Impression of overlapping sounds with long decay:
Effect ofsynchronization of onset and interaural level difference

Teruo Yamasaki[*1], Sonoko Kuwano[*2] and Seiichiro Namba[*2]
**1 Faculty of Human Science, Osaka Shoin Women's University, Nara, Japan*
e-mail: yamasaki@gamelans.org
**2 Osaka University, Osaka, Japan*

Abstract

In our former studies, it was found that even if there was relatively long overlapping between two successive sounds with long decay, the overlapping wasn't detected. We related this phenomenon to self-masking, that is, masking of the former portion of a sound to the latter portion of the same sound. At the same time, this phenomenon might be caused by cognitive factor like whether two sounds are grasped as single sound event. In this study, the localization of each sound was controlled by varying interaural level difference of each sound and it was investigated whether this phenomenon was caused by perceptual factor like self-masking or cognitive factor like cognition of sound event.

A piano sound has a short attack and a fairly long decay (referred to as decaying sounds in this paper). Our former study revealed that a short melody played by decaying sounds needed fairly long overlapping in order to be perceived as being marginally connected (Kuwano, Namba, Yamasaki, and Nishiyama, 1994). This result suggested that such a long overlapping was not perceived in the melody. Another study showed that such a long overlapping made the sounds smooth and beautiful. While the long overlapping like this was not perceived as described above, the sounds were perceived as impure in case that the former portion of the first sound was cut off and two sounds after the onset of the second sound were presented (see figure 1-a). If the appropriate part of former portion of the first sound was added to the sounds as shown in figure 1-b, the impression of the sounds changed to pure one (see figure 1-b) (Namba, Kuwano, Yamasaki, and Nishiyama, 1993, in Japanese).

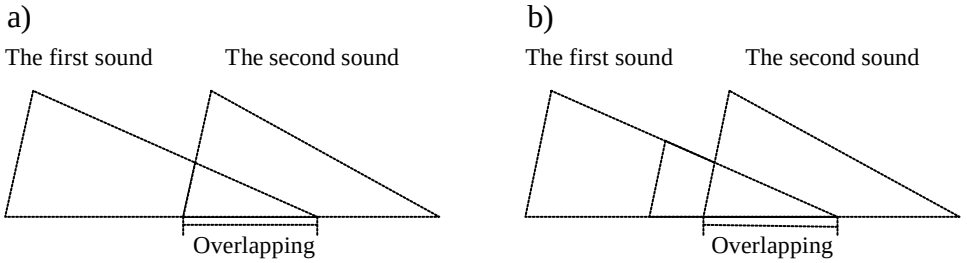

Figure 1 Patterns of the stimuli used by Namba et al.(1993)

We explained these phenomena in terms of auditory masking (Yamasaki, Namba, Kuwano, and Arakawa, 2002). Because the intensity of the second sound was greater than that of the first sound in the portion of overlapping, the second sound masked the first sound. However, this simultaneous masking was not enough to mask the first sound

completely, since the overlapping was detected in the sounds as shown in Figure 1-a. On the basis of results of some perceptual experiments, we pointed out the presence of forward masking that the former portion of the sound interferes the latter portion. We referred to it as self-masking (Yamasaki, Namba, Kuwano, and Arakawa, 2002, in Japanese). The first sound in the portion of overlapping was masked completely, only if simultaneous masking by the second sound and self-masking by the first sound occurred.

When we listen to music in daily life, sounds are located in the space. For example, first violins are located in the left, woodwind instruments are located in the middle, and contrabasses are located in the right in the orchestra. Consequently, the sounds of each instrument reaches each ear with interaural intensity difference and interaural phase difference. When sound played by one instrument follows sound played by the other instrument in orchestra performance, how long should the two sounds overlap? Results of former researches might expect that fairly long overlapping between the two sounds makes the sounds beautiful. Here, to simplify the problem, we would consider only interaural intensity difference. When two sounds are located in different directions, the interaural different intensity of each sound has the contrary inclination. That is, if the first sound is located in the right and the second sound is located in the left, the intensity of the sound reaching right ear is greater than that of the sound reaching left ear in the case of the first sound and vice versa in the case of the second sound. The greater the interaural intensity difference is, the farther the sound is located from center. When successively presented sounds are located in opposite directions, the difference of intensity between two sounds in the portion of overlapping increases in one ear and decreases in the other ear. Therefore, while the amount of simultaneous masking increases in one ear, the amount of simultaneous masking decreases in the other ear. The decrease of the amount of simultaneous masking facilitates the detection of the overlapping and as a result, this makes the impression of the sounds ugly. Consequently, we can expect as follows: the greater the interaural intensity difference, the more ugly the sounds were perceived as.

On the other hand, we can consider the interaural intensity difference as cognitive component. The greater the interaural intensity difference, the more the two sounds are recognized as different sound events. If the two sounds are recognized as quite different sound events, the interference between the sounds might decrease and as a result, the impression of the sounds might be better. This expectation is opposite to what is expected in terms of masking. The aim of this study is to investigate these expectations.

Experiment

When two decaying sounds are presented successively with the interaural intensity difference, it was examined how the interaural intensity difference of the sounds influence the impression of beauty of the sounds.

To add to the interaural intensity difference, two other variables were taken up as independent variables. Those were the overlapping duration and the duration of portion of the first sound before overlapping portion. The experiment examined their effect on the impression of beauty of the sounds.

Method

Subjects. Five subjects participated: 4 females and 1 male, between 20 and 47 years of age, with normal hearing ability.

Stimulus conditions. The stimuli were produced by a sound editing software (Sound Edit 16). The first sound was 523.3 Hz (C5) pure tone and the second sound was 392 Hz (G4) pure

tone. These sounds were produced on different tracks of the sound editing software. Their inter onset interval was fixed at 600 msec. The rise time of these sounds was fixed at 10 msec. The duration of the second sound was fixed at 600 msec, and the fall time was fixed at 590 msec. The overlapping duration was fixed at 120 msec and 240 msec, Therefore, the duration of the first sound was 720 msec and the fall time was 710 msec for the 120 msec overlapping condition, and the duration of the first sound was 840 msec and the fall time was 830 msec for the 240 msec overlapping condition. After that, the previous portion of the first sound before overlapping portion was varied from 0 msec to 120 msec by 30 msec steps by cutting off the former portion of the first sound. Figure 2 shows the stimulus for the 240 msec overlapping condition.

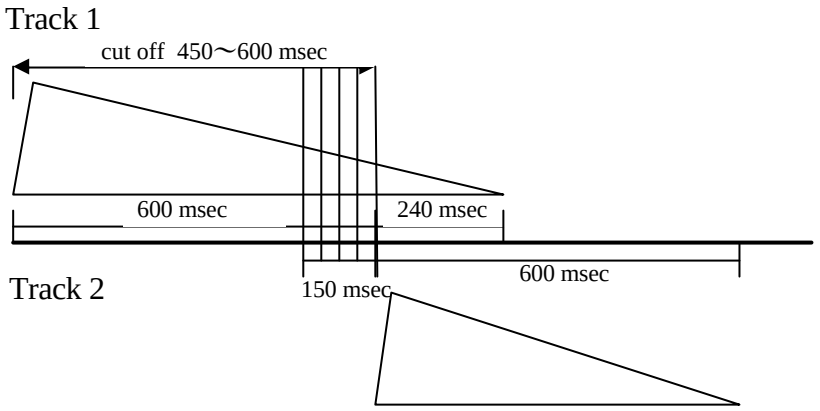

Figure 2 The stimulus for the 240 msec overlapping condition.

Finally, the two sounds were presented to right and left channels as follows: In Monaural condition, the two sounds were presented to both channels at equal level. In Dichotic condition, each sound was presented separately to different channel. In one of the interaural intensity difference condition (abbreviated as IID-12 condition), both sounds were presented with 12 dB greater intensity in one ear than the other ear in order to be located in different directions. In the other of the interaural intensity difference condition (abbreviated as IID-24 condition), both sounds were presented with 24 dB greater intensity in one ear than the other ear in order to be located in different directions. Table 1 shows the stimulus conditions with reference to the interaural intensity difference.

Table 1 Intensity level of the sounds presented to each ear in each condition of the interaural intensity difference.

Condition	Intensity level of the sound			
	The first sound		The second sound	
	Left ear	Right ear	Left ear	Right ear
DichoticLR	- 0 dB	-	-	- 0 dB
IID-24LR	- 0 dB	- 24 dB	- 24 dB	- 0 dB
IID-12LR	- 0 dB	- 12 dB	- 12 dB	- 0 dB
Monaural	- 0 dB	- 0 dB	- 0 dB	- 0 dB
IID-12RL	- 12 dB	- 0 dB	- 0 dB	- 12 dB
IID-24RL	- 24 dB	- 0 dB	- 0 dB	- 24 dB
DichoticRL	-	- 0 dB	- 0 dB	-

Note: The value indicates the attenuation from original intensity. "–" means that sound was not presented.

Because the number of levels of the independent variables was two for the overlapping duration, five for the duration of previous portion of the first sound, and seven for the interaural intensity difference, seventy stimuli were used in this experiment.

Procedure. All stimuli were recorded in two CD-Rs with two kinds of random order and presented to subjects through headphones (STAX RSD-X) in a soundproof room. Subjects were instructed to judge the beauty of each stimulus using seven categories: "very beautiful" to "very ugly". After judging the beauty of all stimuli twice, subjects were asked to judge the degree of separateness of the sounds using seven categories: "very separate from side to side" to "not separate from side to side". Again, subjects conducted this judgments twice.

Results

Judgments of separateness. The results are shown in Figure 3. The left graph shows the results of the 120 msec condition and the right graph shows the results of the 240 msec condition. the vertical axis indicates the separateness. Upper direction means more separate. The horizontal axis indicates the IID condition. Each line indicates each condition of the duration of previous portion of the first sound. Three-way analysis of variance was conducted in order to test the effects of three variables on the separateness judgments. As a result, all main effects were statistically significant ($F(1, 630)=10.41$, $p<0.005$; $F(6, 630)=21.58$, $p<0.0001$; $F(4, 630)=56.78$, $p<0.0001$).

The value of separateness judgments in the 120 msec overlapping condition were greater than that in the 240 msec overlapping condition. Fisher's LSD test was conducted in order to examine the effects of the duration of previous portion of the first sound and the interaural intensity difference. The results were summarized as follows: Generally speaking, the longer the duration of previous portion of the first sound, the more separate the two sounds were perceived as. Similarly, the greater the interaural intensity difference, the more separate the two sounds were perceived as.

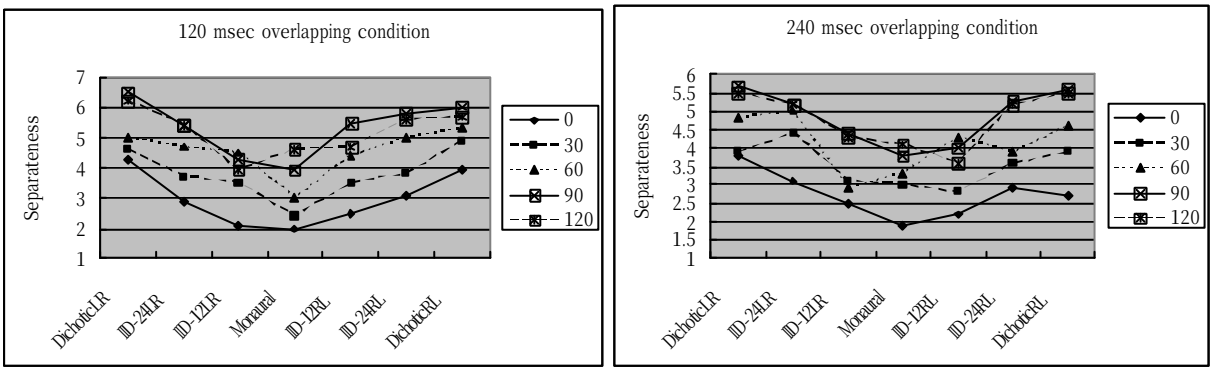

Figure 3 The judgments of the separateness of the sounds.

Judgments of beauty. The results are shown in Figure 4. The left graph shows the results of the 120 msec condition and the right graph shows the results of the 240 msec condition. The vertical axis indicates the beauty. Upper direction means more beautiful. The horizontal axis indicates the IID condition. Each line indicates each condition of the duration of previous portion of the first sound before overlapping portion. Three-way analysis of variance was conducted in order to test the effects of three variables on the beauty judgments. As a result, all main effects were statistically significant ($F(1, 630)=39.64$, $p<0.0001$; $F(6, 630)=8.37$, $p<0.0001$; $F(4, 630)=45.74$, $p<0.0001$).

The stimuli in the 240 msec overlapping condition were judged as more beautiful than those in the 120 msec overlapping condition.

Fisher's LSD test was conducted in order to examine the effects of the duratio n of previous portion of the first sound. The results were as follows: Until the duration reached 60msec, longer duration made the sounds more beautiful. However, among 60 msec and more conditions, there were no statistical difference.

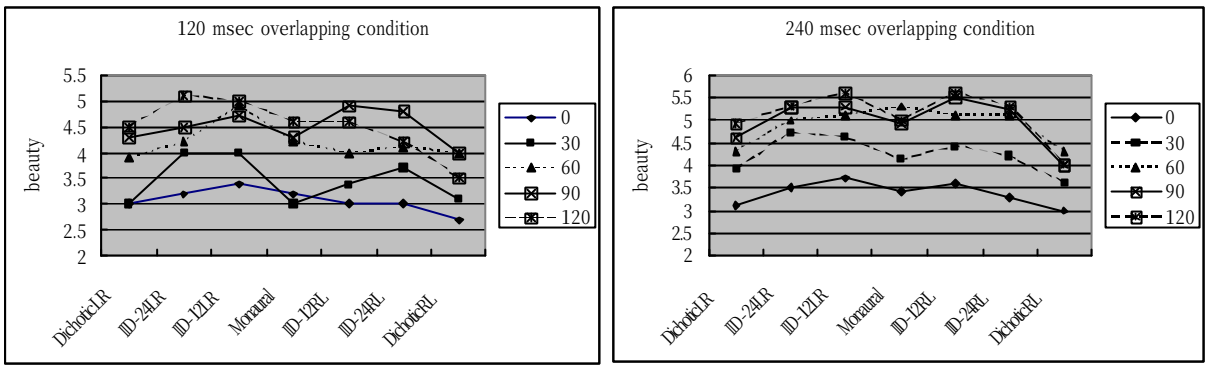

Figure 4 The judgments of the beauty of the sounds.

Figure 5 shows the overall effect of the interaural intensity difference on the judgments of the beauty. The results of Fisher's LSD test were as follows: The stimuli in the IID-12LR condition were judged statistically more beautiful than those in the IID-24LR condition, the Monaural condition, the DichoticLR condition, and the DichoticRL condition. The stimuli in the IID-24LR condition were judged statistically more beautiful than those in the DichoticLR condition and the Dichotic RL condition, and tended to be judged more beautiful than those in the Monaural condition. The stimuli in the Monaural condition was judged statistically more beautiful than those in the DichoticRL condition. The stimuli in the DichoticLR condition tended to be judged more beautiful than those in the DichoticRL condition. Generally speaking, the stimuli in the four IID conditions were judged more beautiful than those in the two Dichotic conditions and the judgments of the stimuli in the Monaural condition were medium.

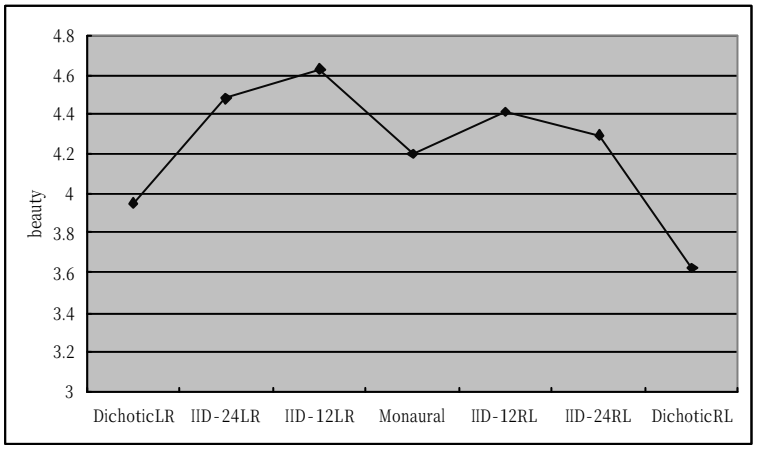

Figure 5 The overall effect of the interaural intensity difference on the judgments of the beauty.

Discussion

The result of the judgments of the separateness agreed with the expectation. That is, the greater the interaural intensity difference, the more spatially separate the two sounds were perceived as.

There were two expectations in reference to the judgments of the beauty. One expectation was as follows: When two sounds are located in opposite directions, difference of intensity between two sounds in the portion of overlapping increases in one ear and decreases in the other ear. As a result, the amount of simultaneous masking increases in one ear and the amount of simultaneous masking decreases in the other ear. It is assumed that the decrease of the amount of simultaneous masking makes the detection of the overlapping easier. Because former studies suggested that the detection of the overlapping had negative effect on the impression of the sounds, it is expected that the beauty of the sounds are related to the interaural intensity difference in the form of an inverted U-shaped curve. On the other hand, the other expectation was as follows: If the two sounds were recognized as different sound events occurring in different locations due to the interaural intensity difference, the interaural intensity difference might reduce the interference between the two sounds and consequently make the impressions of the sounds better. From this view point, it can be expected that the beauty of the sounds are related to the ineteraural intensity difference in the form of an U-shaped curve contrary to the previous expectation.

The result differed from both expectations. The beauty of the sounds was related to the ineteraural intensity difference in the form of a M-shaped curve. The Dichotic conditions, where sounds reaching each ear are quite different from each other, are very artificial. The Monaural condition is not so natural too, because each ear usually does not catch the identical sound in the "real" environment. The results showed that the impressions of the sounds were bad in these "unnatural" conditions. This might suggest that the judgments of the beauty of the sounds are related to, not only the perceptual process like auditory masking and the cognitive process like capturing each sound as different sound event, but also a cognitive process like evaluation of ecological validity on the sound localization. Further investigation is necessary for confirming this interpretation.

References

Kuwano, S., Namba, S., Yamasaki, T., & Nishiyama, K. (1994). Impression of smoothness of a sound stream in relation to legato in musical performance. *Perception & Psychophysics, 56,* 173-182.

Namba, S., Kuwano, S., Yamasaki, T., & Nishiyama, K. (1992). Relation between dynamic characteristics of hearing and "legato" in music performance. *Studies in the Humanities and Social Sciences, College of General Education, Osaka University, 41,* 15-35. (in Japanese).

Yamasaki, T., Namba, S., Kuwano, S., & Arakawa, K. (2002). On the relation between temporal characteristics of hearing and the overlapping impression of successive sounds: The effects of envelope pattern on temporal masking. *Journal of Music Perception and Cognition, 8,* 5-18. (in Japanese).