 1990; Deutsch, 1983; McAdams & Bregman, 1979). The formation of auditory streams is governed by a number of principles (Bregman, 1990; Darwin & Carlyon, 1995) that describe how several auditory events—or sometimes a single event—can constitute an auditory stream (Handel, 1989). Nakajima and Sasaki (1996) proposed that the formation of each auditory event in a stream can be described through the perceptual connection of even smaller perceptual units—auditory subevents—such as sound onsets and offsets. They developed their ideas by using an auditory illusion called the *gap transfer* illusion (Nakajima et al., 2000). They found that when a long glide with a temporal gap in the middle crossed with a shorter, continuous glide at the temporal midpoint, the gap was perceived in the shorter pitch trajectory (Figure 1). Although physically present in the long glide, the gap was thus perceptually allocated to the shorter glide, thereby dividing it into two successive short tones.

Nakajima et al. (2000) explained the gap transfer illusion as follows (see also Remijn & Nakajima, 2005). They assumed that the auditory system deals with onsets and offsets of sounds ("sound edges") as independent entities that can be utilized to construct auditory events (see also Kubovy & Van Valkenburg [2001] on the importance of sound edges in defining "auditory objects"). The idea is that an onset and an offset that are close to each other in time and frequency can be perceptually integrated, even if they are parts of physically different sounds. In the gap transfer illusion, proximity between the onset of the short glide and the offset of the first long glide before the gap causes their perceptual integration. This results in the perception of the first short tone. In a similar vein, the onset of the second long glide after the gap is perceptually coupled with the offset of the short glide, constituting the second short tone (Figure 2). With the offset of the first long glide component and the onset of the second long glide component allocated to the two short tones, the long glide is no longer perceived as delimiting a gap and can hence be perceived as continuous.

In the present study, we tested the validity of this edge integration explanation against three alternative explanations for the gap transfer illusion. These three alternative explanations, respectively, state that the two subjectively divided tones in the gap transfer illusion could also be the result of the perception of combination tones or acoustic

Y. Nakajima, nakajima@design.kyushu-u.ac.jp

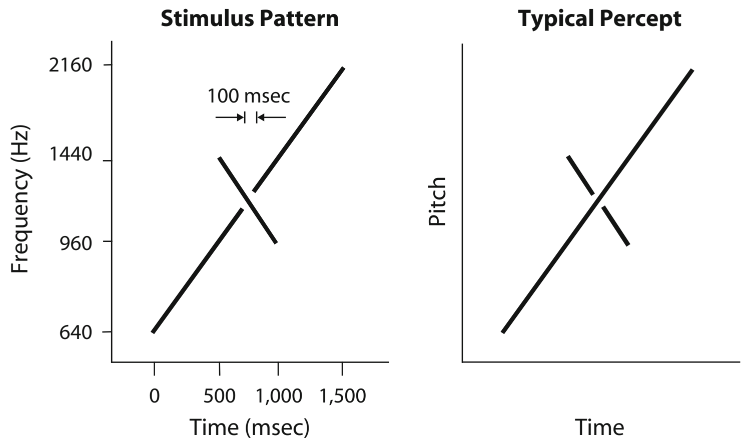

Figure 1. The gap transfer illusion. When listening to the stimulus pattern (left), the gap is perceived in the short pitch trajectory (right), although it is physically present in the long glide.

beats, as well as the result of processes related to auditory induction (Warren, 1999). These alternatives all concern rather low-level phenomena and they needed consideration in light of the results of an informal experiment. The informal experiment showed that the gap transfer illusion occurred robustly when both the long and the short glide components were presented to the same ear. When these glides were presented to different ears, however, 9 out of 10 observers perceived the gap transfer stimulus almost veridically, in that the short glide was judged as more continuous than the long glide. We thus cannot discard the possibility that the gap transfer illusion occurs before information from both ears is combined and that one or more of the three low-level phenomena contributes to the illusion.

In order to test the explanations, we obtained points of subjective equality (PSEs) of the duration and pitch of the two short tones that can be perceived when the gap is transferred from the temporal middle of the long glide to that of the short glide. The edge integration explanation would predict that the duration PSEs are similar to the durations of the parts where the short and the long glide components overlap. The perceived pitches of the short tones could correspond to the frequencies of the relevant sound edges, because it is known that onsets and offsets can carry pitch information (see, e.g., Bregman, Ahad, & Kim, 1994; Brugge, 1992). We will look into the details of the three alternative, low-level explanations for the gap transfer illusion and discuss what kind of duration and pitch PSEs we could expect to obtain if the explanations were valid.

Next to the edge integration explanation, the first alternative explanation for the gap transfer illusion is related to the perception of combination tones (Figure 3, left). Because of nonlinear distortion in the auditory system, combination tones can become audible during the pre-

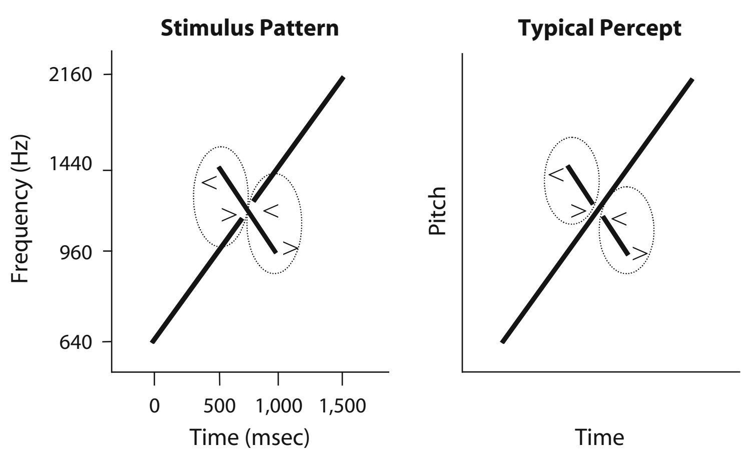

Figure 2. The “edge integration” explanation with regard to the gap transfer illusion. Because of their proximity in frequency and in time, the onsets (<) and offsets (>) of the ascending and descending glide components are perceptually coupled. The gap perceptually transfers, and two short tones are perceived in the temporal middle of the pattern, separated by a temporal gap.

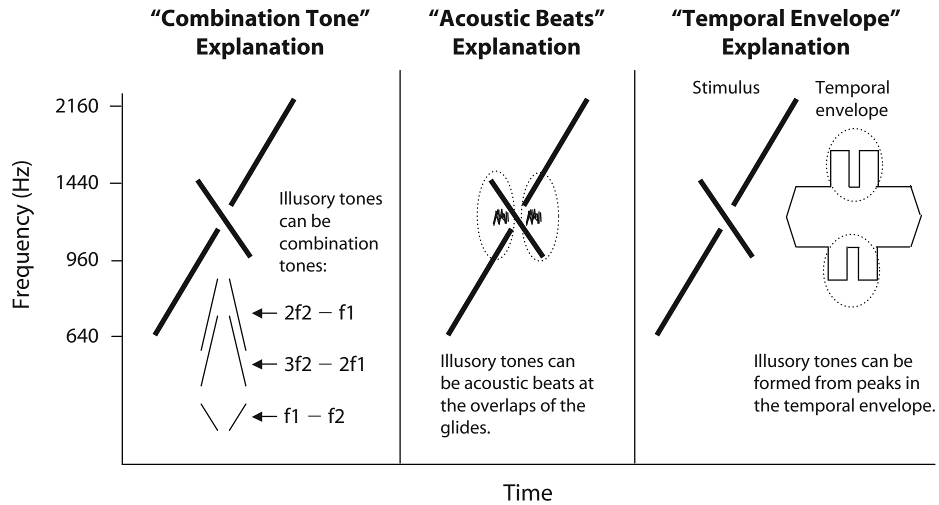

Figure 3. Schematic illustrations of alternative explanations for the gap transfer illusion. The "combination tone" explanation states that the illusory tones in the gap transfer illusion can be combination tones that appear during the overlaps of the glides. f1 and f2 in the left panel of the figure represent the lower and the higher glide component that constitute a two-tone complex when the glides overlap. The "acoustic beats" explanation states that the illusory tones can be caused by acoustic beats that appear during the overlaps of the glides, when the frequencies of the glide components are close toward each other. The "temporal envelope" explanation states that the auditory system makes use of the temporal envelope of the gap transfer stimulus pattern. Similar to processes behind auditory induction, the energy peaks can constitute tones that accompany a continuous tone.

sentation of a two-tone complex. With f1 being the lower and f2 being the higher frequency of a two-tone complex, the most easily audible combination tones have pitches at frequencies $f2 - f1$, $2f1 - f2$, and $3f1 - 2f2$ (Plomp, 1965; Smoorenburg, 1972). In theory, combination tones may appear in any stimulus pattern that causes the gap transfer illusion, because when the shorter glide starts and joins the longer glide, their overlaps make up two two-tone complexes, one before and one after the gap. These two-tone complexes could give rise to the appearance of combination tones. Consequently, the two short tones that listeners hear in the gap transfer illusion might be combination tones. In this study, we examined this hypothesis by obtaining pitch PSEs of the two short tones heard in the gap transfer illusion in two experiments. It is known that the most easily audible combination tones have a pitch that is substantially lower than their primaries f1 and f2 (Plomp, 1965). If the perceived pitches of the short tones in the gap transfer illusion indeed are lower than their primaries (the parts of the short and the long glide that overlap before and after the gap of the long glide), then an explanation of the gap transfer illusion in terms of combination tones must strongly be considered.

A second alternative explanation for the gap transfer illusion is that peripheral processes related to acoustic beats (roughness) cause the perception of the short tones (Figure 3, middle). Acoustic beats may appear perceptually when the frequencies of the two overlapping glides are very close to each other—that is, within a critical bandwidth. This occurs during the overlaps of the crossing glides. Interaction of the glides within a critical bandwidth may cause the overlaps to fuse and cause the perception of two successive tones. When the overlapping glide components are very close to each other in frequency, they can even make up a single tone that fluctuates ("beats") in amplitude (Moore, 1997). In this study, we therefore investigated whether the perceived short tones in the gap transfer illusion are caused by acoustic beats. The acoustic beats explanation dictates that the pitches of the two perceived tones in the gap transfer illusion are fairly similar and close to the frequency centers of the overlapping parts. If the pitch PSEs obtained in the following two experiments show pitch similarity between the two successive tones in the gap transfer pattern, then an explanation in terms of acoustic beats must be considered.

In addition to the combination tone and the acoustic beats explanations, a third alternative explanation for the gap transfer illusion is that the auditory system utilizes the temporal envelopes of amplitude or intensity of the whole stimulus pattern for the formation of the two short tones (Figure 3, right). For example, it is possible that the perception of the two short tones resulted from a process similar to that underlying auditory induction (Warren, 1999). In its simplest form, auditory induction can be obtained when a series of sounds that only differ in intensity are alternated ("homophonic induction"). When listening to the series, often the fainter sound can be heard as continuing along with the more intense sounds, even though it physically is not present behind the more intense sounds (Warren, Obusek, & Ackroff, 1972). Since the amplitude and intensity in the gap transfer stimulus pattern temporally increase during the overlap of the short glide and the long glide, a

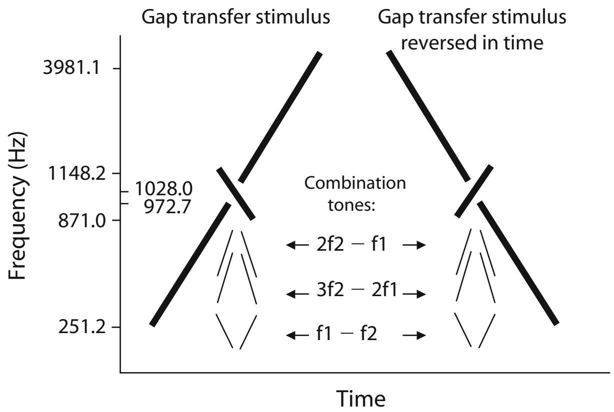

Figure 4. Two of the stimulus patterns used in Experiment 1. Both the gap transfer stimulus pattern and the gap transfer pattern reversed in time should give rise to similar combination tones.

process similar to auditory induction may give rise to the perception of the subjectively divided tones. In the second experiment of this study, we tested this temporal envelope explanation by examining whether the listeners could hear gap transfer in stimulus patterns with completely flat temporal envelopes. If listeners cannot perceive two short divided tones in these patterns, then the gap transfer illusion might have its ground in the intensity envelope or amplitude envelope of the crossing glides.

EXPERIMENT 1

In this matching experiment, participants were asked to judge the pitches and the duration of the subjectively divided tone components in the gap transfer illusion, if perceived. Although the pitch matching tasks and the duration matching tasks were done in separate sessions, we chose to do both pitch matching and duration matching in a single experiment in order to get a full picture of the perception of the two short tones in the gap transfer illusion. As was mentioned in the introduction, PSEs of their pitch and duration can provide strong evidence that the short tones were actually perceived, even though they did not exist physically.

By obtaining information about the pitches of the illusory tones, the combination tone explanation for the gap transfer illusion can be tested. The possibility that combination tones underlie the gap transfer illusion increases if the obtained pitches of the short tones are close to f1 − f2, 2f2 − f1, and/or 3f2 − 2f1, with f1 and f2 being the frequencies of the two overlapping glide components before and after the gap of the longer glide. A second way to test the combination tone explanation was made possible by using stimuli in which a typical gap transfer stimulus pattern was reversed in time. Because the overlaps in the original and the reversed gap transfer stimulus patterns are very similar, the illusory tones in the original stimulus pattern and their counterparts in the reversed pattern should have similar pitches if they are combination tones (Figure 4).

By obtaining information about the pitches of the illusory tones, the acoustic beats explanation for the gap transfer illusion can also be tested. In each gap transfer pattern, the two overlaps are each other's mirror images. If the illusory tones in the gap transfer illusion were the result of acoustic beats or roughness, we may expect that the first tone and the second tone heard in each gap transfer stimulus have similar pitches that are close to the centers of the overlaps.

Method

Participants. Seven students of the Kyushu Institute of Design, Fukuoka, Japan, participated. They were 4 males and 3 females, including author K.K. They were 23–26 years of age and had normal hearing. All had received basic training in music and training in technical listening for acoustic engineers. The reason for asking these trained participants was to reduce possible (inter)subject variability in the data.

Apparatus. The stimulus patterns were generated digitally (12 bits; sampling frequency of 20000 Hz) by a personal computer (Hewlett-Packard 362) and a D/A converter (TEAC PS-9353), that were connected to a low-pass filter with a cutoff frequency of 7 kHz (NF DV-8FL). Through an amplifier (Sansui AU-607XR) and headphones (Rion AD-02), the stimulus patterns were monaurally presented to the participant's left ear in a soundproof room. The sound level of the stimulus patterns—measured with an artificial ear (Brüel & Kjær 4153) and a sound level meter (Brüel & Kjær 2209)—was 77–82 dBA (fast peak).

Stimulus patterns. Eleven stimulus patterns were employed (Figure 5). The first and the second stimulus pattern consisted of a long ascending glide that crossed with a short descending glide. The long ascending glide was 5,000 msec, moving from 251.2 to 3981.1 Hz, and the short descending glide was 500 msec, moving from 1148.2 to 871.0 Hz. Each glide moved linearly on the logarithmic frequency scale at a speed of 0.80 oct/sec (0.24/sec on the common logarithmic scale). In the temporal middle of the stimulus pattern—at $t =$ 2,500 msec from the beginning of the long ascending glide—the long and the short glide crossed each other at 1000 Hz. At the crossing point, the no gap transfer stimulus (Stimulus 1) had a temporal gap of 100 msec in the short descending component, and the gap transfer stimulus (Stimulus 2) had a gap of the same duration in the long ascending component. The rise time and the fall time at the beginning and the end of the long glide were 750 msec. The rise and fall time of the short glide and the rise and fall time at the boundaries of the gap in either the short or the long glide were 3 msec.

The third stimulus pattern consisted of only the short glide with the temporal gap from the no gap transfer stimulus (Stimulus 3). It had frequencies of 1148.2 Hz to 1028 Hz before the gap and frequencies of 972.7 Hz to 871 Hz after the gap. The fourth stimulus pattern (Stimulus 4) also consisted of two short glide components, but differed from the third stimulus. In the fourth stimulus pattern, the first glide component had frequencies of 1148.2 Hz to 972.7 Hz, similar to the starting frequency of the short glide and the end frequency of the long glide—before the gap—in the gap transfer pattern. The second glide component ran from 1028 Hz to 871 Hz, similar to the starting frequency of the second long glide and the end frequency of the short glide—after the gap—in the gap transfer stimulus.

The next four stimulus patterns (Stimuli 5–8) were made by reversing the first four stimuli in time. The remaining three stimulus patterns (Stimuli 9–11) consisted of sinusoidal tones of 500 msec with a temporal gap of 100 msec in the temporal middle (from $t =$ 200 through 300 msec from the start of the tone). The frequencies of these control stimulus patterns were 800, 1000, and 1200 Hz, respectively.

Procedure. The participant was asked to judge the pitches or the durations of the two successive short tones around the temporal middle of each stimulus pattern, if perceived. A standard stimulus pattern, chosen from the eleven stimulus patterns described previ-

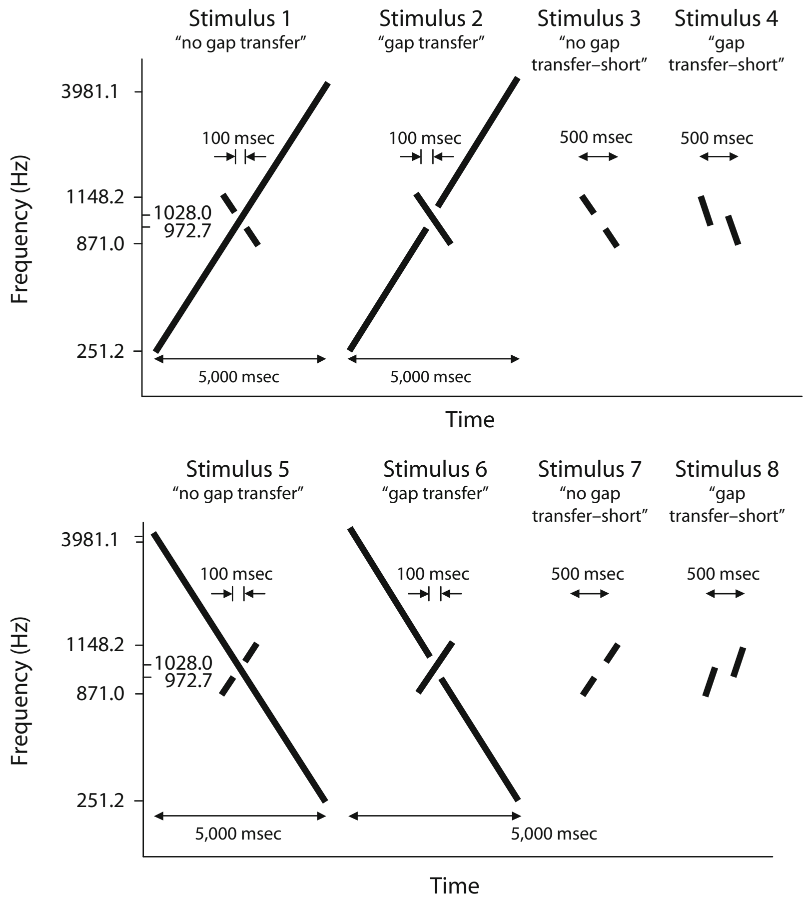

Figure 5. Stimulus patterns employed in Experiment 1. Three control stimuli—each consisting of two steady-state tones separated by a gap—are not shown in the figure.

ously, a silent interval, and a comparison stimulus were presented in this order. The task of the participant was to adjust the comparison stimulus in such a way that it matched one of the following aspects ("targets") of the standard stimulus pattern: (1) the initial pitch of the first short tone, (2) the initial pitch of the second short tone, (3) the duration of the first short tone, and (4) the duration of the second short tone. The only reason for asking the participants to match the initial pitch was to avoid the use of other matching strategies, such as matching the overall pitch or the terminal pitch of each of the short tones. Each session concerned only one of these targets of the standard stimulus pattern. The participant was requested to decline any trial if he/she could not hear "two successive short tones" clearly.

When the target was the initial pitch of the first or the second short tone, the comparison stimulus was a pure tone of 200 msec, including a rise time and a fall time of 20 msec. The sound level of the comparison stimulus was always 78 dBA. When the target was the initial pitch of the first short tone, the comparison stimulus started 4,000 msec after the onset of the short glide/tone of the stimulus patterns. When the target was the initial pitch of the second short tone, the comparison stimulus started 4,000 msec after the next onset in the stimulus pattern (the onset of the second long glide component in the gap transfer stimulus patterns or the onset of the second part of the short glide/tone in the other stimuli). The participant operated a keyboard in order to adjust the frequency of the comparison stimulus to the perceived pitch of the target.

When the target was the duration of the first or the second short tone, the comparison stimulus was a white noise below 7 kHz whose rise and fall were instantaneous. The sound level of the comparison was about 57 dBA. When the target was the duration of the first short tone, the comparison started between 4,000 and 5,000 msec after the onset of the short glide/tone. The time interval between the standard stimulus and the comparison was randomized in order to avoid biasing the participant's duration judgment. When the target was the duration of the second short tone, the comparison started randomly between 4,000 and 5,000 msec after the next onset in the stimulus pattern (the onset of the second long glide component in the gap-transfer stimulus patterns or the onset of the second part of the short glide/tone in the other stimuli). The participant operated a keyboard in order to match the duration of the comparison stimulus to the perceived duration of the target. In this duration matching task, only the stimulus consisting of a 1000 Hz steady-state tone with a 100-msec gap was used as a control, since the other two steady-state stimulus patterns were expected to give almost the same results (see Burghardt, 1973).

In both the pitch matching and the duration matching task, the participant could listen to the standard and the comparison as many times as he/she wanted, and the final frequency or duration of the comparison in each trial was recorded as the PSE. An ascending series and a descending series were assigned to each of the eleven stimulus patterns so that one block consisted of 22 randomized conditions. Each participant finished four blocks, with each block concerning one of the four targets. The participant received two training sessions, one consisting of a pitch matching task and one consisting of a duration matching task. In the experiment, each session was preceded by two warm-up trials. The participant took a break after each session. One session lasted about 45 min. The experiment lasted about 6 h in total.

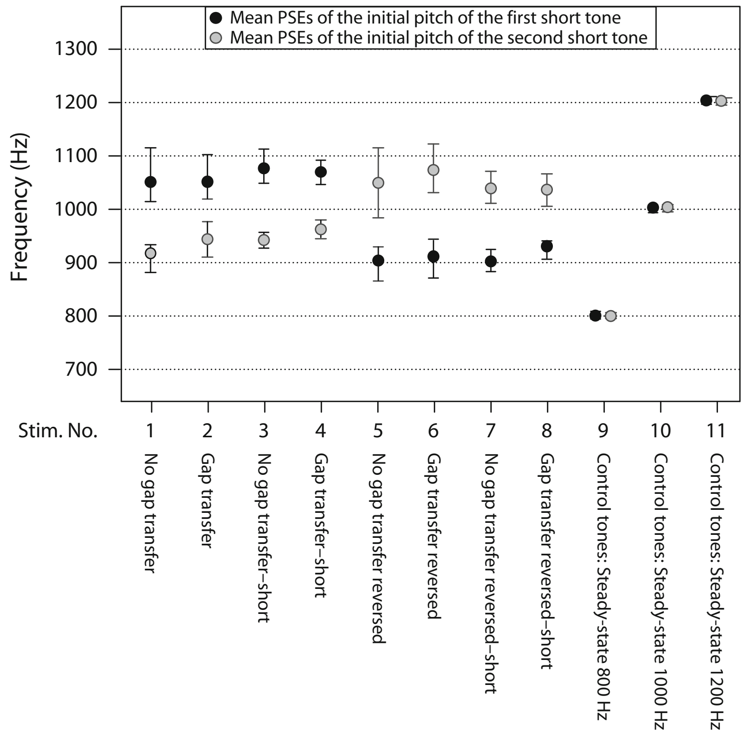

Figure 6. Results of Experiment 1: The means and the 95% confidence interval error bars of the pitch PSEs (points of subjective equality) for the two tones in each stimulus pattern.

Results

Since the subjective properties of the two illusory short tones around the temporal middle in the gap transfer stimulus patterns were going to be examined, it was necessary to check whether the participants could actually perceive these tones. A preliminary phenomenological experiment showed that 6 of the 7 participants clearly heard two successive short tones in the gap transfer stimulus patterns regardless of their listening attitude. One participant could hear two short tones in the stimuli only when he did not attend to the long ascending pitch trajectory. Nevertheless, this participant's percepts were clear enough to perform the present psychophysical task. The participants could clearly hear two short, successive tones in all the other stimulus patterns.

Figure 6 shows the means and the 95% confidence interval error bars of the pitch PSEs of the first short tone and of the second short tone in all stimulus patterns. The figure shows that the pitch judgments were fairly accurate for the control stimuli consisting of a single steady-state tone with a gap (Stimuli 9–11). The 95% confidence interval error bars for these control stimulus patterns ranged just about 9 Hz, indicating that the participants had no difficulties in manipulating the keyboard or performing the matching tasks. Figure 6 shows that the participants could judge the pitches of two tones in all the stimulus patterns with glide components, including the gap transfer stimuli.

The PSEs of the pitches of the short tones in the stimulus patterns with glide components could be classified into two groups, corresponding to the frequency movement of the short glide (components). The first group included the four stimulus patterns with a descending short glide (Stimuli 1–4). The PSEs corresponded to the descending frequency movement of the short glide (components) in these stimuli. The second group included the four stimulus patterns with an ascending short glide (Stimuli 5–8). Here, the PSEs reflected the ascending frequency movement of the short glide (components). The mean PSEs of the pitches of the tones heard in the stimuli were calculated for each of the seven participants and subjected to two ANOVAs in a two-way within-subjects design. The first ANOVA was performed on the mean PSEs of the pitch of the first tone in Stimuli 1–8. The first independent variable was stimulus type (four levels: gap transfer, gap transfer–short tones only, no gap transfer, and no gap transfer–short tones only). The second independent variable was direction of movement of the short glides or glide in Stimuli 1–8 (two levels: descending vs. ascending). An effect of stimulus type was not found [$F(3,18) = 0.31, p = .82$] but the results showed a significant effect of direction of movement [$F(1,18) = 33.25, p < .01$]. That is, the pitch of the first tone heard in Stimuli 1–4 differed from that heard in Stimuli 5–8. The short glide (components) was descending in Stimuli 1–4 and ascending in Stimuli 5–8. A second 4 $\times$ 2 ANOVA with the same independent variables on the mean pitch PSEs of the second tone in Stimuli 1–8 showed similar results. The effect of stimulus type was not significant [$F(3,18) = 1.36, p = .29$] but the

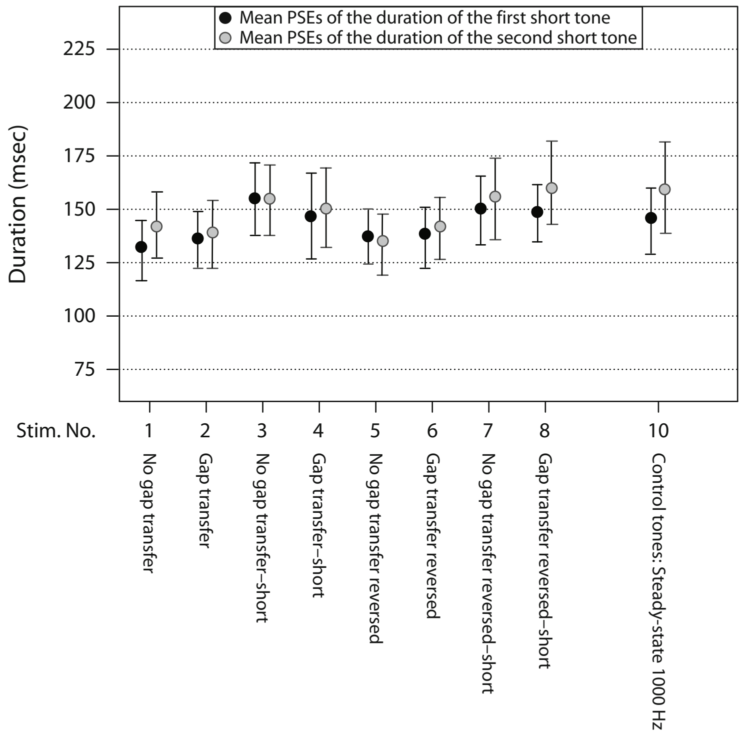

Figure 7. Results of Experiment 1: The means and the 95% confidence interval error bars of the duration PSEs (points of subjective equality) for the two tones in each stimulus pattern.

effect of direction of movement was [$F(1,18) = 14.48$, $p < .01$]. No significant interaction effects were found in either ANOVA.

Figure 7 shows the means and the 95% confidence interval error bars of the duration PSEs of the first short tone and of the second short tone in the stimulus patterns. The figure shows that the duration PSEs of all stimuli—including that of the control stimulus—showed a considerable underestimation. Similar to the pitch PSEs, the duration PSEs were subjected to two ANOVAs in a two-way within-subjects design. The first ANOVA was performed on the mean PSEs of the duration of the first tone in Stimuli 1–8, and the second ANOVA was performed on the mean duration PSEs of the second tone in Stimuli 1–8. Again, the first independent variable was stimulus type (four levels: gap transfer, gap transfer–short tones only, no gap transfer, and no gap transfer–short tones only) and the second independent variable was direction of movement of the short glides or glide in Stimuli 1–8 (two levels: descending vs. ascending). No significant main or interaction effects were found in either ANOVA.

Discussion

First of all, this experiment showed that all participants heard two tones in the gap transfer stimuli, even though these tones did not exist physically. Together with the pitch PSEs, the duration PSEs suggest that the illusory tones corresponded to real tones or real parts of tones in the gap transfer stimulus patterns. The duration PSEs—including those for the control condition—were considerably smaller than the physical duration of the tones, which was 200 msec. We do not have a plausible explanation for the underestimation of the duration PSEs, but perhaps the fact that we used a white noise for the comparison had something to do with the underestimation. In spite of this underestimation, the duration of the illusory tones in the gap transfer stimuli did not differ from the duration of the tones in the other stimulus patterns, where two tones were actually physically present with durations equal to the overlap durations in the gap transfer stimuli.

By obtaining pitch PSEs, we tested whether the illusory tones in the gap transfer stimulus patterns could be combination tones or the results of acoustic beats. The pitch PSEs seem to rule out both explanations. The possibility that the two tones were combination tones seems inappropriate for the following reasons. First, the pitches of the illusory tones in the gap transfer stimulus patterns were not in the pitch ranges in which combination tones would have appeared. Rather, because no effect of stimulus type was found either for the first or the second tone, we can infer that the illusory tones in the gap transfer stimulus patterns were similar to the tones in the stimuli that physically contained two short tones. Second, the pitch-matching results showed a significant effect of direction of movement of the short glide in the stimuli. The results for the gap transfer stimuli in Figure 6 show an example of this effect. The 95% confidence intervals show that there was a difference between the pitch of the first tone in one

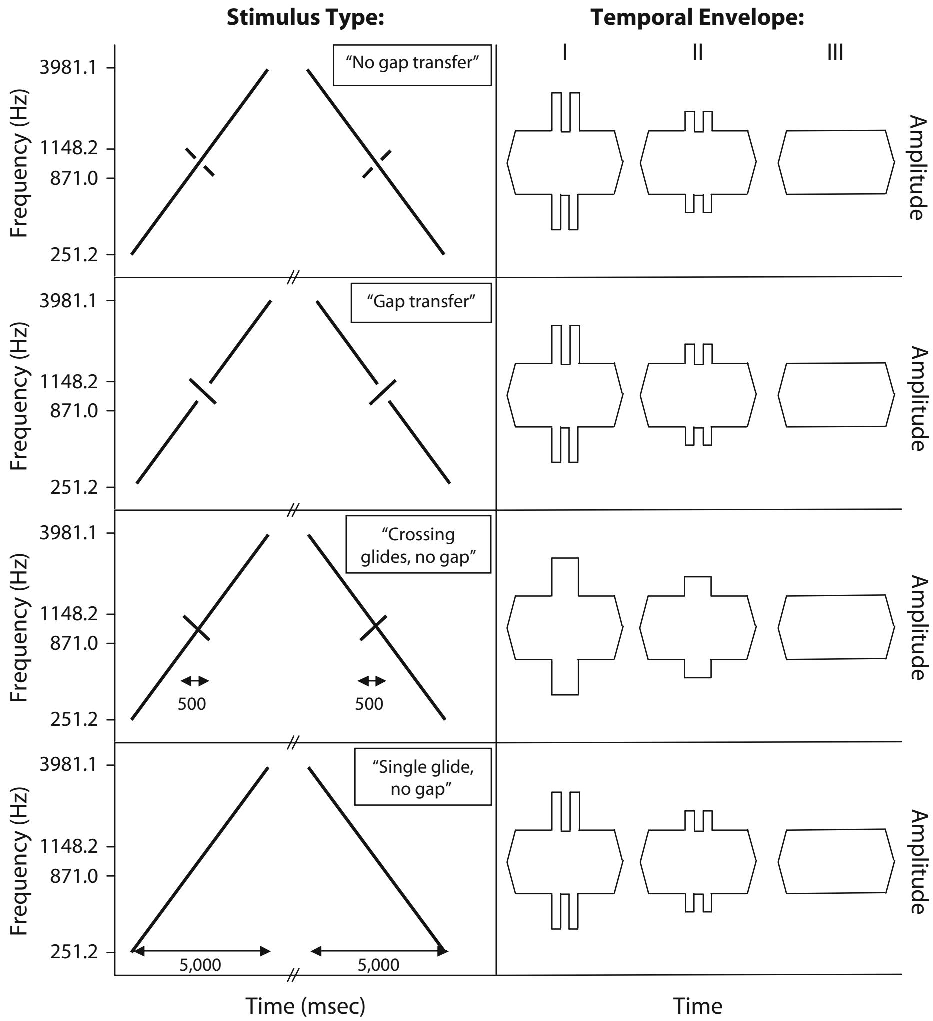

Figure 8. Schematic representations of the stimulus patterns used in Experiment 2 and their temporal envelopes.

gap transfer stimulus (Stimulus 2) and the pitch of the first tone in the corresponding gap transfer stimulus reversed in time (Stimulus 6). The pitch of the second short tone heard in Stimulus 2 also differed from that heard in Stimulus 6. If the illusory tones had been combination tones, then these differences might not have occurred, because the overlapping parts in these stimuli are almost identical (Figure 4).

The results strongly suggest that the acoustic beats explanation for the illusory tones in the gap transfer stimuli is incorrect for the following reason. If the illusory tones had been the results of acoustic beats, then—in theory—the first and the second tone within each gap transfer pattern should have had similar pitches. However, the statistical results show that the pitches of the first and second tone of each gap transfer pattern differed. The above argument on the effect of direction of movement of the short glide also goes against the acoustic beats explanation.

EXPERIMENT 2

In Experiment 2, we tested the temporal envelope explanation as a possible explanation for the illusory tones in the gap transfer illusion. In this experiment, the appearance of the gap transfer illusion was investigated in stimulus patterns in which the temporal envelopes were manipulated. The temporal envelope of the gap transfer stimuli used in Experiment 1 is indicated in Figure 8 (second pane from top, Envelope Condition I). If the temporal envelope belonged to a pure component, then the envelope pattern would result in auditory induction (Warren, 1999), causing a percept consisting of a continuous tone and two successive short tones. It is possible that the same perceptual mechanism that underlies auditory induction caused the gap transfer illusion. This may not fully explain why in Experiment 1 the pitches of the two short tones in one gap transfer stimulus differed from those heard in the gap transfer pattern reversed in time, but we still have to consider this possibility, because auditory induction is a well-established phenomenon. In this experiment, the rise and fall times of the glide components were increased to 25 msec in order to eliminate any possibility of spectral splatter at the onset or offset of a sound (Miyasaka & Sakai, 1980).

In Experiment 2, music students with (acquired) absolute pitch were asked to judge the pitches of the two short tones in the gap transfer illusion in stimulus patterns in which the temporal envelope was completely flat. If the two short tones in the gap transfer illusion were to appear also in such stimulus patterns, then the temporal envelope hypothesis would lose a lot of its explanatory power. We asked the participants to draw their percept of each stimulus pattern and assign musical note names to the perceived pitch trajectories in the same trial. A point of criticism toward Experiment 1 might be that we could not ascertain whether the participants judged the pitches of independent and new short tones or just the beginnings of glide components. By asking participants to give a verbal and graphical description of the stimulus patterns as well as pitch judgments of the two short tones, we could obtain clearer information regarding the matter. In contrast with Experiment 1, the pitches of the first and the second illusory tone were judged in the same trial, which was possible only by employing absolute pitch possessors.

Method

Participants. Six female students of Miyagi Gakuin Women's College, Sendai, Japan, 21–23 years old, participated. They had normal hearing and majored in music. Five of them had acquired absolute pitch.

Apparatus. The stimulus patterns were generated by a personal computer (IBM ThinkPad 560) and converted into audio signals by a personal computer (IBM PS/V 2411) with a D/A converter (Maxam TE410). Each stimulus pattern was recorded on DAT tape by a DAT deck (Pioneer D-C88) through a low-pass filter with a cutoff at 7 kHz (NF DV-8FL). All the stimulus patterns were played back with another DAT deck (Sony TCD-10) whose output was connected to an amplifier (JVC AX-S900) and headphones (STAX SR-Lambda professional). Each stimulus pattern was presented to the left ear of the participant in a soundproof room. The sound level of the stimulus patterns was 70 dBA (fast peak).

Stimulus patterns. Twenty-four stimulus patterns were generated and divided in three groups of 8 stimuli according to the shape of their temporal envelopes (Figure 8). In each group of 8 stimuli, 3 stimuli consisted of a long ascending glide of 5,000 msec (251.2 to 3981.1 Hz) that crossed with a short descending glide of 500 msec (1148.2 to 871.0 Hz). Each glide ascended or descended at a speed of 0.80 oct/sec (0.24/sec on the common logarithmic frequency scale). In all stimulus patterns, the long and the short glide crossed each other at 1000 Hz in the temporal middle of the stimulus pattern, at $t = 2{,}500$ msec from the onset of the long glide. As in Experiment 1, there was a no gap transfer and a gap transfer stimulus pattern. The no gap transfer stimulus pattern had a temporal gap of 100 msec in the short descending glide and the gap transfer pattern had a gap of the same duration in the long ascending glide. In the 3rd stimulus pattern, no temporal gaps were inserted in the crossing glides. As in Experiment 1, these 3 stimulus patterns were also generated backward in time. The 2 remaining stimulus patterns were a control stimulus consisting of only a single glide of 5,000 msec, moving from 251.2 to 3981.1 Hz, and the same stimulus reversed in time.

In each stimulus pattern, the long glide had a rise and a fall time of 750 msec. The short glide component(s) that appeared in 18 of the 24 stimulus patterns had a rise and a fall time of 25 msec. When a gap was inserted in one of the glides, the boundaries around the gap had a rise and fall time of 25 msec. The rise and fall times were increased to 25 msec in the present experiment in order to minimize any undesirable spectral changes in the procedures to shape temporal envelopes, which are described below.

The temporal envelopes of the eight stimulus patterns were varied in the following three ways (Figure 8). In Envelope Condition I, each glide had a fixed amplitude, as was the case in Experiment 1. This means that the total peak amplitude doubled during the overlaps of the two glides. In Envelope Condition II, the total intensity of each stimulus pattern was kept constant (except for the beginning and the end of the whole stimulus pattern). This means that the total peak amplitude increased by a factor of 1.41 ($= \sqrt{2}$) during the overlaps of the two glides. In Envelope Condition III, the total peak amplitude of each stimulus pattern was kept constant (except for the beginning and the end of the whole stimulus pattern). This resulted in stimulus patterns with a completely flat temporal envelope. The temporal envelopes of the control stimuli with a single glide were also varied in these three ways, so that the total peak amplitude changed in the same manner as in the other stimulus patterns.

Procedure. Each stimulus pattern was played repeatedly to the participant with silent intervals of 4.2 sec between presentations, so the participant could listen to each stimulus pattern as many times as she wanted. The participant was instructed to describe the perceptual impression of each stimulus pattern both graphically and verbally. In each trial, the participant first drew a freehand figure of her perceptual impression of the stimulus pattern with a pencil on a sheet of graph paper. The participant was instructed to consider the horizontal dimension as corresponding to time and the vertical dimension as corresponding to pitch. This kind of graphical representation of auditory percepts was used also in Nakajima et al. (2000). After drawing the percept of the stimulus pattern, the participant gave a verbal report of her perceptual impression. If a short perceptual component could be perceived in the stimulus pattern, the participant was asked to draw the precise shape of the short perceptual component and name its pitch if possible, corresponding to the music scale. Each participant described the 24 stimulus patterns twice. A randomized block of 24 stimulus patterns was divided in two sessions of 12 trials. The participant completed the first session of each block after three warm-up trials. The second session of each block was preceded by one warm-up trial, and the participant was required to take a break after each session. One session lasted about 30 min. The experiment lasted about 2 h in total.

Results

The graphical descriptions of the 6 participants' perceptual impressions showed that they perceived all stimulus patterns as consisting of a long ascending or descending pitch trajectory and a short pitch trajectory or trajectories (tone or tones). The note names that were assigned by the five absolute pitch possessors to the ascending no gap transfer and the ascending gap transfer stimulus patterns are indicated in Figure 9. The degrees of discontinuity of the perceptual components heard in the stimuli could be classified into three categories (1) completely discontinuous, (2) partially discontinuous (a small unevenness in pitch or in time could be perceived in the pitch trajectory, but it could not be perceived as clearly consisting of two sounds separated in pitch or by a temporal gap), and (3) completely continuous. These categories made an ordinal scale and, for $n = 12$ (6 participants $\times$ 2 replications), two-tailed sign tests were conducted between the continuity scores of the long pitch trajectory and those of the short pitch trajectory for each stimulus pattern

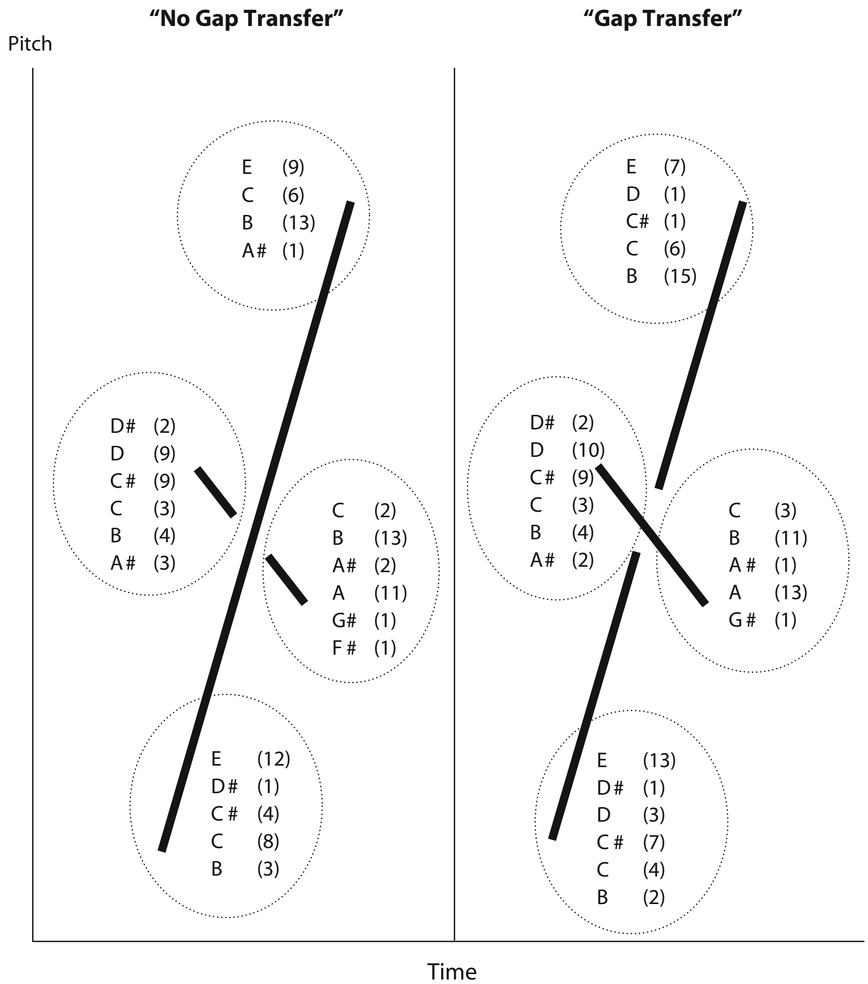

Figure 9. Note names assigned to the ascending "no gap transfer" stimulus and the ascending "gap transfer" stimulus, by absolute pitch possessors in Experiment 2. Each of the five absolute pitch possessors made two judgments for each of the three envelope conditions of the stimulus patterns, resulting in 30 pitch judgments in total. The number of times a tone was heard as corresponding to a particular musical note is indicated in between brackets behind that note. Some unclear pitch judgments were left out of the figure.

(Table 1). Table 1 shows that in the no gap transfer and the gap transfer stimuli, the short glide was judged as significantly more discontinuous than the long glide in all envelope conditions ($p < .01$). Five of the 6 participants even heard discontinuity in the short glide in crossing-glide stimuli without a gap in all envelope conditions.

A long ascending or descending tone and two successive short tones (crossing percepts) were also perceived in two of the control stimuli, even though they consisted of just a single long glide. In Envelope Condition I, the short pitch trajectory was significantly more discontinuous than the long pitch trajectory in the stimulus with a single ascending glide ($p < .05$) and a single descending glide ($p < .01$). In Envelope Condition II, the short pitch trajectory was significantly more discontinuous than the long pitch trajectory in the stimulus with a single ascending glide ($p < .05$). In Envelope Condition III, only a single pitch trajectory was perceived.

Discussion

The experiment showed that discontinuity was also heard in the short pitch trajectory in the control stimulus consisting of crossing glides without a gap in all envelope conditions. Although each individual glide in the stimulus was physically continuous, acoustic beating seemed to have caused two short dips (of 13 msec at 3 dB below peak level) around the glides' crossing point (see Nakajima et al., 2000). These dips seem to have been allocated only to the shorter glide. The reason why only the short glide was heard as interrupted—even though the dips were shared by both the short and the long glide—can tentatively be explained by exclusive allocation of the dips to the attended auditory stream and impaired acuity of the auditory system when it has to compare two auditory streams at the same time (Remijn et al., 2007).

As for the gap transfer stimuli, the results of this experiment showed that the short continuous glide was

Table 1
Number of Times That the Long Glide Was Judged As More Continuous Than the Short Glide in Experiment 2

Temporal Envelope Type	Stimulus Type							
	No Gap Transfer	No Gap Transfer (Reversed)	Gap Transfer	Gap Transfer (Reversed)	Crossing Glides No Gap	Crossing Glides No Gap (Rev.)	Single Glide No Gap	Single Glide No Gap (Rev.)
I	12**	12**	12**	12**	7**	8**	12**	12**
II	12**	12**	12**	12**	8**	8**	9**	7*
III	12**	12**	11**	12**	8**	8**	nst	nst

Note—Each score represents the $N+$ value in a two-tailed sign test over 12 repetitions of each stimulus. $N-$ was zero in all conditions. $^*p < .05$; $^{**}p < .01$; nst, no short tone was perceived in the stimulus.

subjectively divided into two short tones in all envelope conditions. The appearance of the short tones in the gap transfer stimulus patterns cannot be related to spectral splatters at the onset or offset of the glides, since this was greatly reduced by increasing the glides' rise and fall time (Miyasaka & Sakai, 1980). More importantly, since the gap transfer illusion was perceived in all envelope conditions, we may infer that the temporal envelope explanation cannot fully account for the illusion's appearance. The basis of the temporal envelope explanation is that sudden change in amplitude may signal an onset or an offset of a sound (see Bregman, 1990). The experiment showed that such amplitude changes in the two control stimuli indeed could have caused the perception of two short tones, even though the stimuli consisted of just a single glide. The sudden amplitude changes may have facilitated the perception of the two tones in the gap transfer illusion in a similar way. The results of the stimuli with Envelope Conditions II and III, however, indicate that local amplitude changes were not a requisite for the gap transfer illusion. In the gap transfer stimulus patterns, except for the first 30 msec at the onset of the short glide, the short glide and the long glide move in the same critical bandwidth (1 equivalent rectangular bandwidth at the 1000 Hz reference frequency in the temporal middle of the stimuli is 132.6 Hz, following Glasberg & Moore, 1990). Therefore, in Envelope Condition II, the participants must have perceived each stimulus pattern as having the same overall loudness. In Envelope Condition III, the loudness of the stimulus during the overlaps may have even been softer than the nonoverlapping parts. In spite of this fact, the gap transfer stimulus patterns gave rise to the gap transfer illusion in all the envelope conditions, thus including the flat patterns. The temporal envelope explanation centered around a mechanism of auditory induction (Warren, 1999) thus cannot fully explain the gap transfer illusion.

A point to strengthen the claim that the gap transfer illusion cannot be explained fully with the temporal envelope hypothesis comes from the finding that the envelope conditions did not have any influence upon the note names assigned to the short tones in the gap transfer stimuli. With regard to the note names, Experiment 2 seems to suggest that the pitches assigned to the two perceived short tones in the stimuli with crossing glides corresponded to the frequency ranges of the short glide component(s) in the temporal middle of the stimuli. Figure 9 shows that in the gap transfer stimuli with the long ascending glide, the musical notes assigned most frequently to the pitch of the first short tone were the notes D and C♯. These notes correspond to frequencies of 1174.7 and 1108.7 Hz, respectively. These frequencies are close to the frequency range covered by the first half of the short physical glide (1148.2–1000 Hz). A similar tendency was found in Experiment 1, where the mean PSE of the first short tone in the ascending gap transfer stimulus pattern was 1059.9 Hz (Figure 6), and the frequency range covered by the first half of the short glide was 1148.2–1000 Hz. Figure 9 shows that the musical notes assigned most frequently to the second short tone in the gap transfer stimuli with the long ascending glide were the notes B and A, which correspond to frequencies of 987.8 and 880 Hz, respectively. These frequencies fall in the frequency range covered by the last half of the short glide in the ascending gap transfer stimuli (1000–871 Hz). This result was also found in Experiment 1, where the mean PSE of the second short tone in the ascending gap transfer pattern was 944.4 Hz. The pitch names assigned to the short tones in the gap transfer stimuli with the long descending glide showed the same tendency.

GENERAL DISCUSSION

In the gap transfer illusion, a physically continuous short glide is subjectively divided into two short tones (Figure 1). Nakajima et al. (2000) described that the short tones in the gap transfer illusion are formed through the integration of sound edges that physically belong to different sounds (Figure 2). By obtaining subjective measures of the pitches and durations of the two short tones in the gap transfer illusion, in the present study, we tested the validity of this "edge integration" explanation against three alternative low-level explanations for the illusion.

With regard to the first alternative explanation, the results show that the two short tones heard in the gap transfer illusion are not the result of combination tones that could have appeared during the overlaps of the crossing glides in the stimulus patterns. The main reason for this is that the pitch PSEs of the illusory tones were not in the frequency ranges in which combination tones could have occurred. The second alternative explanation, based on the possible appearance of acoustic beats in the gap transfer stimulus patterns, also could not explain the results that followed from the pitch matchings in the experiments. If the illusory tones in the gap transfer illusion were the results of acoustic beats, then the pitches of the first and the second short tones in each gap transfer pattern should have been very close to each other. However, their pitches differed systematically. With regard to the third alternative explanation, Experiment 2 showed that the gap transfer illusion appeared even when the intensity or peak amplitude of the gap transfer stimuli remained constant. This result showed that a change of the total amplitude or the total intensity is not a requisite for the illusion. However, we cannot simply discard the temporal envelope explanation in view of the fact that even a single glide with local increases in amplitude gave rise to a percept that was similar to that heard in the gap transfer stimuli. The perception of the two short tones in the gap transfer illusion, therefore, may have been facilitated by the detection of amplitude increases in the stimulus patterns.

The temporal envelope hypothesis, however, has the disadvantage that on its own, it cannot explain why the short tones in the gap transfer illusion were perceived as having different pitches. The edge integration explanation, on the other hand, can account for the perceived pitches or pitch movements of the short illusory tones reasonably well. Physiological and psychophysical studies have shown that onsets and offsets of sounds can carry pitch information (see, e.g., Bregman et al., 1994; Brugge, 1992; Heil & Irvine, 1998). According to Kubovy and Van Valkenburg (2001), pitch is one of the strongest cues to define the auditory edges that delimit auditory objects. Pitch information, for example, can be used by the auditory system to segregate soft sounds from a noisy background. According to the ideas of Nakajima et al. (2000), the short tones in the gap transfer illusion would incorporate the pitch information of the perceptually coupled onsets and offsets detected in the stimulus patterns. As with the temporal envelope explanation, increments in sound intensity are important for the edge integration explanation as well, because a sudden increase and decrease in the amplitude of a sound may be interpreted by the auditory system as the appearance of an onset and an offset, respectively (see also Bregman, 1990; Fishbach, Nelken, & Yeshurun, 2001). However, in order to detect sound edges, the auditory system does not always need a change in amplitude or intensity. The advantage of the edge integration explanation is that a sudden change in pitch itself can signal an onset or offset of a sound, without a change in amplitude or intensity at the pitch change. The edge integration explanation could therefore account for the finding that the gap transfer illusion appears when the stimulus has a flat temporal envelope. Because the pitch changes are close to each other in time, the perceived onset and offset can be integrated, with an illusory auditory event that contains pitch information of the onset and offset as a result. This is how the short tones in the control patterns with just a single glide with amplitude changes may have been constructed as well.

Handel (1989) pointed out that, "Hearing must bridge the interval between beginning and end. It is almost impossible to take a 'snapshot' of auditory events, so that the totality is available at one time" (p. 163). The edge integration explanation—as a part of "auditory grammar" (Nakajima & Sasaki, 1996; Nakajima et al., 2000)—can be used to draw a rough sketch of how the auditory system connects "beginning and end." This rough sketch may be the starting point of the construction of auditory events. In another study (Remijn & Nakajima, 2005), we have already demonstrated how the edge integration model can predict the illusory appearance of an auditory event in stimuli consisting of two partly overlapping glides, whereas attempts to connect these ideas to language perception are being made (Wang & Nakajima, 2005).

AUTHOR NOTE

All authors contributed equally to this article and are listed in alphabetical order. The research was supported by grants to Y.N. from the Japanese Ministry of Education (07459017 in the fiscal years 1995–1997), the Japan Society for the Promotion of Science (14101001 in the fiscal years 2002–2006), and the "Center of Excellence" program of the Japanese government, entitled "Design of Artificial Environments on the Basis of Human Sensibility." Chris Darwin, Stephen Handel, Gert ten Hoopen, and Daigoh Suetomi gave us valuable suggestions. Tsuyoshi Yoshizumi and Yoko Fujishima are thanked for their technical support. All correspondence should be addressed to Y. Nakajima, Kyushu University, Department of Acoustic Design, 4-9-1 Shiobaru, Minami-ku, Fukuoka 815-8540, Japan (e-mail: nakajima@design.kyushu-u.ac.jp).

REFERENCES

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

Bregman, A. S., Ahad, P. A., & Kim, J. (1994). Resetting the pitch-analysis system. 2. Role of sudden onsets and offsets in the perception of individual components in a cluster of overlapping tones. *Journal of the Acoustical Society of America*, **96**, 2694-2703.

Brugge, J. F. (1992). An overview of central auditory processing. In A. N. Popper & R. R. Fay (Eds.), *The mammalian auditory pathway: Neurophysiology* (pp. 1-33). New York: Springer.

Burghardt, H. (1973). Die subjektive Dauer schmalbandiger Schalle bei verschiedenen Frequenzlagen. *Acustica*, **28**, 278-284.

Darwin, C. J., & Carlyon, R. P. (1995). Auditory grouping. In B. C. J. Moore (Ed.), *Hearing* (Handbook of Perception and Cognition, 2nd ed., pp. 387-424). San Diego: Academic Press.

Deutsch, D. (1983). Auditory illusions, handedness, and the spatial environment. *Journal of the Audio Engineering Society*, **31**, 607-618.

Fishbach, A., Nelken, I., & Yeshurun, Y. (2001). Auditory edge detection: A neural model for physiological and psychoacoustical responses to amplitude transients. *Journal of Neurophysiology*, **85**, 2303-2323.

Glasberg, B. R., & Moore, B. C. J. (1990). Derivation of auditory filter shapes from notched-noise data. *Hearing Research*, **47**, 103-138.

Handel, S. (1989). *Listening: An introduction to the perception of auditory events*. Cambridge, MA: MIT Press.

Heil, P., & Irvine, D. R. F. (1998). The posterior field P of cat auditory cortex: Coding of envelope transients. *Cerebral Cortex*, **8**, 125-141.

Kubovy, M., & Van Valkenburg, D. (2001). Auditory and visual objects. *Cognition*, **80**, 97-126.

McAdams, S., & Bregman, A. S. (1979). Hearing musical streams. *Computer Music Journal*, **3**, 26-43.

Miyasaka, E., & Sakai, H. (1980). Detectability of switching transients. *Journal of the Acoustical Society of Japan*, **36**, 244-252.

Moore, B. C. J. (1997). *An introduction to the psychology of hearing* (4th ed.). London: Academic Press.

Nakajima, Y., & Sasaki, T. (1996). A simple grammar of auditory stream formation [Abstract]. *Journal of the Acoustical Society of America*, **100**, 2681.

Nakajima, Y., Sasaki, T., Kanafuka, K., Miyamoto, A., Remijn, G., & ten Hoopen, G. (2000). Illusory recouplings of onsets and terminations of glide tone components. *Perception & Psychophysics*, **62**, 1413-1425.

Plomp, R. (1965). Detectability threshold for combination tones. *Journal of the Acoustical Society of America*, **37**, 1110-1123.

Remijn, G. B., & Nakajima, Y. (2005). The perceptual integration of auditory stimulus edges: An illusory short tone in stimulus patterns consisting of two partly overlapping glides. *Journal of Experimental Psychology: Human Perception & Performance*, **31**, 183-192.

REMIJN, G. B., NAKAJIMA, Y., & TANAKA, S. (2007). Perceptual completion of a sound with a short silent gap. *Perception*, **36**, 898-917.

SMOORENBURG, G. F. (1972). Audibility region of combination tones. *Journal of the Acoustical Society of America*, **52**, 603-614.

WANG, H., & NAKAJIMA, Y. (2005). Illusory Chinese syllables perceived in stimulus patterns consisting of two partly overlapping harmonic glides. *Report for the Acoustical Society of Japan*, **H-2005-116**, 681-684.

WARREN, R. M. (1999). *Auditory perception: A new analysis and synthesis* (4th ed.). Cambridge: Cambridge University Press.

WARREN, R. M., OBUSEK, C. J., & ACKROFF, J. M. (1972). Auditory induction: Perceptual synthesis of absent sounds. *Science*, **176**, 1149-1151.

(Manuscript received March 20, 2006;
revision accepted for publication September 11, 2006.)