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 5: Psychophysical investigation of factors involved in the perception of the pitch trajectory in double-tone stimulus patterns.

5.1 General purpose

In Experiments 1 and 2, the perceptual modes showed that only when a middle tone was perceived, a long pitch trajectory could be perceived. This was in line with the hypothesis of Nakajima et al. (2000), discussed in the Introduction. They hypothesized that the perception of the long pitch trajectory depends on the perceptual construction of the middle tone. The perceptual integration of the first and second glide may be facilitated when the offset of the first glide and the onset of the second glide are perceptually integrated to constitute the middle tone. This hypothesis implies that the factors that influence the perception of the middle tone, as found in Experiment 5, should also have an influence on the perception of the long pitch trajectory in the double-glide stimulus patterns. The nature of the long pitch trajectory is investigated in the three following psychophysical experiments, in which participants were asked to judge if, and how clearly, they could hear a single, continuous pitch trajectory in stimulus patterns consisting of two partly overlapping frequency components.

5.2 Experiment 6: Judging the continuity of the pitch trajectory.

5.2.1 Purpose

Variations in the instantaneous frequency separation, the slope, and the overlap duration of two partly overlapping frequency components were made in order to investigate the influence of these factors on the perception of the long pitch trajectory in the double-glide stimulus patterns.

5.2.2 Method

Stimulus patterns

All stimulus patterns consisted of two partly overlapping, ascending glides of 1200 ms each. The rise time of the first glide was 200 ms, and the fall time 4 ms, with linearly shaped ramps. Conversely, the rise time of the second glide was 4 ms, and the fall time 200 ms. The glides could overlap each other for 100, 200, 400, or 800 ms, making the total duration of the stimulus patterns 2300, 2200, 2000, and 1600 ms, respectively. The slope of the glides was 0.5, 1.0, or 1.5 octaves per second.

The frequency separation between the overlapping glides was 0.5, 0.75, 1.0, 1.5, or 3.0 times a critical bandwidth of the reference frequency, from which the stimulus patterns were calculated. The reference frequency between the two components at the temporal midpoint of each stimulus pattern was fixed at 1600 Hz. The critical bandwidth at this point is 240 Hz, as interpolated from the data reported by Zwicker and Fastl (1999). The ERB at this reference frequency is 197.4 Hz (Glasberg & Moore, 1990). In this experiment, the value of 240 Hz was used to calculate the frequency separation between the overlapping glides. For each stimulus pattern, the frequency separation value was divided by two, and this value was both added to and subtracted from the reference frequency. This resulted in two frequency points at the temporal middle of the stimulus pattern, from which the first and the second glide were calculated linearly on a log scale. Combining the variations in overlap duration, slope, and frequency separation, a total of 4 x 3 x 5 = 60 double-glide stimulus patterns was generated.

Apparatus

The stimulus patterns were generated by a computer (16 bit, sampling frequency 22050 Hz) with a D/A converter (TEAC PS 9353), low-pass filtered at 7 kHz (NF Electronic Instruments DV-04), and recorded on a DAT (Digital Audio Tape) with a DAT deck (Sony 500-ES) in random order. The stimulus patterns were presented to the participant in a sound proof booth via a DAT deck (Sony 500-ES) and an amplifier (Sansui 607 KX) through headphones (Rion AD 02), monaurally to the left ear. The sound level of the stimulus patterns maximally reached 74 dBA, as measured (fast-peak) with a sound level meter (Brüel & Kjær 2209), and an artificial ear (Brüel & Kjær 4152), mounted with a microphone (Aco 7023).

Participants

Six students, two females and four males, took part in the experiment. They had received basic training in music, and training in technical listening for acoustic engineers. They were 21 - 25 years of age. The participants had normal hearing.

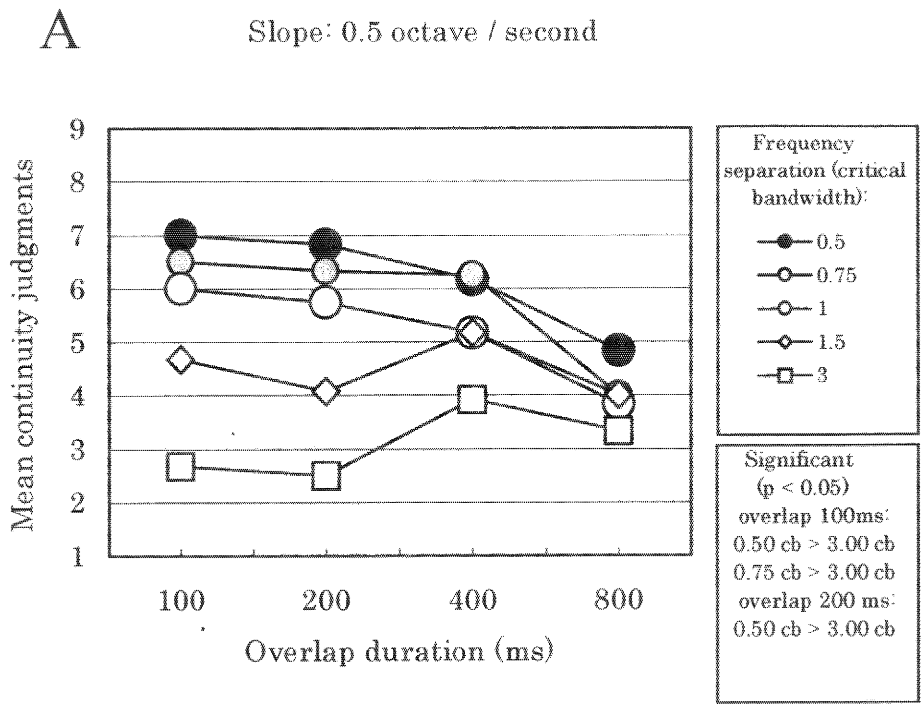
Procedure

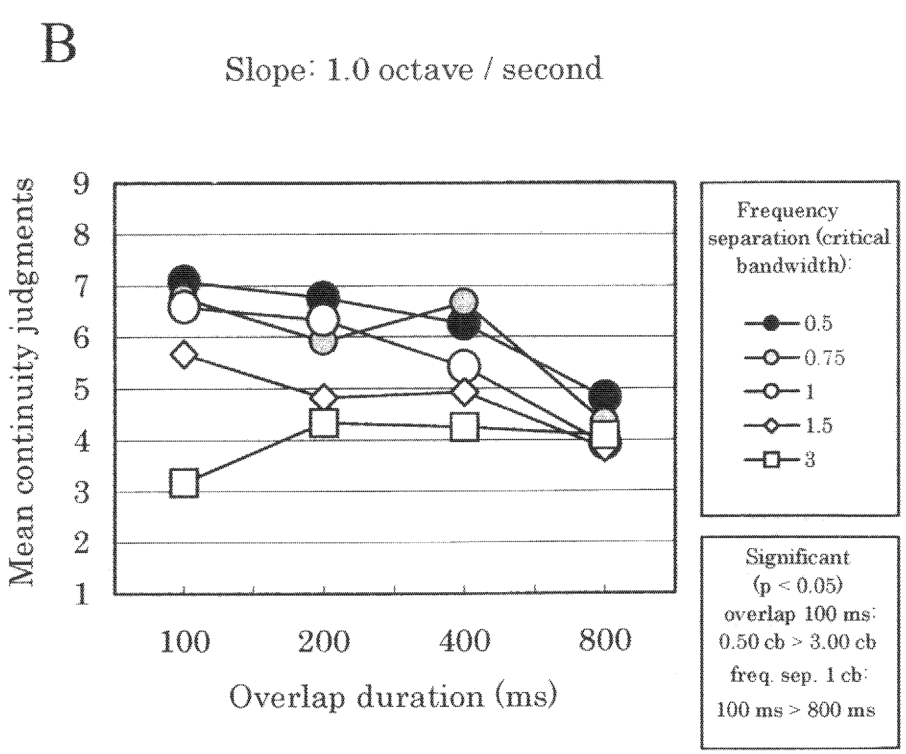
The participant was asked to judge the continuity, in pitch and in time, of the ongoing pitch trajectory of each stimulus pattern. The judgments were made on a 9-point rating scale. The right extreme of the scale, the number '9', corresponded to an ongoing pitch trajectory that was completely continuous, i.e., did not stop to exist in time and was coherent with respect to pitch and loudness. The left extreme of the scale, the number '1', indicated that the pitch trajectory was discontinuous, i.e., consisted of two successive sounds. The participant received two differently randomized sessions, each consisting of the 60 stimulus patterns in random order, and performed four warm-up trials before starting each session and after each break. In each trial, a stimulus pattern was presented three times, with silent intervals of 3 s between each presentation. The silent time between consecutive trials was 10 s. The participant was asked to give his/her judgment of the continuity of the pitch trajectory during the silent intervals between presentations or during the 10 s before the start of the next trial. The participant was allowed to take a break at any time he/she wanted during the task, and was required to take a break between the two sessions. The whole experiment lasted about one hour.

5.2.3 Results and discussion

Figure 18 (A-C) depicts the means of the judgments of the perceptual continuity of the pitch trajectory as obtained in Experiment 6. The data for the 60 stimulus patterns, which were judged twice by each of the six participants (12 times in total), were subjected to the Friedman analysis of variance. For $n = 12$ and $df = 60 - 1 = 59$, the obtained $F = 229.7$ (not corrected for ties) was significant ($p<0.05$), and multiple comparisons followed.

No significant differences existed between the stimulus patterns with slopes of 0.5, 1.0, and 1.5 octaves per second, under equal values of overlap duration and instantaneous frequency separation. As for the variations in instantaneous frequency separation, though, a number of significant differences was found. In five cases, the stimulus patterns with frequency separations of 3 times the critical bandwidth were judged as yielding less continuous pitch trajectories than the stimulus patterns with frequency separations of 0.75 or 0.5 time the critical bandwidth. The significant cases are indicated in Figure 18 (A-C). The comparisons between the perceptual continuity scores for stimulus patterns in which the glides were separated by 0.5, 0.75, 1, and 1.5 times the critical bandwidth, on the other hand, showed no significant differences at all. As can be seen from Figure 18 (A-C), the mean continuity scores in general showed a gradual decline in perceptual continuity with increasing instantaneous frequency separation between the glides, for patterns with overlap durations of 100 ms through 400 ms.





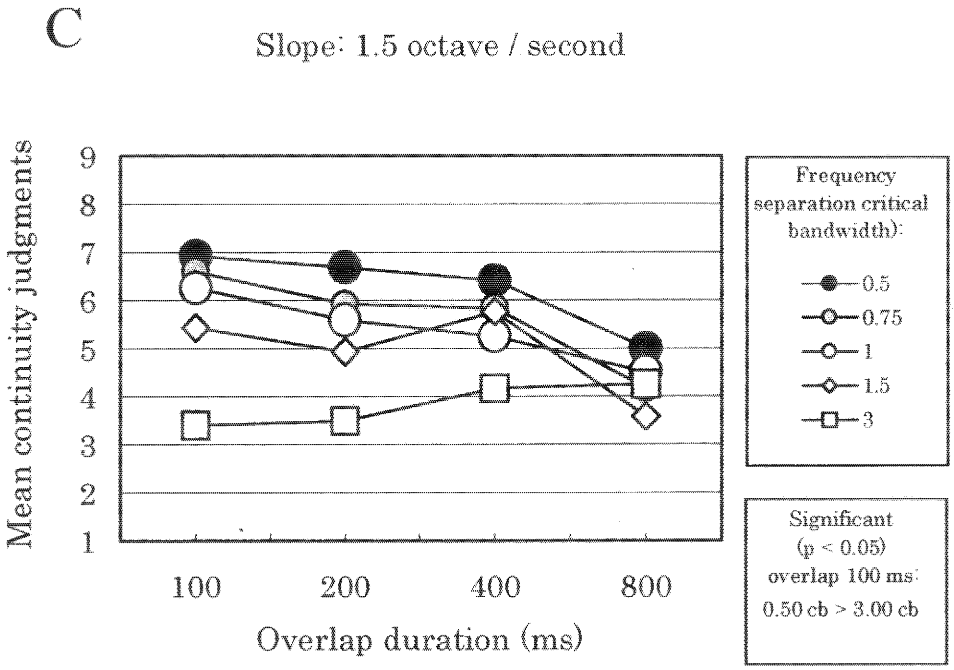


Figure 18. *Mean continuity judgments obtained in Experiment 6. The graphs (A-C) show the mean continuity judgments regarding the long pitch trajectory perceived in double-glide stimulus patterns. The reference frequency of the stimulus patterns was 1600 Hz, and the corresponding critical bandwidth 240 Hz, as interpolated from data described by Zwicker and Fastl (1999).*

Similar to the perception of the middle tone as discussed earlier, the perceptual continuity of the pitch trajectory did not disappear suddenly when the glides were separated by more than one critical bandwidth at the temporal midpoint of the stimulus patterns. The notions of the critical band or ERB do not seem to play a decisive role in explaining the mechanism behind the continuity perception of the long pitch trajectory.

The judgment scores in Figure 18 (A-C) seem to suggest that the continuity of the pitch trajectory weakened when the overlap duration was 800 ms. Closer inspection of the data revealed that this was the case for five of the six participants and that this trend appeared in stimulus patterns with instantaneous frequency separations of 0.5 through 1.5 times the critical bandwidth. The stimulus patterns in which the glides were separated by 3 critical bandwidths, on the other hand, did not show this trend. As discussed before, the judgment scores of these stimulus patterns already showed a relatively low perceptual continuity compared with the other stimulus patterns. Regarding the possible influence of overlap duration on perceptual continuity, therefore, the stimulus patterns with the largest frequency separation did not seem to provide much information, and they were left out of the following analysis. Three separate Friedman tests were done, corresponding to each variation in slope (0.5, 1, and 1.5 octaves per second). The scores obtained from the stimulus patterns with frequency separations of 0.5, 0.75, 1, and 1.5 critical bandwidths were averaged for each of the four overlap durations (100, 200, 400, and 800 ms), for each of the six participants. The three Friedman tests ($n = 6$, $df = 4 - 1 = 3$) all showed a significant difference ($p<0.05$) in the perceptual continuity between stimulus patterns with a 100 ms overlap and an 800 ms overlap. Next to an increase in instantaneous frequency separation, an increase in overlap duration thus seems to deteriorate the perceptual continuity of the pitch trajectory. The matter is further investigated in Experiment 7.

5.3 Experiment 7: Judging the continuity of the pitch trajectory.

5.3.1 Purpose

In this experiment, the influence of the factors that were varied in Experiment 6 on the perception of the pitch trajectory perceived in the stimulus patterns was further investigated. The stimulus patterns, apparatus, and participants were the same as in Experiment 5, in which the clarity of the middle tone was judged. By using the same stimulus patterns as used in Experiment 5, the correlation between the continuity of the perceived pitch trajectory and the clarity of the middle tone in the double-tone stimulus patterns could be investigated.

5.3.2 Method

The participants of Experiment 5 also participated in this experiment. They were asked to judge the (dis-)continuity, in pitch and in time, of the ongoing pitch trajectory. Judgments were made on a 7-point rating scale. The left extreme represented ‘discontinuity’, corresponding to the perception of two separate pitch trajectories. The right extreme, the number ‘7’, represented ‘complete continuity’ in pitch and in time of a single pitch trajectory. The stimulus patterns, apparatus and further experimental procedure were the same as used in Experiment 5. The total task lasted about three hours.

5.3.3 Results and discussion

The nine single-tone stimulus patterns were all judged as completely continuous and they were left out of the statistical analysis. The distribution of the raw scores of the 45 double-glide stimulus patterns was somewhat skewed to the right (it was found that in the raw scores of four of the five participants the median was larger than the mean, and the mean was larger than the mode of the scores). As described by Howell (2002), a distribution that is skewed to the right can be normalized by subjecting the raw data to a square-root transformation. Similar to Experiment 5, the transformed scores of the 45 double-tone stimulus patterns were then analyzed by using the multivariate analysis of variance (MANOVA) model in a three-way, within-subjects design. The slope of the overlapping tones (3 levels), frequency separation between the tones (5 levels), and their overlap duration (3 levels) were the independent measures. The mean (dis-)continuity judgments of the long trajectory can be seen in Figure 19. Four significant effects on the continuity perception of the long trajectory were found and they were the following.

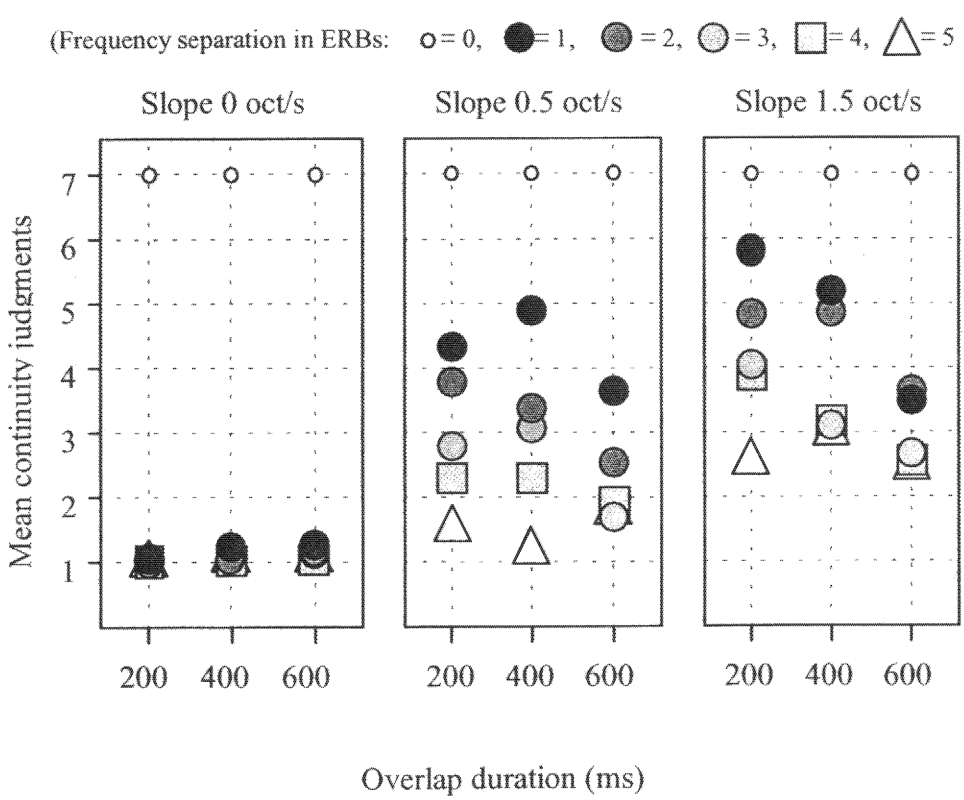


Figure 19. *The mean continuity judgments obtained in Experiment 7.*

Although Experiment 6 indicated that the overlap duration of the frequency components could be important for the perception of the single pitch trajectory, no significant main effect of overlap duration was found here. In this experiment, however, the longest overlap duration was 600 ms, compared with the 800 ms overlap duration used in Experiment 6. A significant effect of overlap duration was found in an additional analysis, in which the scores of the stimulus patterns with the steady-state tones (zero slope) were not included. In these stimulus patterns, as can be seen in Figure 19, a single pitch trajectory could not be perceived. When the frequency components were gliding, however, the continuity of the pitch trajectory improved when the overlap duration of the glides became shorter [$F(1,22) = 18.85$, $p<0.0003$].

The main effect of slope was significant [$F(1,37) = 116.60$, $p<0.0001$]. In Figure 19, it can be seen that continuity increased when the slope of the frequency components increased. This effect of slope was not found in Experiment 6, in which the differences in the slopes of the stimulus patterns used were relatively small compared with the conditions of the present experiment. Two interaction effects were found, related to the slope of the stimulus patterns. The interaction slope * frequency separation [$F(1,37) = 4.60$] and slope * overlap duration [$F(1,37) = 4.17$] were both significant at the 5% level. In an analysis in which the steady-state results were left out of the set of data, these interaction effects disappeared, but the main effect of slope firmly remained [$F(1,22) = 38.99$, $p<0.0001$].

The main effect of instantaneous frequency separation was significant [$F(1,37) = 34.66$, $p<0.0001$]. Figure 19 shows that when the instantaneous frequency separation between the overlapping tones became smaller, perceptual continuity became more compelling. Similar to the results of Experiment 6, though, a single pitch trajectory could also be perceived when the two overlapping frequency components were separated by more than a critical bandwidth, and their slope was larger than zero.

Experiment 7 shows that the factors that are important in the perception of the middle tone, as found in Experiment 5, are also important in the perception of the pitch trajectory. An additional analysis on the means of the 45 stimulus double-tone patterns as used in Experiments 5 and 7 showed a significant correlation between the clarity of the middle tone and the continuity of the long pitch trajectory [$r=.86$, ($t=11.11$), $p<0.0001$]. This may indicate that the long pitch trajectory results from the perception of the middle tone, as suggested by Nakajima et al. (2000). A perceptually salient middle tone seems to cause better perceptual continuity of the pitch trajectory than a perceptually weak middle tone.

5.4 Experiment 8: Judging the perceptual continuity in 'step-up' and 'step-down' stimulus patterns

5.4.1 Purpose

The perceptual modes found in Experiments 1 and 2 showed that a single pitch trajectory in the double-glide stimulus patterns could only be perceived when a middle tone was perceived. Furthermore, in Experiment 7, a significant correlation was found between the clarity of the middle tone and the continuity of the pitch trajectory in the double-glide stimulus patterns. The perception of the pitch trajectory, therefore, seems to depend on the perception of the middle tone. However, two alternative explanations of the experimental results must be considered. First, the middle tone may be the result of the perception of the pitch trajectory, and not the other way around. Perceptual mode B as found in Experiments 1 and 2, however, indicated that a middle tone could be perceived even when two pitch trajectories were perceived. A second alternative explanation is related to the perception of typical auditory continuity, as discussed in the Introduction. In studies of typical auditory continuity, the frequency proximity between the weaker sound components, just before and after the more intense sound, can promote the perceptual continuity of the softer pitch trajectory (Bregman, 1999; Bregman, 2000). Since the overlap of the double-glide stimulus patterns is 3.01 dB more intense than the glide components outside the overlap, it may be argued that the double-glide stimulus patterns resemble those that render typical auditory continuity. Similar to typical auditory continuity, the frequency proximity between the (weaker) glide components just outside the overlap may promote the perception of a continuous pitch trajectory in the double-glide stimulus patterns. For example, the results of Experiment 7 showed that a decrease in the instantaneous frequency separation and an increase in the slope of the glides caused a more continuous pitch trajectory.

An explanation of the results of Experiment 7 may be that the pitch trajectory became smoother because the decrease in frequency separation and increase in slope gave rise to a more robust middle tone. This robust middle tone facilitated the perception of a continuous pitch trajectory. However, an alternative to this thus may be that the decrease in instantaneous frequency separation and the increase in slope simply improved the frequency proximity between the glide components outside the overlap themselves. Frequency proximity may have influenced the perception of the continuity of the pitch trajectory directly, rather than indirectly via the middle tone. The correlation between the perception of the middle tone and the perception of the pitch trajectory could have appeared because the same factors (instantaneous frequency separation and slope) influenced the clarity of the middle tone and the continuity pitch trajectory directly and independently. In this experiment, the matter is investigated by using the 'step-up' and 'step-down' stimulus patterns, similar to those used in Experiment 4.

As found in Experiment 4, the 'step-up' stimulus patterns yielded a significantly more robust middle tone than the 'step-down' stimulus patterns. It was argued that this resulted from the difference between the 'step-up' and 'step-down' stimulus patterns with regard to the frequency proximity between the stimulus edges that delimited the overlap in both sets of stimuli. If the perception of the pitch trajectory depended on the perception of the middle tone, then the 'step-up' stimulus patterns should render a more smoothly continuing pitch trajectory because they give rise to a more robust middle tone. The generally weak middle tone as heard in the 'step-down' stimulus patterns, on the other hand, should accompany a less smooth pitch trajectory. However, as can be seen in Figure 20, the frequency proximity between the glide components just outside the overlap is much better in the 'step-down' stimulus patterns than in the 'step-up' stimulus patterns, under conditions of equal instantaneous frequency separation between the glides.

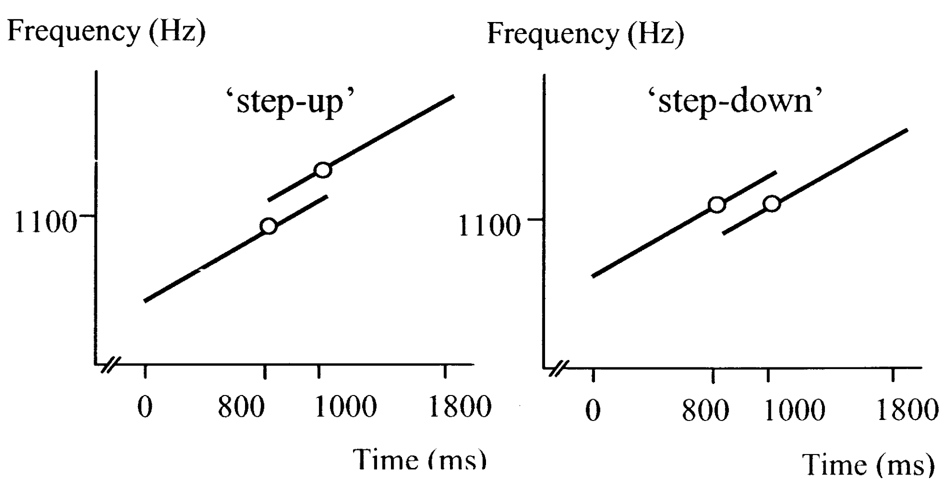


Figure 20. *Example of the difference in frequency proximity between the glide components outside the overlap in 'step-up' and 'step-down' stimulus patterns. The frequency proximity between the glide components outside the overlap (white dots) is better in the 'step-down' stimulus patterns (right) than in the 'step-up' stimulus patterns (left), under conditions of equal instantaneous frequency separation between the glides.*

If frequency proximity between the glide components just outside the overlap had an effect on the pitch trajectory, then the 'step-down' stimulus patterns should render a smoothly continuing pitch trajectory, despite a weak middle tone. The correlation between the clarity of the middle tone and the continuity of the pitch trajectory may thus disappear when 'step-down' stimulus patterns are used.

The purpose of Experiment 8 was to investigate whether the frequency proximity between the glide components outside the overlap is important for the perception of a continuous pitch trajectory in the double-glide stimulus patterns. If this is the case, regardless of the robustness of the middle tone, then it may be necessary to reconsider the idea that the perception of the pitch trajectory depends on the perception of the middle tone.

5.4.2 Method

Stimuli

Twelve double-glide stimulus patterns were generated. They consisted of two partly overlapping, ascending glides of 1200 ms each. The glides had a slope of one octave per second. The rise time of the first glide was 200 ms and the fall time 4 ms, whereas the rise time of the second glide was 4 ms and the fall time 200 ms. The rise and fall times had linearly shaped ramps. Six stimulus patterns had an overlap duration of 100 ms. The other six stimulus patterns overlapped each other for 400 ms. The stimulus patterns with the 100 ms overlap lasted 2300 ms in total, and the stimulus patterns with the 400 ms overlap had a total duration of 2000 ms.

Three stimulus patterns with a 100 ms overlap and three with a 400 ms overlap had the second glide starting at a frequency higher than that of the first glide at that point in time ('step-up' stimulus patterns). The remaining six stimulus patterns had the second glide starting at a frequency lower than that of the first glide at that point in time ('step-down' stimulus patterns). The stimulus patterns had a reference frequency of 800 Hz, and were calculated in the same way as the double-glide stimulus patterns used in the former experiments. The instantaneous frequency separation between the glides was 0.5, 1.0, or 1.5 times the critical bandwidth of the reference frequency (148 Hz, as interpolated from data described by Zwicker and Fastl, 1999).

Apparatus

The same apparatus as used in Experiment 6 was used. The sound level of the double-glide stimulus patterns ranged from 70.5 through 74 dBA (fast-peak).

Participants

Four participants, aged 20 through 22 years of age, participated in the experiment. They were one female and three males, and were studying in the field of auditory perception. All had received basic training in music, and training in technical listening for acoustic engineers.

Procedure

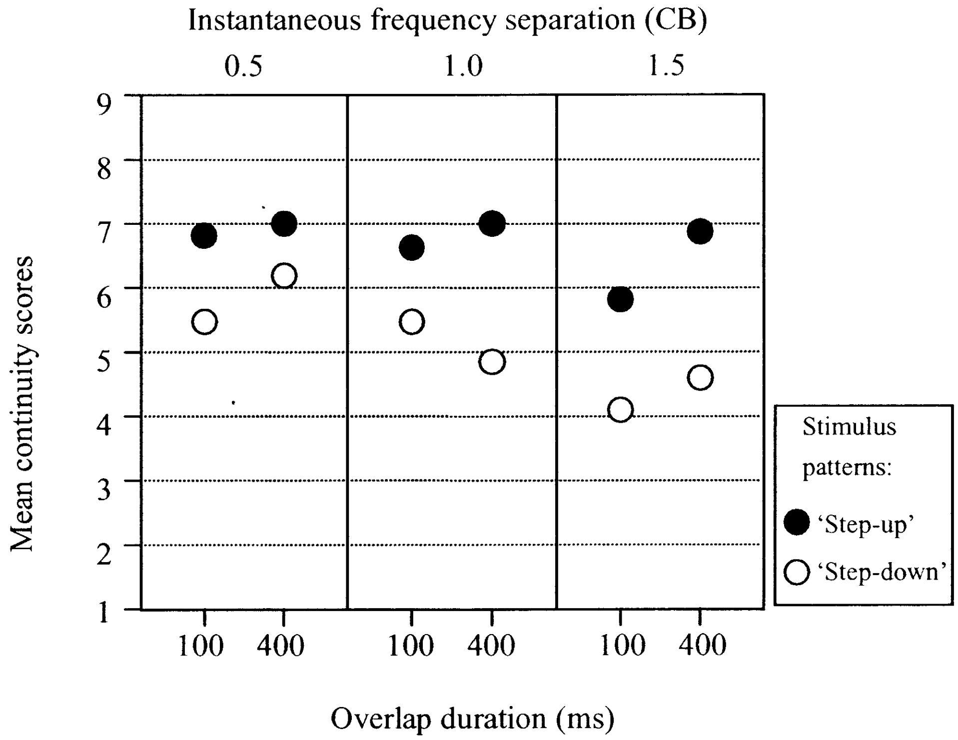
The procedure was the same as that of Experiment 6, in which the participants were asked to judge the continuity of the ongoing pitch trajectory on a 9-point rating scale. The experiment lasted about 30 minutes.

5.4.3 Results and discussion

The mean continuity scores of the pitch trajectory as perceived in the 'step-up' and 'step-down' stimulus patterns are depicted in Figure 21. The scores were subjected to a sign-test (two-tailed, $p<.05$), and the significant cases are indicated below the figure. Generally, the instantaneous frequency separation did not have a large effect on the continuity of the pitch trajectory. Only the 'step-down' stimulus patterns with the 100 ms overlap showed two significant cases. The 'step-down' stimulus patterns with the instantaneous frequency separations of 0.5 or 1 critical bandwidth gave rise to a significantly more continuous pitch trajectory compared with the 'step-down' stimulus patterns with the instantaneous frequency separation of 1.5 times the critical bandwidth.

Figure 21 shows that the pitch trajectory as perceived in the 'step-up' stimulus patterns was generally more continuous than that of the 'step-down' stimulus patterns. This tendency was general, and four significant cases were found. The 'step-up' stimulus patterns rendered a significantly more continuous pitch trajectory than the 'step-down' stimulus patterns when the overlap was 100 or 400 ms and the instantaneous frequency separation was 1 or 1.5 times the critical bandwidth in both sets of stimulus patterns. The relative smoothness of the pitch trajectory of the 'step-up' stimulus patterns indicates that frequency proximity between the glide components just outside the overlap is not mainly responsible for the perception of the pitch trajectory.

As an example, when the overlap duration is 400 ms and the glides are separated by 1.5 times a critical bandwidth, the differences between the frequency points just outside the overlap is only 1.62 Hz (793.07 through 791.45 Hz) in the 'step-down' stimulus pattern. However, the difference between the glide components just outside the overlap in the corresponding 'step-up' stimulus pattern is as large as 446.65 Hz (599.81 through 1046.46 Hz). Nonetheless, this 'step-up' stimulus pattern was significantly more continuous than the 'step-down' stimulus pattern.



Experiment 8: Significant cases ($p<.05$)

Overlap (ms):	Stimulus pattern:	Inst.Freq. sep.(CB):		Stimulus pattern:	Inst. Freq. sep. (CB):
100	step-up	0.5	more continuous than	step-down	1.5
100	step-up	1.0	more continuous than	step-down	1.0, 1.5
100	step-up	1.5	more continuous than	step-down	1.5
100	step-down	0.5	more continuous than	step-down	1.5
100	step-down	1.0	more continuous than	step-down	1.5
400	step-up	1.0	more continuous than	step-down	1.0
400	step-up	1.5	more continuous than	step-down	1.5

Figure 21. *Results of Experiment 8*

A preliminary experiment with the same stimulus patterns was done to obtain judgments regarding the clarity of the middle tone, if perceived, in the double-glide stimulus patterns. An analysis of the mean judgments of the clarity of the middle tone and the continuity of the pitch trajectory showed a significant correlation between the two [$r = .73$, ($t = 3.41$), $p<.01$]. The perception of a continuous pitch trajectory is thus facilitated when a robust middle tone is perceived, even when the frequency separation between the glide components that lie outside the overlap is large, as in the 'step-up' stimulus patterns. On the other hand, frequency proximity between the glide components outside the overlap does not facilitate the perception of a continuous pitch trajectory when the middle tone is weak, as in the 'step-down' stimulus patterns.

The results of the present experiment indicate that frequency proximity between the glide components outside the overlap itself is not essential for the perception of a continuous pitch trajectory in double-glide stimulus patterns. With regard to the results of Experiment 7, it can therefore be argued that because frequency proximity influences the perception of the middle tone, it influences the continuity of the pitch trajectory.

5.5 Summary

In Experiments 6, 7, and 8, the perceptual continuity of the pitch trajectory as perceived in double-glide stimulus patterns was investigated. In Experiments 6 and 7, it was found that, similar to the perception of the middle tone, a continuous pitch trajectory could be heard even when the instantaneous frequency separation between two partly overlapping glides was larger than a critical bandwidth, or an ERB. It was found that two factors that influenced the perception of the middle tone also influenced the perception of the pitch trajectory. The factors were the instantaneous frequency separation between the glides, and the slope of the glides. Moreover, in Experiment 7, a significant correlation was found between the clarity of the middle tone and the continuity of the long pitch trajectory. This provided support for the hypothesis of Nakajima et al. (2000), which says that the pitch trajectory that is perceived in the double-glide stimulus patterns results from the perception of the middle tone. In terms of this hypothesis, because a change in the instantaneous frequency separation between the glides, or the slope of the glides, influences the perception of the middle tone, the perception of the pitch trajectory is influenced as well.

An alternative to this is that a change in instantaneous frequency separation, or slope, influenced the perception of the pitch trajectory directly. A change in the instantaneous frequency separation or slope, for example, could have influenced the frequency proximity between the glide components outside the overlap. Similar to typical auditory continuity, this could have had consequences for the perceptual continuity of the pitch trajectory. In Experiment 8, however, it was found that frequency proximity between the glide components outside the overlap did not influence the perception of the pitch trajectory directly. Even when the frequency proximity between the glide components outside the overlap was small, the perception of continuity was not promoted in 'step-down' stimulus patterns that rendered a relatively weak middle tone. The significant correlation between the clarity of the middle tone and the continuity of the pitch trajectory that was found in Experiment 7, therefore, indicates that the perception of the continuous pitch trajectory indeed results from the perception of the middle tone, as suggested by Nakajima et al. (2000).