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 2: Phenomenological descriptions of the perception of two partly overlapping glides.

2.1 General purpose

The general purpose of the following two experiments is to investigate whether stimulus patterns consisting of two partly overlapping glides can be perceived as consisting of a long pitch trajectory, accompanied by a short tone in its temporal middle. Stimulus patterns are investigated with the use of phenomenological methods. Phenomenological observation was the experimental approach of the original Gestalt psychologists. The method is simple, in that the observer is asked to view or listen to a stimulus pattern and to describe its apparent organization, either verbally or graphically. The observers is free to comment on any aspect of the stimulus pattern, and is encouraged to report his/her perception of the stimulus pattern as detailed as possible (Pomerantz & Kubovy, 1986). Observers asked to express their perception of the stimulus pattern can comment on the global nature of the stimulus pattern, as well as aspects of it, and can give descriptions of more than one percept in case they can perceive the stimulus pattern in more than one way. Because just three observers described their perception of the stimulus patterns in Nakajima et al. (2000), at first, it seemed necessary to gather more detailed information about the perception of the stimulus patterns. For this purpose, the phenomenological method seemed more appropriate than a psychophysical method, since the latter are often restricted to the investigation of only certain aspects of a stimulus pattern.

The main parameter changed in the stimulus patterns in the following two experiments is the instantaneous frequency separation between the overlapping glides. The instantaneous frequency separation refers to the frequency separation between the glides at any given point in time of the overlap. Stimulus patterns with an instantaneous frequency separation of 0.3 octave, for example, consist of two overlapping glides that are separated by 0.3 octave during any given point in time of their overlap. Of special interest is how the stimulus patterns are perceived when the instantaneous frequency separation between the glides is smaller than a critical bandwidth. Two frequency components that are placed within a critical bandwidth interact (Moore, 1997). The interaction of two overlapping glides that are moving in one critical bandwidth may therefore influence the perception of a middle tone and a long pitch trajectory.

2.2 Experiment 1: Verbal descriptions of the perception of stimulus patterns

consisting of two partly overlapping glides.

2.2.1 Purpose.

The purpose of Experiment 1 was to get detailed reports about the perception of stimulus patterns consisting of two partly overlapping glides, and to investigate the role of the instantaneous frequency separation between the glides on the perception of the stimulus patterns.

2.2.2 Method.

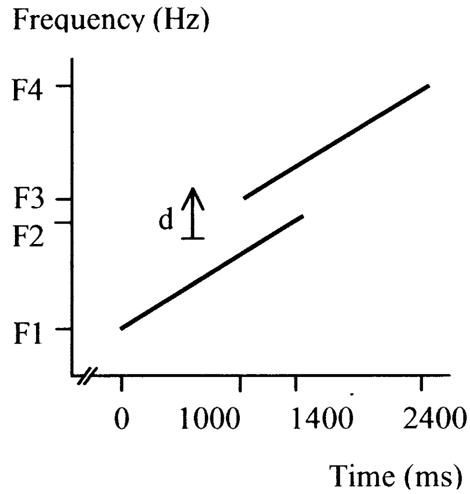
Stimulus patterns

A total of eight stimulus patterns, two single-glide and six double-glide stimulus patterns, was generated. The double-glide stimulus patterns consisted of two pure frequency components that moved in logarithmic frequency, and partly overlapped each other (Figure 8). The duration of each glide was 1400 ms. Since the second glide started 400 ms before the first glide's offset, the glides overlapped each other from t = 1000 ms through t = 1400 ms from the onset of the first glide (t = 0). The rise time of the first glide and the fall time of the second glide was 500 ms. The fall time of the first glide and the rise time of the second glide was 4 ms. The rise and fall times of the glides had linearly shaped ramps. The total duration of the six double-glide stimulus patterns was 2400 ms.

In order to categorize the phenomenological descriptions, only a single type of stimulus patterns was used. It consisted of two ascending glides, with the second glide starting at a frequency that was higher than the frequency of the first glide at that point in time. Descending stimulus patterns and patterns in which the second glide started at a lower frequency than that of the first glide at that point in time were not used.

At t=1200 ms from the onset of the first glide, a frequency of 1000 Hz was chosen from which the instantaneous frequency separation between the glides was calculated. The instantaneous frequency separation between the glides was 0.1, 0.3, 1.3, and 2.3 octaves, respectively. The slope of the glides, or the increase or decrease in frequency by time, was one octave per second. This resulted in glides spanning middle range frequencies, implicating that each stimulus pattern had, approximately, an overall equal loudness.

Two single-glide stimulus patterns were generated. One stimulus pattern comprised of the first glide, and one of the second glide of the double-glide stimulus pattern with the 0.1-octave frequency separation. The duration of these control stimulus patterns was 1400 ms.



St.	d (oct.)	F1	F2	F3	F4
1	0.1	420.5	1109.6	901.3	2378.4
2	0.3	392.3	1035.7	965.9	2549.1
3	0.8	329.9	870.5	1148.7	3031.4
4	1.3	277.4	732.0	1366.0	3605.0
5	1.8	233.6	615.6	1624.5	4287.1
6	2.3	196.2	517.6	1931.9	5098.2

Figure 8. *Stimulus patterns employed in Experiment 1. The horizontal axis represents time, whereas the vertical axis represents frequency on a logarithmic scale. The instantaneous frequency separation (d) in the six double-glide stimulus patterns) was 0.1, 0.3, 0.8, 1.3, 1.8, and 2.3 octaves. The frequencies of the double-glide stimulus patterns are depicted in the table on the right.*

Apparatus

The stimulus patterns were generated by a computer (Teac PS 9000-216), low-pass filtered at 7 kHz (NF DV-04), and recorded on DAT (Digital Audio Tape) with a DAT deck (Sony DTC-M100). This experiment was partly conducted in Fukuoka, (Japan), and in Leiden (the Netherlands). In Fukuoka, the stimulus patterns were presented off-line via a DAT deck (Sony 500 ES), an amplifier (Sansui AU 607 KX), via headphones (Rion AD-02) to the observer's left ear. In Leiden, the Netherlands, the stimulus patterns were presented via a DAT deck (Sony XD-Z505), and headphones (Telephonics TDH 39). The sound level of the stimulus patterns was 67 dBA on average both in Japan and in the Netherlands, as measured (fast-peak) by a sound level meter (Brüel & Kjær 1613), and an artificial ear (Brüel & Kjær 4152), mounted with a microphone (Brüel & Kjær 4134).

Observers

Three males of 23, 39, and 52 years old, all with normal hearing, participated in the experiment. Two of the observers were Dutch and one was Japanese. They were all studying in the field of auditory perception and had previous experience in participating in auditory experiments.

Procedure

The eight stimulus patterns were randomized and monaurally presented to the observer's right ear in a sound proof booth. Each stimulus pattern was presented 45 times, with a silence of 5 seconds between each presentation. The observer was asked to give verbal descriptions of his perception of the stimulus pattern during the repeated presentation of each stimulus pattern, and was able to hear the same stimulus pattern as many times as necessary. The observer was asked to describe his perception of each stimulus pattern as detailed as possible, and to describe different modes of perception of the same stimulus pattern if perceived. Care was taken not to impose any restraints on the observer's responding. When the observer indicated that he was satisfied with his description of a stimulus pattern, the next 45 presentations of a following stimulus pattern were presented. In Japan, the reports were given in English. The observer participating in the Netherlands verbalized his perception of the stimulus patterns in Dutch. The observer could take a break whenever wanted. The experiment lasted about one hour.

2.2.3. Results and discussion

The two single-glide stimulus patterns were described exactly according to their physical properties by all observers. The stimulus pattern in which the instantaneous frequency separation between the glides was 0.1 octave, however, was not perceived veridically. Rather, the stimulus pattern consisted, according to all observers, of one long, pure, ascending pitch trajectory and a relatively short tone. The long pitch trajectory was heard as continuous. The short tone was perceived as rough and steady-state, and it was perceived in or near the temporal middle of the long pitch trajectory. In the following, this tone will therefore be referred to as the 'middle tone'. The stimulus pattern with the 0.3-octave frequency separation between the glides was also perceived as consisting of a long, pure, ascending pitch trajectory, accompanied by a middle tone. The long pitch trajectory was heard as continuous. The middle tone was judged as less rough compared with the 0.1-octave condition. Two of the three observers also heard the middle tone at least partially ascending in pitch. According to the same observers, this pitch was lower than the pitch of the temporally corresponding part of the long ascending pitch trajectory.

The stimulus pattern with the 0.8-octave frequency separation between the glides was perceived differently by all three observers. One observer reported that the stimulus pattern consisted of a long, ascending pitch trajectory and a middle tone. The long pitch trajectory was heard as continuous, and the middle tone was said to be close to a pure tone. One observer also perceived a middle tone and a long pitch trajectory. The ascending pitch trajectory, however, was not judged as completely continuous. The third observer did not perceive a continuous long pitch trajectory, but perceived two ascending pitch trajectories. The beginning of the second pitch trajectory was heard as non-pure, and could, with some effort, be heard as an independent short tone.

The stimulus patterns with an instantaneous frequency separation of 1.3 octaves between the glides was perceived as consisting of two ascending pitch trajectories, according to all observers. The onset of the second long pitch trajectory was perceptually standing out, according to two observers after listening to the stimulus patterns many times, in that it was more salient than the remainder of the second pitch trajectory, and the first pitch trajectory. This onset could not be heard as an independent short tone.

Quite the same description was given to the stimulus patterns with the 1.8-octave and the 2.3-octave frequency difference between the glides. Two pure, long, ascending pitch trajectories were heard by all observers. According to two observers, the onset of the second pitch trajectory was perceived as more salient than the remainder of the second pitch trajectory, and the first pitch trajectory, when the frequency separation between the glides was 1.8 octaves. The third observer, however, could hear the offset of the first pitch trajectory as more salient than the remainder of the first pitch trajectory, and the second pitch trajectory. In the 1.8-octave condition, and especially in the 2.3-octave condition, the two long pitch trajectories could both be scanned from beginning to end. Observers reported that they could hear the beginning of the second pitch trajectory while the first still continued. In these conditions, a middle tone could not be perceived.

According to the phenomenological reports, the perceptual impressions of the double-glide stimulus patterns can be categorized as follows (Figure 9):

(A) A long ascending pitch trajectory was heard, accompanied by a short tone in or near the temporal middle of the pitch trajectory. The long pitch trajectory was perceptually continuous. The short tone was heard as an independent tone, with an onset and an offset.

(B) Two successive or partly overlapping, ascending pitch trajectories were heard, accompanied by a short tone at or near the shift from the first pitch trajectory to the second pitch trajectory. The short tone was heard as an independent tone, with an onset and an offset.

(C) Two partly overlapping, ascending pitch trajectories were heard. The onset of the second pitch trajectory was perceived as more salient than the remainder of the percept. A middle tone was not perceived.

(D) Two partly overlapping, ascending pitch trajectories were heard. The offset of the first pitch trajectory was perceived as more salient than the remainder of the percept. A middle tone was not perceived.

(E) Two partly overlapping, ascending pitch trajectories could be perceived. They had no particularly salient parts. A middle tone was not perceived.

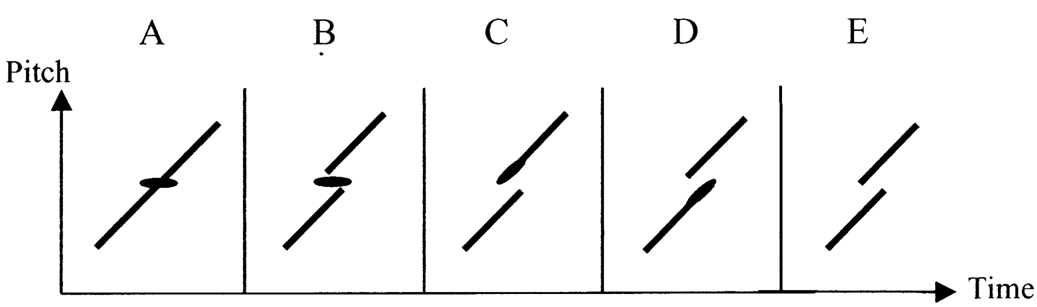


Figure 9. *Schematic representations of the perceptual impressions of the double-glide stimulus patterns obtained in Experiment 1. The perceptual impressions were categorized in five perceptual modes A-E.*

The appearances of the five perceptual modes obtained in Experiment 1 are shown in Table 1. All the observers indicated that their perception changed while listening to particular stimulus patterns for a long time, or that they could switch between two perceptual modes. A change from one perceptual mode to another, in which the latter mode became the dominant one, is indicated by '=>'. For example, 'A=>B' indicates a case in which Mode A was perceived at first, but Mode B became the dominant percept after repeated listening to the same stimulus pattern. Cases in which an observer could switch between perceptual modes, with both modes equally prominent, are indicated by '/'. For example, 'A/B' indicates a case in which the observer could perceive both modes A and B, without being able to report the dominant mode.

Mode A represents the perception of two partly overlapping glides as described by Nakajima et al. (2000). The appearance of Mode A shows that two acoustically different glides of equal duration can be perceived non-veridically. Instead of two pitch trajectories of equal duration, a single, long pitch trajectory is heard, along with a short tone in or near the temporal middle of the pitch trajectory. No mode was mentioned in which a long, continuous pitch trajectory could be perceived without a middle tone. This may suggest that the appearance of the single pitch trajectory depends on the perception of the middle tone, as proposed by Nakajima et al. (2000).

Table 1. *Appearances of the five perceptual modes obtained in Experiment 1. A change from one perceptual mode to another, with the latter one becoming the dominant one, is indicated by '=>'. Cases in which an observer could perceive two perceptual modes, and switch between them without being able to mention a dominant mode, is indicated by '/'. In the table, 'd' indicates the instantaneous frequency separation (in octaves) between the glides.*

St.	d (oct.)	Observer 1	Observer 2	Observer 3
1	0.1	A	A	A
2	0.3	A	A	A
3	0.8	A	A=>B	C
4	1.3	E=>C	C	E
5	1.8	E/C	C	E/D
6	2.3	E	E	E

The perception of a single pitch trajectory and a middle tone disappeared when the instantaneous frequency separation between the glides became larger. The instantaneous frequency separation between the glides thus influenced the qualitative nature of the percept. The interaction of the two overlapping components within a critical bandwidth, however, did not seem essential for the perception of the middle tone and the long pitch trajectory. Although the perception of the stimulus patterns changed when the instantaneous frequency separation between the glides became larger, Mode A was also perceived when the glides were separated by more than a critical bandwidth. At a reference frequency of 1000 Hz, the critical bandwidth is 160 Hz, interpolated from the data described by Zwicker and Fastl (1999). The equivalent rectangular bandwidth (ERB), another measure of a critical bandwidth, is even smaller (133 Hz) at the reference frequency of 1000 Hz (Glasberg & Moore, 1990). The results of Experiment 1 show that Mode A could be perceived when the instantaneous frequency separation between the glides was 0.1 or 0.3 octave. Two observers could hear Mode A even when the glides were separated by 0.8 octave. When separated by an instantaneous frequency separation of 0.3 octave or more, the glides are separated by 208 Hz, which is larger than a critical bandwidth, or ERB. A frequency separation of 0.8 octave is considerably larger (561 Hz) than a critical bandwidth, or ERB, of the reference frequency. The matter is further investigated in Experiment 2.

2.3 Experiment 2: Graphical descriptions of the perception of stimulus patterns consisting of two partly overlapping glides.

2.3.1 Purpose

In Experiment 2, the appearance of the five perceptual modes of the double-glide stimulus patterns as used in Experiment 1 was further investigated. The role of the instantaneous frequency separation between the glides on the perception of the stimulus patterns was investigated with a phenomenological method, in which participants were asked to draw their perceptual impressions of the stimulus patterns on paper.

2.3.2 Method

Stimulus patterns and apparatus

The stimulus patterns and apparatus used in Experiment 2 were the same as those used in Experiment 1.

Participants

Five participants, one female and four males, took part in the experiment. All participants had normal hearing, and were studying in the field of auditory perception. Three participants were Japanese, and two were Dutch. The Japanese participants had received basic training in music, and technical listening training for acoustic engineers.

Procedure

The eight stimulus patterns were each recorded 45 times on DAT tape, with a silence of five seconds between each presentation. In a sound proof booth, the stimulus patterns were randomly presented to the participant's left ear. The participant was asked to draw his/her perceptual impression of each stimulus pattern on a paper sheet, and to give written explanations of his/her drawing. The participant could listen to the same stimulus pattern as many times as he/she wanted and adjust his/her drawing of the percept until satisfied with the result. The only restriction was to draw the percept in a small cadre, and to consider the horizontal axis as 'time' and the vertical axis as 'pitch'. If a stimulus pattern was heard in different perceptual modes, the participant was instructed to draw one percept, give written explanations of the other(s), and indicate the order of dominance between the percepts, if any. At the beginning of the experiment, the participant was instructed just to listen to the eight randomized stimulus patterns four times. After that, three warm-up trials were given. Instructions were given in English. The experiment lasted about one hour.

2.3.3 Results and discussion

The two stimulus patterns that consisted of a single glide were depicted as such by all observers. The drawings of the double-glide stimulus patterns showed that no other perceptual modes were perceived than those reported in Experiment 1. Table 2 shows the frequency of appearance of Modes A, B, C, and E in Experiment 2. Mode D did not appear in Experiment 2. Since the drawings of Experiment 2 rendered additional information to the descriptions of the perceptual modes of Experiment 1, the perceptual modes are discussed in more detail in the following.

The results show that, similar to the phenomenological reports of Experiment 1, Mode A could be perceived when the instantaneous frequency separation between the glides was 0.1, 0.3, or 0.8 octave, but was not perceived when the frequency separation between the glides was larger. The instantaneous frequency separation between the glides thus influenced the perception of a middle tone and a long, continuous pitch trajectory. The results show, however, that Mode A was not mentioned at all by two observers. On the other hand, three observers could perceive a middle tone and a continuous (Mode A) or discontinuous pitch trajectory (Mode B) when the instantaneous frequency separation was even larger than a critical bandwidth (0.3 or 0.8

octave).

Table 2. *Frequencies of appearance of the perceptual modes obtained in Experiment 2 with regard to the double-glide stimulus patterns. A change from one perceptual mode to another, with the latter one becoming the dominant one, is indicated by '=>'. Cases in which an observer could perceive two perceptual modes, and switch between them without being able to mention a dominant mode, are indicated by '/'.*

St.	d (oct.)	Obs1	Obs2	Obs3	Obs4	Obs5
1	0.1	A	A	C	B	A
2	0.3	A	A	C/B	B	A
3	0.8	A=>B	A	C	C	A
4	1.3	C	B	E	B/C	E
5	1.8	C	E	E	C	C=>E
6	2.3	E	E	E	E	E

Mode E corresponds to the veridical solution of the stimulus pattern and was, for every observer, the dominant percept when the glides were separated by the largest frequency separation. Mode C also consists of two separate pitch trajectories of equal duration, but differs from Mode E in that the beginning of the second pitch trajectory was perceptually more salient than the remainder of the second pitch trajectory, and the first pitch trajectory. Mode D, mentioned in Experiment 1, did not appear in Experiment 2. Mode D differs from Mode C in that the offset of the first pitch trajectory was perceived as more salient than the remainder of the first pitch trajectory, and the second pitch trajectory. Mode D was mentioned by one observer for one condition in Experiment 1.

The frequency of appearance of Mode C in Experiment 1 (Table 1) and especially in the present experiment shows that the onset of the second pitch trajectory was relatively salient in quite a number of cases. Only one observer did not report Mode C. The appearance of Mode C may be related to research done by Bregman and Ahad (1994). They studied the effects of suddenness of onsets on the discrimination of the order of pitches of individual tones in a 1 s, 4-tone cluster of overlapping pure tones. They found that when the tones had asynchronous onsets and were turned on with sudden onsets, the onset of the tone that was turned on would stand out perceptually from the cluster. They argued that the auditory system restarts a pitch-computing process when a sound joins an ongoing spectrum with a sudden onset, and that the onset of the added sound would contribute dominantly to the new pitch. Mode C seems to represent a similar case in which the onset of the second glide, joining the first glide with a sudden onset, contributes dominantly to the pitch of the percept at that point in time. Bregman and Ahad (1994) also found that when the onsets of the 4-tone cluster were synchronous but the offsets asynchronous, the effects of suddenness on the saliency of the offsets were relatively small. Discrimination of the pitches of the offsets in general was more difficult compared with discrimination of onsets. Pitches of offsets seem less likely to stand out perceptually from a cluster of tones and to contribute to a process of pitch computation than onsets. The finding that Mode C appeared much more often than Mode D in Experiments 1 and 2 may be in line with this.

Mode B represents a percept in which a middle tone is perceived along with a pitch trajectory that has a discontinuity in pitch and/or in time. According to the hypothesis of Nakajima, the perceptual construction of a single, long pitch trajectory is facilitated when the onset of the second and the offset of the first glide perceptually form the middle tone. The two glides are left without an offset and an onset, respectively, which should promote their perceptual integration. Mode B, however, shows that the perception of a middle tone does not guarantee the perception of a continuous pitch trajectory. The perception of the long pitch trajectory will be further investigated and discussed in Chapter 5 of this thesis.

2.4 Summary

The results of Experiments 1 and 2 showed that two partly overlapping glides can be perceived as consisting of a long pitch trajectory, accompanied by a middle tone. Although two of the eight participants did not perceive a long, continuous pitch trajectory, and one of them did not perceive a middle tone at all in the stimulus patterns, six participants perceived the percept robustly. When the instantaneous frequency separation between the glides became smaller, the frequency of appearance of this percept increased. However, the percept appeared also when the instantaneous frequency separation between the glides was rather large, i.e., more than a critical bandwidth. This shows that the interaction of the two overlapping glides within a critical bandwidth is not essential for the perception of a middle tone and a long pitch trajectory.

The perceptual modes showed that a middle tone could be perceived without a single, long pitch trajectory. The perception of a single pitch trajectory, therefore, cannot be a prerequisite for the perception of a middle tone. On the other hand, a long pitch trajectory was never perceived without a middle tone. The middle tone therefore seems more robust, and may be a prerequisite for the appearance of the long pitch trajectory. Because of this, the nature of the middle tone is investigated in Chapter 3. Although the critical band concept does not seem to play an essential role in the perception of the middle tone, other peripheral processes may. It seems necessary to check the possibility that the middle tone may be the result of a combination tone, or tones, or even the result of spectral splatter at the onset of the second glide, or the offset of the first glide, due to the short, linearly raised rise and fall times that were used in Experiments 1 and 2.