THE SUBJECTIVE CONTINUITY OF FREQUENCY-GLIDE TONES IN THE GAP TRANSFER ILLUSION

Tsuyoshi Kuroda[1], Yoshitaka Nakajima[2]
Department of Acoustic Design, Graduate School of Design, Kyushu University
4-9-1 Shiobaru, Minamiku, Fukuoka 815-8540 Japan
[1]*kuroda@gsd.design.kyushu-u.ac.jp,* [2]*nakajima@design.kyushu-u.ac.jp*

Abstract

The gap transfer illusion is an auditory illusion in which a very short temporal gap inserted in a long frequency glide is perceived as if it were in a crossing short glide. We examined whether the subjective continuity of the long glide in this illusion is caused by reallocation of sound energy from the short glide to the long glide. We found that a glide of 4000 ms with a gap of 100 ms could be perceived as continuous even when a crossing glide of 400 ms was 9 dB weaker, but a single glide of 4000 ms with an intensity dip of 100 ms was perceived as discontinuous even when the dip was only 3 dB. Thus, the subjective continuity of the long glide in the gap transfer illusion cannot be attributed to reallocation of sound energy.

The gap transfer illusion is an auditory illusion which takes place typically in stimulus sounds consisting of two frequency-glide components (Nakajima, Sasaki, Kanafuka, Miyamoto, Remijn, & ten Hoopen, 2000; Figure 1). One glide is around 1500 ms or longer, and the other glide is around 500 ms or shorter. The two glides cross each other at their temporal middles; the longer glide ascends or descends in frequency, and the shorter glide moves in the opposite direction. A temporal gap is inserted at the temporal middle of the longer glide, but the gap is perceived as if it were in the shorter glide, that is, the longer glide is perceived as continuous, and the shorter glide is perceived as discontinuous. The gap transfer illusion is explained by introducing auditory subevents such as onsets and terminations (Kanafuka, Nakajima, Remijn, Sasaki, & Tanaka, 2007; Nakajima et al., 2000). When the gap transfer illusion takes place, onsets and terminations are detected at the temporal edges of the glides (Figure 1). The principle of proximity, one of the Gestalt principles, is applicable to these auditory subevents; an onset and a termination that are close to each other in frequency and time are likely to be connected perceptually to form an auditory event. In the stimulus sound indicated in Figure 1, the onset and the termination preceding the gap are close to each other, and thus, they are connected with each other. The onset and the termination succeeding the gap are also connected. The residual onset and termination form a long continuous tone. Thus, the gap transfer illusion has been explained.

There is another explanation on the subjective continuity of the long glide in the gap transfer illusion (Nakajima et al, 2000). In the stimulus sounds that cause the gap transfer illusion, the sound energy of the short glide at the temporal middle is sufficient to restore the missing part of the long glide, and thus, a portion of this energy is reallocated to the long glide. This explanation, which we call the energy-reallocation model, may be accepted widely in the studies of the auditory continuity illusion (Bregman, 1990; Miller & Licklider, 1950; Warren, Obusek & Ackroff, 1972), in which a sound with a temporal gap is perceived as continuous when the gap is filled with a more intense sound.

However, it is difficult for us to accept the above idea that the subjective continuity of the long glide in the gap transfer illusion is attributed to the reallocation of sound energy. Kuroda, Nakajima, Tsunashima, and Yasutake (2009) found that the gap transfer illusion can take place even when the short glide is 6 dB weaker than the long glide. If a portion of the sound energy of the short glide, in this case, were reallocated to the long glide, there should be a temporal dip of 6 dB around the temporal middle of the long glide. The dip is likely to cause the glide to be perceived as discontinuous, and the energy-reallocation model would not work. We never examined, however, the perception of dips inserted in glides systematically, and thus, we conducted an experiment.

Method

Observers

Eight observers with normal hearing participated. They were acoustic engineering students at Kyushu University. All observers had received basic training in music and seven observers had received training in technical listening for acoustic engineers.

Stimulus sounds and apparatus

We employed two sets of stimulus sounds (Figure 2). One set included crossing-glide patterns consisting of a glide of 4000 ms and a glide of 400 ms that crossed each other at their temporal middles. The other set included single-glide patterns of 4000 ms. The rise time and the decay time of amplitude were 20 ms with cosine-shaped ramps.

There were two “direction” conditions in the crossing-glide patterns: In the ascending condition, the long glide ascended and the short glide descended, and vice versa in the descending condition. These two conditions were distinguished also in the single-glide patterns on depending on whether the glide ascended or descended. The frequency gliding rate was 1/3 oct/s. The crossing frequency of the two glides in the crossing-glide patterns and the instantaneous frequency at the temporal middle of the glide in the single-glide patterns were both 1000 Hz.

In the crossing-glide patterns, a temporal gap of 100 ms was inserted at the temporal middle of the long glide. The level of the short glide was varied from -18 dB to +18 dB in steps of 3 dB. Thus, there were 13 “short-glide-level” conditions. The level of the long glide was always 0 dB.

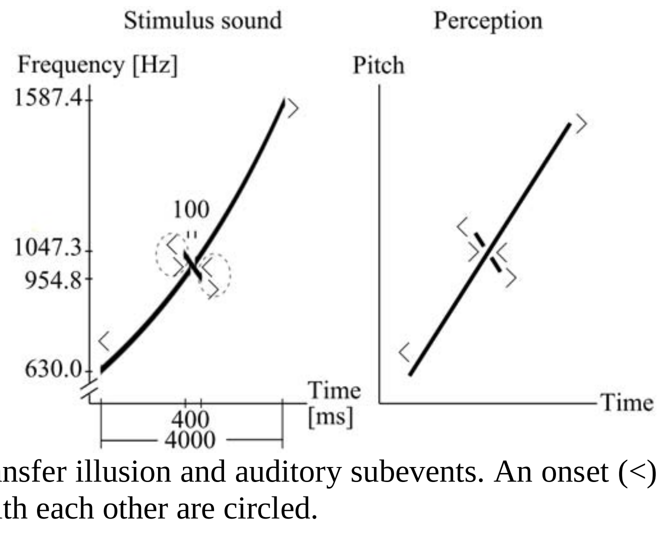

Figure 1. The gap transfer illusion and auditory subevents. An onset (<) and a termination (>) that are connected with each other are circled.

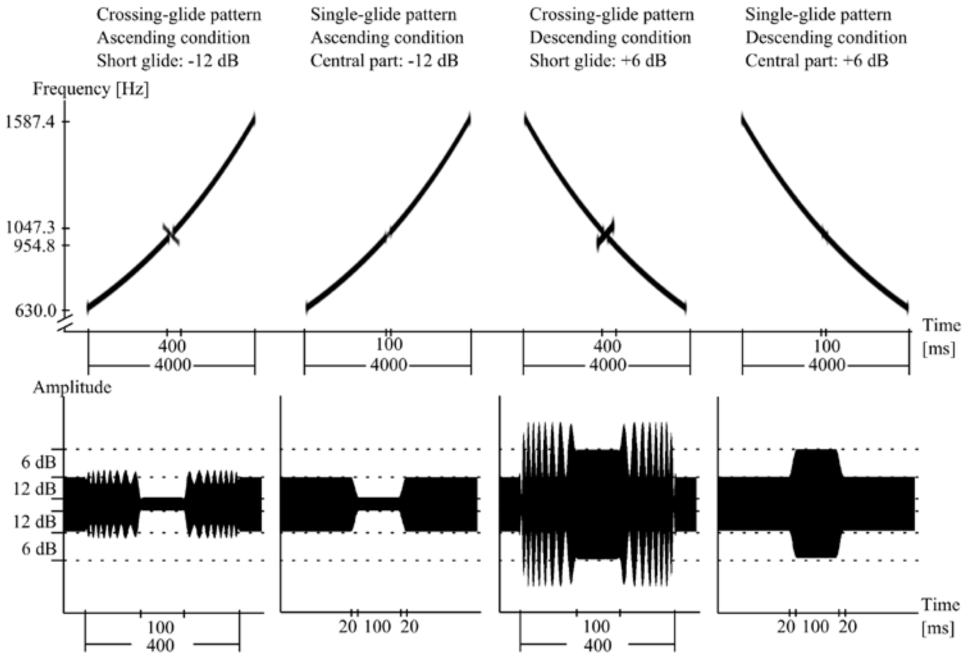

Figure 2. Stimulus sounds. The spectrograms are shown at the top. The wave forms at the temporal middles are shown at the bottom.

In the single-glide patterns, there were two “middle-part” conditions: In the gap condition, a temporal gap of 100 ms was inserted at the temporal middle of the glide; in the level-difference condition, the level at the temporal middle portion of 100 ms was varied from -18 dB to +18 dB in steps of 3 dB (Figure 2). Thus, there were 13 “central-part-level” conditions. There were amplitude transitions of 20 ms just before and after the central part. A cosine-shaped envelope was applied to these transitions. The other part of the glide was always at 0 dB.

The level of 0 dB was calibrated to 70 dB SPL both in the crossing-glide patterns and in the single-glide patterns. Thus, the sound pressures at the temporal middles of the crossing-glide patterns and the single-glide patterns were equal if the level of the short glide in the crossing-glide patterns and the level of the central part in the single-glide patterns were of the same value (Figure 2).

There were 26 crossing-glide patterns (2 directions × 13 short-glide levels), and 28 single-glide patterns (1 gap × 2 directions + 1 level-difference × 2 directions × 13 central-part levels). Thus, there were 54 stimulus sounds in total.

We used a computer application (Microsoft Visual Basic 2005) to make a program steering the course of the experiment. We used an application (Jsoftware J504b) to generate sound files containing the digital signal of each stimulus sound. The signals were sampled at 44100 Hz and quantized to 16 bits.

The stimulus sounds were presented to the observer in a sound-attenuating chamber, where background noise was below 30 dBA. The sound file of each stimulus sound was played once after the observer clicked on a pane on the computer display. All sound files contained a silent interval of 2000 ms before the onset of the stimulus sounds. The digital signals of the stimulus sounds were sent out from an audio card (E-MU 0404) installed in a computer (Frontier KZFM71/N), which was located outside the chamber. The digital signals were converted into analogue signals in an audio processor (Onkyo SE-U55GX). The signals were led through an active low-pass filter with a cutoff frequency of 15000 Hz (NF DV8FL), a graphic equalizer (Roland RDQ-2031), an amplifier (Stax SRM-313), and headphones (Stax SR-303). The low-pass filter was for anti-aliasing, and the equalizer was for keeping the frequency response of the system flat. The SPLs were measured with a sound level meter (Nagano Keiki 2072) and an artificial ear (Brüel & Kjær 4153).

Procedure

The observer and the experimenter stayed in the sound-attenuating chamber. The observer was allowed to listen to each stimulus sound as many times as he/she wanted. The observer was instructed to write words explaining his/her perception with a pencil on a sheet of paper. The observer was also instructed to draw a figure (figures) representing his/her perception, using the horizontal axis for time and the vertical axis for pitch. These axes were drawn in advance on each paper sheet. When there was more than one type of perception, the observer reported all types and indicated the order of dominance if possible. The observer handed this sheet to the experimenter when he/she finished the observation of each stimulus sound. When there was an ambiguous or contradictory part in the report, the experimenter asked the observer to make this part clearer, but the experimenter did not mention anything about the content of the report.

Before the experiment, the observer listened to all stimulus sounds once in random order by clicking on a pane on the display. Training observations were carried out in the same way as in the experimental observations. The training sounds included both stimulus sets, but the levels of the short glide in the crossing-glide patterns and the central part in the single-glide patterns were varied from -18 dB to +18 dB in steps of 9 dB instead of 3 dB. Thus, there were 10 crossing-glide patterns (2 directions × 5 short-glide levels) and 12 single-glide patterns (1 gap × 2 directions + 1 level-difference × 2 directions × 5 central-part levels) for training. There were 22 training sounds in total, and they were presented in random order.

The experimental observations of the 54 stimulus sounds were divided into 3 blocks, and thus, each block contained 18 observations. The stimulus sounds were presented in random order. At the beginning of each block, a warm-up of two observations was carried out in the same way as in the experimental observations, and the warm-up sounds were the same as to be presented in the last two observations of the block. Breaks of few minutes were inserted between the blocks. Each block took about 20 min.

Results and Discussion

Typical examples of drawing and writing descriptions are shown in Figure 3. The observers' reports indicated the perception of a long tone and a short tone in the crossing-glide patterns. One report indicated only a long tone when the short glide was at -18 dB in the ascending condition. This report was omitted from further analyses.

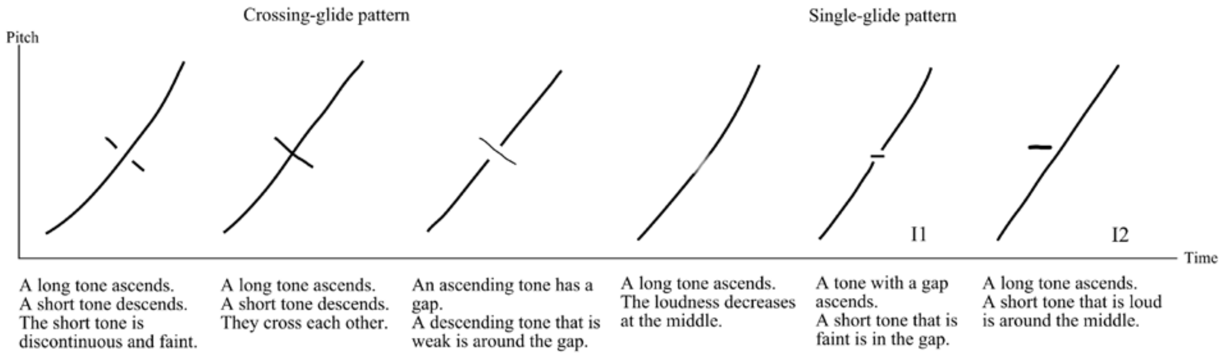

Figure 3. Typical examples of drawing and writing descriptions.

The observers' reports indicated the perception of a single tone in the single-glide patterns, but some reports indicated the appearance of an illusory short tone around the temporal middle of the percept of the glide. The illusory tone was accompanied by the subjective discontinuity of the glide (Figure 3, I1) or the subjective continuity of the glide (I2). When the central part was above +3 dB, the illusory tone appeared in seven out of the eight observers with the subjective continuity of the glide. This perception may have been a special case of the homophonic continuity illusion (Warren et al, 1972).

We classified the continuity-discontinuity of the tone into three categories: The tone was classified as "discontinuous (D)" if the report indicated clear discontinuity of the tone, as "partially discontinuous (P)" if the report indicated a weak effect related to the discontinuity like a transient decrease in loudness, as "continuous (C)" if the report indicated no discontinuity. When the illusory tone appeared in the single-glide patterns, we classified only the percept of the glide. When more than one type of percepts appeared in a single observation, only the most dominant percept was classified. The frequencies of the reports in these categories are shown in Table 1.

We compared the subjective continuities of the long tone and the short tone in the crossing-glide patterns statistically. We checked, in each report, whether the long tone or the short tone was more continuous, and conducted a sign test (two-tailed) for each stimulus sound. The sign test was based on 8 comparisons for the eight observers if no report had been omitted. When both tones belonged to the same category, the comparison was regarded as a tie and was not included in the calculation. The results are shown in Table 1. When the long glide and the short glide were at the same level, the percept of the long glide was significantly more continuous than the percept of the short glide, that is, the gap transfer illusion took place. The gap transfer illusion also tended to take place when the short glide was weaker.

The continuity category of the long glide in the crossing-glide patterns changed from "continuous" into "partially discontinuous" or "discontinuous" when the level of the short glide became low. We defined the continuity threshold as a mean value between two levels where this change of the category took place. This threshold was calculated in each observer. The mean threshold (N = 8) in the ascending condition was -8.6 dB (SD = 5.8), and the mean threshold in the descending condition was -9.8 dB (SD = 4.2). We also calculated the continuity threshold in the single-glide patterns in the same way. The mean threshold in the ascending condition was -3.4 dB (SD = 1.5), and the mean threshold in the descending condition was -2.6 dB (SD = 1.5). We conducted a two-way ANOVA with repeated measures to establish the significant difference between the mean continuity thresholds of the crossing-glide patterns and the single-glide patterns. The main effect of the stimulus sets was significant, $F(1, 7) = 14.2$, $p < .01$.

The long glide could be perceived as continuous even when the crossing short glide was about 9 dB weaker. If this continuity of the long glide had been caused by reallocation of sound energy from the short glide to the long glide, there should have been a dip of 9 dB, which must be detected clearly. Thus, the subjective continuity of the long glide in the gap transfer illusion cannot be attributed to reallocation of sound energy. The temporal configuration of auditory subevents and the temporal distribution of sound energy may be utilized independently for determining the subjective continuity-discontinuity of sound.

Acknowledgments

The research was supported by the Japan Society for the Promotion of Science (20330152 in the fiscal years 2008-2012; 20653054 in the fiscal years 2008-2009).

Table 1. Frequency distribution of categorized continuities

Single-glide pattern				Crossing-glide pattern							
Level of central part [dB]	D	P	C	Level of short glide [dB]	Long tone D	Long tone P	Long tone C	Short tone D	Short tone P	Short tone C	Sign test
Ascending condition											
-18	5	3		-18	6	1		1		6	
-15	4	4		-15	7	1		3		5	
-12	5	3		-12	4	1	3	6		2	
-9	2	6		-9	2	1	5	7		1	
-6	5	3		-6	1	1	6	8			>
-3	1	2	5	-3		1	7	8			>>
0			8	0		1	7	8			>>
+3			8	+3		1	7	5		3	
+6			8	+6			8	1		7	
+9			8	+9			8			8	
+12			8	+12		1	7			8	
+15			8	+15			8			8	
+18			8	+18			8			8	
Gap	8										
Descending condition											
-18	4	4		-18	8			3		5	
-15	4	4		-15	6	1	1	4		4	
-12	6	2		-12	4	1	3	6		2	
-9	5	3		-9	4		4	6		2	
-6	5	3		-6	1	1	6	8			>
-3	3	2	3	-3			8	8			>>
0			8	0			8	8			>>
+3			8	+3			8	4		4	
+6			8	+6			8	1		7	
+9			8	+9	2		6			8	
+12			8	+12	2		6			8	
+15			8	+15	2		6			8	
+18			8	+18	2		6			8	
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." The sum of classifications in a pattern could be below 8 since some reports were omitted. >: The long tone is significantly more continuous than the short tone at the 5% level (>>: at the 1% level).

References

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

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.

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., 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.

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