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

CHAPTER 7: Summary and conclusions

Intracellular recordings of neurons in the auditory cortex of mammals have indicated that, similar to the visual system, the auditory system decomposes an auditory scene into features, and groups these features in order to construct auditory events. The grouping of auditory features, however, cannot be demonstrated easily with methods other than neurophysiological ones. In this thesis, an attempt was made to demonstrate the integration of auditory onsets and offsets that belong to acoustically different sounds into auditory events, and the factors that influence their integration, with phenomenological and psychophysical methods. Auditory onsets and offsets are important auditory features, and a demonstration of the perceptual connection of stimulus edges of acoustically different sounds could support the neurophysiological findings.

A stimulus pattern consisting of two partly overlapping glides was used to study the perceptual grouping of onsets and offsets. The main changes in the spectral content of the stimulus pattern occur at the onset of the second and the offset of the first frequency component. The changes in the spectrum are sudden. Whereas a gradual increase or decrease in a spectrum may indicate change in a single sound, a sudden increase or decrease in a spectrum typically signals that a 'new' sound joins the ongoing sound or that one of the sounds has terminated (Bregman, 1990). Here, the sudden increase in level at the onset of the second frequency component suggests to the auditory system that a new sound is joining the ongoing sound. The sudden decrease in level at the offset of the first sound suggests to the auditory system that one of the sounds has terminated. Then, the task of the auditory system is to decide whether the terminated sound was the one that joined the ongoing sound, or the ongoing sound itself. It thus has to decompose a mixture of two overlapping sounds at the overlap of the stimulus pattern. Remarkably, often the auditory system did not resolve the overlap veridically. Rather than a percept of two overlapping pitch trajectories of equal duration, the stimulus pattern was perceived as consisting of a long pitch trajectory, accompanied by a short tone in the temporal middle of the pitch trajectory. This percept appeared typically when the stimulus pattern has the following physical parameters.

(1) The overlapping glides had a slope of 0.25 through 1.5 octaves per second. The percept became more stable when the slope of the glides got steeper. The slope of the glides had a significant effect on the clarity of the middle tone and the continuity of the long pitch trajectory (Experiments 5 and 7).

(2) The overlap duration of the glides was 100 through 400 ms. The percept became weaker when the overlap duration became longer. The overlap duration of the glides had a significant effect on the clarity of the middle tone (Experiment 5). Some significant differences were found between the continuity of the long pitch trajectory in stimulus patterns with short overlaps of 100 ms and relatively long overlaps of 800 ms (Experiment 6).

(3) The instantaneous frequency separation between the glides was 0.5 through 1.5 times the critical bandwidth, or ERB, of the reference frequency at the temporal midpoint of the stimulus patterns. The percept could appear when the glides were separated by more than a critical bandwidth, or ERB, but became more robust when the instantaneous frequency separation became smaller. The instantaneous frequency separation between the glides had a significant effect on the clarity of the middle tone and the continuity of the long pitch trajectory (Experiments 5 and 7).

Since the subjective duration of the middle tone corresponded roughly to the duration of the overlap (Experiment 3), it can be assumed that the overlap plays an important role in the perception of the middle tone. In view of this, it was first investigated whether or not the middle tone could be a combination tone, or the result of spectral splatters occurring at the onset and the offset that delimited the overlap of the glides. With regard to the middle tone, the following was found:

(4) The middle tone was not the result of spectral splatter at the onset and/or the offset that delimit the overlap (Experiments 3 and 4).

(5) The middle tone was not a combination tone that could have become prominent during the overlap of the glides (Experiment 3 and 4).

In view of the perceptual continuity of the long pitch trajectory in the double-glide stimulus patterns, one has to assume that a single pitch trajectory can be perceived only when the offset of the first and the onset of the second frequency component are not perceived as such. One possibility is that the offset of the first and the onset of the second glide are masked by the middle tone. This, however, seems unlikely. The overlap of the glides seems to play an important role in the perception of the middle tone. The overlap, however, is only 3 dB more intense than the glide components outside the overlap. Even if all the energy of the overlap was allocated to the middle tone, and the middle tone were regarded as 3 dB more intense than the remaining glide components, then still its masking potential is low, compared with the masking potential of a louder sound necessary for typical auditory continuity of a softer sound to occur. One of the prerequisites of typical auditory continuity is that the louder sound should have the potential to mask the softer sound had this softer sound actually been present behind the louder sound (Warren, 1999). In stimulus patterns that render typical auditory continuity, therefore, often the intensity difference between the interrupting sound and the weaker sounds is made much larger than 3 dB, i.e.10 dB or more (Warren, 1999). A reason why the intensity difference between the interrupting sound and the weaker sounds should be large in studies of typical auditory continuity is, according to Bregman (1990), because the interrupting sound should mask the offset of the weaker sound before the more intense sound, and the onset of the weaker sound after the more intense sound. In the double-glide stimulus patterns, however, such masking may not be necessary. Rather than masking of the onset of the second and the offset of the first glide, the perception of the long pitch trajectory may become facilitated when these stimulus edges are allocated to the middle tone, and, therefore, are not perceived as belonging to the long glides.

Nakajima et al. (2000) hypothesized that the middle tone was the perceptual result of the perceptual connection of the offset of the first glide to the onset of the second. Proximity between these stimulus edges would facilitate their perceptual integration. The results of the present experiments support this hypothesis.

(6) The perception of the middle tone was facilitated when the stimulus edges that delimited the overlap were close together in frequency (Experiments 4 and 5).

This occurred when the slope of the glides became steeper, and the instantaneous frequency separation between the glides decreased. The robustness of the middle tone in ascending ‘step-up’ stimulus patterns, as compared with the corresponding ‘step-down’ stimulus patterns showed the same influence of frequency proximity between the stimulus edges that delimited the overlap.

(7) Next to proximity in frequency, the perception of the middle tone seemed to be facilitated when the stimulus edges that delimited the overlap were close together in time (Experiment 5).

The middle tone as perceived in the double-glide stimulus patterns became more robust when the overlap duration became shorter. Since a shortening of the overlap duration did not influence the frequency proximity between the stimulus edges that delimited the overlap in a systematic way, the results suggest that proximity in time between the stimulus edges may also promote the perception of the middle tone.

With regard to the perception of the long pitch trajectory, the following was found.

(8) A significant correlation was found between the clarity of the middle tone and the continuity of the long pitch trajectory (Experiments 5 and 7).

(9) Frequency proximity between the glide components just before and after the overlap did not improve the perceptual continuity of the pitch trajectory when the middle tone was weak, as in 'step-down' stimulus patterns. A relatively large frequency separation between the glide components just before and after the overlap did not deteriorate the perceptual continuity of the pitch trajectory when the middle tone was salient, as in 'step-up' stimulus patterns (Experiment 8).

The long pitch trajectory thus depended on the saliency of the middle tone. This supports the hypothesis of Nakajima et al. (2000), which says that after the attribution of the offset of the first and the onset of the second glide to the middle tone, the remaining components of the glides could be integrated perceptually.

In studies of typical auditory continuity, frequency proximity between softer sounds that flank a louder sound can improve the perceptual integration of the softer sounds. In the double-glide stimulus patterns as investigated here, however, frequency proximity between the glide components before and after the overlap did not seem to improve their perceptual integration into a long pitch trajectory. Rather, the results seemed to show a so-called 'trajectory effect'. Listeners may have based their judgments of continuity on an internal representation or idea about 'continuity' of a pitch trajectory, based on a simple rule. Since all the stimulus patterns were ascending, and repeated several times, the listeners may have judged the pitch trajectory according to whether or not it was consistent in its ascending pitch movement. Since the start of the glide component after the middle tone in the 'step-down' stimulus patterns was lower in some cases than the end of the glide component before the overlap, a small dip in the ascending movement of the pitch trajectory may have been heard. The 'step-up' stimulus patterns, on the other hand, all rendered a constantly ascending pitch trajectory that may have fitted better with the listeners' idea of continuity. Bregman (1990) described a superiority of consistency in pitch trajectory, or 'trajectory effect', as found in Ciocca and Bregman's (1987) experiments, as a result of more central processes.

An alternative view on the perception of the double-glide stimulus patterns could be that the middle tone resulted from the perceptual construction of the long pitch trajectory, rather than the other way around. The auditory system may have detected that the two glide components outside the overlap moved on the same pitch trajectory, and may have concluded that they constituted the same sound. The middle tone could be some kind of residue of the integration process of the two glides. The appearance of Mode B in Experiments 1 and 2, however, showed that a middle tone could appear also when no continuous, long pitch trajectory was perceived. Furthermore, the finding that frequency proximity is involved in the formation of the middle tone, and more central processes seem involved in the perception of the pitch trajectory, makes it unlikely that the middle tone resulted from the perception of the long pitch trajectory. The formation of auditory events under the influence of frequency proximity is supposed to be a pre-attentive and innate process (Bregman, 1990). It seems therefore unlikely that the formation of the middle tone occurs *after* the formation of the long pitch trajectory. This would, for example, implicate that the auditory system processes the stimulus pattern only after the offset of the second frequency component (at the end of the stimulus pattern), and that spectral changes such as the onset and offset that delimit the overlap have to be kept in auditory memory for a long time, before they are interpreted.

(10) The continuity of the long pitch trajectory was found to be as convincing as the continuity of a pitch trajectory heard in stimulus patterns consisting of two glides that flanked an intense noise band. This stimulus pattern caused typical auditory continuity (Experiment 8).

Since the masking potential of the overlap is low, the most plausible way in which the perceptual continuity of the pitch trajectory can be as good as that heard in stimulus patterns that can render typical auditory continuity is by the attribution of the offset of the first and the onset of the second glide to the middle tone. The assumption that, after the attribution of the onset and the offset, more central processes are involved in creating the 'best solution' for the remaining glide components is supported by the following:

(11) Two successive glides that moved on a corresponding pitch trajectory and were separated by a short temporal gap could be heard as constituting a relatively continuous pitch trajectory when the offset of the first and the onset of the second frequency component were attributed to different sounds (Experiment 10).

Thus, even when no excitation patterns are present fully between two successive glides that perceptually lack an offset and an onset, respectively, they can be heard as relatively continuous. The perceptual integration of the remaining glide components in the double-glide stimulus patterns, after the attribution the onset and offset that delimit the overlap, should therefore be possible, since at least some excitation patterns between the remaining glide components are provided by the overlap.

(12) The support for the hypothesis that the middle tone is perceptually constructed of the onset and the offset of two acoustically different sounds shows that an auditory scene can be decomposed into features such as onsets and offsets. The middle tone can be evidence of the perceptual integration of these features ('binding') by the auditory system, obtained without neurophysiological methods.

In order to construct an auditory event of some duration, such as the middle tone, by integrating an onset and an offset of two acoustically different sounds, a 'filling part' is necessary in between the onset and the offset. In order to perceptually construct the middle tone, energy of the overlap seems necessary for the 'filling part' between the onset of the second glide and the offset of the first. This may result, for example, in audible changes in the timbre of the middle tone when the instantaneous frequency separation between the glides increases. The phenomenological reports obtained in Experiments 1 and 2 showed that some observers heard a difference in the timbre of the middle tone when the glides were separated by more than a critical bandwidth and when they moved within one critical bandwidth. However, in order to get more understanding of the conditions in which auditory onsets and offset can be perceptually integrated, more research with regard to the nature of the 'filling part' of the middle tone needs to be done. For example, the relative intensity of the overlap that is necessary for the perception of a middle tone may be important for further investigation on the matter.