# Aspects of the gap transfer illusion

PACS: 43.66.Ba

Yoshitaka Nakajima

Department of Acoustic Design
Kyushu Institute of Design
4-9-1 Shiobaru, Minami-ku
Fukuoka 815-8540
Japan
Tel: +81 92 553 4558
Fax: +81 92 553 4520
E-mail: nakajima@kyushu-id.ac.jp

## ABSTRACT

One of my colleagues and I found a new auditory illusion. In a typical situation, an ascending frequency glide of 2500 ms with a temporal gap of 100 ms in the middle and a continuously descending frequency glide of 500 ms cross each other at their temporal centers. The gap in the long glide is perceived as if it were in the short glide. Recently, we are investigating the illusion in a broad context including speech perception. We hypothesize that onsets and terminations of sounds are detected independently, and that they are organized perceptually to form auditory events obeying the principle of proximity.

## INTRODUCTION

Attempts to apply Gestalt principles to auditory perception are now recognized as comprising an indispensable part of auditory psychology (e.g., Handel, 1989; Moore, 1997). This was not the case when auditory psychology started to be a solid branch of modern science. Leading researchers in this field (e.g., Stevens and Davis; 1938) paid no attention to Gestalt psychology even when its importance to perceptual psychology was being revealed widely (e.g., Koffka, 1935). Most Gestalt psychologists, on the other hand, directed their efforts mainly to studying vision when they dealt with perception. They established a 'general' theory of perception doing little empirical research on auditory perception. The auditory modality and the visual modality, however, play different roles in our everyday life, and they must have different principles. Although it is often necessary to import traditional Gestalt notions into auditory psychology, it does not seem always productive to try to find a systematic analogy between audition and vision. Auditory organization must be governed by its own principles. My aim is to establish Gestalt principles for the auditory modality.

The majority of the researchers investigating auditory perception seem to agree with each other that the concept of auditory streams plays an important role when we investigate auditory organization (e.g., Bregman, 1990). An auditory stream is a linear string of auditory events and silences. A phrase in speech, a melody in music, a sequence of footsteps, and a continuous

fan noise are all typical examples of auditory streams. Strictly speaking, the percepts of these sequences are auditory streams. Our ears are often exposed to a disorganized flood of sound energy, and to distinguish auditory streams is the first step to construct the subjective world of sound. Auditory streams are similar to 'figures' in vision, but a great difference is that auditory streams may stay behind other auditory streams without being hidden completely. We sometimes begin to pay attention to an auditory stream (for example, a cricket's chirps) noticing that it has stayed there, behind other sounds, for some time. This means that the same auditory stream can change from a part of the ground into a figure without losing its identity.

The definition of an auditory event can diverge among researchers and contexts. I will use this word in a strictly basic sense in this report. A syllable in speech, a note in music, and a single footstep are typical auditory events. In some cases, a consonant at the end of a syllable, a short spoken word consisting of a few syllables, an attack of a musical note, or a note together with grace notes can be a single auditory event. The definition fluctuates a little.

## THE GAP TRANSFER ILLUSION

One of my colleagues and I found a new auditory illusion (Nakajima & Sasaki, 1993), which was reported in English later (Nakajima et al, 2000). We take up stimulus patterns where a long ascending frequency glide of 1500 ms from 640 to 2160 Hz and a short descending frequency glide of 500 ms from 1440 to 960 Hz cross each other at 1175.8 Hz at their temporal centers. These glides move at the same speed in logarithmic frequency in opposite directions, and their rise and fall times were always 10 ms. In one of the patterns, the short descending glide has a temporal gap of 100 ms in the middle, and the long ascending glide is continuous. The perception of this pattern is veridical. A typical percept is a long ascending component accompanied by a short descending component with a temporal gap in the middle. This may not be a very impressive result, but it should be noted that bouncing perceptual components, which often appear in this kind of stimulus patterns of crossing frequency glides (Halpern, 1977; Bregman, 1990), do not appear.

When a temporal gap of the same duration is introduced not into the short descending glide, but into the middle of the long ascending glide, the typical percept remains almost the same. Although the physical temporal gap is in the long ascending component, it is perceived as being in the short descending component. This phenomenon is what we call the *gap transfer illusion*. The same kind of illusory transfer of a temporal gap takes place also when this stimulus pattern is reversed in time.

An aspect of this illusion seems to have something in common with temporal induction (the illusion of continuity) reported by Miller and Licklider (1950). Systematic discussions on auditory induction appeared in Warren (1999) and Bregman (1990). When the gap transfer illusion takes place, the long glide component has a physical temporal gap that is long enough to be perceived, but the subjective continuity of this component is restored by the presence of the short glide component. It seems as if a portion of sound energy of the short component around the crossing point was reallocated to the long component. This might be interpreted as an example in which the neural response to the inducer is partly used to restore the inducee (Warren, Bashford, Healy, & Brubaker, 1994). It is theoretically important that the real portion necessary for restoration was specified clearly.

However, I do not think that the gap transfer illusion is just a variation of temporal induction. The continuity of the short component, which would be solid if this component were presented alone, is broken subjectively by introducing the long component. If the principle of good continuity worked here, this would mean that the subjective continuity of one component is sacrificed for that of the other component. We would have to explain, then, why the continuity of the long component has priority. It does not seem productive to try to explain the gap transfer illusion just by a simple analogy to auditory induction.

Tsunashima and Nakajima (2002) used harmonic frequency glides consisting of three

components in a related experiment, and showed that the gap transfer illusion took place robustly also in their paradigm unless there were amplitude differences between corresponding components of the long and the short glide, which crossed each other.

They also attempted to relate the temporal gap perceived in the gap transfer illusion to the perception of stop consonants in speech. They demonstrated a stimulus pattern where a long harmonic frequency glide of 4000 ms from 126 to 317 Hz and a short harmonic frequency glide of 400 ms from 209 to 191 Hz crossed each other at 200 Hz at their temporal centers. These glides had three formants each to make them perceived as vowels /a/, and move at the same speed in logarithmic frequency in opposite directions. The long glide had a temporal gap of 100 ms in the middle, and a very short noise to begin a consonant /k/ was placed in the middle of this gap. A formant transition for /k/ was also introduced immediately after the gap. If the long glide with the temporal gap and the very short noise had been presented alone, a stop consonant /k/ flanked by two vowels /a/ would have been perceived. Note that the temporal gap must have been crucial to cause the perception of /k/. When the short frequency glide was added, however, the consonant /k/ was perceived not in the long component, but in the short component, where there was no physical clue about the consonant except for the very short noise, whose physical allocation was not clear. The short component, which would have been perceived as a single vowel /a/ when presented alone, was perceived as /aka/ in the present context. The most plausible explanation is that the temporal gap in the long component was transferred into the short component subjectively, and that this gap was utilized by the perceptual system as a clue of /k/. An interesting point is that the temporal gap often does not appear in the percept of this pattern when the transferred consonant /k/ is perceived.

## THEORETICAL IMPLICATIONS

I take up the very first example where frequency glides of 500 and 1500 ms cross each other. Suppose that the ascending glide, whether it is longer or shorter, has a temporal gap in the middle. If this glide is shorter (500 ms), the gap is perceived veridically as being in the corresponding, i.e., shorter, subjective component. If this glide is longer (1500 ms), the gap is perceived as being in the other component, which is again shorter. The gap is almost always perceived as being in the shorter component. In the present example, these two stimulus patterns are almost the same within the 500 ms range around the crossing point. This means that what happens more than 200 ms before or after the temporal gap seems to determine whether the gap is perceived in the veridical or in the illusory position. It is unlikely, therefore, that the illusion is explained only at a peripheral level.

Nakajima et al (2000) reported a closely related phenomenon. They used a stimulus pattern where a long ascending glide of 2500 ms from 421.7 to 2371.3 Hz and a short descending glide of 500 ms from 1188.5 to 841.4 Hz crossed each other at 1000 Hz in opposite phases at their temporal centers. The glides caused acoustic beats around the crossing point, where they had opposite phases. Thus, the temporal envelope of the whole stimulus pattern had a gap of 38 ms in the middle (if the boundaries of a gap can be located at the points at which the sound pressure level is 3 dB below the peak level). This gap belongs to neither of the glides physically, but the general tendency is that clearer discontinuity is perceived in the shorter glide component. It seems that our perceptual system allocates temporal gaps more easily to shorter glide components than to longer glide components.

We proposed a simple perceptual model, i.e., the event construction model, in order to understand the mechanism of the gap transfer illusion (Nakajima et al, 2000; see also Nakajima & Sasaki, 1996). Our basic idea is that onsets and terminations could behave as if they were independent 'subevents'. A temporal gap is accompanied by, or includes, a termination of a sound and an onset of another sound in this order. If the termination at the beginning of the gap is connected perceptually to an onset which took place before its occurrence, then we get a new auditory event. If the onset at the end of the gap and a termination are connected perceptually

in this order, then we get another auditory event. Probably, an onset and a termination can be connected with each other more easily when they are closer to each other in time and frequency. This means that the onset and the termination of the shorter glide should be connected more easily to the termination and the onset bounding the temporal gap than their counterparts of the longer glide. Thus, the gap is more likely to be perceived in the shorter glide component.

We introduced one of the Gestalt principles, the proximity principle, in order to understand how onsets and terminations are connected perceptually to construct auditory events. We will also need the good continuity principle in order to understand the perception of the glide components in our experimental paradigm. We believe Gestalt principles are still very useful conceptual tools if we try to introduce them in a reserved manner.

## ACKNOWLEDGMENTS

This report is based on my collaboration with Takayuki Sasaki and Gert ten Hoopen. The present study was supported by the Japan Society for the Promotion of Science (14101001 in the fiscal year 2002).

## BIBLIOGRAPHICAL REFERENCES

Bregman, A. S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, MA: MIT Press.

Halpern, L. (1977). The effect of harmonic ratio relationships on auditory stream segregation. Unpublished research report, McGill University, Psychology Department.

Handel, S. (1989). *Listening: An Introduction to the Perception of Auditory Events.* Cambridge, MA: MIT Press.

Koffka, K. (1935). *Principles of Gestalt Psychology*. Harcourt, Brace & World, New York.

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

Moore, B. C. J. (1997). *An Introduction to the Psychology of Hearing*. (Fourth Edition) London: Academic Press.

Nakajima, Y., & Sasaki, T.(1993). Perceptual transfer of onsets and offsets between crossing glide tone components [in Japanese, abstract in English](Tech. Rep. No.SP92-146, pp.73-80).Tokyo: Institute of Electronics, Information and Communication Engineers.

Nakajima, Y., & Sasaki, T.(1996). A simple grammar of auditory stream formation [Abstract]. *Journal of the Acoustical Society of America*, **100**(4, Pt.2), 2681.

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

Stevens, S. S. & Davis, H. (1938). *Hearing*. Wiley, New York.

Tsunashima, S., & Nakajima, Y. (2002). Demonstrations of the gap transfer illusion. *Proceedings of the 7th International Conference on Music Perception and Cognition, Sydney, 2002*.

Warren, R. M. (1999). *Auditory Perception: A New Analysis and Synthesis.* Cambridge: Cambridge University Press.

Warren, R. M., Bashford, J. A., Jr., Healy, E. W., & Brubaker, B. S.(1994). Auditory induction: Reciprocal changes in alternating sounds. *Perception & Psychophysics*, 55, 313-322.