

Roles of temporal proximity between sound edges in the perceptual organization of veridical and illusory auditory events

Nakajima, Yoshitaka
Department of Human Science, Kyushu University

Remijn, Gerard Bastiaan
Department of Human Science, Kyushu University

Kamimura, Yuki
Graduate School of Design, Kyushu University

Kanafuka, Kyoko
Nippon Telegraph and Telephone West Corporation

<https://hdl.handle.net/2324/7153271>

出版情報 : Hearing Research. 422, pp.108546-, 2022-09-01. Elsevier
バージョン :
権利関係 : Creative Commons Attribution 4.0 International





Roles of temporal proximity between sound edges in the perceptual organization of veridical and illusory auditory events

Yoshitaka Nakajima^{a,b,*}, Gerard B. Remijn^a, Yuki Kamimura^c, Kyoko Kanafuka^d

^a Department of Human Science/Research Center for Applied Perceptual Science, Kyushu University, Shiobaru 4-9-1, Minami-ku, Fukuoka-shi 815-8540, Japan

^b Sound Corporation, Japan

^c Graduate School of Design, Kyushu University, Japan

^d Nippon Telegraph and Telephone West Corporation, Osaka, Japan

ARTICLE INFO

Article history:

Received 27 December 2021

Revised 20 May 2022

Accepted 24 May 2022

Available online 26 May 2022

Keywords:

auditory organization
gap transfer illusion
proximity principle
Auditory Grammar
crossing and bouncing

ABSTRACT

The gap transfer illusion is an auditory illusion in which a temporal gap in a long glide is perceived as if it had transferred to a physically continuous shorter glide. The illusion typically occurs when the long and the shorter glide cross each other at their temporal midpoints, where the long glide is physically divided by the gap. The occurrence of the gap transfer illusion was investigated in stimuli in which the duration and the slope of the long glide were 5000 ms and ~ 0.8 oct/s. The shorter glide was given different frequency ranges and different temporal ranges, and thus its time-frequency slope was also varied. The overlap configuration of these crossing glides was varied as well. As control stimuli, we used stimuli in which a continuous long glide crossed a shorter glide with a gap, i.e., the opposite configuration of the gap-transfer stimuli as above, as well as stimuli in which both crossing glides were continuous. The perception of two crossing tones tended to be facilitated when the glides differed in duration and/or slope. When the glides were relatively similar in duration and slope, however, bouncing percepts appeared more often. Similarity between the crossing tones thus promoted auditory bouncing, while dissimilarity between them facilitated the crossing percept. If the crossing percept dominated in gap-transfer stimuli, the gap transfer illusion took place in a typical manner, but the illusory transfer of the gap could occur even when the crossing percept was not dominant. When the shorter glide was as short as 500 ms, the crossing percept and the gap transfer illusion were robust. The mechanism of the illusion was examined in terms of factors that can influence the perceptual integration of auditory stimulus edges, i.e., onsets and offsets, of physically different sounds. Much like the perceptual construction of speech units, we suggest that the auditory system utilizes a rough time window of several hundreds of milliseconds to construct an initial skeleton percept of auditory events. The present data indicated the importance of the temporal proximity, rather than the frequency proximity, between sound edges in the illusory tone construction.

© 2022 The Author(s). Published by Elsevier B.V.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)

1. Introduction

This article focuses on an auditory illusion, the *gap transfer illusion*, which functions as a tool to investigate basic principles of auditory organization, in particular those related to gestalt psychology. Our immediate purpose is to clarify the mechanism of this illusion, thereby aiming at clarifying how certain gestalt principles relate to auditory organization. Since the early 1900s, researchers of visual perception share a solid understanding that gestalt principles

and the organization of figures and ground underlie the general framework of visual organization (e.g., Koffka, 1935). It took until the end of the last century, however, before these ideas were thought to be applicable to auditory organization as well. Handel (1993) and Bregman (1990) in particular systematically argued that the perceptual formation of auditory events and auditory streams can be bases of auditory organization, and their arguments lead the present research.

In order to study auditory organization, Nakajima et al. (2000) reported the gap transfer illusion. Their most important idea, developed further by Kanafuka et al. (2007), is that an illusory auditory event can be constructed by perceptually connecting temporal sound edges to generate a framework for an auditory event. In the gap transfer illusion, the onset of a glide tone and the

* Corresponding author at: Department of Human Science/Research Center for Applied Perceptual Science, Kyushu University, Shiobaru 4-9-1, Minami-ku, Fukuoka-shi 815-8540, Japan.

E-mail address: nakajima.yoshitaka.310@m.kyushu-u.ac.jp (Y. Nakajima).

offset of a different glide tone are perceptually connected, obeying the proximity principle: If an onset and an offset are proximate to one another, they can be integrated to construct an auditory event, even when they physically belong to different tones. It was not clearly stated, however, whether they meant proximity in time, in frequency, or both. Since there were no empirical data on this issue, the present experiments were performed.

The stimulus patterns in our previous studies (Nakajima et al., 2000; Kanafuka et al., 2007) were first created to further investigate how organizational principles in hearing relate to those in visual motion perception (see also Wang et al., 2015). In particular, we wanted to see how crossing sounds, consisting of two glide tones, are perceptually organized. In the visual modality, it is known that two identical, visual objects that move towards each other, overlap, and then move away from each other commonly yield bistable ‘streaming’ or ‘bouncing’ percepts (Metzger, 1934; a closely related phenomenon was discovered by Benussi, 1917). In the streaming percept, each object moves towards the opposing object’s starting position. In the bouncing percept, the objects bounce off of each other at their contact point, and retrace their movement back to their own starting position. Ongoing research has shown that the bouncing percept is facilitated by similarity between the physically crossing objects. Dissimilarity, such as in object shapes, speed, and brightness, as well as the active tracking of one object over the other, promotes the streaming percept (e.g., Metzger, 1934; Burns and Zanker, 2000).

A similar phenomenon exists in the auditory modality (Halpern, 1977, as quoted by Bregman, 1990; McPherson et al., 1994; see also Tougas and Bregman, 1985). An ascending and a descending frequency glide that cross each other at their temporal midpoint can either be heard as streaming through each other, i.e., the ‘crossing’ percept, or as ‘bouncing’ off of each other at the crossing point. In the latter bouncing case, two pitch trajectories can be heard at the same time, but not exactly as bouncing at a particular point: as a U-shaped pitch trajectory together with an inverted U-shaped pitch trajectory, often apart from each other. Each trajectory is composed of portions of two different glides. The glide components in the top half of the (logarithmic) frequency range are integrated according to the frequency proximity principle (Bregman, 1990), and the glide components in the bottom half as well. Research has shown that, likewise to the influence of object similarity on visual bouncing perception, physical similarity of slope or timbre between the different glides promotes auditory bouncing (Bregman, 1990; McPherson et al., 1994).

While trying to disambiguate crossing and bouncing in auditory stimuli, Nakajima and colleagues came across an illusory stimulus (Nakajima et al., 2000). Their initial idea was to promote auditory crossing in a typical crossing-bouncing stimulus by making one of the glides far shorter than the other, and by making one of the glides discontinuous with a gap at the crossing point. When the gap was in the shorter glide, indeed a long continuous pitch trajectory was heard as crossing with a shorter, discontinuous pitch trajectory. When the gap was inserted in the longer glide, however, the percept remained the same: The gap was heard in the shorter pitch trajectory, while the longer pitch trajectory, which physically contained the gap, was heard as continuous. Hence, the effect was called the *gap transfer illusion*.

It was argued that the gap transfer was caused by the perceptual reintegration of acoustic cues that physically belonged to different sounds. Under the header of “Auditory Grammar” (Nakajima et al., 2014), which describes a number of heuristics regarding how sound components, such as *onsets*, *offsets*, *fillings*, and *silences*, are perceptually integrated into auditory events and auditory streams, it was argued that the first portion of the short discontinuous trajectory in the illusory percept results from the perceptual reintegration of the onset of the short glide with the offset

of the long glide component before the gap, obeying the proximity principle. Likewise, the second portion of the short trajectory is caused by the reintegration of the onset of the long glide component after the gap and the offset of the short glide. Research has shown that the two illusory short tones cannot be the result of acoustic beats nor combination tones occurring during the overlap of the glide components (Kanafuka et al., 2007); their pitches were basically determined by the glide portion between the onsets and the offsets. So far, research on the occurrence of such illusory gap transfer has been extended to stimuli with manipulations in the glides’ intensity and spectra (Kuroda et al., 2009, 2010).

In the present study, we investigated the occurrence of the gap transfer illusion in stimuli in which the duration (temporal range) and the frequency range of the short glide (hence, its slope) and the overlap configuration were varied in order to manipulate the temporal and the frequency proximity between the onsets and the offsets to be reintegrated for the illusion to occur. Differences in slope between the crossing glides might promote crossing perception (McPherson et al., 1994), making the illusion more likely to occur.

2. General method

Three very similar experiments were conducted, and their methods will be described together in this section except for the stimuli used in Experiments 2 and 3, which were introduced as new ideas developed. The stimuli used in Experiment 1 will be described in the present section in order to give a general idea of our paradigm. Procedure uniquely related to the stimuli used in a particular experiment will be described in the section of this experiment.

2.1. Participants

Twelve/13/8 participants of 20–24/20–25/22–30 years of age, joined Experiment 1/2/3. They were students of the Department of Acoustic Design or the Department of Human Science, Kyushu University, Japan, who all had received listening training for acoustic engineers (Iwamiya et al., 2003). All had heard basic demonstrations of the gap transfer illusion in classes. None of the participants of Experiment 2 had participated in Experiment 1. Three of the 8 students that participated in Experiment 3, had participated in Experiment 2; there was at least 6 months between the experiments. All participants had provided written, informed consent as to their participation. The procedures of the experiments were approved by the Ethics Committee of the Faculty of Design, Kyushu University.

2.2. Stimuli (in Experiment 1)

Three basic stimulus types were employed as in Fig. 1: (1) *gap-transfer* stimuli, (2) *gap-transfer control* stimuli, and (3) *no-gap* stimuli. All stimuli consisted of two gliding tones that moved in frequency with the same speed of $0.24 \text{ }^1/\text{s}$ on the common logarithmic scale, approximately 0.80 oct/s , as we will refer to this glide speed for convenience. The two gliding tones moved in opposite directions, so that they crossed each other at a (logarithmic) center frequency of 1000 Hz. In each stimulus type, the long glide was 5000 ms and moved from 251.2 to 3981.1 Hz or back (3.99 oct). The short glide was 500 ms and traversed a frequency range of 871.0–1148.2 Hz (0.40 oct). The gliding tones in each stimulus type had a rise and a fall time of 20 ms, with cosine-shaped amplitude-envelope ramps. In cases where one of the glides contained a silent gap, the rise and the fall time to delimit the gap were also 20 ms, with cosine-shaped ramps, which

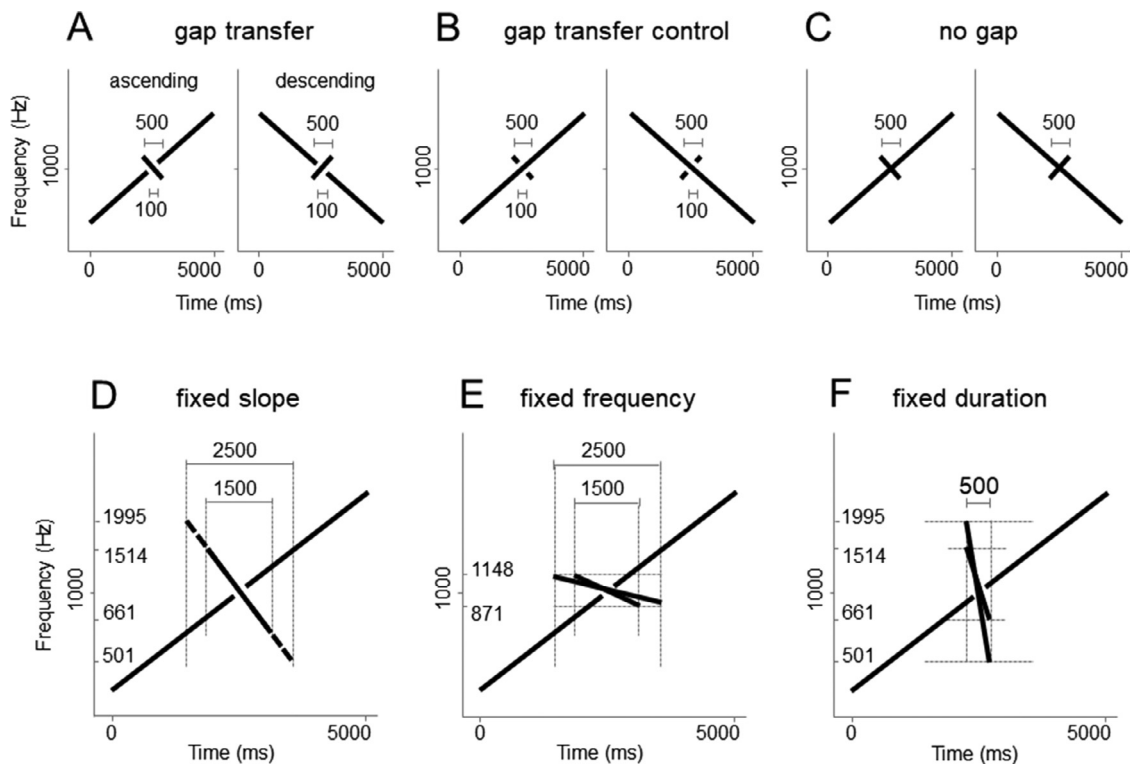


Fig. 1. Schematic impression of the stimuli used in Experiment 1. (A): the *gap-transfer* stimuli with a long ascending glide (left) and with a long descending glide (right), (B): the *gap-transfer control* stimuli with a long ascending glide (left) and with a long descending glide (right), (C): the *no-gap* stimuli with a long ascending glide (left) and with a long descending glide (right), (D): the *fixed-slope* gap-transfer stimuli depicted with a long ascending glide (the mirror-image stimulus with a long descending glide, the fixed-slope gap-transfer control stimuli, and the fixed-slope no-gap stimuli are not depicted), (E): the *fixed-frequency* gap-transfer stimulus depicted with a long ascending glide (the mirror-image stimulus with a long descending glide, the fixed-frequency gap-transfer control stimuli, and the fixed-frequency no-gap stimuli are not depicted), and (F): the *fixed-duration* gap-transfer stimulus depicted with a long ascending glide (the mirror-image stimulus with a long descending glide, the fixed-duration gap-transfer control stimuli, and the fixed-duration no-gap stimuli are not depicted).

were not included in the gap duration. When both glides were continuous, they shared the same phase at the crossing point (see also Nakajima et al., 2000; Remijn et al., 2007).

(1) A gap-transfer stimulus (Fig. 1A) included a long glide of 5000 ms with a 100-ms gap in its temporal middle. The long glide crossed with a short glide of 500 ms, so that both glides crossed at their shared temporal midpoint, and the short glide thus moved through the gap in the long glide. The long glide was either ascending with a descending short glide, or vice versa. Although earlier research on gap-transfer stimuli had not revealed any particular effect of glide direction on perceptual content (e.g., Nakajima et al., 2000), we included the ascending and descending conditions here in order to balance the conditions.

(2) A gap-transfer control stimulus (Fig. 1B) also included a 5000-ms long glide crossing with a 500-ms short glide. Here, though, a 100-ms gap was not in the long glide but in the short glide. In one version, the long glide was ascending with a descending short glide, and in the other, vice versa. (3) In a no-gap stimulus, a continuous 5000-ms glide crossed with a continuous 500-ms glide (Fig. 1C). The two glides crossed each other at the temporal midpoint at the same phase. Here too, the long glide was either ascending with a descending short glide, or vice versa.

These basic stimuli were expanded in three different ways as follows. As mentioned, we had hypothesized that proximity between stimulus edges of physically different glides was important for the perceptual organization that led to the gap transfer illusion. Our purpose was to investigate how the integration of onsets and offsets mainly would concern proximity in time, proximity in frequency, or both. Proximity in both time and frequency could be investigated with *fixed-slope* stimuli (Fig. 1D). In the fixed-

slope stimuli, the long and the short glide crossed each other with the same slope of $0.24 \text{ }^1/\text{s}$ on a common logarithmic scale ($\sim 0.80 \text{ oct/s}$). The duration of the short glide in the expanded stimuli was either 1500 or 2500 ms. When the short glide was 1500 ms, it traversed a frequency range of 660.7–1513.6 Hz ($\sim 1.20 \text{ oct}$). The 2500-ms short glide traversed 501.2–1995.3 Hz ($\sim 1.99 \text{ oct}$). Twelve (expanded) fixed-slope stimuli were made (2 short-glide durations \times 3 basic stimulus types \times ascending/descending direction of the long glide). Including the basic stimuli, of which the short glide had the same fixed slope, the short glide duration was thus varied in 3 steps: 500, 1500, and 2500 ms.

In order to investigate the role of proximity exclusively in time between stimulus edges in auditory event organization, *fixed-frequency* stimuli were used (Fig. 1E). In the fixed-frequency stimuli, the duration of the short glide was also either 1500 or 2500 ms. Here, though, the short glide traversed the same frequency range as in the basic stimuli, where the 500-ms glide moved from 871.0 to 1148.2 Hz ($\sim 0.40 \text{ oct}$). The short glide spanned this frequency range both when it was 1500 ms and when it was 2500 ms. The slope of the short glide in the expanded stimuli thus became shallower, and was $\sim 0.27 \text{ oct/s}$ when it was 1500 ms, and $\sim 0.16 \text{ oct/s}$ when it was 2500 ms. Twelve (expanded) fixed-frequency stimuli were made (2 short-glide durations \times 3 basic stimulus types \times ascending/descending direction of the long glide). The short glide duration was varied in 3 steps including the basic stimuli.

Finally, in order to investigate the role of proximity exclusively in frequency between stimulus edges in auditory event organization, *fixed-duration* stimuli were used (Fig. 1F). In these stimuli, the duration of the short glide was always 500 ms. It moved, though,

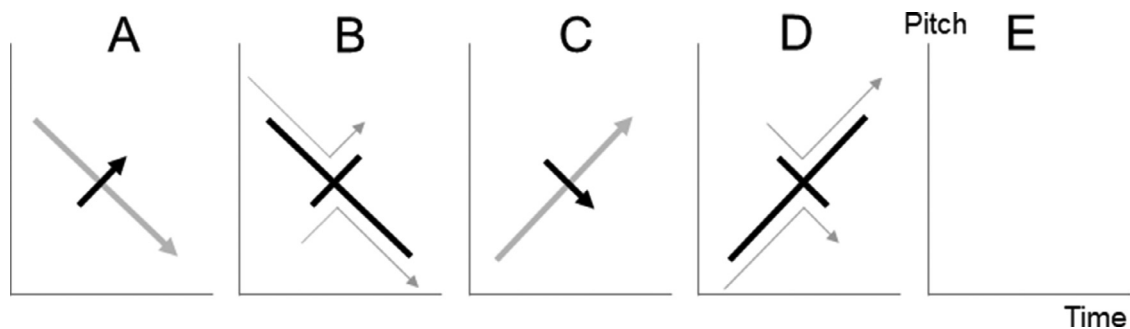


Fig. 2. Alternatives of percept patterns. These illustrations for 1) crossing (A and C), 2) bouncing (B and D), or 3) else (E) were printed on each response sheet with explanations in Japanese. The 'else' panel also worked to show what the horizontal and the vertical axis meant.

through the same frequency ranges as in the fixed-slope stimuli. When the short glide moved from 660.7 to 1513.6 Hz (or vice versa), its slope was ~ 2.39 oct/s. When the glide moved from 501.2 to 1995.3 Hz (or vice versa) its slope was ~ 3.99 oct/s. Twelve (expanded) fixed-duration stimuli were made (2 short-glide frequency ranges \times 3 basic stimulus types \times ascending/descending direction of the long glide). Including the basic stimuli, the frequency range of the short glide was varied in 3 steps.

The stimuli were presented to the participant at 70 dB SPL, referenced from an ongoing 1000-Hz pure tone of the same amplitude as that of each glide component. The presentation levels were measured with a precision sound level meter (Naganokeiki 2071), mounted with an artificial ear (Brüel & Kjær 4231).

2.3. Equipment

The participants sat in front of a computer screen in the sound-attenuating booth. The stimuli were presented monaurally to the participant by means of a computer (Frontier KZFM71/N) and an audio processor (ONKYO SE-U55GX). The generated sound signals passed through a low-pass filter (NFDV-04 DV8FL; cut-off frequency = 15,000 Hz), a graphic equalizer (Roland RDQ-2031), and a headphone amplifier (STAX SRM-313), before the stimuli reached the participant through headphones (STAX SR-303). The audio processor was used as a DA converter, and the equalizer was used to maintain flat-shaped frequency characteristics of the whole audio system. The low-pass filter was used for anti-aliasing.

2.4. Procedure

The first task of the participants was to indicate the type of percept they heard for each stimulus pattern: (1) crossing, (2) bouncing, or (3) else (Fig. 2). They then rated the continuity/discontinuity of the pitch trajectories in the percept (corresponding to one of the arrows in the selected panel in Fig. 2 except in panel E) on a 4-point scale, in which the two extremes were 'continuous' and 'discontinuous'. They were allowed to listen to each pattern repeatedly until they were sure about their judgment. This procedure was introduced in order to relate the present data directly to our previous data (Nakajima et al., 2000; Kanafuka et al., 2007), in which we first collected phenomenological reports of the participants utilizing new auditory stimulus patterns. This was necessary to prove the existence of a new illusion itself first without implying how the illusion could appear in the experimental procedure or instructions. We then conducted psychophysical experiments utilizing 4-point scales, because it turned out that 3-point or 4-point scales were the only possible alternatives to be directly connected to the contents of the phenomenological reports. Since our aim was to determine in which conditions the new illusion could occur, the participants were allowed to

listen to each pattern as many times as they wanted, as is necessary to separate different perceptual categories in phenomenological experiments in general. As an indication of how many times a given stimulus was heard, the data from Experiment 3, which had the largest number of stimuli, showed that in 97% of the trials the participants listened to the stimuli just up to 5 times. This implies that the participants' perceptual impressions stabilized after a few presentations of the same stimulus. Earlier research had shown that the gap transfer illusion occurs in a very stable manner over multiple presentations for each listener, for example, enabling pitch- and duration-matching of the illusory short tones (Kanafuka et al., 2007). The illusory gap-transfer tones could also affect visual apparent motion stimuli, as if they were real tones (Wang et al., 2015).

The participants heard each stimulus two seconds after they clicked a "Play" button on the computer screen. For each stimulus trial, the participants received a sheet of paper showing five panels to choose between (Fig. 2). Crossing sounds (Fig. 2A and C) were depicted (1) by a long ascending arrow crossing a shorter descending arrow at the arrows' shared geometrical middle, or (2) by a long descending arrow crossing a short ascending arrow at the arrows' geometrical middle. Bouncing sounds (Fig. 2B and D) were schematically depicted by two lines that resembled a checkmark (\checkmark) or its mirror image touching an inverted checkmark, i.e., the original or mirrored checkmark, rotated 180° around the bottom. In one bouncing pattern, (3) the long, aligned ends of the checkmarks were ascending, while in the other, (4) the long ends were descending. For each trial, the participants were asked to indicate what they had heard by marking one of the four panels by pencil. If the participants' percept differed from what was indicated on any of these panels, their task was to draw the percept by pencil in an "else" panel. For each stimulus, a new response sheet was prepared. Different order of the four panels was used for each block; 16 different response sheets were randomly selected from the 24 variations that could be made by arranging the four panel patterns.

After the participants indicated the perceived pattern by marking one of the four panels (in most cases), they judged the perceptual (dis)continuity of the pitch trajectories in the percept, by marking 4-point rating scales depicted below each panel. A rating scale was provided for each of the longer and the shorter trajectory, or for each of the checkmark trajectories. After the participants finished judging for a stimulus, they clicked on a 'Next' button on the computer screen to listen to the next stimulus.

3. Experiment 1

The purpose of Experiment 1 was to investigate how differences in the frequency ranges and in the temporal ranges of the glides in the gap-transfer stimulus affect the occurrence of the gap transfer illusion, for which a crossing percept is a requisite. Besides the

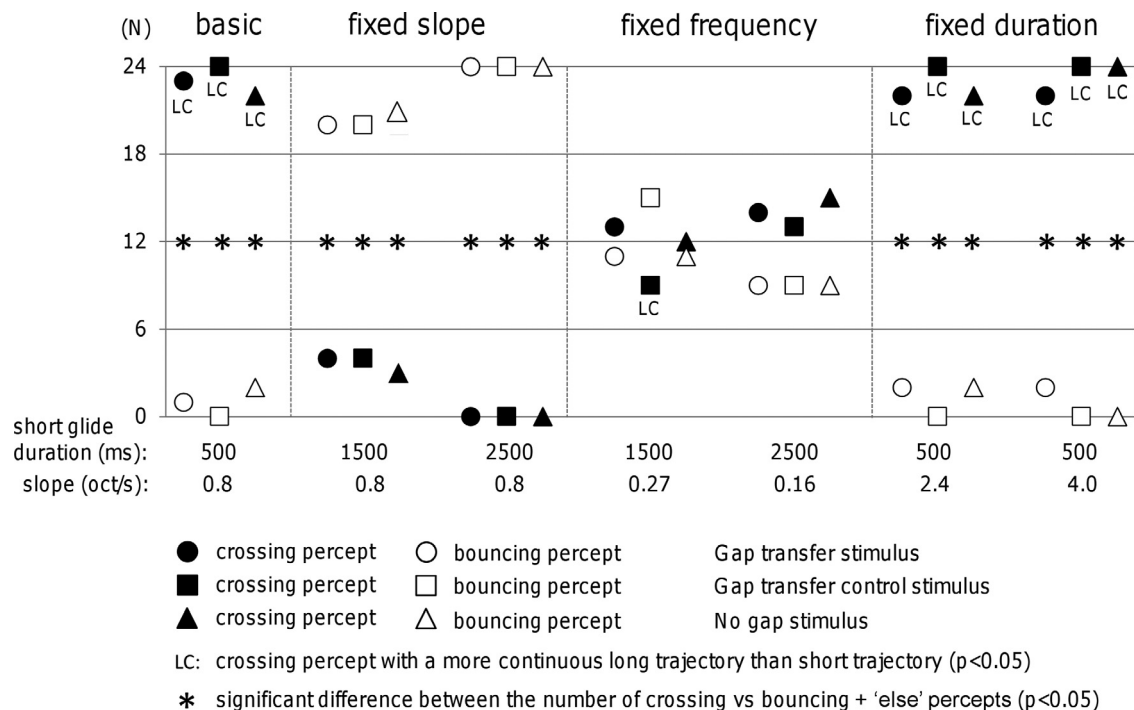


Fig. 3. The results of Experiment 1 for the gap-transfer stimuli, the gap-transfer control stimuli, and the no-gap stimuli. Black symbols indicate the crossing percept, and white symbols the bouncing percept; the frequencies of these percepts are compared on the y-axis. The asterisks show cases where crossing was significantly more frequently perceived than bouncing + 'else' percepts ($p < 0.05$). 'LC' near a black (crossing-percept) symbol indicates that the longer pitch trajectory in the percept was heard significantly as more continuous than the short trajectory ($p < 0.05$).

gap-transfer stimulus, two types of control stimuli were employed as described in General Method (Section 2).

3.1. Stimuli and procedure

Combining (1) the 6 basic stimuli with (2) the 12 fixed-slope stimuli, (3) the 12 fixed-frequency stimuli, and (4) the 12 fixed-duration stimuli made a total of 42 stimuli. After receiving instructions, the participants first listened to all of the 42 stimulus patterns once, in random order. They then performed a training session of 16 trials, randomly selected from the 42 stimuli. The purpose of the training session was to familiarize the participants with the type of stimuli and the task utilizing the rating scales. No feedback to their responses was given, however, since we did not want to bias their responses. The actual experiment consisted of 3 blocks of 14 stimulus patterns, randomly assigned from the 42 stimulus patterns. Each block started with two additional warm-up trials, for which the stimulus patterns were the same as for the last two trials in the block.

4. Results

Since our immediate purpose was to construct a theoretical framework to understand how gestalt principles could work in the time-frequency domain, we were obliged to deal with qualitative rather than quantitative natures of auditory percepts, and this led us to obtain and analyze data based on rudimentary non-parametric statistics. This should be a necessary step to clarify the mechanism of an auditory illusion taking place robustly in an unbelieveably simple context.

The results of Experiment 1 are depicted in Fig. 3. For the 504 perceptual impressions in total (42 stimulus patterns each judged by 12 listeners), only on four occasions, for different stimuli, did

the listeners report a different percept than a crossing or bouncing one.

For each stimulus, we counted the number of crossing percepts and the number of alternative percepts, i.e., bouncing percepts and 'else' percepts. We then performed sign tests (two-tailed, $p < 0.05$) to see whether crossing percepts were heard more/less frequently in stimuli with a long ascending glide than in otherwise the same stimuli with a long descending glide. The same was done to see whether the number of bouncing percepts was influenced by the movement direction of the long glide. The tests over crossing percepts in the 21 stimulus pairs ($N^+ = 7$, $N^- = 4$, ties = 10, $p = 0.549$) and over bouncing percepts in the same 21 stimulus pairs ($N^+ = 4$, $N^- = 7$, ties = 10, $p = 0.549$) showed that the selected percept type did not vary significantly with longer-glide or longer-part direction (ascending or descending). For further analyses, data for both glide directions were combined and results were analyzed with $n = 24$ (with a null hypothesis that two potential percepts could occur with the same probability).

To see whether any of the stimuli facilitated the crossing percept or an alternative percept (bouncing or 'else'), two-tailed sign tests were performed between "crossing percepts" and "bouncing + else percepts" ($n = 24$). The basic gap-transfer stimuli (Fig. 1A) was almost always perceived as crossing ($N^+ = 23$, $N^- = 1$; $p < 0.001$). The basic gap-transfer control stimulus (Fig. 1B) was always perceived as crossing ($N^+ = 24$, $N^- = 0$; $p < 0.001$), and the basic no-gap stimulus (Fig. 1C) dominantly as crossing ($N^+ = 22$, $N^- = 2$; $p < 0.001$); in all these basic stimuli, the long glide was judged as more continuous than the short glide ($p < 0.01$).

The expanded fixed-slope stimuli (Fig. 1D) were more often perceived as bouncing than as crossing. Both in the gap-transfer stimulus and in the gap-transfer control stimulus, bouncing was perceived in 20 cases ($p < 0.01$) when the short glide was 1500 ms, and in 24 cases ($p < 0.001$) when the short glide was 2500 ms. The no-gap stimulus was perceived as bouncing in 21 cases

($p < 0.001$) and 24 cases ($p < 0.001$), respectively. No significant difference in perceived (dis)continuity was found between the higher (checkmark-shaped) and the lower (inverted checkmark-shaped) pitch trajectory in these bouncing percepts.

The expanded fixed-frequency stimuli with the 1500-ms or the 2500-ms short glide (Fig. 1E) did not show a significantly dominant percept. Bouncing or crossing percepts occurred roughly in similar proportions. If bouncing percepts appeared, the (dis)continuity judgments of the higher (checkmark-shaped) and the lower (inverted checkmark-shaped) pitch trajectories did not significantly differ. As for crossing percepts, in the 9 (out of the 24) cases in which the gap-transfer control stimulus with a short glide of 1500 ms was heard as crossing, the long glide trajectory was significantly more continuous than the short one ($p < 0.01$). No other such difference in (dis)continuity between crossing glide trajectories was observed.

The expanded fixed-duration stimuli (Fig. 1F), in which a 500-ms short glide with a steeper slope crossed the long glide, were robustly perceived as crossing regardless of stimulus type. In the gap-transfer stimulus, crossing was dominant in both slope conditions ($N^+ = 22$, $N^- = 2$; $p < 0.001$), with a long glide significantly more continuous than the short glide ($p < 0.01$). In the gap-transfer control stimulus, crossing was dominant as well in both slope conditions ($N^+ = 24$, $N^- = 0$; $p < 0.001$), with a long glide significantly more continuous than the short glide ($p < 0.001$). Both the no-gap stimuli with the slope of 2.39 oct/s and with the slope of 3.99 oct/s were heard as crossing too ($N^+ = 22$, $N^- = 2$; $p < 0.001$; and $N^+ = 24$, $N^- = 0$; $p < 0.001$; respectively), again with a significantly more continuous longer glide ($p < 0.01$ and $p < 0.001$, respectively).

5. Discussion

The results of Experiment 1 (Fig. 3) confirmed the occurrence of the gap transfer illusion in the basic gap-transfer stimuli, in which the 5000-ms glide with a gap crossed with the continuous 500-ms glide. Although only the long glide physically contained the gap, it was consistently perceived as more continuous than the short glide. The basic gap-transfer control stimuli were perceived veridically, keeping the long glide more continuous. Also the basic no-gap stimuli showed a significantly more continuous long pitch trajectory as compared with the short trajectory, with which it crossed. Nakajima et al. (2000), utilizing similar no-gap stimuli, found this tendency, and indicated that the perceptual discontinuity in such cases could be attributed to the short acoustic gap(s) observed in the acoustic beats around the crossing point.

The fact that bouncing percepts were dominant only in the expanded fixed-slope stimuli suggests that, for bouncing to occur stably in stimuli with physically crossing glides, both glides need to be not too different from each other in length as well as in slope. Analogous to stream-bounce perception in vision (Burns and Zanker, 2000), a certain degree of similarity in sensory information thus may facilitate bouncing in audition as well. The present results, including those for the gap-transfer stimuli, follow earlier findings on auditory bouncing (Tougas and Bregman, 1985; McPherson et al., 1994).

The increase in bouncing percepts (or in rare cases including an 'else' percept) for stimuli with a relatively longer short glide may be also due to the following. As established here and earlier (Nakajima et al., 2000; Kanafuka et al., 2007; Kuroda et al., 2010), the gap is allocated to the short pitch trajectory in a typical gap-transfer stimulus with a short and a long glide. The idea has been that this occurs because the onset of the short glide and the offset of the long glide before the gap are in close proximity in frequency and in time. Their perceptual connection makes

a coherent entity—the first short tone. The second short tone in the percept results from the perceptual connection of the onset of the long glide component after the gap and the offset of the short glide. These sound edges are close in frequency and in time as well, and therefore are likely to be perceptually integrated into another short tone. Thus the gap transfer illusion occurs, and crossing inevitably takes place. If the short glide is not short enough in this case, however, the proximity is lost, and such a process to cause inevitable crossing does not take place. This should favor bouncing.

No percept became dominant in the expanded fixed-frequency stimuli, in which the short glide was either 1500 or 2500 ms, while traversing the same frequency range as in the basic 500-ms condition. As compared to the percepts for basic stimuli with a short glide of 500 ms, the amount of bouncing percepts increased for short-glide durations of 1500 and 2500 ms. However, in any of these conditions, bouncing did not become the dominant percept. One plausible reason for this is that the temporal proximity between onsets and offsets was lost, and crossing was not facilitated, whereas the slope difference between glides made their similarity weaker, and bouncing was not facilitated, either.

In all fixed-duration stimuli, including the corresponding basic stimuli, crossing was the dominant percept. Here the short glide was fixed at 500 ms, yet even when it traversed a relatively wide frequency range, bouncing was not observed. This leads to the idea that crossing is always facilitated by the temporal proximity between onsets and offsets, and that bouncing is facilitated only when both glides have similar slopes and are not too different in duration.

With regard to the edge-integration account for the gap transfer illusion, from the fixed-duration stimuli, we see that even when the frequency separation between the onsets and offsets increased markedly, as in the 2.39- and 3.99-oct/s conditions, crossing and gap transfer are perceived stably as long as the onsets and offsets are in close temporal proximity. The present data indicate that temporal proximity between sound edges causes perceptual integration of the edges of physically different sounds, but that frequency proximity does not. If we assume that the temporal proximity of onsets and offsets is solely important to cause the gap transfer illusion, the fact that the illusion stably took place in the basic stimuli and in the fixed-duration stimuli – and not in others – can be explained.

Strong temporal proximity as described above should make the percepts of two short tones and one long continuous glide compulsory, making a crossing percept inevitable. The very stable crossing percept in the fixed-duration stimuli, including the basic stimulus, is likely to be caused by such a mechanism. Additionally, it is possible that the proximity between the onset and the offset of the shorter glide, for example, works in the same way in no-gap stimuli. The effect of temporal proximity between stimulus edges turned out to be important, and was further investigated in Experiment 2.

6. Experiment 2

We employed only fixed-slope stimuli of the three basic stimulus types as in Experiment 1. We increased the overlap duration of the short glide asymmetrically, and thus lengthened the short glide duration only before, or only after the glides' crossing point. We thus examined whether the increase in overlap duration would favor bouncing percepts; if one side of the shorter glide was lengthened, it would be perceptually similar to the temporally opposite side of the longer glide located in the same frequency range. Another purpose of this experiment was to examine whether the temporal proximity of an onset and an offset facilitates them to be connected perceptually also in these asymmetric stimuli.

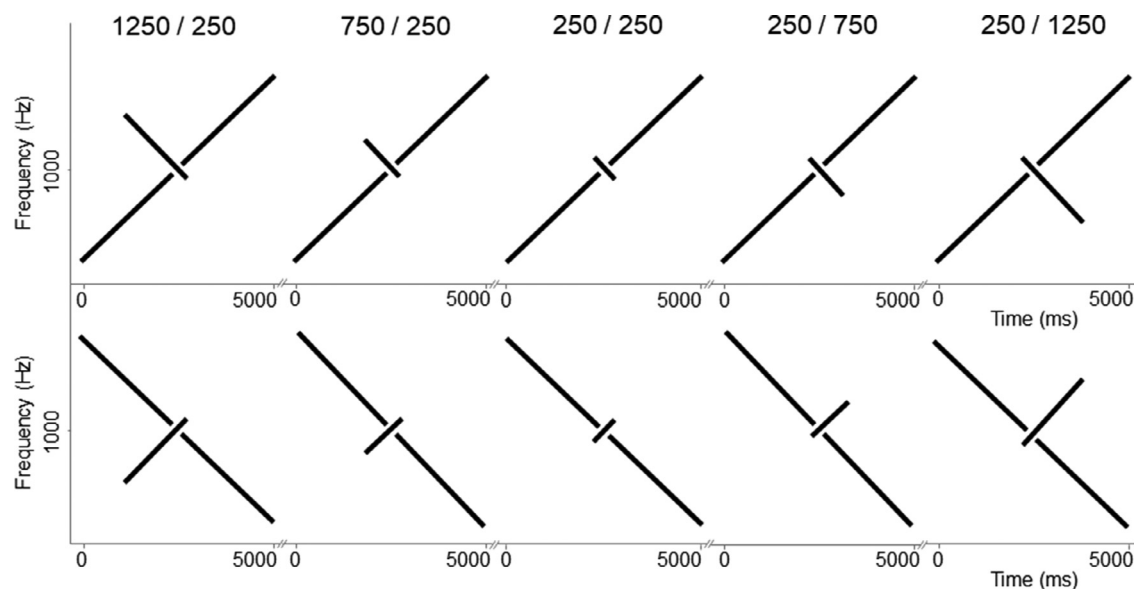


Fig. 4. Schematic impression of a subset of the stimuli used in Experiment 2. Gap-transfer stimuli are indicated with a long ascending glide overlapping with a short descending glide in 5 variations in overlap configuration (indicated in milliseconds at the top). The mirror-image gap-transfer stimuli with a long descending glide, the gap-transfer control stimuli, and the no-gap stimuli are not indicated.

6.1. Stimuli and procedure

The 30 stimuli used in Experiment 2 were all fixed-slope stimuli comprising the same three stimulus types as used in Experiment 1, along with the same two variations in glide direction. We thus employed gap-transfer stimuli (indicated in Fig. 4 representatively), gap-transfer control stimuli, and no-gap stimuli, in each case with an ascending long glide and a descending short glide, or vice versa. These six subtypes of stimuli were made with five short-glide overlap configurations. In the first configuration, the short glide was 1500 ms, and it overlapped the long glide for 1250 ms (or 1200 ms when there was a 100-ms gap) before the glides' crossing point, and 250 ms (or 200 ms) after it. In the second configuration, the short glide was 1000 ms, and it overlapped the long glide for 750 ms (or 700 ms) before the crossing point, and 250 ms (or 200 ms) after it. In the third configuration, the short glide was 500 ms, and crossed the long glide in its temporal middle with a 250-ms (or 200-ms) overlap on both sides of the crossing point. In the fourth configuration, the short glide was 1000 ms overlapping the long glide before and after the crossing point for 250 (or 200 ms) and 750 ms (or 700 ms), respectively, while in the fifth configuration the short glide was 1500 ms, overlapping the long glide for 250 (or 200 ms) and 1250 ms (or 1200 ms), respectively, before and after the crossing point. From here on, we refer to these five overlap configurations as the 1250/250-, 750/250-, 250/250-, 250/750-, and 250/1250-ms conditions.

The participants first listened to the 30 stimulus patterns once in random order. After that, the stimuli were presented to the participants in three blocks of 10 stimuli each, with two warm-up trials added to each block.

7. Results

The results of Experiment 2 are depicted in Fig. 5. Out of 390 perceptual impressions (30 stimuli judged by 13 listeners), 35 were 'else' percepts. As in Experiment 1, stimuli with a long ascending glide and with a long descending glide were combined for each stimulus type and each overlap configuration, and a sign test

($p < 0.05$) was performed to see whether the crossing percept or an alternative percept (bouncing + 'else') was dominant. In the 250/250-ms condition, the gap-transfer stimuli (18 out of 26 judgments, $p = 0.08$) were not significantly more perceived as crossing than as bouncing + 'else'. The gap-transfer control stimuli (20 out of 26 judgments, $p < 0.01$) were significantly more perceived as crossing than as bouncing. In the no-gap stimuli, the crossing percept was perceived in 18 out of 26 judgments ($p = 0.08$).

Similar to the analysis of Experiment 1, sign-tests were performed to see whether continuity differences existed between the long and the short pitch trajectory in crossing percepts or between the high and the low pitch trajectory in bouncing percepts. The results showed that when the gap-transfer stimuli and the gap-transfer control stimuli were perceived as crossing, listeners heard the long trajectory in the percept as significantly ($p < 0.05$) more continuous than the short trajectory, for every overlap configuration. When the no-gap stimulus was perceived as crossing, only the stimulus with the 250/250-ms overlap yielded a significantly better continuity of the long trajectory ($p < 0.01$). No significant difference in perceived continuity between the checkmark-shaped pitch trajectories in bouncing percepts was observed.

8. Discussion

The gap-transfer stimuli and the gap-transfer control stimuli with the shortest, symmetrical overlap were typically perceived as consisting of a continuous long pitch trajectory accompanied by a shorter pitch trajectory with a gap: For both stimulus types, this crossing percept appeared more frequently than the bouncing or 'else' percept. None of the stimuli with an asymmetrical overlap gave rise to a clearly dominant percept, although the number of bouncing percepts increased as the overlap duration in these stimuli became longer. Lengthening just one side, i.e., not lengthening the other side, of the overlap thus seemed not to facilitate perceptual organization into bouncing percepts. Whereas in Experiment 1 bouncing was significantly more often perceived than crossing in the gap-transfer stimuli, and in the gap-transfer control stimuli with the symmetrical 1500-ms overlap, the comparable stimuli shaped here with an asymmetric 1500-ms overlap did not show

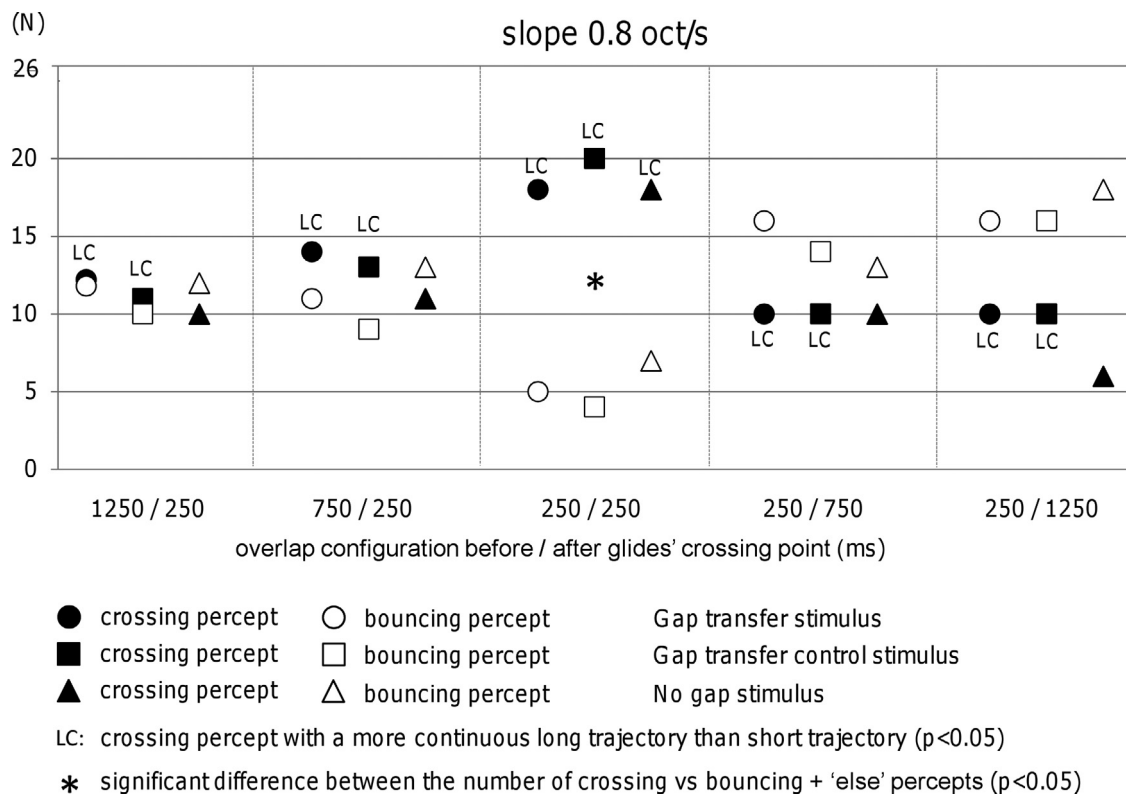


Fig. 5. The results of Experiment 2 for the gap-transfer stimuli, the gap-transfer control stimuli, and the no-gap stimuli. Black symbols indicate the crossing percept, and white symbols the bouncing percept. The asterisks show cases where crossing was significantly more frequently perceived than bouncing + 'else' percepts ($p < 0.05$). 'LC' near a black (crossing-percept) symbol indicates that the longer pitch trajectory in the percept was heard significantly as more continuous than the short trajectory ($p < 0.05$).

significantly more bouncing percepts. It will be a critical step to clarify what hampers bouncing percepts in such asymmetric stimulus patterns.

The most important issue to be noticed is that, once a crossing percept took place, the longer glide tended to be perceived as more continuous than the shorter glide (as indicated by "LC" in Fig. 5). This was the case for all gap-transfer stimuli. Even when crossing percepts were not dominant, the gap transfer illusion, in a broader sense, took place for crossing percepts. This is in line with our explanation of the illusion, in which the temporal proximity between the onset of a glide and the offset of another glide is considered to construct an illusory tone between them (with supporting data on the natures of the illusory tones in Kanafuka et al., 2007).

In the stimuli with the asymmetric overlap, one overlapping part is notably short in absolute terms (i.e., 250 or 200 ms). The sound edges delimiting the notably short part fall within a temporal window of "compulsory" auditory integration (for a discussion see Yabe et al., 1997; Remijn and Nakajima, 2005), facilitating the perception of a short tone either before or after the gap. If a crossing percept takes place in the present paradigm, this should cause the gap transfer illusion, and, if a bouncing percept takes place, this should become a separate part of a bouncing trajectory.

9. Experiment 3

The stimuli as used in Experiment 2 were employed with three different slope configurations. We wished to examine the results of Experiment 1, which showed that a difference in slope between the two crossing glides facilitated a crossing perceptual organization, and that the temporal proximity between an onset and an offset was essential to connect them illusorily.

9.1. Stimuli and procedure

Stimuli with the same overlap configurations as used in Experiment 2 were employed, each with three variations in the slope of the short glide. The long glide in each stimulus moved with a speed of ~ 0.80 oct/s, while the short glide was either ~ 0.27 oct/s, ~ 0.80 oct/s, or ~ 2.39 oct/s (0.08 $1/s$, 0.24 $1/s$, or 0.72 $1/s$, exactly, on a common logarithmic scale). In total, 90 stimuli were employed (3 stimulus types \times 2 long-glide directions \times 5 overlap configurations \times 3 slope configurations).

The participants first listened to the 90 stimulus patterns once in random order, and they performed a training session of 10 trials in order to get familiar with the stimuli and the task. No feedback to their responses was given. After that, the stimuli were presented to the participants in 5 blocks of 18 stimuli each, with two warm-up trials added to each block.

10. Results

From the 720 judgments in total (90 stimuli judged by 8 participants), in all but 3 cases a crossing or a bouncing percept was heard. Two-tailed sign tests ($p < 0.05$) were performed over the combined ascending and descending conditions to test for dominance basically between these percepts. The results are shown in Fig. 6, in which the asterisks show cases where crossing was significantly more frequently perceived than bouncing and 'else' percepts (combined). The indication 'LC' near a black symbol indicates that, within a crossing percept, the long pitch trajectory was judged as significantly more continuous than the short pitch trajectory ($p < 0.05$).

The results obtained with the 250/250-ms configuration (with the shortest overlaps) were as described below. When the slope of

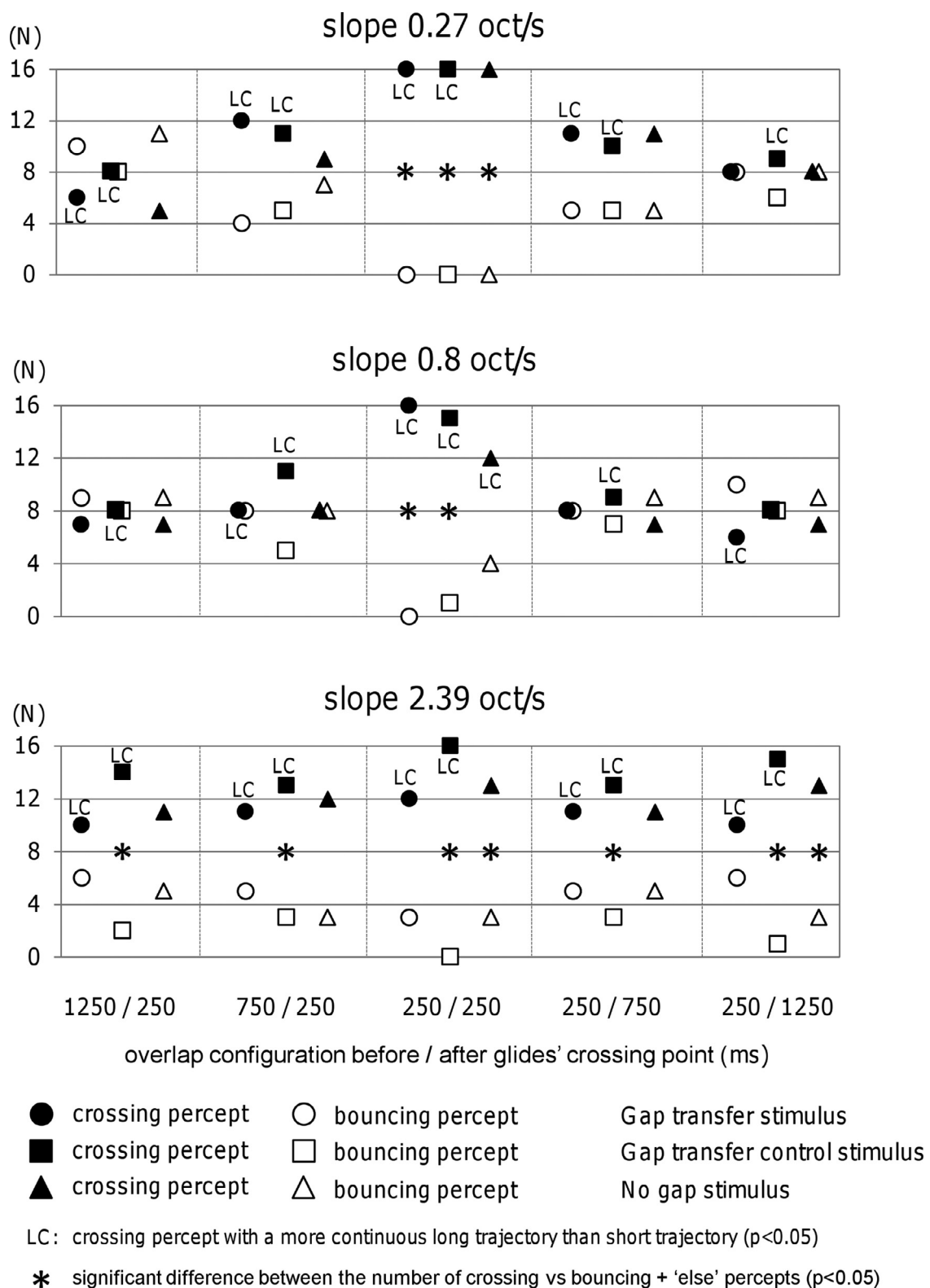


Fig. 6. The results of Experiment 3 for the gap-transfer stimuli, the gap-transfer control stimuli, and the no-gap stimuli. Black symbols indicate the crossing percept, and white symbols the bouncing percept. The 'asterisks' show cases where crossing was significantly more frequently perceived than bouncing + 'else' percepts ($p < 0.05$). 'LC' near a black (crossing-percept) symbol indicates that the longer pitch trajectory in the percept was heard significantly as more continuous than the short trajectory ($p < 0.05$).

the short glide was 0.27 oct/s, significantly more crossing percepts than bouncing and 'else' percepts were heard in all stimuli. This was also the case in the 0.80-oct/s condition but only in the gap-transfer stimuli and in the gap-transfer control stimuli. When the short glide moved with 2.39 oct/s, the dominance of crossing percepts occurred as well, though not significantly for the gap-transfer stimulus. In all gap-transfer stimuli with the 250/250-ms configuration, the crossing percepts were generally dominant, and the long trajectory was more continuous: The gap transfer illusion appeared in a typical manner.

Also in many gap-transfer stimuli in which the crossing percepts were not significantly dominant, the longer trajectory in the crossing percept was significantly more continuous: The gap transfer illusion often appeared in a weaker manner. When the slope of the short glide was the steepest, 2.39 oct/s, the gap transfer illusion including such weaker cases took place in all gap-transfer stimuli (Fig. 6).

Since the amount of bouncing percepts reported by the participants was not small, we had an opportunity to have a look at the (dis)continuity of bouncing pitch trajectories too. Overall, 84 bouncing percepts were reported in gap-transfer stimuli (see the white circles in Fig. 6) with an unequal overlap configuration (i.e., 1250/250 ms, 750/250 ms, 250/750 ms, and 250/1250 ms), thus with either a shorter upper trajectory or a shorter lower trajectory ("shorter" means 250/2500 ms or 2500/250 ms compared, for example, with 2500/750 ms or 1250/2500 ms). (Perfect or partial) discontinuity was perceived in 78 among the 84 bouncing cases: only in the longer trajectory in 9 cases, only in the shorter trajectory in 62 cases, and in both trajectories in 7 cases. Typically (in 71 among all 84 bouncing cases) discontinuity appeared in one of the two trajectories, but not in both, and mostly (in 62 cases) in the shorter trajectory associated with the shortest (250- or 200-ms) overlap.

11. Discussion

Fig. 6 shows that the stimuli with a relatively steep short glide (2.39-oct/s condition), against the long glide, facilitated crossing better than a stimulus with a less pronounced slope difference between the long and the short glide. In terms of the current explanation for the gap transfer illusion, the issue is important since the slope influences the frequency proximity between the sound edges that are potentially supposed to be integrated into short, illusory tones. When the short glide is relatively shallow, its onset frequency is closer to the offset frequency of the first long glide component before the gap. A steep short glide has an onset frequency that is farther away in frequency from the offset frequency of the first long glide component. Because frequency proximity between sound edges improves with a shallower slope, one might expect more compulsory gap transfer to occur in shallow-sloped glide stimuli. However, the results of Experiment 1 (the fixed-duration stimuli) and Experiment 3 indicate that stimuli with a relatively steep short glide give rise to robust gap transfer, in the broader sense, despite the relatively weak frequency proximity between the relevant sound edges.

One plausible explanation for this robust gap transfer is that fast-moving glides are perceptually less coherent. The edge-integration account for the gap transfer illusion maintains that sound edges from physically different sounds break off from their carrier sound and connect to one another. If the slope of a glide is shallow, the glide as a whole makes a relatively strong perceptual entity making its edges less prone to segregation. Whatever explanation may be given, it has become clear that the frequency proximity between an onset and an offset does not play an important role in the auditory organization related to the gap transfer illusion.

In our earlier research, the effect of slope on the perception of two partly overlapping gliding tones has been explained in the same way. Depending on their frequency separation, two partly overlapping glide tones of the same duration can be perceived as consisting of a continuous pitch trajectory which is as long as the whole pattern, accompanied by a short tone in its temporal middle (Remijn and Nakajima, 2005). The short tone is assumed to be the result of the perceptual integration of the onset of the second glide to the offset of the first. The perception of the short tone, however, becomes less compelling when the slope of the glides becomes shallower, or zero, i.e., when the stimulus consists of two partly overlapping steady-state tones. Here too, it is feasible that the sound edges delimiting the overlap are less likely to split off of their carrier, because the carrier sounds themselves make strong perceptual entities (see Remijn et al., 2008, for detailed discussion).

12. General discussion

Three experiments were performed to further investigate factors that promote the perception of two crossing pitch trajectories in gap-transfer stimuli and related stimuli. Proceeding from the fact that temporal sound edges (i.e., onsets and offsets) have a strong neurophysiological representation (e.g., Petkov et al., 2007; Ramamurthy and Recanzone, 2017) and constitute the framework of auditory events, we hypothesized that the proximity between onsets and offsets is vital to cause the gap transfer illusion (Nakajima et al., 2000). In order to investigate this, we therefore varied the proximity (distance) systematically. In a series of conditions in Experiment 1, the proximity between onsets and offsets was varied only in time, in another series it was varied only in (logarithmic) frequency, and there was also a series in which the proximity was varied together both in time and in frequency. Inevitably, also the slope of the shorter glide was varied. The most important finding was that the proximity between onsets and offsets in time, but not the proximity in frequency, is a key factor to construct the two illusorily divided tone components in the gap transfer illusion. Another important finding, for which we may need another series of studies, was that bouncing percepts can be dominant only when the ascending and descending glides have the same slopes, and only when the shorter glide is not too short.

In order to confirm that the temporal proximity plays an important role, we started from stimuli in which the shorter glide was as short as 500 ms, and lengthened only the preceding or the succeeding half of it in Experiments 2 and 3. This sometimes made the occurrence of the crossing percepts less frequent, especially when the ascending and the descending glide had the same slope, but, once crossing percepts took place, the longer glide tended to be perceived as more continuous. In this sense, the gap transfer illusion, in a broader sense, turned out to be very robust, and the temporal proximity between onsets and offsets explains the situation well. The gap transfer illusion in this broader sense was not observed in the expanded fixed-frequency stimuli in Experiment 1, in which the temporal proximity between onsets and offsets were weak both before and after the gap.

Getting into details, Experiment 1 showed that illusory gap transfer occurred in gap-transfer stimuli in which a 5000-ms long glide crossed with a 500-ms short glide. When the short glide duration (= the overlap duration plus the gap duration) increased to 1500 or 2500 ms, typical gap transfer subsided, and the stimuli were significantly more often perceived as bouncing than as crossing. In bouncing percepts, a checkmark-shaped pitch trajectory was heard along with an inverted checkmark-shaped trajectory, roughly speaking. It seems that, when the physically crossing sounds become more similar in duration, more frequently the bouncing percept appears. This should be connected to visual streaming-bouncing (Metzger, 1934), in which similarity between

(physically crossing) objects is a prerequisite for the perception of bouncing.

It is remarkable in the obtained auditory bouncing percepts that the gap was exclusively allocated, in most cases, either to one or the other checkmark-shaped trajectory—but not to both. The results of Experiment 3 clearly indicated that the gap tends to be allocated to the shorter trajectory, but we would need more data to take up this issue systematically. A similar perceptual mechanism to the mechanism to cause the gap transfer illusion is likely to have worked as will be discussed below.

With regard to the explanation of the gap transfer illusion in terms of Auditory Grammar (Nakajima et al., 2014), the present data clearly show that the illusory coupling of sound edges occurs when they are notably close in time (200 ms = 250 – 50 ms, subtracting a half of the gap duration). When the short glide duration increases, either in both directions or in one direction, i.e., when the temporal proximity between the relevant sound edges decreases, the number of bouncing percepts increases. The temporal proximity between the onset of the short glide and the offset of the long glide component before the gap, and – likewise – the temporal proximity between the onset of the long glide component after the gap and the offset of the short glide, thus, are important for the crossing percept to occur compulsorily in the present paradigm: Once the onsets and the offsets are interpreted for auditory organization, they are not to be interpreted again, thus leaving the long trajectory continuous (Nakajima et al., 2014). The temporal gap in the long glide in a gap-transfer stimulus was almost never perceived in its physical place in spite of many stimulus variations.

Experiments 2 and 3 particularly probed whether the occurrence of the gap transfer illusion would change if stimuli had an asymmetrical overlap. In the stimuli with overlap configurations of 1250/250, 750/250, 250/750, or 250/1250 ms, the crossing percept still appeared, although not dominantly, and the gap tended to be perceived in the shorter pitch trajectory including a portion emerging from the 200-ms overlap. Experiment 3 confirmed that, when crossing was heard in stimuli with an asymmetrical overlap, the gap was typically attributed to the shorter pitch trajectory. It is remarkable that, even when bouncing was heard, the gap tended to be in the shorter pitch trajectory – the trajectory that emerged from the 200-ms overlap. These results again indicate that temporal proximity between sound edges is an essential determining factor in auditory event formation.

The results of Experiment 3 suggest that slope difference between the longer and the shorter glide facilitated crossing. As an extension of Nakajima et al.'s theoretical framework (Auditory Grammar), this indicates that auditory event formation based on sounds with relatively swift changes in frequency, which should make the *filling* cue between the onset and the offset of a sound less stable as a single perceptual unity, relies more on building a perceptual skeleton from the temporal edges of sounds. As pointed out in Kuroda et al. (2010), Bregman and Dannenbring (1977) already mentioned that frequency-modulated tones cause a special mode of perception when determining sound continuity or discontinuity. Compared to steady-state or slow-moving sounds, fast-moving sounds that traverse a large frequency range stimulate many neurons with different frequency response characteristics whose excitation persists only briefly. In determining sound (dis)continuity, for the auditory system to rely upon these brief excitation patterns may result in incoherent auditory organizations: It would be difficult to judge whether there are many cues or just one. Instead, building an auditory-event framework initially based on neural excitations patterns in response to sound onsets and offsets may be more efficient.

In fact, research on speech and music has shown that the auditory system can make use of both of the following temporal as-

pects: the temporal fine structure of sound, with short duration cues of some tens of milliseconds (30–50 Hz range; Giraud and Poeppel, 2012), along with information that falls within a larger time scale of several hundreds of milliseconds (2–5 Hz; e.g., Ding et al., 2017). The auditory system thus has the capacity to register information by means of different temporal windows, and information regarding different sound edges falling within the same time window of several hundreds of milliseconds can be subject to mandatory integration in an initial auditory-event framework as suggested earlier (e.g., Yabe et al., 1997; Remijn and Nakajima, 2005). This integration might occur even if a large frequency separation exists between the relevant sound edges as in the case of a large slope difference, such as in Experiment 3 when the short glide was 2.39 oct/s and the long glide 0.80 oct/s.

As a step to verifying our findings, we made a new auditory demonstration (Fig. 7). It was confirmed in the present series of experiments that the onset of a glide and the offset of a glide are more likely to be connected perceptually if they are closer in time. Perceptual integration of such temporal sound edges indeed occurred whether or not they belonged to the same glide component, often compulsorily if the temporal distance between the sound edges was relatively short, i.e., 200 ms. This introduced us to a new version of an auditory illusion we reported previously (Nakajima et al., 2000; Remijn et al., 2001; Remijn and Nakajima, 2005).

We made an example of such an illusory stimulus: A glide component moved from 500.0 to 1148.7 Hz taking 1200 ms, and another glide component of the same amplitude and duration moved from 957.6 to 2200.0 Hz (Fig. 7A). They both had rise and fall times of 15 ms. These components were presented in this order with an overlap of 200 ms at the temporal center, and thus comprised a stimulus pattern of 2200 ms. Typically, a long ascending tone and a short tone in the middle were perceived, thus the physical configuration of the stimulus pattern was not preserved: What we call the *split-off effect* appeared. Our argument goes that the onset of the second glide component and the offset of the first glide component, away from each other just by a short distance of 200 ms, are recoupled perceptually constructing an illusory auditory event: the short tone in the middle of the long tone.

In a new variation of this stimulus, we introduced a silent gap of 10 ms in the middle of the pattern (Fig. 7B), in both glide components. Since there were a fall time and a rise time of 15 ms before and after this gap, the effective gap duration measured as the distance between the –3-dB points was 29 ms. A typical percept of this stimulus consisted of a long ascending glide and two successive short tones in the middle. If our previous argument can be extended, the onset of the second glide component is perceptually connected to the common offset to begin the gap, and the common onset to end the gap to the offset of the first glide component. Two short auditory events are thus constructed. It is remarkable that a long ascending glide without a gap was perceived; the offset and the onset to begin and to end the gap had been interpreted already to construct the percepts of the two short tones, and thus it was not necessary to interpret them again because the gap was not long enough as a cue of a silence (see Remijn et al., 2007).

Interestingly, almost the same percept appeared even when the gap in the middle remained only in one of the two overlapping components (Fig. 7C and D). The configuration of the onsets and offsets did not change much in this case, and it seems that the silent part in the middle that appeared when the gap was put in both components did not play a substantial role perceptually to construct a temporal skeleton (see Nakajima et al., 2014, for further theoretical details; the perception of the first offset may be suppressed finally leaving two temporally adjacent short tones, but this is beyond the scope of this paper). The present argument held also when the stimulus patterns as described above were reversed

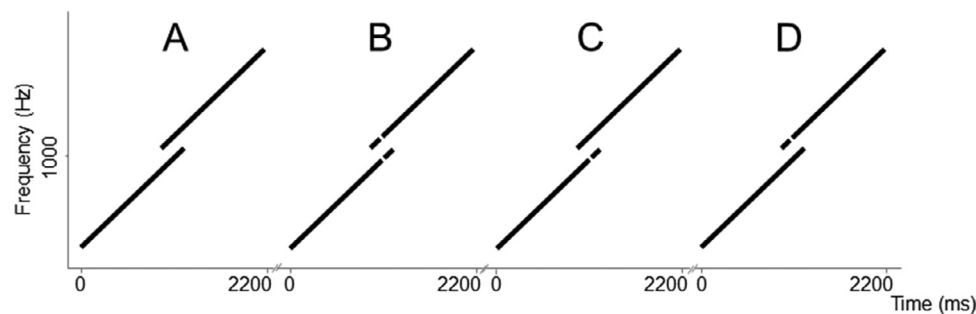


Fig. 7. Schematic impression of stimuli consisting of two partly-overlapping glide tones. In the temporal middle of the stimuli the frequency glides were either physically continuous (A) or discontinuous, with a 10-ms gap delimited by 15-ms rise and fall times in both glides (B), only in the lower glide (C), or only in the upper glide (D).

in time comprising stimulus patterns of descending glide components.

Taken together, the above-described demonstration shows that, within a certain temporal window, mandatory integration of sound edges occurs, and that stronger temporal proximity has priority. Once the edges have been assigned to an auditory event, they are not allocated to another event (unless required by the context), which in this case leads to illusory continuity of a sound with a physical gap. The brief lack of sound energy in between the temporal sound edges thus does not much affect the auditory (dis)continuity, and only the temporal sound edges play certain roles.

In the stimuli described above, as well as in the stimuli used in the present experiments, the perceived gap is often included in the perceptual organization of the most recent sound that enters the listener's ear, the short pitch trajectory. The gap is seldom allocated to the starting, ongoing long pitch trajectory in crossing percepts in the present paradigm. It is feasible that the auditory system, in order to reduce processing load, puts new sensory information into the most recently formed auditory organization, unless its spectrotemporal content fits undeniably into the “old” auditory organization (Remijn et al., 2007). This should be closely related to what Bregman (1990) calls the “old-plus-new” heuristic. Although perceptual roles of onsets in related situations have been emphasized, the roles of offsets have not yet been investigated systematically. For an auditory event to be constructed, however, an offset cue presented without an onset cue can even induce an illusory onset (Sasaki et al., 2010). If the auditory system detects a new offset, very probably it is obliged to check whether there is, or was, an onset to be connected to this offset, and we are inclined to think from the present results that the auditory system prefers to connect an onset and an offset that are close to each other, and as far as possible within the same temporal window of several hundreds of milliseconds: An offset prefers a new onset. This mechanism should work more easily if an onset and an offset are not connected strongly to the carrier sound as in a steep glide. Contrarily, a steady-state filling between an onset and an offset is likely to be preserved as a single perceptual entity. To think how onset cues and offset cues work together in time will be an important part of future research.

Author statement

We wish to explicitly mention that our study was approved by the Ethics Committee of Kyushu University, Japan, and that the procedures were followed according to the regulations established in the Helsinki Declaration.

The work described here has not been published or submitted for publication, nor is being considered – in whole or in part – for publication elsewhere, and all the authors have read the paper and have agreed to have their names listed as authors. We further

declare that this work is free from any limitations by financial or any other relationships that might lead to a conflict of interest.

Declaration of Competing Interest

None.

Acknowledgments

We thank Gert ten Hoopen, Takayuki Sasaki, Tsuyoshi Kuroda, and Kazuo Ueda for their comments on Auditory Grammar. Yuko Yamashita encouraged us in our regular seminars. This study was supported by the Japan Society for the Promotion of Science (25242002 in FYs 2013–2017 and 17H06197 in FYs 2017–2019 to YN) and the Kyushu University Center for Clinical and Translational Research (FY 2017).

References

- Benussi, V., 1917. *Psychologie Der Zeitauffassung*. Carl Winter's Universitätsbuchhandlung, Heidelberg.
- Bregman, A.S., 1990. *Auditory Scene analysis: The perceptual Organization of Sound*. MIT Press, Cambridge, MA.
- Bregman, A.S., Dannenbring, G.L., 1977. Auditory continuity and amplitude edges. *Can. J. Psychol.* 31, 151–159.
- Burns, N.R., Zanker, J.M., 2000. Streaming and bouncing: observations on motion defined objects. *Clin. Exp. Ophthalmol.* 28, 220–222.
- Ding, N., Patel, A.D., Chen, L., Butler, H., Luo, C., Poeppel, D., 2017. Temporal modulations in speech and music. *Neurosci. Biobehav. Rev.* 81, 181–187.
- Giraud, A.-L., Poeppel, D., 2012. Cortical oscillations and speech processing: emerging computational principles and operations. *Nat. Neurosci.* 15, 511–517.
- Halpern, L., 1977. *The Effect of Harmonic Ratio Relationships on Auditory Stream Segregation*. McGill University, Psychology Department Unpublished research Report.
- Handel, S., 1993. *Listening: An Introduction to the Perception of Auditory Events*. The MIT Press, Cambridge, MA.
- Iwamiya, S., Nakajima, Y., Ueda, K., Kawahara, K., Takada, M., 2003. Technical listening training: improvement of sound sensitivity for acoustic engineers and sound designers. *Acoust. Sci. Technol.* 24, 27–31.
- Kanafuka, K., Nakajima, Y., Remijn, G.B., Sasaki, T., Tanaka, S., 2007. Subjectively divided tone components in the gap transfer illusion. *Percept. Psychophys.* 69, 641–653.
- Koffka, K., 1935. *Principles of Gestalt Psychology*. Harcourt, Brace & World, New York.
- Kuroda, T., Nakajima, Y., Eguchi, S., 2010. Illusory continuity without sufficient sound energy to fill a temporal gap: examples of crossing glide tones. *J. Exp. Psychol.* 38, 1254–1267.
- Kuroda, T., Nakajima, Y., Tsunashima, S., Yasutake, T., 2009. Effects of spectra and sound pressure levels on the occurrence of the gap transfer illusion. *Perception* 38, 411–428.
- McPherson, L.M.P., Ciocca, V., Bregman, A.S., 1994. Organization in audition by similarity in rate of change: evidence from tracking individual frequency glides in mixtures. *Percept. Psychophys.* 55, 269–278.
- Metzger, W., 1934. Beobachtungen über phänomenale Identität. *Psychol. Forsch.* 19, 1–60.
- Nakajima, Y., Sasaki, T., Kanafuka, K., Miyamoto, A., Remijn, G., ten Hoopen, G., 2000. Illusory recouplings of onsets and terminations of glide tone components. *Percept. Psychophys.* 62, 1413–1425.
- Nakajima, Y., Sasaki, T., Ueda, K., Remijn, G.B., 2014. Auditory grammar. *Acoust. Aust.* 42, 97–101.

- Petkov, C.I., O'Connor, K.N., Sutter, M.L., 2007. Encoding of illusory continuity in primary auditory cortex. *Neuron* 54, 153–165.
- Ramamurthy, D.L., Recanzone, G.H., 2017. Spectral and spatial tuning of onset and offset response functions in auditory cortical fields A1 and CL of rhesus macaques. *J. Neurophysiol.* 117, 966–986.
- 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. *J. Exp. Psychol.* 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.
- Remijn, G.B., Nakajima, Y., ten Hoopen, G., 2001. Continuity perception in stimulus patterns consisting of two partly overlapping frequency glides. *J. Music Percept. Cognit.* 7, 77–92.
- Remijn, G.B., Pérez, E., Nakajima, Y., Ito, H., 2008. Frequency modulation facilitates (modal) auditory restoration of a gap. *Hear. Res.* 243, 113–120.
- Sasaki, T., Nakajima, Y., ten Hoopen, G., van Buuringen, E., Massier, B., Kojo, T., Kuroda, T., Ueda, K., 2010. Time stretching: illusory lengthening of filled auditory durations. *Attention Percept. Psychophys.* 72, 1404–1421.
- Tougas, Y., Bregman, A.S., 1985. Crossing of auditory streams. *J. Exp. Psychol.* 11, 788–798.
- Wang, Q., Guo, L., Bao, M., Chen, L., 2015. Perception of visual apparent motion is modulated by a gap within concurrent auditory glides, even when it is illusory. *Front. Psychol.* 6, 564. doi:10.3389/fpsyg.2015.00564.
- Yabe, H., Tervaniemi, M., Reinikainen, K., Näätänen, R., 1997. Temporal window of integration revealed by MMN to sound omission. *Neuroreport* 8, 1971–1974.