AN EFFECT OF AMPLITUDE MODULATION ON
PERCEPTUAL SEGREGATION OF TONE SEQUENCES

Mamoru Iwaki, Hiromi Seki

Niigata University
Graduate School of Science and Technology
8050 Ikarashi-2, Niigata, Japan
iwaki@bc.niigata-u.ac.jp

ABSTRACT
Auditory stream segregation is known as a process whereby sound elements are separated and integrated into some perceptual objects as a coherent whole, in the auditory scene analysis. Such a perceptual faculty is considered to be dependent on some factors in sounds such as similarity, good continuation, common fate, disjoint allocation, closure, and so on. For example, when we listen to fast alternative sinusoidal tone sequences, a number of separate auditory objects may be perceived, according to their acoustical attributes. It is known that such perception, fusion or fission, has close relation to frequency difference, time pattern and harmonic components in sounds. In this paper, we investigate an effect of amplitude modulation on sequential stream segregation for two close sinusoidal tones in frequency domain. As a result, it is shown that amplitude modulation improves the sequential stream segregation performance.

1. INTRODUCTION
Auditory stream segregation[1, 2] is a psychoacoustical phenomenon which has been investigated even now. Each perceived stream is considered to correspond to an auditory object. The auditory stream segregation is a function to reflect the auditory attributes of different components to different sources. Configuration of stimuli in frequency, time, and space is considered as an important factor for the segregation. A relation between streaming and perceived rhythm were reported by L. P. A. S. van Noorden[3, 4]. When two tone bursts, A and B with different frequencies, were arranged as 'ABA ABA ...' with shorter silent gaps separating the successive bursts, listeners heard the fainter tone burst regardless of the presence of the louder burst. The perception depended on the difference of tones in frequency. For the shorter difference, the two tone bursts made a fusion and one auditory stream like galloping sound was perceived. The frequency difference increasing, two streams, one for a series of A and the other for B, were perceived. The rate in time of stream A was double than that of B. M. R. Jones et al[5, 6] reported the importance of the presence of temporally equi-spaced series of sound for auditory segregation. This indicated that the larger predictability of the series should make auditory segregation easier because the relationship between the preceding and following sounds expectedly became clear. J. Vliegen et al[7] treated the dependency on harmonics in stimuli. Although it is said that components with similar harmonics tends to perceived in one auditory stream, there were only small difference comparing to the case without harmonics. As another configuration of stimuli in temporal and frequency domains, amplitude of the stimuli is able to be considered as a factor of auditory segregation. M. M. Rose[8] reported such influence on the auditory segregation by varying the whole amplitude of one stimulus against the other. However, an effect of amplitude modulation on the auditory stream segregation has not been investigated although it is known that correlated changes in amplitude and frequency are one of important cues for the sound source separation.

In this paper, we evaluate an influence of amplitude modulation on the auditory stream segregation for tone sequences. The temporal configuration of stimuli is 'ABA ABA ...' as is in others. The B-sound has higher frequency than A, which is amplitude-modulated with some modulation frequencies and modulation indices. For such stimulus, the perception and the frequency distance of A and B are measured. As a result, the amplitude modulation has an effect on the auditory streaming, and the amount of the effect depends on the value of the modulation parameters. In addition, on the modulation frequency axis, the most efficient modulation frequency for the auditory stream segregation is obtained.

2. EXPERIMENT
In this section, configurations of the psychoacoustical experiment in this paper is described. First, the design of stimuli
is shown with the settings of their parameters. Next, the protocol of the experiment is explained. The task to be asked for subjects is also described.

2.1. Stimuli

Figure 1 shows the design of stimuli in our experiment. The stimuli consisted of two kinds of sinusoidal tones called A-sound and B-sound. As depicted in Fig. 1, the frequency of B-sound was different from that of A-sound. They were presented like ‘ABA ABA ...’ and distinct in both of time and frequency domains. In the unit of ABA, both sounds were separated with 20[ms] silence. The unit was repeatedly presented twelve times in one stimulus with 120[ms] long silence portions. The A-sound was an intermittent sinusoidal tone expressed as follows:

\[ A(t) = \begin{cases} 
\sin 2\pi f_A t, & \text{if } t \text{ in A-range} \\
0, & \text{otherwise} 
\end{cases} \]

where the frequency was fixed as \( f_A = 100[\text{Hz}] \). Such a low frequency of 100[Hz] was selected for eliminating fatigue in exposure to simple sound for a long time. The B-sound was one of amplitude modulated sinusoidal tones represented by

\[ B(t) = \begin{cases} 
(1 + m \sin 2\pi g t) \sin 2\pi f_B t, & \text{if } t \text{ in B-range} \\
0, & \text{otherwise} 
\end{cases} \]

The distance of A-sound and B-sound in the frequency axis was one of 1, 2, 2.5, 3, 3.5, 4, 4.5, 5, 6, 7, and 11 semitone. Then the frequency \( f_B \)’s of B-sound were 105.95, 112.25, 115.55, 118.92, 122.41, 125.99, 129.68, 133.48, 141.42, 149.83, and 188.77[Hz], respectively, shown in Table 1. In this amplitude modulation, \( g \) and \( m \) mean the modulation frequency and the modulation index, respectively. The modulation frequency was selected from a set of 2, 4, 16, and 32[Hz], and the modulation index was one of 0, 0.3, 0.6, and 0.8. The B-sound is identical to a pure tone in case of \( m = 0 \). The duration of both sounds was 100[ms]. They were digitized with sampling frequency of 20[kHz] and quantization of linear 16[bit], and tapered linearly by 10[ms]. All stimuli were presented to both ears through a closed headphone in a silent room.

![Fig. 1. Design of stimuli](image)

2.2. Protocol

In the experiment, one trial consisted of 143 stimulations and judgments to them. The 143 stimuli were constructed with 11 \( f_B \)’s for one non-amplitude-modulated stimuli in \( m = 0 \) and 12 combinations of 4 modulation frequencies and 3 modulation indeces (i.e., \( 143 = 11 \cdot (1 + 4 \cdot 3) \)). Each stimulus was presented at random one by one.

The subjects in this experiment were three males and a female of age from 22 to 24 with normal audibility. They had trained with some examples to understand what they were going to be heard, beforehand. They were asked to judge whether the stimuli were perceived a single stream like galloping (fusion), or two streams of A and B separately (fission). It was a two-alternative forced-choice procedure. All subjects carried out 15 trials in total.

In this experiment, the effect of amplitude modulation is going to be considered large if the distance of \( f_B \) and \( f_A \) for fission is small. On the other hand, the effect of the modulation is considered small if the distance does not get small.

3. RESULTS

In this section, the results of the experiment are summarized. Since the results show almost common tendency, results for one subjects are illustrated in Fig. 2 as a typical example. Figure 2 depicts the segregation ratio, which is the ratio of the number of judgements where the subject answered that two separate auditory streams were perceived. The abscissa in these graphs represents the frequency distance of \( f_B \) from \( f_A = 100[\text{Hz}] \) in semitone. Figures 2(a), (b), (c), and (d) correspond to the results for the modulation frequencies \( g \) of 2, 4, 16, and 32[Hz], respectively. In each graph, four segregation ratios are illustrated according to the modulation index \( m \). As a result of these graphs, the segregation ratio has a dependency on frequency distance, modulation frequency, 

### Table 1. Configuration of the frequency of B-sound: \( f_B \)

<table>
<thead>
<tr>
<th>( f_B - f_A ) [semitone]</th>
<th>( f_B ) [Hz]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>105.95</td>
</tr>
<tr>
<td>2.0</td>
<td>112.25</td>
</tr>
<tr>
<td>2.5</td>
<td>115.55</td>
</tr>
<tr>
<td>3.0</td>
<td>118.92</td>
</tr>
<tr>
<td>3.5</td>
<td>122.41</td>
</tr>
<tr>
<td>4.0</td>
<td>125.99</td>
</tr>
<tr>
<td>4.5</td>
<td>129.68</td>
</tr>
<tr>
<td>5.0</td>
<td>133.48</td>
</tr>
<tr>
<td>6.0</td>
<td>141.42</td>
</tr>
<tr>
<td>7.0</td>
<td>149.83</td>
</tr>
<tr>
<td>11.0</td>
<td>188.77</td>
</tr>
</tbody>
</table>
Fig. 2. Segregation ratio (i.e. The stimuli is perceived as if it were two streams of tones: fission or segregation, while one stream: fusion)

Fig. 3. Relationship between modulation index and frequency distance

Fig. 4. Relationship between modulation frequency and frequency distance

4. DISCUSSIONS

As shown in Fig. 2, all subjects heard two separate auditory streams nearly at 100% when the frequency distance was larger than 7[semitone], even if modulation index \( m = 0 \). This almost matches with the conventional research[7]. Furthermore, in our results, the promotion effect of amplitude modulation of B-sound on auditory stream segregation is demonstrated. For example, Figure 3 shows that the segregation in the perception is promoted by increasing the modulation index. The increase of the modulation index will
contribute to the increase of similarity or good continuation in the series of B-sound. Generally, it is said that sounds with similar acoustical properties tend to be grouped into a single auditory stream. From this viewpoint, it is natural that the larger modulation promotes the segregation performance.

In Fig. 4, it is observed that the segregation becomes easy whenever modulation frequency \( g = 4 \text{[Hz]} \). The frequency matches the value that the detectability of amplitude modulation becomes maximal in the range of this experiment, as reported by E. Zwicker[9].

There were some differences in the effect, depending on subjects. For example, some of the subjects can perceive two auditory streams for frequency distance of about 5 semitone. However the difference is not enough for another subjects. For such a subject insensitive to amplitude modulation, as a tendency, the promotion of segregation by amplitude modulation was not observed clearly.

5. CONCLUSIONS

We evaluated an effect of amplitude modulation for auditory stream segregation in perception of tone sequences by psychoacoustical experiments. In this experiment, the higher frequency sinusoidal tone had an amplitude-modulated envelope with good continuation. As a result, it was shown that the amplitude modulation had an effect on the auditory stream segregation while conventional investigations have been suggested that temporal and frequency patterns should have notable relation to the segregation characteristics. As the modulation index increased from 0.3 to 0.8, the effect on the auditory stream segregation became large. Furthermore, for the modulation frequency of 4[Hz], the effect of amplitude modulation became large, comparing to the other frequencies of 2, 16, and 32[Hz]. To show more general characteristics and relations to the audibility for amplitude modulated sound, we need more experiments with more subjects, which still remains for future works.

6. ACKNOWLEDGMENT

This work was supported by the Japan Science and Technology Corporation(JST) under a program of the Core Research for Evolutional Science and Technology(CREST).

7. REFERENCES