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：

CHAPTER 4: Further psychophysical investigation of factors involved in the perception of the middle tone in double-tone stimulus patterns.

4.1 Experiment 5: Judging the clarity of the middle tone

4.1.1 Purpose

In Chapter 3 it was discussed that frequency proximity between the onset and the offset that delimit the overlap of two partly overlapping glides may be important in the perception of the middle tone. This was further investigated in the present experiment. Stimulus patterns were used in which factors that influenced the frequency proximity between the onset and the offset that delimit the overlap were varied. These factors were the slope of the glides, and the instantaneous frequency separation between the glides. An increase in the slope of the glides in this experiment decreases the frequency separation between the onset and the offset that delimit the overlap. A decrease in the instantaneous frequency separation has the same effect. Double-glide stimulus patterns with variations in the overlap duration of the glides were employed to study the influence of temporal proximity between the onset and the offset that delimit the overlap on the perception of the middle tone. In the present experiment, next to stimulus patterns with partly overlapping glides, the perception of a middle tone in stimulus patterns with partly overlapping steady-state tones ('zero slope' stimulus patterns) was also investigated. Similar to Experiment 4, a rating scale was used to obtain data with regard to the clarity of the middle tone, if perceived.

4.1.2 Method

Stimulus patterns

A total of 45 double-tone patterns was generated. The stimulus patterns consisted of two partly overlapping pure frequency components with rise and fall times of 20 ms, with cosine shaped ramps. In all stimulus patterns, the onset of the second frequency component started before the offset of the first, at a frequency higher than that of the first component at that point in time ('step-up'). The components therefore overlapped each other and were separated in frequency. The overlap duration was 200, 400, or 600 ms. In stimulus patterns with an overlap duration of 200 ms, each frequency component lasted 1200 ms. In stimulus patterns with a 400 ms overlap, each frequency component had a duration of 1300 ms. The frequency components in the stimulus patterns with the 600 ms overlap lasted 1400 ms each. Thus, the total duration of each stimulus pattern was kept fixed at 2200 ms.

The frequency components were separated by 1, 2, 3, 4, or 5 ERBs of the 800 Hz reference frequency at the temporal midpoint of the stimulus patterns (at t = 1100 ms from the onset of the first component). At the reference frequency, 1 ERB equals 111.05 Hz (Glasberg & Moore, 1990). The frequency separation was further calculated as in the former experiments. The slope of the overlapping frequency components was 0, 0.5 or 1.5 octaves per second. The overlapping components with the slope of 0 octave per second were consequently steady-state tones. The overlapping components with slopes of 0.5 or 1.5 octaves per second were ascending glide tones.

Nine single-tone stimulus patterns ('zero' frequency separation) were generated. They consisted of a 2200 ms pure tone, moving through the 800 Hz reference frequency at the temporal midpoint (at t = 1100 ms from the onset of the tone). The three variations in slope were applied to these single-tone patterns, and each pattern was replicated three times. The replications were included to make a full factorial design of the stimulus set. Combining the three variations in overlap duration (200, 400, and 600 ms), the three variations in slope (0, 0.5, and 1.5 octaves per second), and the six variations in frequency separation (0 – 5 ERBs) resulted in a full factorial design of 54 stimulus patterns.

Apparatus

The stimulus patterns were generated and controlled by a computer (16 bit, sampling frequency 44100 Hz). Via a DAT recorder (Tascam DA-30 MKII), that was used as a D/A converter, a low-pass filter at 8.3 kHz (NF Electronic Instruments DV-04), an amplifier, and headphones (STAX Lambda Nova), the stimulus patterns were monaurally presented to the participant's preferred ear in a sound proof booth. The sound level averaged 68 dBA for the single-tone stimulus patterns, whereas the level for the double-tone stimuli averaged 71 dBA. The levels were measured (fast-peak) with a sound level meter (Brüel & Kjær 2209) and an artificial ear (Brüel & Kjær 4153), mounted with a microphone (Aco 7013).

Participants

Five students of auditory perception, two females and three males, participated in the experiment. They were 20 - 25 years of age, and had normal hearing. They had received basic training in music, and training in technical listening for acoustic engineers.

Procedure

The procedure was the same as that of Experiment 4. Each participant received three randomized sessions, consisting of the 54 randomized trials, and two warm-up trials before each session. The experiment lasted about two hours.

4.2 Results and discussion

The nine single-tone stimulus patterns were perceived as having no middle tone in the temporal middle of the sound. They were left out of the statistical analysis. The raw data of the 45 double-tone stimulus patterns were normalized by a square-root transformation, and analyzed by using the multivariate analysis of variance (MANOVA) model in a three-way, within-subject design. The independent measures were the slope of the overlapping tones (3 levels), the frequency separation between the tones (5 levels, without the single-tone stimulus patterns), and overlap duration (3 levels). The mean judgments regarding the clarity of the middle tone are depicted in Figure 17. Three significant effects were found.

A significant main effect of frequency separation was found [$F(1,37) = 59.17$, $p<0.0001$]. Figure 17 shows that, similar to the results of the former experiments, a middle tone was perceived at smaller values of instantaneous frequency separation, but could be perceived also at frequency separations larger than a critical bandwidth, or one ERB. The middle tone became less robust, however, when the frequency separation between the frequency components became larger. The results of Experiment 4 suggested that the frequency proximity between the stimulus edges that delimit the overlap could be important for the perception of the middle tone. In the present experiment, the robustness of the middle tone at smaller values of instantaneous frequency separation between the glides may also have been caused by frequency proximity between the onset and the offset that delimited the overlap of the glides. By taking only the main effect of instantaneous frequency separation into account, however, it is not possible to determine whether the instantaneous frequency separation between the overlapping frequency components, the frequency separation between the stimulus edges that delimit the overlap, or both these factors had an influence on the perceptual construction of the middle tone.

The variations in slope gave some more information regarding the matter. In this experiment, the frequency separation between the onset of the second frequency component and the offset of the first became smaller when the slope of both components became steeper. For example, when both components overlapped each other for 200 ms and were separated by 2 ERBs, the frequency separation between the onset and the offset that delimit the overlap was 222 Hz when both components were steady-state tones, 157 Hz when the slope of the components was 0.5 octave per second, and only 47 Hz when their slope was 1.5 octaves per second. An increase in slope thus enhanced the frequency proximity between the onset and the offset that delimited the overlap, independent of the instantaneous frequency proximity between the frequency components. The main effect of slope was significant [$F(1,37) = 68.55$, $p<0.0001$]. In Figure 17, it can be seen that the saliency of the middle tone decreased when the slope of the overlapping tones became shallow. The significant effect of slope thus supports the hypothesis of Nakajima et al. (2000), and the findings of Experiment 4, that frequency proximity between the stimulus edges that delimit the overlap facilitates the perception of the middle tone.

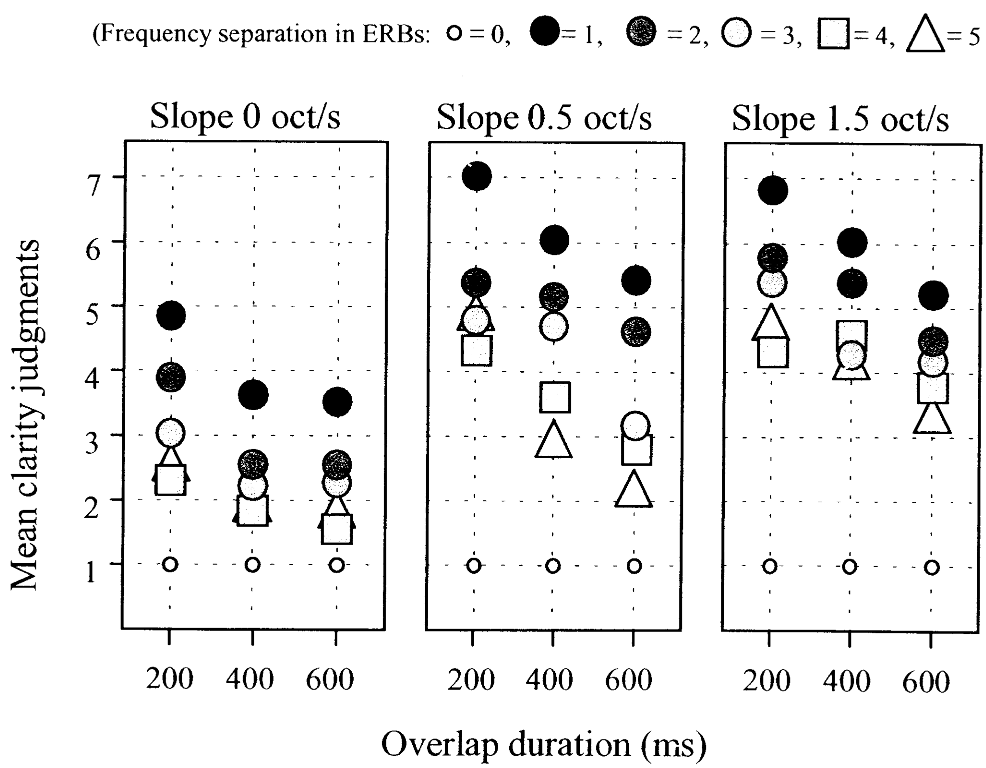

Figure 17. *Results of Experiment 5. Mean clarity judgments of the middle tone in double-glide stimulus patterns.*

The third significant effect was the main effect of overlap duration [$F(1,37) = 23.00$, $p<0.0001$]. The results show that at larger overlap durations, the middle tone became less salient. Just like a change in the slope, a change in the overlap duration between the overlapping components affects the frequency proximity between the onset and offset that delimit the overlap, when both components are glides. At a fixed frequency separation between the glides, though, the frequency proximity between the stimulus edges that delimit the overlap does not change in a systematic way. When the slope of the glides was 0.5 octave per second in the present experiment, the frequency proximity between the stimulus edges even became smaller when the overlap duration increased. Nevertheless, a relatively weak middle tone was perceived in these cases. Furthermore, the effect of duration also appeared when the overlapping tones were steady-state tones (Figure 17, left plane). In these patterns, at a fixed instantaneous frequency separation between the tones, the frequency separation between the stimulus edges that delimited the overlap remained the same regardless of the overlap duration. When the overlap duration increased, though, the robustness of the middle tone deteriorated. This was statistically confirmed by a two-way MANOVA on the fifteen stimulus patterns with overlapping steady-state tones, with overlap duration and frequency separation as independent measures. The effect of overlap duration was significant [$F(1,11) = 14.00$, $p<0.0033$]. In general, the effect of overlap duration may indicate that the temporal proximity between the stimulus edges that delimit the overlap is important as well for the appearance of the middle tone. Perceptual integration of both stimulus edges may worsen when they become further separated in time, regardless of their frequency relationship.

Some relevant research regarding the temporal integration of auditory features was done in experiments in which 'event related potentials' were measured, which reflect the activity of neurons in the auditory cortex in response to different types of auditory stimuli. One of the components of the event related potential is a negative-going wave termed 'mismatch negativity'. Mismatch negativity is typically related to pre-attentive detection of change in sound features (e.g. Remijn & Sugita, 1995). For example, when a series of identical 'standard' sounds is followed by a sound that deviates in a particular feature from the series, mismatch negativity can occur with a latency of 100-200 ms from the onset of the deviant sound. When two different deviant sounds were presented successively within a repetitive series of standard sounds, each of the two deviant sounds elicited a separate mismatch negativity when the temporal interval between the onsets of the two deviant sounds was 300 ms. However, only a single mismatch negativity was elicited when the temporal separation between the onsets of the two deviant sounds was 150 ms (Sussman et al., 1999). Similar results were obtained by using two temporally separated deviations carried by a single sound (Tervaniemi et al., 1997). Based on this, the authors suggested that there might be a general temporal integration mechanism in the formation of auditory events with a ca. 200 ms window. The duration of the temporal integration window is about 150 ms according to Yabe et al. (1997). Applied to the present results, the finding that a more robust middle tone appeared when the overlap duration of the overlapping glides or steady-state tones was short may be related to the existence of a temporal frame in which the auditory system integrates sound features. As an extension of this theoretical speculation, if voluntary attention switches can be made only after this time span, attention can be directed to the pitches of the overlapping components only when the overlap duration is longer than 200 ms. An increase in the overlap duration may enable the participants to hear out both overlapping components more easily. It is reasonable to assume that a separate middle tone is not likely to be heard when the participant is able to register the partial pitches of the components that physically constitute the overlapping part. More research needs to be done to investigate the matter.

This experiment showed that the three factors that were varied, namely, the instantaneous frequency separation between the frequency components, the slope, and the overlap duration of the frequency components all influenced the perception of the middle tone. From the results, it can be argued that not only frequency proximity, but also temporal proximity between the stimulus edges that delimit the overlap could be important in the perception of the middle tone.