The perceptual integration of auditory onsets and offsets in stimulus patterns of two partly overlapping frequency glides

Remijn, Gerard Bastiaan

https://doi.org/10.15017/458553

出版情報：Kyushu Institute of Design, 2002, 博士（芸術工学）, 課程博士
バージョン：
権利関係：

CHAPTER 1: Introduction and theoretical background

1.1 Gestalt principles in perception

A famous anecdote regarding the life in the studio of the Dutch painter Rembrandt van Rijn (1606-1669) tells that his students used to paint coins on the floor for the pleasure of watching their teacher bend down in a vain effort to pick them up. They followed a tradition in visual art that later, from the middle to the end of the nineteenth century, has become known as 'Trompe l'oeil', a French term that literally means 'to deceive the eye'. The purpose of 'Trompe l'oeil' artists is to make two-dimensional representations of objects that look so realistic that the viewer is fooled into thinking that actual three-dimensional objects exist (Dars, 1979). The artists deceive our eyes, confronting us with something that makes us question our perception of reality. People have been intrigued by 'deceptions of the eye' throughout history. 'Trompe l'oeil', for example, dates back as far as 400 B.C. as a part of the Greco-Roman culture, and the drawings and paintings were used for entertainment and decoration.

Many years later, a revival of 'Trompe l'oeil' at the end of the nineteenth century inspired the creation of a great number of 'deceptions of the eye', visual illusions, that not only provided entertainment, but also promoted the study of our perception of the physical world. An important group of psychologists who studied the perceptual system with the aid of diverse illusory figures and phenomena were the Gestalt psychologists. The German word 'Gestalt' means 'shape' or 'form'. A central question to their movement, and to the study of the perceptual system in general, was how the perceptual system could integrate sensory elements, corresponding to immediate responses in the brain that are caused by excitation of a sensory organ (James, 1892), into a coherent perceptual organization (Köhler, 1947). The ideas of the Gestalt psychologists can best be explained by a visual example, like Figure 1. The figure consists of two identical white line segments (A and C) that flank a gray, square-shaped plane B.

Although the figure is two-dimensional, it is possible to perceive it as a three-dimensional picture consisting of a continuous white bar that is obstructed by a gray plane. The Gestalt psychologists asked themselves questions such as how and why our perceptual system can organize the white line segments A and C into the same perceptual organization, and how the impression of depth in the picture is obtained. Figure 1 is just an example of what our visual system is confronted with in our daily life.

In order to interpret and understand the world around us, our visual system integrates sensory elements such as edges, lines, and colors into larger, meaningful perceptual organizations. Although this process is complicated and cannot be described fully by just mentioning the work of the Gestalt psychologists, their ideas are valuable in our understanding of perceptual organization in general.

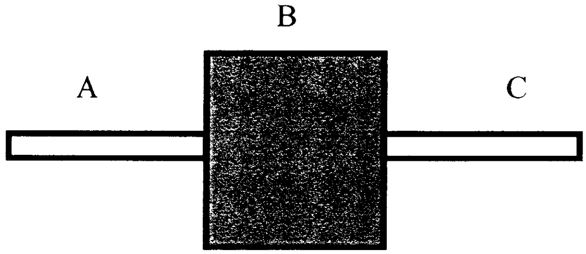

Figure 1. *Two-dimensional figure that shows the perceptual integration of white line segments A and C. Together, A and C can form a continuous line segment that is occluded by gray plane B.*

The perceptual integration of sensory elements, according to the Gestalt psychologists, involves a number of principles. One of these principles is the principle of similarity: sensory elements that are similar are likely to become part of the same perceptual organization. In Figure 1, according to the Gestalt psychologists, the fact that both line segments A and C are white, and physically are of the same width, promotes their perceptual integration. Because the line segments have the same widths, another principle that may be important for the perception of the figure is that of 'good continuity'. If we would extend one of the white line segments A or C, they physically would blend and become one. If the angle and slant of one of the white line segments would be different, however, the perceptual integration of both segments becomes more difficult. Another principle of perceptual integration, according to the Gestalt psychologists, is that of proximity. Sensory elements that are close to each other are more liable to become part of the same perceptual organization. In the figure, the two white line segments A and C are close to each other. If the two line segments A and C were farther removed from each other, for example, when separated by a much wider gray plane B, their perceptual integration would become more difficult.

Similarity, proximity, and good continuity are thus some of the principles that underlie the perceptual integration of sensory elements, according to the Gestalt psychologists. Their work has been very influential to the understanding of our perceptual system, and research based on their work is thriving even today (e.g. Spillmann, 1999; Hermann & Bosch, 2000). Moreover, Gestalt psychology is not only important for the study of the visual perceptual system, but also became important for the understanding of other sensory systems. Nowadays, the field of auditory perception also benefits from Gestalt psychology (Deutsch, 1982; Handel, 1989; Bregman, 1990). In fact, Figure 1 could very well be a typical example of an auditory stimulus pattern that yields the so-called 'auditory continuity effect' (Warren, 1999). In the auditory modality, a stimulus pattern that is made up of two weaker sounds that flank a more intense sound can often be heard as consisting of a single, long pitch trajectory that continues behind a louder sound. Although physically no sound is present behind the more intense sound, a continuous pitch trajectory can be heard. In Figure 2, the two white line segments A and C that flank the gray plane B represent two weaker sounds, whereas the gray plane B represents a more intense sound. The gray plane B and the white line segments A and C are depicted with the x-axis corresponding to time, and the y-axis to frequency. Time (x-axis), frequency (y-axis), and relative intensity of sounds in stimulus patterns (by means of dark and light texturing) will be depicted in this way throughout this thesis.

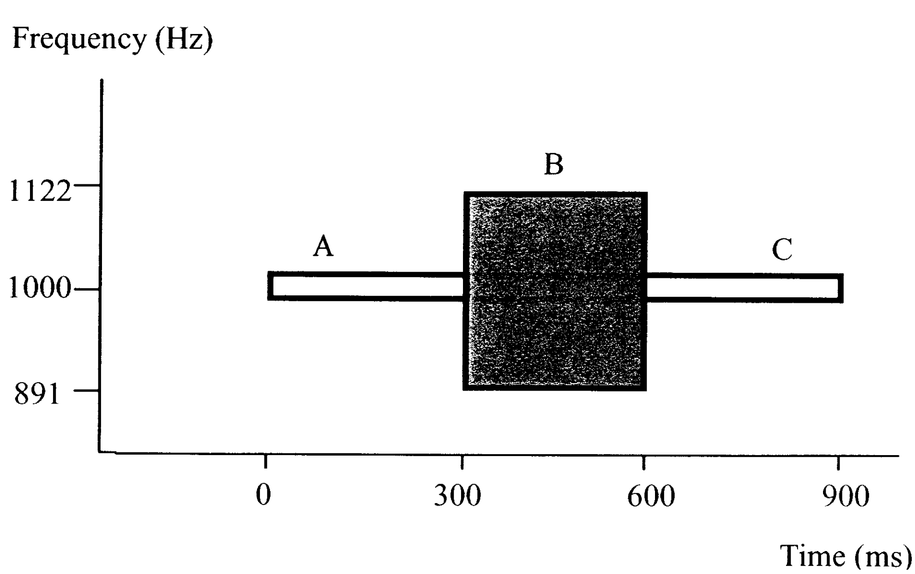

Figure 2. *An example of a stimulus pattern that renders typical auditory continuity. The weak sounds (the white line segments) can be heard as a single sound continuing through a more intense sound (the gray plane), even though the weaker sounds are physically interrupted by the more intense sound, and do not continue behind it.*

Similar to the visual modality, sounds A and C can become part of the same perceptual organization, and together they can form a relatively long, soft sound that appears to continue behind louder sound B. The auditory continuity effect can occur when the more intense sound is a noise burst and the weaker sounds are glides or steady-state frequency components (Ciocca & Bregman, 1987; Dannenbring, 1976). The auditory continuity effect can also appear when both the weaker and the more intense sounds are alternating with each other, and both the more intense sounds and weaker sounds are noise bursts, or when both are pure tones (see reviews of Bregman, 1990; Warren, 1999).

The same kind of Gestalt principles that are involved in the perception of continuity in the visual modality, for example in regarding the white line segments A and C as a single bar in Figure 1, can be involved in the perception of the auditory continuity effect (Bregman, 1990). Similarity of the two weaker sounds, in intensity and frequency, may promote the perception of a continuous pitch trajectory. Also, the proximity between the end frequency of weaker sound A and the starting frequency of weaker sound C, just before and after the more intense sound B, could promote the auditory continuity effect (Bregman, Colantonio, & Ahad, 1999; Bregman, Ahad, Crum, & O'Reilly, 2000). Finally, when weaker sounds A and C align on the same frequency trajectory, often they are more likely to become integrated perceptually into the same perceptual organization (Ciocca & Bregman, 1987). The Gestalt principles of similarity, proximity, and good continuity thus can describe the perceptual integration of auditory sensory elements as well.

When we look closely at Figure 1, we see that a part of a single line separates white line segment A and gray plane B. Although this borderline can be perceived as the end of white line segment A, often the borderline is perceived as belonging to gray plane B. In these cases, segment A appears to lack an end, and seems to be occluded by gray plane B. The same can be said about the borderline between B and white line segment C. Since the borderline perceptually belongs to B, the white line segment C seems to appear from behind B. This shows the Gestalt principle of 'belongingness' of sensory elements to a single object, or a single perceptual organization. Sensory elements, such as the borderline, are often included only into a single perceptual organization. According to Bregman (1990), this 'belongingness' has an ecological validity, in that in our natural world only a negligible chance exists that the touching edges of two objects (perceptual organizations) have exactly the same shape. Similarly, it is highly unlikely that a sound stops exactly at the moment that another sound begins. Our eyes and ears are therefore used to regard the edge that separates two visual or auditory perceptual segments as belonging to only one perceptual organization. If, for example, an edge could be part of two objects perceptually, then it would be impossible to see or hear which object is nearer. A famous visual illusion that demonstrates this is the face/vase illusion by Rubin, made in 1915, depicted in Figure 3. The borderline that separates the white and black surfaces can be perceptually allocated to either the white surface or the black surface. When the borderline is allocated to the white surface, we perceive a vase against a black background. When it is allocated to the black surface, however, the area for the vase becomes the background of two faces confronting each other. The borderline belongs to only a single perceptual organization at the same time, and that perceptual organization appears to be closer to us.

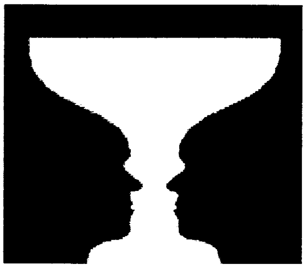

Figure 3. *The face/vase illusion, designed by Rubin in 1915. The line that separates the white and the black surface can be perceptually allocated to either the white surface (the vase) or the black surface (two faces).*

One of the famous slogans of the Gestalt psychologists was: the whole is not the sum of its parts. With this, they basically meant that the sensory elements interact nonlinearly in perception, and that the particular arrangement of sensory elements determines the identity of a perceptual organization (Pommerantz & Kubovy, 1986). The face/vase illusion shows this 'non-linearity' by demonstrating that our perceptual system can organize the same sensory data in more than one way. Because of the perceptual allocation of a line segment to one or the other potential perceptual organization in the face/vase illusion, we know that a visual scene is decomposed into sensory elements, such as stimulus edges. However, since the Gestalt psychologists mainly dealt with the holistic properties of visual or auditory scenes, they were basically concerned with patterns of arrangements of sensory elements. It was not until the middle of the twentieth century that a number of neurophysiological findings, and the development of the 'feature' theory of perception, generated more interest into the decomposition of visual and auditory scenes into sensory elements.

1.2 Feature detection

The 'feature theory' of visual perception proposes that objects in a visual scene are analyzed into various 'features', such as colors, textures, lines, and edges, which are then used as the basis for identifying objects (Barlow, 1999). These features are regarded as the building blocks of perceptual organizations. Selfridge and Neisser (1960) proposed the existence of a hierarchical system of 'feature detectors' that look for progressively more complex features, and combinations of features, in an initial image of external signals. Simple feature detectors would, for example, first detect an edge in a visual scene. Higher order feature detectors would be able to detect the orientation of the edge, and its length. This information would be passed on to feature detectors that could identify conjunctions of edges, making up shapes and sizes of certain objects, until a detector is able to identify a certain object as a whole.

Feature theorists regarded perception as the result of a series of processes that are achieved in simple neural circuits, and can be described in computational terms (Barlow, 1999). Their way of thinking was mainly driven by the findings of neurophysiologists who could identify the kind of information that neurons in the visual and auditory system extract from the signals arriving at the sense organs. Neurophysiologists found that almost all neurons in the visual nervous system respond to stimulus edges in the visual field. Some neurons respond more strongly than others, however, and the first feature detectors that could be identified were neurons that specifically responded to stimulus edges. Early research by Barlow (1953) showed the existence of simple and complex line and edge detectors in the nervous system of frogs. He showed the existence of neural compositions that function as simple edge detectors in the frog's eye, with detectors responding to borders between light and dark regions. He also found far more complex edge detectors that respond to small, dark objects moving into the field of vision of the frog. These so-called 'bug detectors' enable the frog to catch its food. The existence of line and edge detectors in the visual cortex of cats was demonstrated by a famous study of Hubel and Wiesel (1965). Studies on the visual cortex of monkeys have also shown that the decomposition of a visual scene is initiated by the detection of edges (Lamme et al., 1999). Line and edge detectors are also present in the human visual system (Burr et al., 1989). Nowadays, rather than trying to identify neural units that function as feature detectors, researchers of the visual system are concentrating themselves more on investigating how local edge elements can be grouped in global contours of objects in a visual scene and how such grouping can be predicted (Geisler et al., 2001). Moreover, researchers are investigating how the visual system is able to integrate ('bind') visual features if more objects are present in the visual field (Usher & Donnelly, 1998). Synchronization of the activity of interacting populations of neurons is assumed to be involved in a mechanism for feature integration in vision (Gray et al., 1992).

Feature detection theory also became recognized in the field of auditory perception. It was found that most neurons in the auditory nervous system respond strongly to onsets or offsets of sounds (Heil, 1997). Again, differences in the magnitude of responses were found, and the identification of the first specific feature detectors of the auditory system were neural units responding exclusively to onsets and offsets of sounds. Earlier work by Whitfield and Evans (1965) provided evidence for neural units in the cat's auditory cortex that respond to 'edges' of sounds, both in frequency and in time, with special neurons responding only to onsets or offsets of sounds, or both. They also found more complex neural compositions that respond to frequency-modulated tones, many with preferences for a certain upward or downward direction of the movement. Recent research has shown that neurons in the primary auditory cortex of the awake primate have complex patterns of sound-feature selectivity that indicate sensitivity to stimulus edges in frequency or in time, stimulus transitions in frequency or in intensity, and feature conjunctions (DeCharms & Merzenich, 1996; DeCharms et al., 1998). According to the authors, the data suggest 'that the cortex decomposes an auditory scene into component parts using a feature processing system reminiscent of that used for the cortical decomposition of visual images'. Researchers of the auditory system are also investigating how auditory features can be integrated into a unified auditory organization (Hall et al., 2000). Suggestions are made that, just like in the visual nervous system, the integration of auditory features involves synchronization of neural activity (Eggermont, 2000). However, the decomposition of an auditory scene and the integration of auditory features into auditory perceptual organizations cannot be demonstrated easily with methods other than neurophysiological ones. Until recently, no auditory stimulus patterns have been designed that clearly show that an auditory scene can be decomposed into auditory features, such as stimulus edges, and that these features can be integrated to a perceptual organization. By using auditory stimulus patterns consisting of crossing glide tones, however, Nakajima et al (2000) demonstrated a possible example of an auditory illusion that shows the integration of auditory features, the 'gap transfer' illusion.

1.3 The gap transfer illusion

Figure 4 shows the gap transfer illusion (Nakajima et al., 2000). The stimulus pattern consists of two successive, long frequency glides on a single continuous trajectory, that cross with a shorter, physically continuous glide. The two longer frequency glides are separated by a temporal gap of 100 ms, where they cross with the shorter glide. Contrary to what one might expect, listeners typically do not perceive a gap between the two longer frequency glides, but perceive a continuous long pitch trajectory. The physically short glide, on the other hand, is heard as a short pitch trajectory with a gap. The transfer of the gap can be perceived both when the long frequency glides ascend and the short glide descends, and when the directions of the glides are reversed.

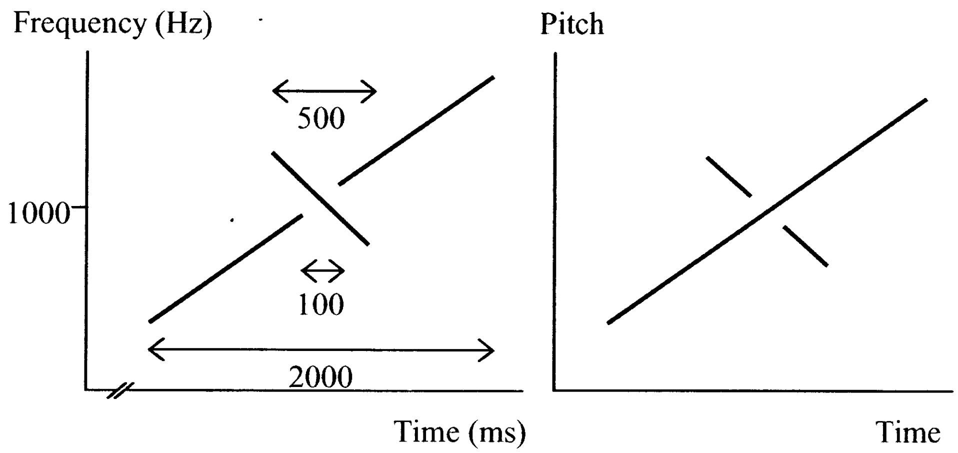

Figure 4. *Stimulus pattern that renders the 'gap transfer illusion' (left plane). Although a short temporal gap is physically present between two long, ascending glides, a single, continuous ascending pitch trajectory is heard. On the other hand, the physically continuous, short glide is perceived as a descending pitch trajectory with a short gap (right plane).*

The mechanism behind the appearance of the gap transfer from the long to the short pitch trajectory was described in terms of the 'event construction model' of auditory organization as proposed by Nakajima et al. (2000). This model is based on the rudimentary idea that the auditory system treats onsets and offsets of sounds as independent auditory elements, and that these elements are perceptually organized according to their proximity in time and frequency. The model regards onsets and offsets of sounds as the main 'features' of sounds, and presumes that these features can be perceptually integrated according to the Gestalt principle of proximity.

Evidence for the integration of stimulus edges according to their proximity can be found in visual perception. Research on the computation of contour information has shown that the integration of stimulus edges into global contours and visual objects can be influenced by the spatial proximity of the stimulus edges (Boucart et al., 1994). Furthermore, it is known that when two visual patterns are presented in rapid succession, their contours may be combined into a single, unified percept. The proximity of the stimulus edges, amongst others, affects the temporal integration of the contours (Visser & Enns, 2001). The temporal proximity between sensory elements may also influence their grouping into a single perceptual organization. Contour integration, for example, is facilitated when sensory elements that perceptually constitute a contour are presented simultaneously, and are temporally separated from sensory elements comprising a background (Usher & Donnolly, 1998; Beaudot, 2002).

Just like in vision, according to Nakajima et al. (2000), stimulus edges of sounds can also be integrated according to their proximity in time and frequency. The gap transfer illusion shows that proximity may cause 1) the perceptual connection of the onset of the short glide to the offset of the first long glide, and 2) the perceptual connection of the onset of the second long glide to the offset of the short glide This results in the perception of the two short tones (Figure 5).

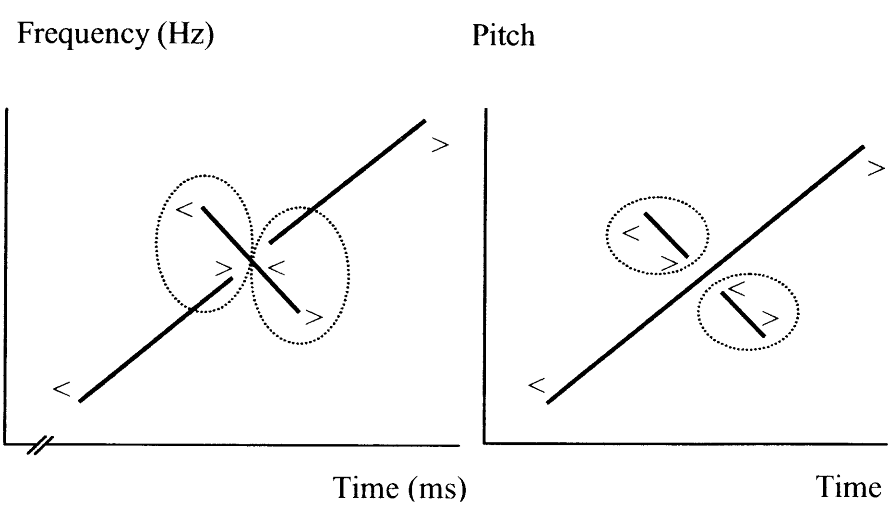

Figure 5. *The event construction model. In the gap transfer illusion, the perception of a gap in the short descending pitch trajectory (right plane) is caused by the perceptual connection of the onset (<) and the offset (>) of the short descending glide to, respectively, the offset of the first, and the onset of the second ascending glide (left plane). It was suggested that the perceptual integration of the onsets and offsets of acoustically different sounds occurs because of their proximity in time and frequency (Nakajima et al., 2000).*

As a result of the perceptual construction of the two short tones, the long glides around the temporal middle are left without an offset or an onset. This facilitates the perception the long, continuous pitch trajectory. If this hypothesis were correct, the gap transfer illusion could demonstrate the decomposition of an auditory scene into features such as stimulus edges, and the perceptual integration of these stimulus edges into perceptual organizations. This grouping can occur even when the stimulus edges belong to different sounds acoustically.

The idea that stimulus edges such as onsets and offsets can perceptually break off from the stimuli to which they physically belong was demonstrated earlier by Steiger (1980). He demonstrated that when a repetitive series of frequency glides is presented, the onsets of the glides perceptually segregate, and together form an auditory organization of repetitive short clicks (Figure 6). Bregman (1990), in explaining this demonstration, mentioned that proximity in time and frequency between the onsets could have caused their perceptual segregation.

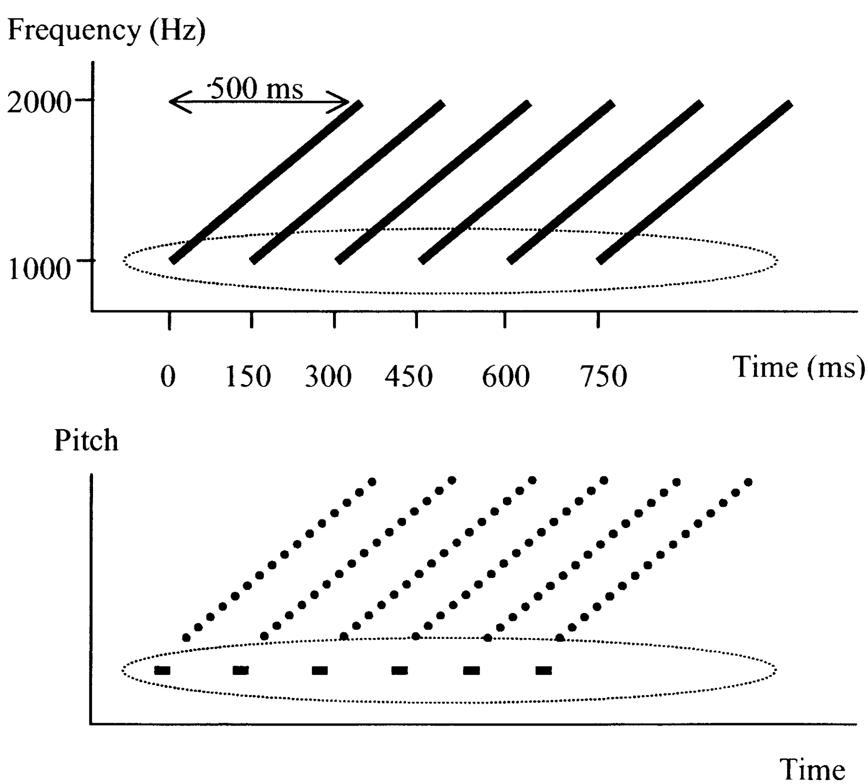

Figure 6. *An example of a stimulus pattern investigated by Steiger (1980), consisting of a series of identical glides played in rapid succession (upper plane). Often, the onsets of the glides perceptually segregate from the glides to which they physically belong, and a series of rhythmic short tones can be perceived (lower plane).*

The gap transfer illusion differs from Steiger's demonstration, however, in that it seems to show the perceptual segregation of both onsets and offsets. Moreover, it may show their integration into a perceptual organization consisting of only a single auditory event, a short tone with a salient beginning and a salient end. The gap transfer illusion may also demonstrate that, if an onset or an offset perceptually segregates from a sound, the remainders of the sounds can be perceptually integrated into a perceptual organization as well. The perception of the remainders of the glides in Steiger's demonstration (1980) is not clearly described, with regard to whether or not the remaining glide components still have a perceptual beginning, and whether or not the remaining glide components are also grouped into a single perceptual organization.

Following the demonstration of the gap transfer illusion, the hypothesis that proximity in time and frequency between onsets and offsets of sounds could result in their perceptual integration, even though they are acoustical parts of different sounds, was tested in a different paradigm. A typical example of the paradigm is depicted in Figure 7. The stimulus pattern consists of two parallel glides of 1400 ms, both either ascending or desceding, that partly overlap each other for 200 ms in the temporal middle of the stimulus pattern.

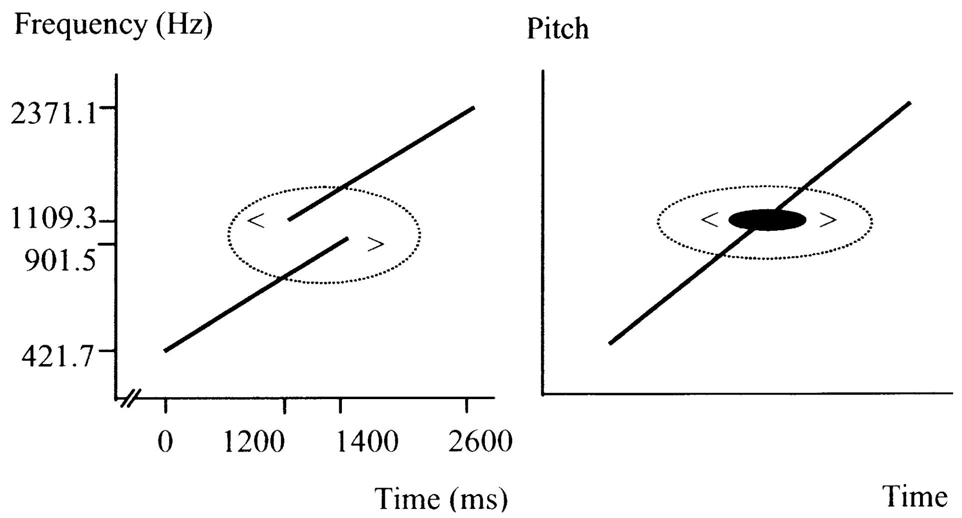

Figure 7. *A stimulus pattern consisting of two partly overlapping glides (left plane). Proximity in frequency and time between the onset of the second glide (<) and the offset of the first (>) leads to their perceptual connection, and the perception of a short tone. This 'middle tone' can be perceived in the temporal middle of a long, continuous pitch trajectory, that is perceptually composed of the remaining components of the glides (Nakajima et al., 2000).*

According to the hypothesis of Nakajima et al. (2000), proximity between the onset of the second glide and the offset of the first glide could lead to their perceptual connection, and the perception of a short tone. This short tone would be illusory, since the stimulus pattern consists of two physically separated, relatively long glides. They further hypothesized that the remaining components of the glides, as a result of the perceptual connection of the offset of the first glide to the onset of the second glide, would be perceptually connected as well. Left without an offset and an onset, respectively, the remaining components together could form a perceptual organization consisting of a single, long pitch trajectory. The results of a phenomenological experiment, in which three observers verbally described their perception of the stimulus pattern, indeed showed that a single long pitch trajectory could be perceived, along with a short tone in the temporal midlde of the long pitch trajectory (Nakajima et al., 2000). Throughout this thesis, the short tone is referred to as the 'middle tone'.

1.4 The aim of this thesis

The aim of this thesis is to investigate, by use of phenomenological and psychophysical methods, whether auditory stimulus edges of physically different sounds, such as onsets and offsets, can be perceptually integrated to construct auditory organizations. The perceptual integration of auditory stimulus edges of physically different sounds would show that an auditory scene is decomposed into auditory features such as stimulus edges, in a similar way the visual system decomposes a visual scene into stimulus edges. The matter will be investigated by using stimulus patterns that consist of parallel glides that partly overlap each other.

The appearance of illusory auditory events composed of onsets and offsets of physically different sounds can be investigated more easily with these simplified stimulus patterns, than with those that render the gap transfer illusion. The stimulus patterns that render the gap transfer illusion also consist of partly overlapping glides, which, however, are crossing each other. In stimulus patterns that cause the gap transfer illusion, often some simultaneous components of the two crossing glides move within the same critical bandwidth. The critical bandwidth is 'a range of frequencies that surrounds the frequency of a designated pure tone. When other pure tones whose frequencies are within this bandwidth are played at the same time as the designated tone, the auditory system does not hear the two completely independently' (Bregman, 1990). This interaction may influence the perceptual connection of stimulus edges that are close together in frequency and in time, assumed to underlie the perception of the gap transfer illusion and the perception of the two partly overlapping parallel glides (Nakajima et al., 2000). In the latter stimulus patterns, however, the influence of the interaction of the overlapping glides at the peripheral level of the auditory system on the perception of the stimulus patterns can be investigated more easily. The frequency separation between the parallel glides can be made in such a way that the two glides do not move within the same critical band. Next, in these stimulus patterns, only one overlap exists, delimited by one onset and offset, which enables the investigation of the frequency relation between these stimulus edges more easily. Manipulations of the overlap duration of the glides can also clarify the possible effect of temporal proximity between the stimulus edges on the perception of the stimulus pattern. The stimulus patterns further allow simple manipulations of the slope of the glides.

1.5 Contents of this thesis

In Chapter 2, phenomenological descriptions of stimulus patterns consisting of two partly overlapping parallel glides are described. In two experiments, participants gave verbal and graphical descriptions of their perception of the stimulus patterns. In these experiments, the influence of critical bands on the perception of the stimulus patterns was investigated by presenting stimulus patterns with increasing frequency difference between the overlapping glides. In Chapter 3, psychophysical experiments are described that investigated whether peripheral processes other than those concerned with critical bands could be responsible for the perception of the stimulus patterns. Two psychophysical experiments were done in order to investigate whether or not the perception of spectral splatter at the onset or offset that delimit the overlap influenced the perception of the stimulus patterns. In the same experiments, it was investigated whether the perception of a combination tone, or tones, influenced the perception of the stimulus patterns. Two psychophysical experiments are described in Chapter 4, which had the aim of investigating the influence of the temporal and frequency proximity of the stimulus edges that delimit the overlap on the perception of the stimulus patterns. In Chapter 5, rather than concentrating on the overlap of the stimulus patterns, the perception of the remainders of the glides was investigated in three psychophysical experiments. The perception of stimulus patterns as investigated here was compared with those that render a typical auditory continuity effect (see Chapter 1.1) in Chapter 6. Chapter 7 summarizes the findings and contains the conclusions of this thesis.