A hysteresis in a rhythm perception of tone-bursts by a gradual change in the incremental level

PACS REFERENCE: 43.66.+y

SUZUKI Yôiti; NISHIMURA Ryouichi; KUROSAWA Tomoyuki
Research Institute of Electrical Communication, Tohoku University
2-1-1 Katahira, Aoba-ku,
Sendai 980-8577,
JAPAN
Tel: +81.22.217.5460
Fax: +81.22.217.5535
E-mail: yoh@ais.riec.tohoku.ac.jp

ABSTRACT

Grouping in rhythm perception based on the difference of intensity among tone bursts was examined in terms of the sequential context effect. In the perception of a rhythm consisting of a set of four tone bursts, three bursts of equal intensity following a more intense one easily results in perception of a single group consisting of the four. With an increase in the intensity of the third burst, the set may be perceived as two groups, and as with a decrease, perception as a single group is reestablished. Experimental results show that the critical level at which splitting or merging occurs depends on the path of the change of intensity, namely, whether it is incremental or decremental. Once a rhythm is perceived, listeners seem to keep the rhythm, resulting in this newly found hysteretic contextual perception.

INTRODUCTION

Double interpretation illusion has been long studied in the visual domain. For example, a picture titled “a wife and her mother” introduced by E. G. Boring in 1930 is very famous; this picture can be seen as either a young or an old woman. Whether the young or the old woman appears at first glance depends on characteristics of the observer such as age, gender, familiarity with young and old woman, and so on. If you know well about this picture, however, you may attempt to see it as the other figure and you will probably be able to do so. Leeper examined whether the latter perception can be forced by external factors instead of the observer’s intention [1]. He made two other pictures by slightly modifying “a wife and her mother” to have them seen distinctively as a young woman and an old woman, respectively. If subjects are asked to view the original picture after looking at one of the modified pictures for a while, they tend to see the original figure as the same woman which they first saw. Fischer showed a similar phenomenon using a number of pictures gradually changing from a face of a man to the shape of a woman with her knee bent [2]. Looking at them successively, observers seemed to keep the image of picture they saw first even for the ambiguous pictures located midway between the two typical pictures located at the ends.

This kind of hysteretic perception is also found in the human auditory system. The Békésy-type audiometer, which employs the up-down method to measure the threshold of hearing, may be a good example. In this method, the sound pressure level of stimuli is gradually increased until it becomes audible, and vice versa. The transient level from audible to inaudible in a downward sequence is usually several decibels lower than that from inaudible to audible in an upward

sequence. After several upward and downward sequences, the audible threshold level is given as an average of the turning points. Though such hysteretic effect could cause difficulties in psychophysical experiments, research directly relating to hysteretic perception in the human auditory system seems to be very limited.

The ultimate goal of this study was to investigate the mechanism of utilizing preliminary information other than the current physical stimuli. In the first stage of this study, we focused on a grouping of sequential tone burst signals, i.e., rhythm perception [3]. This is because sequential tone bursts, unlike music or speech, are free from any semantic effects. Moreover, rhythm is easy for subjects to understand.

SOUND STIMULI FOR PSYCHOACOUSTIC EXPERIMENTS

A set of four tone bursts (TB1, TB2, TB3, and TB4), as depicted in Fig. 1, was used as the basic unit of the sound stimuli. This is called a set, hereafter. Each tone burst was a 500-Hz pure tone with a duration of 50 ms, including 5 ms raised cosine rise and fall times. The inter-burst-interval was 200 ms as shown in the figure. In a set, the sound pressure levels of TB1, TB2 and TB4 were fixed at 66, 60 and 60 dB, respectively. In contrast, that of TB3 was varied and ranged from 60 dB to 66 dB for each set. The level of TB3 is expressed as L3 hereafter. The possible perceived rhythm for a set was twofold: a group consisting of four bursts or two groups consisting of two bursts. Which rhythm appeared depended on L3. This set was presented repeatedly as a sequence as shown at the bottom of Fig. 1.

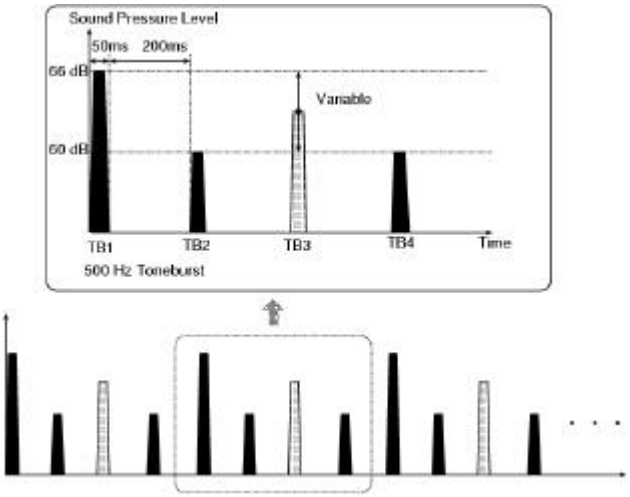

Fig. 1 Time pattern of tone burst sequence used in the experiment

In the first experiment, L3 was fixed throughout a sequence, and thus no contextual effect could be introduced. In this set, basic rhythm perception without any contextual effect could be observed as a function of L3. Another experiment was then conducted to evaluate the effect of context on the perception of rhythm. In this experiment, L3 was incrementally or decrementally changed during a single sequence to examine the effect of hysteresis on the perception of rhythm.

EXPERIMENT 1: RHYTHM PERCEPTION WITHOUT CONTEXTUAL EFFECT

Method

In this experiment, L3 was set at 60, 61, 62, 63, 64, 65, or 66 dB for each sequence. One single sequence lasted 5 s followed by a silent period of 5 s. Subjects were asked to depress a button labeled “2” on a response box while they perceived a rhythm consisting of two bursts, and to depress a button labeled “4” while they perceived a rhythm consisting of four bursts. They were also asked not to push any button when they did not perceive any rhythm or perceived a rhythm other than those consisting of two or four bursts. Fourteen sequences were presented in a single session. Therefore, a certain level of TB3 appeared twice in a single session. The experiment was composed of five sessions for one subject.

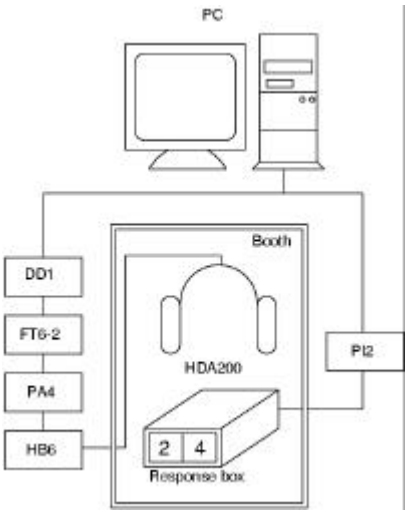

Fig. 2 Apparatus used in the experiments

Apparatus of the experiment is depicted in Fig. 2. System2^{TM} of Tucker-Davis Technologies was employed to generate stimuli as well as to acquire a subject's response. Stimuli were generated by DD1 followed by an anti-aliasing filter with a cut-off frequency of 10 kHz (FT6-2), an attenuator (PA4), and a headphone buffer (HB6). Finally, stimuli were diotically presented to both ears through headphones (HDA200, Sennheiser). Subjects' responses were sent to a host computer via a parallel interface (PI2). The sampling frequency was 32 kHz. Five subjects, all male and in their twenties with normal hearing, participated this experiment.

Results

The time periods of pushing button "2" or "4" are referred to as T_2 or T_4, respectively. The time period of no pushing is referred to as T_0. The ratios of T_2, T_4, and T_0 to the entire presentation time are calculated for each L3. Figure 3 shows the results averaged over the subjects. The individual numerical values are shown in Table 1. The solid line in Fig. 3 indicates the time delay needed to push a button after commencement of a stimulus presentation.

Discussion

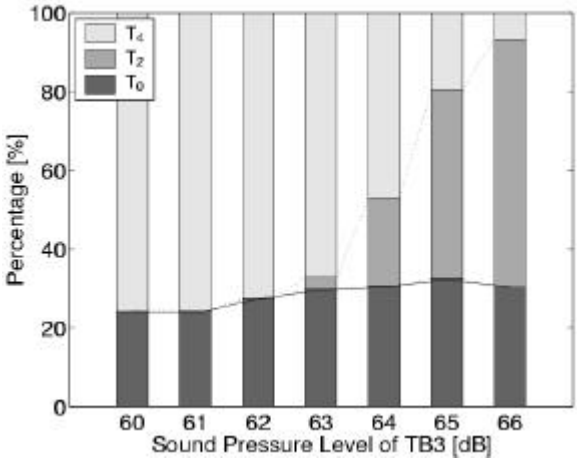

Fig. 3 Ratios of T_0, T_2, and T_4 to the total presentation time

Table 1 Ratios of T_0, T_2 and T_4 to total presentation time for various sound pressure level of TB3 in the non-context condition

Sub.		60	61	62	63	64	65	66
A	T_0	24.1	19.9	28.0	27.5	28.0	35.0	32.3
	T_2	0.0	0.0	0.0	0.4	7.1	22.9	60.3
	T_4	75.9	80.1	72.0	72.0	64.9	42.	7.5
B	T_0	33.7	36.7	42.1	50.0	50.0	50.4	46.5
	T_2	0.0	0.0	0.0	3.1	6.7	49.6	53.5
	T_4	66.3	63.3	57.9	47.0	43.1	0.0	0.0
C	T_0	18.2	20.2	20.7	20.6	17.4	26.1	24.0
	T_2	0.0	0.0	0.0	1.3	4.8	27.6	51.7
	T_4	81.8	79.8	79.3	78.0	77.8	46.3	24.4
D	T_0	20.5	21.9	21.8	25.0	29.5	28.2	29.1
	T_2	0.0	0.0	0.0	8.3	36.1	64.8	69.3
	T_4	79.5	78.1	78.2	66.7	34.4	6.9	1.7
E	T_0	24.6	22.9	24.4	26.3	28.3	23.0	19.9
	T_2	0.0	0.0	0.0	1.9	57.1	74.4	80.1
	T_4	75.4	77.1	75.6	71.8	14.6	2.6	0.0

From Fig. 3 and Table 1, it can be easily seen that the ratio of perceiving the rhythm of the two bursts simply increases as L3 increases. In contrast, the ratio of perceiving the rhythm of the four bursts simply decreases as the sound pressure level of TB3 increases. Meanwhile, T_0 and the solid line that corresponds to the time delay to make the first decision for a sequence are almost constant and identical over the entire range of the L3 level. This implies that once the subjects perceived the rhythm of as either two or four bursts, they tended to keep perceiving that rhythm for the rest of the stimulus set.

EXPERIMENT 2: RHYTHM PERCEPTION WITH A CONTEXT

Methods

In Exp. 1, a rhythm of the two bursts was always perceived for the stimuli for L3 from 60 dB to 62 dB. Then, beyond this level, rhythms of both two and four were perceived. Taking this into account, an incremental change of L3 sequence-by-sequence was introduced to impose a context. That is, L3 for the first sequence in a session was set at 62 dB and then increased by 1 dB for each following sequence, reaching 66 dB at the fifth, i.e., the last sequence. This condition is referred to as "increasing condition" hereafter. Another condition where L3 was decreased by 1 dB for each sequence from 66 to 62 dB was also examined and is referred to as "decreasing condition". Time patterns of L3 for these two conditions are shown in Fig. 4. The subjects' task was the same as that in Exp. 1. Ten sessions were carried out for each condition. The five subjects who participated in Exp. 1 also took part in this experiment.

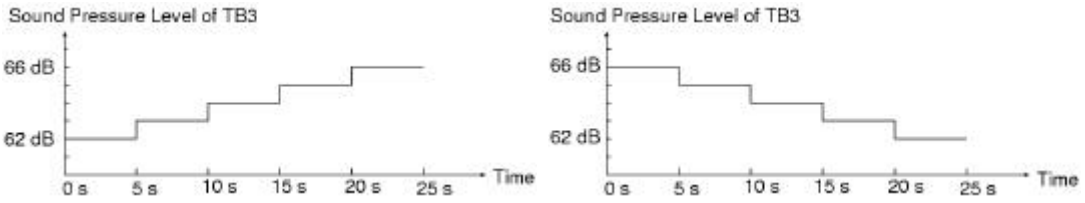

Fig. 4 Time patterns of the sound pressure level of TB3 (left: increasing condition; right: decreasing condition)

Results

For each condition, the obtained data were divided into five parts according to each L3.The ratios of T_2, T_4, and T_0 to the entire presentation time were reduced for each L3. The results averaged over all subjects are depicted in Fig. 6. Moreover, the results are individually shown in Fig. 5.

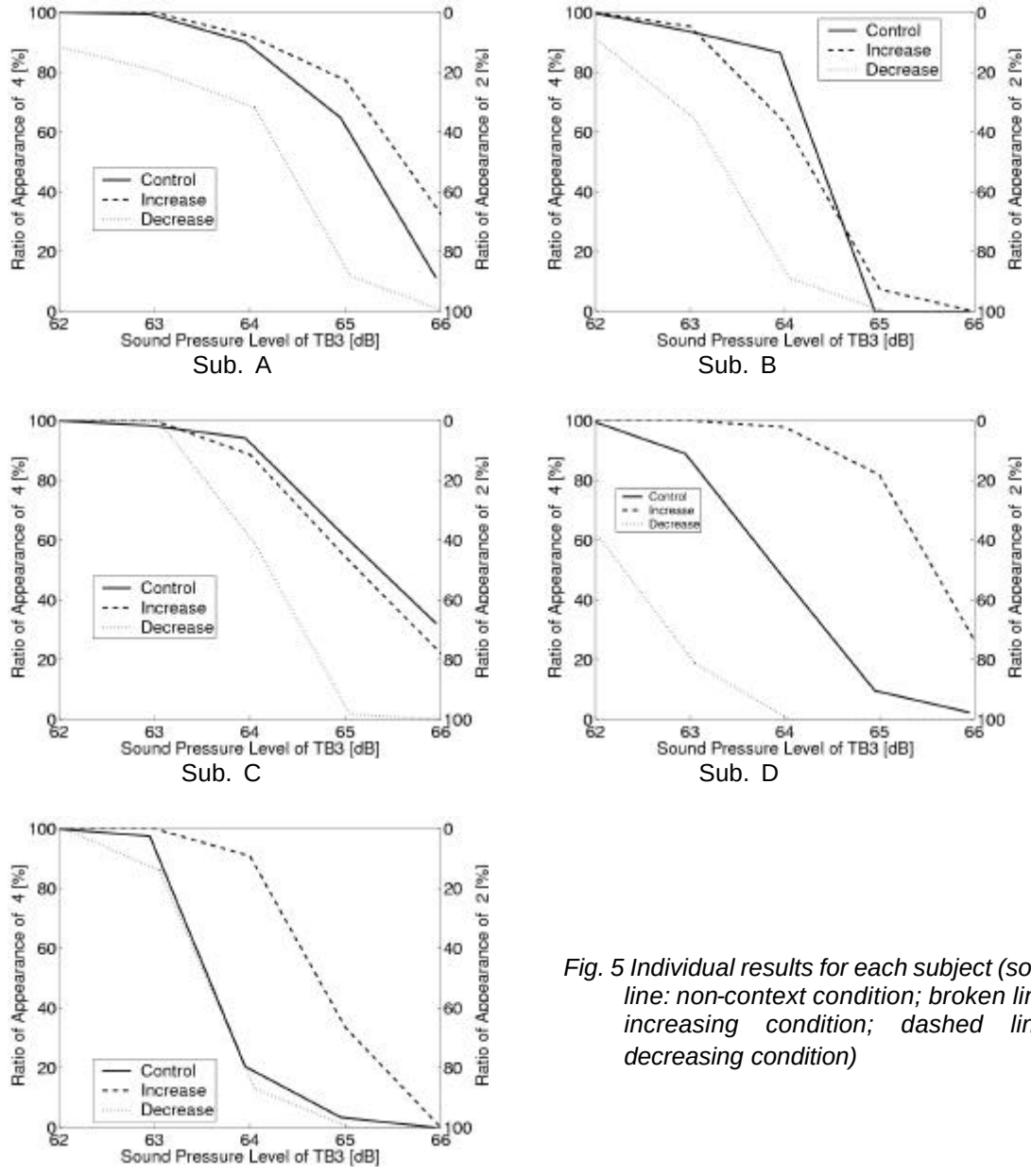

Fig. 5 Individual results for each subject (solid line: non-context condition; broken line. increasing condition; dashed line: decreasing condition)

Hysteretic Context Effect

T_0 was observed almost only for the first sequence in a session, that is, the sequence with L3 of 62 dB in the increasing condition, and 66 dB in the decreasing condition. Therefore, the observed T_0 must have been the time to make the first decision. In addition, since T_0 was merely observed for the sequences other than the first one, listeners were always able to perceive one of the two rhythms once one of the rhythms was perceived.

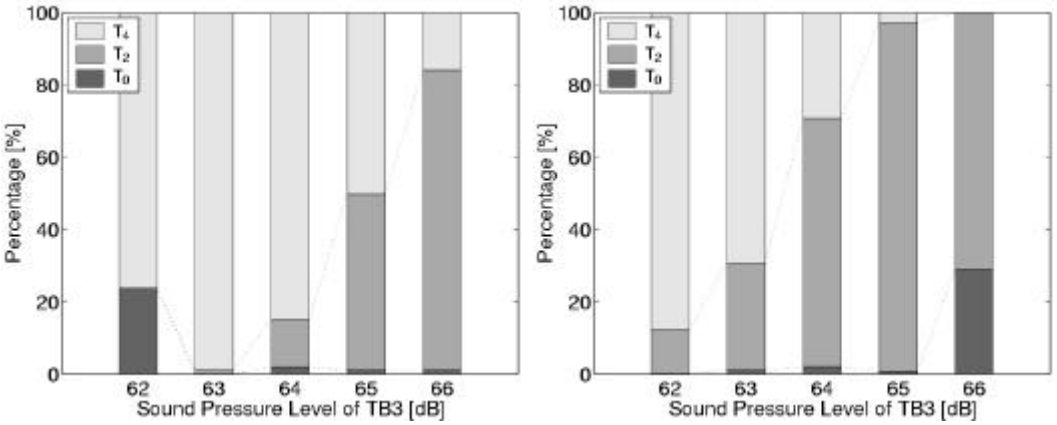

Fig. 6 Ratios of TB4, TB2, and TB0 to the entire presentation time for various sound pressure levels of TB3 (left: increasing condition; right: decreasing condition)

To compare the results of Exp. 2 with that of Exp. 1, the ratios were recalculated by ignoring T_0. The results obtained after this recalculation are shown in Fig. 7. This figure clearly shows that the perceived rhythm hysteretically depends on the context of the stimuli. In the increasing condition, the ratios of perceiving the rhythm of four bursts are significantly higher than those in the non-context condition (Exp. 1). In the decreasing condition, in contrast, the rhythm of two bursts dominates. Considering these tendencies, it can be concluded that a rhythm perceived at a certain moment has an inertial effect on the perception of the following rhythms. This demonstrates that a hysteretic context effect exists in rhythm perception.

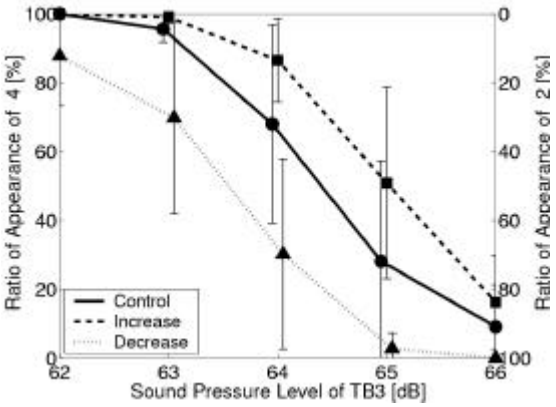

Fig. 7 Ratios of T_2 and T_4 to the entire time excluding T_0 (solid line: non-context condition; broken line: increasing condition; dashed line: decreasing condition)

If individual results are observed, the amount of context effect strongly depends on the subjects. Subjects seem to fall into three groups in terms of their perception tendency. One consisted of three subjects (A, B, C) who showed strong effect only in the increasing condition. Another consisted of only one subject (D) who showed strong effect only in the decreasing condition. The rest (E) showed little effect both in the increasing and decreasing condition. In the previous experiment, subjects A, B, and C tended to perceive the rhythm of four bursts more frequently than the others, and subject E tended to perceive the rhythm of two bursts. Subject D showed an intermediate tendency. It might be that subjects who tend to perceive a specific rhythm would not be strongly influenced by the context. Those who did not show any tendency as to the rhythm perception are rather easily influenced by the context.

EXPERIMENT 3: INFLUENCE OF THE STEP SIZE OF L3

Methods

The incremental/decremental difference of 1 dB for L3 between two consecutive sequences is much higher than the threshold of intensity discrimination of a pure tone at 60 to 70 dB. As this could cause some unexpected effects in the perception of the rhythm of the stimuli, another experiment was conducted in which the level change of L3 was set so as to be less than the discrimination threshold. In this experiment, L3 was increased or decreased by 0.2 dB for each following set. The initial and the final L3s were the same as those in Exp 2. The time patterns of L3 are shown in Fig. 8. Five subjects participated in this experiment, three of whom had participated in the previous experiments.

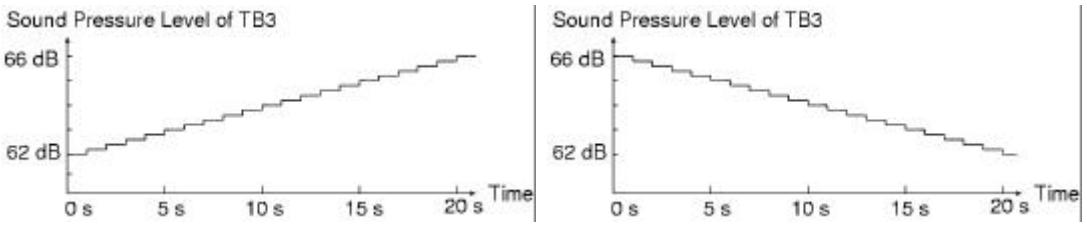

Fig. 8 Time patterns of the sound pressure level of TB3 (left: increasing condition; right: decreasing condition)

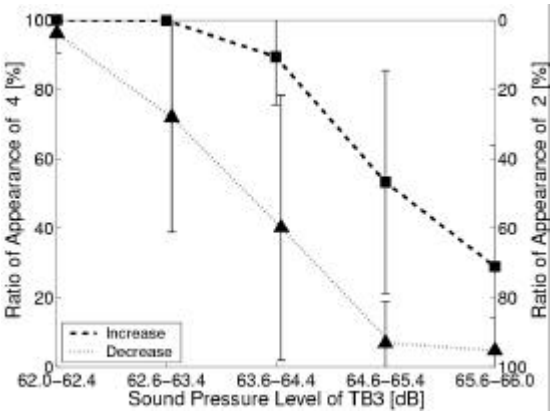

Fig. 9 Ratios of T2 and T4 to the total presentation time excluding T0 (broken line: increasing condition; dashed line: decreasing condition)

Results and Discussion

T_0, T_2, and T_4 were measured for each L3. The ratios of T_2, and T_4 to the entire presentation time excluding T_0 are shown in Fig. 9. To compare the obtained data with the results of Exp. 2, the data of Exp. 3 were classified into five categories of L3: 62.0 to 62.4, 62.6 to 63.4, 63.6 to 64.4, 64.6 to 65.4, and 65.6 to 66.0 dB. The results of the analysis of variance showed no significant difference between Exps. 2 and 3. This shows that the hysteretic context effects are consistently observed for both of the step sizes.

CONCLUSION

Psychophysical experiments using tone bust signals with a certain context were conducted to examine the hysteretic context effect in rhythm perception. Results showed that perception of a rhythm at a certain moment tends to continue to subsequent similar stimuli. This indicates that hysteretic perception exists in our auditory system. For comprehensive understanding of the mechanism causing this phenomenon, further experiments are needed.

REFERENCES

[1] R. Leeper, “A study of neglected portion of field of learning: The development of sensory organization,” *Journal of Genetic Psychology*, vol.46, pp. 41-75, 1935.
[2] G. H. Fisher, “Preparation of ambiguous stimulus materials,” *Perception and Psychophysics*, vol. 52, no. 9, pp. 421-422, 1967.
[3] T. L. Bolton, “Rhythm,” *American Journal of Psychology*, vol. 6, pp. 145-238, 1894.