ILLUSORY UNIFICATION OF SIMULTANEOUS TEMPORAL GAPS OF GLIDE TONES THAT ARE APART FROM EACH OTHER IN FREQUENCY

Yoshitaka Nakajima
Perceptual Psychology Section, Graduate School of Design, Kyushu University,
815-8540 Fukuoka, Japan
nakajima@design.kyushu-u.ac.jp

Tsuyoshi Kuroda and Tomoko Uehara
Department of Acoustic Design, Kyushu University,
815-8540 Fukuoka, Japan
kuroda@gsd.design.kyushu-u.ac.jp, staccato.hop_step-jump@hotmail.co.jp

Abstract

When a temporal gap typically shorter than 50 ms in a frequency glide longer than 1 s is shared by a shorter glide, its perceptual presence in the longer glide can be weakened or lost. In the present study, we conducted an experiment to examine a possibility that this illusion, illusory gap unification, is caused by peripheral processes of the auditory system. A long glide of 3000 ms and two short glides of 300 ms shared an intensity dip of 40 ms in their temporal middles. They all ascended or descended together at a rate of 1 oct/s. The middle frequency of the long glide component was 1000 Hz, and the short glide components were 13 semitones lower and higher. Thus, the influence of peripheral interactions between these components was minimized. The illusion still took place: A majority of eight observers perceived the long glide as completely continuous and the short glides as completely or partially discontinuous.

To investigate whether and how discontinuous temporal points appear in auditory streams is important for understanding the mechanism of auditory perception and communication. *Illusory gap unification* is an illusion in which a tone with a temporal gap long enough to be detected, if in isolation, is perceived as continuous (Remijn, Nakajima, & Tanaka, 2007; Remijn et al., 2008; Figure 1). We also called this phenomenon *illusory auditory completion*. When a glide of 1.5 s or above and a shorter glide of about 400-600 ms cross each other, sharing a temporal gap typically shorter than 50 ms, the gap is often perceived only in the shorter glide. The longer glide is perceived typically as continuous, and, even when some discontinuity is perceived, it is not as clear as in the shorter glide. A similar phenomenon takes place when there is no real gap but two glides cross each other in opposite phases so that there is a common dip of intensity at the crossing point (Nakajima et al., 2000). The same kind of illusion also occurs clearly in a stimulus pattern where synthesized recorder tones of different durations, which have steady pitches, share a gap in their central positions (Nakajima, 2008b). A remarkable point about these phenomena is that there is a gap or dip of sound energy which is unfilled and long enough to be detected, and that the longer tone sharing this gap is still perceived as continuous; an illusory auditory continuity takes place without a slight portion of sound energy to bridge the gap or dip.

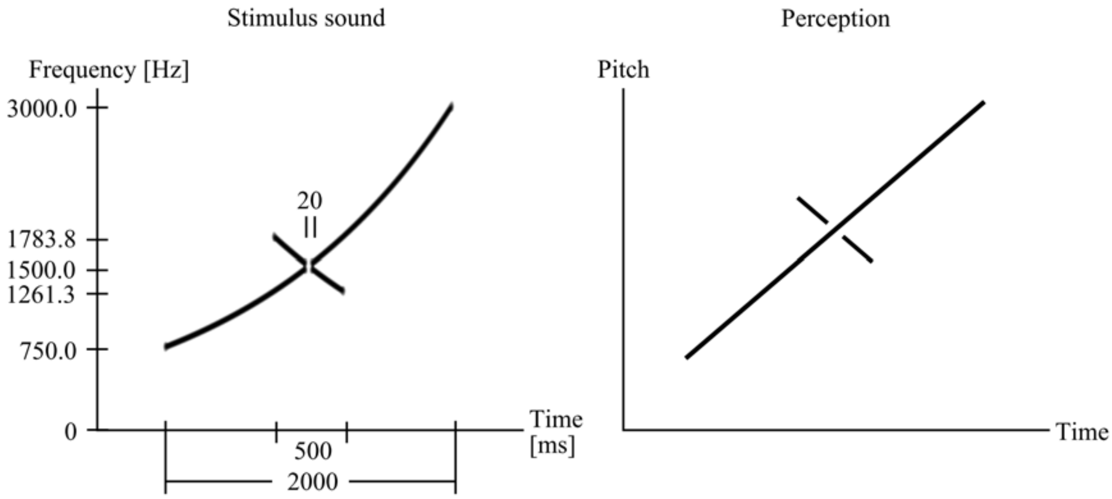

Figure 1 Illusory unification of simultaneous temporal gaps (Remijn, Nakajima, & Tanaka, 2007).

A simple model, *the event construction model* (Nakajima et al., 2000; Kanafuka et al., 2007), gives an explanation. The crossing glides in the above-mentioned stimulus pattern share a very short gap or dip. The decaying parts of these glides that mark the beginning of the shared gap or dip are simultaneous in time and within a critical bandwidth in frequency. It should not be very unreasonable to assume that these decaying parts just give a single cue of an offset. The same kind of argument holds for the rising parts of these glides that mark the end of the gap or dip, and we can assume that they serve as a single cue of an onset.

Because the onset cue at the very beginning of the shorter glide and the integrated offset cue at the beginning of the gap are close to each other in time and frequency, they are likely to be connected with each other perceptually forming a percept of a short tone. In the same way, the integrated onset cue at the end of the gap or dip and the offset cue at the very end of the shorter glide are likely to form another percept of a short tone. Thus, two successive short tones are obtained, and the integrated onset and offset cues need not be interpreted again perceptually. If the offset cue and the onset cue delimiting the gap are very close to each other in time, probably they do not allow a clear cue of a silence, which must require some amount of time, to be inserted between them. A longer glide that shares a very short temporal gap or dip with a shorter glide as in the present context, thus, does not contain an offset cue, an onset cue, nor a clear cue of a silence around its middle; such a glide has every reason to be perceived as continuous if the gap or dip is short enough, despite the fact that there is no sound energy to fill the gap or dip.

The illusory continuity of the longer glide is quite different from the illusory continuity as discovered by Miller and Licklider (1950), and as explained by Bregman (1990) and Warren (2008), whose views on related phenomena are accepted widely. In most reported cases displaying illusory continuity of a sound, there is a portion or portions of sound energy belonging to another sound, and the energy can fill the gap or gaps of the former sound. For example, suppose that a 100-ms portion of a pure tone is replaced with a loud noise, and that one still hears the pure tone as continuous. This continuity may be illusory in the sense that the gap of the pure tone is not heard as such, but is not actually so illusory as one may think because the peripheral excitation pattern in time and frequency, which roughly corresponds to the sound energy distribution, must be showing a fairly continuous extension in time. The physical structure of the stimulus pattern and the corresponding peripheral excitation just do not indicate clearly whether the tone is continuous or discontinuous. In illusory gap unification, however, the peripheral excitation pattern is very likely to be discontinuous in time. This means that the longer glide is likely to have a temporal gap or dip in its peripheral excitation pattern, and this gap or dip should be clear enough to cause a gap to be perceived, for the same shared gap or dip in the excitation pattern is perceived as a gap in the shorter glide. This means that the subjective continuity of the longer glide is really illusory, and must be caused at a higher level of the auditory system. The phenomenon, therefore, should give new insights to our understanding of auditory organization.

There is one weak point, however, in the above argument. Because the longer and the shorter glide cross each other sharing a gap or dip, the total sound intensity just before or after the temporal gap or dip is doubled, compared with the situation where the longer glide is presented alone, and this may have changed the local excitation pattern within the period corresponding to the gap or dip.

The main purpose of the present study was to examine whether the same kind of 'really' illusory continuity appears in conditions where the local peripheral excitation in response to a long glide really remains unchanged.

Experiment

We employed stimulus patterns consisting of three glides that either ascended or descended together. One of them was 3000 ms, and the others were 300 ms. In some cases, they shared a temporal gap or dip in their temporal middles. We were interested in whether illusory auditory continuity or related phenomena take place in such conditions where the peripheral interaction between glides was minimized.

Method

Participants:
Eight undergraduate and graduate students of Kyushu University, four males and four females, took part in the experiment. They had received basic training in music and training in technical listening for acoustic engineers.

Stimulus patterns:
We employed two types of stimulus patterns: complex stimulus patterns and simple stimulus patterns. A complex stimulus pattern consisted of one long glide component and two concurrent short glide components. A simple stimulus pattern was either a long glide component or two concurrent short glide components.

A complex stimulus pattern consisted of three glide components of the same amplitude, among which one was 3000 ms and two were 300 ms. These glide components either ascended or descended together at a rate of 1 oct/s, and they shared a temporal middle. The middle frequency of the longest component was 1000 Hz, and the other components were 13 semitones above and below this glide. There were four gap conditions as follows (Figure 2): 1) No-gap condition: All the glide components were continuous physically; 2) Gap-in-short-glide condition: The shorter components had a common gap in their temporal middle, but the longest component was continuous; 3) Gap-in-long-glide condition: The longest component had a gap in its temporal middle, but the shorter components were continuous; 4) Shared-gap condition: A temporal gap was shared by all the components in their middles.

A simple stimulus pattern was either the longest component or the combination of the two shorter components in one of the complex stimulus patterns. All possible simple patterns were extracted from all the complex patterns.

The duration of the gaps, as defined as the duration of a silent period, was varied from 0 to 160 ms. The rise and fall times of the components were always 20 ms, and were shaped by raised cosine curves. When a gap was 0 ms, the glide component was not continuous, but there were a fall time before and a rise time after the 0-ms gap, which means that there was an intensity dip of 40 ms.

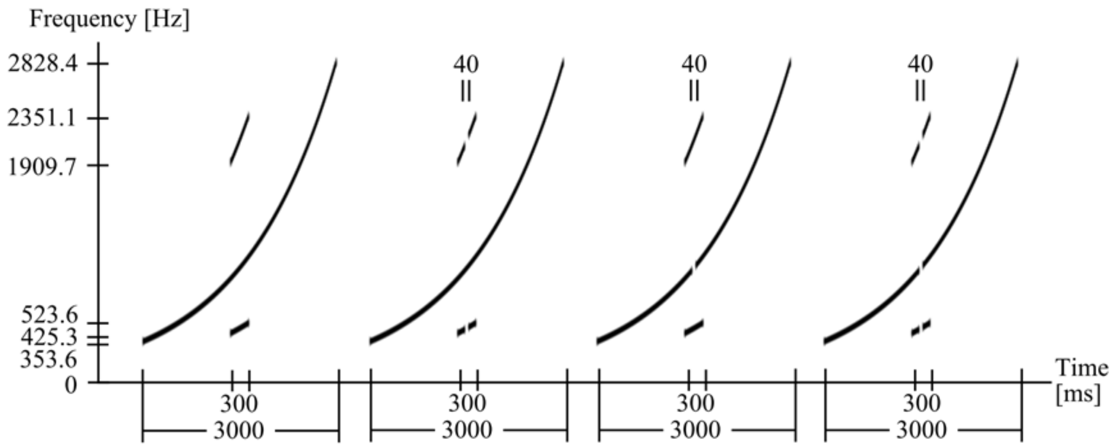

Figure 2 The four gap conditions employed in the experiment. Simplified spectrograms of stimulus patterns in 1) no-gap, 2) gap-in-short-glide, 3) gap-in-long-glide, and 4) shared-gap condition are indicated. Only ascending patterns were chosen in this figure, but there were also descending patterns, which would be represented by the mirror images of these spectrograms.

The stimulus patterns were presented diotically to the participant through headphones (Stax SR-303), whose frequency responses, together with the other parts of the experimental system, were made flat by an equalizer (Roland RAD RDQ-2031). The presentation level was calibrated so that a pure tone of 1000 Hz of the same amplitude as that of each glide component be 65 dB SPL. An artificial ear (Brüel & Kjær 4153) was used to measure sound pressure levels.

Procedure:

The task of the participant was to report the percept of each pattern both verbally (by writing words) and graphically (by drawing). A half of the participants listened to the complex stimulus patterns and the simple stimulus patterns in this order, and the other half in the opposite order. The patterns of each type were presented in randomized order. Each pattern was presented 2 s after the participant pressed a mouse button, and the participant was allowed to listen to the pattern as many times as he/she wanted.

Results and Discussion

The simple patterns were employed basically to check whether the auditory system was capable of utilizing the information of temporal gaps in the complex patterns, and we are going to examine how they were perceived only when necessary in our arguments. Complex stimulus patterns were perceived in two different modes. In the first mode, 1) a long glide tone or two successive long glide tones and 2) a short tone or two successive short tones appeared in the percept. A short tone could be either a glide or a steady-state tone. In the second mode, each short tone in the first mode was represented by two concurrent components as in a musical chord in which two notes are distinguished. Because our main purpose was to examine the subjective continuity of the long glide tone, which was represented by two successive glide tones when discontinuous, we did not differentiate between these two modes in the present analysis. That is, if two concurrent short components were reported, they were interpreted as a short tone consisting of two components. This should be justified by the fact that such two components started and ended simultaneously in the participants' reports.

Roughly speaking, a complex pattern was perceived mostly as consisting of one long glide tone and one short tone (which could be a glide tone). A tone in the percept could be D) discontinuous, P) partially discontinuous, or C) continuous. A discontinuous (D) tone could be described either as one tone with a gap or as two successive tones. A partially discontinuous (P) tone was characterized by a rapid change that could be associated with a gap, e.g., a temporary loudness decrease. A continuous (C) tone was without a rapid change between the onset and the offset.

The stimulus patterns in the no-gap condition, the gap-in-short-glide condition, and the gap-in-long-glide condition were almost always perceived veridically. That is, physically continuous tones were reported without a particular change between the onset and the offset, and physically discontinuous tones were reported as two successive tones or a tone with a temporal gap. The stimulus patterns in the shared-gap condition were not always perceived veridically, however. Table 1 shows the results. When the gap duration was 0 ms, i.e., when the dip duration was 40 ms, illusory gap unification took place in a typical manner; a long continuous glide tone and a short tone with a temporal gap, or two successive short tones, were reported. The illusion disappeared rapidly when the gap duration was increased. If we define the threshold of the gap duration for the occurrence of the illusion, it must be between 0 and 20 ms in the present situation, but, in order to establish such values, we would need a psychophysical experiment in which the presentation of stimulus patterns and the participant's responses are more strictly controlled. The present data show that the illusion could take place, although infrequently, even when the gap duration was as long as 40 ms.

Table 1. Frequency distributions of categorized degrees of subjective continuity-discontinuity

	Shared-gap pattern								Single-glide pattern		
	Long tone			Sign	Short tone			Sign	(Long tone)		
Gap duration [ms]	D	P	C	test 1	D	P	C	test 2	D	P	C
Ascending condition											
0			8	>	5	1	2	>>	7	1	
20	5	1	2		7	1			7	1	
40	6	1	1		8				8		
80	8				8				8		
160	8				8				8		
No gap											8
Descending condition											
0	2		6	>	7	1		>	7	1	
20	5	1	2		8				7	1	
40	6	1	1		8				8		
80	8				8				8		
160	8				8				8		
No gap											8

Note. The subjective continuity-discontinuity of the tones, which was reported by eight observers, was classified into the following categories: D) discontinuous, P) partially discontinuous, and C) continuous. The frequency of each category is indicated. Blanks represent "0." Sign test 1: A sign test was conducted in each condition to compare whether the percept of the long tone was more continuous than the percept of the short tone or vice versa. ">" represents that the long tone was significantly more continuous at the 5% level. Sign test 2: A sign test was conducted in each condition to compare whether the percept of the long tone in the shared-gap pattern was more continuous than the percept of the corresponding single-glide pattern or vice versa. ">" represents that the long tone was significantly more continuous at the 5% level, and ">>" at the 1% level.

Examining the results in the simple stimulus patterns, even the 0-ms gaps in these complex stimulus patterns should be detectable at a certain stage of perceptual processing, and yet the long glide with a 0-ms gap, which is about 4-5 critical-bands apart from the short glide components, is perceived as continuous by all the participants in the ascending pattern, and by 6 out of the 8 participants in the descending pattern. An explanation based on *the event construction model* (Nakajima et al., 2000; Nakajima, 2008a) is that the onset cue given by the simultaneous onsets of the two short components and the offset cue given by the simultaneous offsets of all the glide components just before the gap were connected with each other, because they were close to each other in time and frequency, to construct an auditory event of a short tone just before the gap, and that the onset cue and the offset cue after the gap were connected with each other in the same way to construct another auditory event.

Conclusions

Illusory unification of simultaneous temporal gaps takes place even when a gap in a long glide, which is to be filled perceptually, receives almost no peripheral influence of other acoustic components. This illusory continuity cannot be explained in a conventional manner, but an explanation based on the event construction model works.

Acknowledgements

This work was supported by the Japan Society for the Promotion of Science (19103003 in fiscal 2007-2009; 20653054 in fiscal 2008-2009).

References

Bregman, A. S. (1990). *Auditory Scene Analysis*. Cambridge, MA: MIT Press.

Kanafuka, K., Nakajima, Y., Remijn, G. B., Sasaki, T., & Tanaka, S (2007). Subjectively divided tone components in the gap transfer illusion. *Perception & Psychophysics*, 69, 641-653.

Miller, G. A. & Licklider, J. C. R. (1950). The intelligibility of interrupted speech. *Journal of the Acoustical Society of America*, 22, 167-173.

Nakajima, Y. (2008a). Illusions related to the temporal continuity and discontinuity of sounds. *Proceedings of the 24th Annual Meeting of the International Society for Psychophysics, Toronto, Canada*. 3-6.

Nakajima, Y. (2008b). Illusions related to auditory grammar: Ten demonstrations in musical contexts. *Proceedings of the 10th International Conference on Music Perception and Cognition, Sapporo, Japan*. in press.

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.

Remijn, G., Nakajima, Y., & Tanaka, S. (2007). Perceptual completion of a sound with a short silent gap. *Perception*, 36, 898-917.

Remijn, G. B., Pérez, E., Nakajima, Y., & Ito, H. (2008). Frequency modulation facilitates (modal) auditory restoration of a gap. *Hearing Research*, 243, 113-120.

Warren, R. M. (2008). *Auditory Perception: An Analysis and Synthesis: Third Edition*. Cambridge: Cambridge University Press.