Interdependence between temporal and ambient spatial variations of a successive sound sequence

Ryota Miyauchi (1), Yukio Iwaya (2), and Yōiti Suzuki (2)

(1) School of Information Science, Japan Advanced Institute of Science and Technology,
1-1 Asahidai, Nomi, Ishikawa, 923-1292 Japan
(2) Research Institute of Electrical Communication, Tohoku University,
2-1-1 Katahiraa, Aoba-ku, Sendai, 980-8577 Japan

PACS: 43.66.Mk, 43.66.Qp

ABSTRACT

Time and space are interdependent in perception. The most typical example is that the temporal and spatial patterns of three successive stimuli, defining two inter-onset intervals and two spatial distances, can respectively affect the experience of spatial and temporal variations. In regard to auditory modality, most studies have equated pitch space with ambient space and demonstrated that each of the temporal and pitch intervals can affect the respective perceptual pitch and temporal variations. A very interesting question is whether the interdependence between temporal and ambient spatial variations, not pitch variations, could be found when successive sounds differ in location. To investigate what effect temporal variations in a successive sound sequence had on the experience of spatial variations, we measured the subjective differences between two neighboring distances (Experiment 1) and subjective absolute locations (Experiment 2) when three successive sounds, S₁, S₂, and S₃, were presented to participants from different loudspeakers. The results demonstrated that varying the time interval between S₁ and S₂ (t₁) and between S₂ and S₃ (t₂) affected the perceived spatial distances between the sounds. If t₁ was larger than t₂, the subjective distance between S₁ and S₂ (d₁) was greater than that between S₂ and S₃ (d₂), although the physical distances of d₁ and d₂ were identical. The results indicate that there is a typical Tau effect in auditory modality, and the auditory Tau effect is caused by the mislocalization of S₂. Furthermore, the findings suggest that a perceptual, not physical, temporal pattern affects the experience of the auditory spatial layout. Spatial information on successive sounds may be organized after a sound sequence is temporally structured.

INTRODUCTION

Time and space are interdependent in perception. Temporal and spatial variations in a stimulus sequence can respectively affect the experience of the spatial and temporal layout in the visual and tactile modality. The most typical example is where each temporal variation or spatial variation of three successive stimuli, defining two temporal inter-onset intervals (IOIs) and two spatial distances, can affect the experience of each spatial variation or temporal variation. When two constant spatial distances with variable temporal IOIs are presented, the distances are judged to be different according to the temporal IOIs. Helson (1931) called this phenomenon the "Tau effect." Conversely, when temporal IOIs are constant and spatial distances are variable, temporal judgments vary according to the spatial pattern. This phenomenon was called the "Kappa effect" by Cohen, Hansel, and Sylvester (1953), although the phenomenon had previously been reported by Benussi (1913) and Abe (1935). These effects have been found for various modalities (see Jones and Huang, 1982 for an overview).

Most studies on auditory modality have equated the pitch space with ambient space and confirmed that Tau and Kappa effects occur between perceived temporal and pitch intervals (see ten Hoopen, Miyauchi, and Nakajima, 2008 for an overview). However, auditory-motion perception of a moving sound source does not pertain to pitch space but rather to ambient auditory space. It is thus a very interesting question whether the interdependence between temporal variations and ambient spatial variations, not pitch variations, can be found when successive sounds differ in location. Sarrazin, Giraudo, and Pittenger (2007) recently demonstrated that variable IOIs of a sound sequence affect the reproduction of constant distances (Tau effect) and conversely variable spatial distances of a sound sequence affect constant time intervals (Kappa effect). Grondin and Plourde (2007) found that there are conditions under which increasing the spatial distance between sound sources increases the perceived duration of a temporal interval (Kappa effect). The findings by Sarrazin et al. and Grondin et al. indicate that time and space are interactive in auditory perception, memory, cognition, and reproduction. However, they used complex spatio-temporal sequences of sounds. To obtain direct evidences of perceptual interactions between temporal and spatial information in auditory modality, it is necessary to conduct experiments using a simpler sound sequence.

The purpose of the present study was to investigate whether or not temporal variations affect the experience of spatial variations. In the present experiments, sequences consisting of three successive sound bursts (in order of S₁, S₂, and S₃) presented from different loudspeakers were used as the stimuli. The sequence included two IOIs (t₁ and t₂) and two
spatial distances ($d_1$ and $d_2$) delimited between $S_1$ and $S_2$, and between $S_2$ and $S_3$. In Experiment 1, participants judged the difference between $d_1$ and $d_2$ in various combinations of $t_1$ and $t_2$, and $d_1$ and $d_2$. In Experiment 2, we measured the perceptual location of $S_3$ with various temporal and spatial patterns.

**EXPERIMENT 1**

**Methods**

Participants: Seven males and one female (aged between 21 and 33) took part in the experiment. One of the males was one of the authors and the other participants were naive regarding the configuration of the stimulus and the objective of the experiment. All had normal hearing (confirmed by audiometric testing).

Apparatus: Each participant sat on a chair located at the center of a circular array of loudspeakers with a radius of 1.5 m in an anechoic chamber (noise level of 19.5 dB). The position of their head was fixed straight ahead (0°) by a headrest. The array at eye level in the horizontal plane consisted of a series of 36 loudspeakers (Fostex FE83E) installed in a custom-made cylindrical box separated by 10° intervals. Eleven loudspeakers located between −60° and 60° were used in the experiment (Fig. 1). The participants were blindfolded so that they could not see the loudspeakers during the experiment.

Auditory stimuli were computer generated (sampling frequency of 48 kHz and quantization of 32 bits) and presented via the loudspeakers to the participants through a D/A converter (Pavec MD-8D72-133) and amplifiers (Biamp MCA8050). The levels of the sound bursts from all loudspeakers at the head position of participants were calibrated with a measuring amplifier (Briel & Kjær 2610) and a microphone (Briel & Kjær 2669 and 4165).

Stimulus: A stimulus pattern consisted of three successive sound bursts (i.e., in order to $S_1$, $S_2$, and $S_3$) presented from different loudspeakers. The sound bursts were pink noise low-pass filtered at 12 kHz. Each sound burst had a duration of 50 ms, which included a cosine-shaped rise-and-fall time of 10 ms. The A-weighted sound pressure level was 78 dB (as measured at the head position of participants when sound was presented continuously).

The total duration delimited by $S_1$ and $S_3$ was fixed at 360 ms. The IOI between $S_1$ and $S_2$ was varied from 90 to 270 ms in steps of 30 ms, and consequently, the IOI between $S_2$ and $S_3$ varied from 270 to 90 ms in steps of 30 ms. Thus, the temporal patterns of $t_1$ and $t_2$ ($t_1/t_2$) were /90/270/, /120/240/, /150/210/, /180/180/, /210/150/, /240/120/, and /270/90/ [ms]. The loudspeakers presenting $S_1$, $S_2$, and $S_3$ were horizontally changed from the left to the right or from the right to the left. These three loudspeakers were located on the left of the hemisagittal plane (Fig. 1a), on its right (Fig. 1c), or centered across it (Fig. 1b). The angular distances between the two loudspeakers presenting $S_1$ and $S_2$ ($d_1$), and $S_2$ and $S_3$ ($d_2$) were 20° and 30°, 30° and 30°, and 40° and 20°. Thus, the spatial patterns of $d_1$ and $d_2$ ($d_1/d_2$) were 1.2, 1.1, and 2.1 (see Fig. 1). In total, there were 126 stimulus patterns [7 temporal patterns × 3 spatial patterns × 3 spatial locations (left, center, and right) × 2 directions of movement (left-to-right and right-to-left)].

Procedure: All 126-stimulus patterns were presented in random order in each session and they started with four warm-up trials. Each participant performed in nine sessions over three days. The first session served as training. The data obtained in the training session and the warm-up trials were excluded from the final analyses. A two-alternative forced-choice (2AFC) method was employed. Participants were instructed to push a “left” or “right” key according to which of the neighboring two angular distances, they perceived to be closer.

The proportion of responses starting that the distance between $S_1$ and $S_2$ ($d_1$) was shorter than that between $S_2$ and $S_3$ ($d_2$) (i.e., $P_{d_1<d_2}$) was calculated. Differences of $P_{d_1<d_2}$ between spatio-temporal conditions were evaluated using a four-way repeated-measured analysis of variance (ANOVA) with factors related to temporal pattern ($t_1$/$t_2$), spatial pattern ($d_1$/$d_2$), spatial location (left, center, and right), and direction of movement (right-to-left and left-to-right). The temporal and spatial patterns were found to have significant effects $F(6, 42) = 26.784$ for temporal pattern and $F(2, 14) = 23.654$ for spatial pattern), while spatial location, direction of movement, and interactions had no significant effects. Thus, the factors of spatial location and direction of movement were merged, and $P_{d_1<d_2}$ was plotted as a function of the temporal pattern (i.e., $t_1$/$t_2$) with error bars indicating standard errors (see Fig. 2). These results indicate that $P_{d_1<d_2}$ decreases as $t_1$ increases, and consequently, as $t_2$ decreases. That is, the distance between two sounds closer in time is perceived to be closer in space. The same tendency was found for all spatial patterns and locations. The results clearly demonstrated that the Tau effect between time and ambient space occurred in auditory modality.

**Discussion**

If variations in a physical temporal pattern affect the experience of ambient space, $d_1$ and $d_2$ should be perceived as equal when $t_1$ is equal to $t_2$. However, our results revealed that when $t_1$ was about 60 ms shorter than $t_2$ (i.e., $t_1/t_2 = 5/7$), $d_1$ and $d_2$ were perceived as equal (see the results indicated by the open triangles and solid line in Fig. 2). Nakajima, ten Hoopen, Hilkhuyzen, and Sasaki (1992) demonstrated that when two neighboring empty time intervals ($t_1$ and $t_2$) composed of three brief successive sounds were presented from a single source, $t_2$ was perceived to be much shorter than the physical interval because $t_2$ was perceptually “shrunk” by adding $t_1$ (Nakajima, ten Hoopen, & van der Wilk, 1991; Nakajima et al., 2004). This illusory phenomenon in regard to auditory-time perception, called “time-shrinking,” typically occurs when $t_2$ is longer than $t_1$ and when the difference between $t_2$ and $t_1$ is less than about 100 ms ($0 < (t_2 - t_1) < 100$ [ms]). Under the time conditions to meet this range, $t_1$ and $t_2$ were perceived to be the same duration (ten Hoopen et al., 2006; Sasaki, Nakajima, & ten Hoopen, 1998; Miyauchi &
Nakajima, 2007). The midpoint of this 1:1 temporal category (i.e., \( t_2 - t_1 = 50\) [ms]) and the temporal pattern in which \( d_1\) and \( d_2\) are perceived as equal (i.e., \( t_2 - t_1 = 60\) [ms]) is almost the same. These findings suggest that the experience of auditory space is affected by the perceptual temporal variations consisting of that space.

**EXPERIMENT 2**

The results from Experiment 1 clearly demonstrated that the Tau effect between time and ambient space occurred in auditory modality. However, it remains unclear whether the temporal pattern affects the ability to localize each sound or to organize some sound into a spatial pattern. In Experiment 2, we measured the absolute perceptual locations of \( S_2\) under four conditions: (1) \( S_2\) was presented alone, (2) \( S_1\) was presented before \( S_2\), (3) \( S_3\) was presented after \( S_2\), and (4) \( S_1\), \( S_2\), and \( S_3\) were sequentially presented. The perceptual location of \( S_2\) was only varied according to the temporal pattern under condition (4) if time patterns affect the sound localization.

**Methods**

**Participants:** A male and a female took part in the experiment. They were naive regarding the configuration of the stimulus and the objective of the experiment and they had normal hearing (confirmed by audiometric testing).

**Apparatus:** Each participant individually sat on a chair located at the center of a semicircular array (radius of 1.1 m) of loudspeakers in a sound-proof room. Their heads were fixed in a straight-ahead position (0°) by a chin rest. The array at eye level on the horizontal plane consisted of a series of 19 loudspeakers (Hoshiden 7N101; \( \alpha = 30\) mm), mounted in a custom-made enclosure and separated by 2.5° intervals.

**Stimulus:** All sound bursts used in the experiment were 1-kHz pure-tone bursts and they had a duration of 10 ms, which included cosine-shaped rise and fall time of 5 ms. We prepared four types of stimulus patterns, i.e., \( S_2\), \( S_1\), \( S_2\), \( S_1\), \( S_2\), and \( S_1\)–\( S_2\)–\( S_1\) conditions. Under the \( S_2\) conditions, a single sound burst (\( S_2\)) was presented by itself from one of six loudspeakers located at \( -12.5^\circ\), \(-7.5^\circ\), \(-2.5^\circ\), \( 2.5^\circ\), \( 7.5^\circ\), and \( 12.5^\circ\). Under the \( S_1\)–\( S_2\) conditions, a sound burst (\( S_1\)) was presented from the loudspeaker located at \(-15^\circ\) before \( S_2\) was presented. The time interval between \( S_1\) and \( S_2\) was 120, 180, or 240 ms. Under the \( S_1\)–\( S_2\)–\( S_1\) conditions, a sound burst (\( S_1\)) was presented from the loudspeaker located at \( 15^\circ\) after \( S_2\) was presented. The time interval between \( S_2\) and \( S_2\) was 120, 180, or 240 ms. Under the \( S_1\)–\( S_2\)–\( S_1\) conditions, three sound bursts, \( S_1\), \( S_2\), and \( S_3\), were sequentially presented. The temporal patterns of these sound bursts \( (t_1, t_2)\) were \( 120/240\), \( 180/180\), and \( 240/120\) ms. In total, there were 6 (locations of \( S_2 \times 1\) time variation), 18, 18, and 18 (6 locations of \( S_2 \times 3\) time variations) stimulus patterns under respective \( S_2\), \( S_1\)–\( S_2\), \( S_2\), and \( S_1\)–\( S_2\)–\( S_1\) conditions.

**Procedure:** The participants were instructed to concentrate on a fixation line located at a position that was straight-ahead (0°) of them during the experiment. Each participant’s task was to judge the perceived location of \( S_2\) relative to the fixation line and to push a “left” or “right” key according to whether \( S_2\) was located at left or right of the fixation line. Each participant took part in 14 trials for each stimulus pattern.

**Results**

We first calculated the proportion of judgments in which the participant judged the sound as being presented from the left of the fixation line. We then fitted logistic functions, by using the maximum-likelihood method (Wichmann and Hill, 2001), to the response distributions as a function of the location of \( S_2\) relative to the fixation line under all temporal and spatial conditions. The results obtained under the \( S_2\)–\( S_2\)–\( S_1\) and \( S_1\)–\( S_2\)–\( S_1\) conditions are in Fig. 3. The results under \( S_2\) conditions have been given in all panels as a reference.

The point of subjective equality (PSE), defined as the 0.5-point on the psychometric function, was estimated from the fitted psychometric functions for each individual participant. The PSE indicates the relative position of the loudspeaker at which the sound was perceived at the same location as the fixation line. The PSE under \( S_2\) conditions, in which \( S_2\) was presented alone, was \( -1.4^\circ\). We defined the PSE under \( S_2\) conditions as the reference perceptual location of participants without the influence of other sounds, and calculated the disparities in PSEs between \( S_2\) and the other conditions (Table 1).
location of S₂ had slightly shifted to left under the influence of S₁ presented at right of S₂. These results indicate that S₂ was mislocated in the direction opposite to the preceding or following sound. In other words, the spatial disparity between two sequential sounds was underestimated.

Table 1. Disparity in PSEs between S₂ and other conditions.

<table>
<thead>
<tr>
<th>Condition</th>
<th>PSE (°)</th>
</tr>
</thead>
<tbody>
<tr>
<td>/120/240/</td>
<td>-0.7°, -2.7°, -3.4°</td>
</tr>
<tr>
<td>/180/180/</td>
<td>1.4°, 2.2°, 0.2°</td>
</tr>
<tr>
<td>/240/120/</td>
<td>8.5°, 2.9°, -10.1°</td>
</tr>
</tbody>
</table>

Under S₁–S₂–S₃ conditions (Fig. 3(c)), the perceptual location of S₂ strongly depended on the temporal patterns. When t₁ was shorter than t₂ (/120/240/), the PSE drastically shifted to the right as 8.5°. This means that the perceptual location of S₂ shifted to the left under the influence of S₁ and S₃ (Fig. 4(a)). Under these conditions, the time interval consisting of S₁ and S₂ was shorter than that consisting of S₂ and S₃. Thus, the results indicate that the spatial disparity presented with shorter time intervals was underestimated, and/or conversely, the spatial disparity presented with longer time intervals was overestimated. This tendency was the same as that in the findings in Experiment 1.

However, when t₁ was longer than t₂ (/240/120/), PSE drastically shifted to the left as -10.1°. This means that the perceptual location of S₂ shifted to the left under the influence of S₁ and S₃ (Fig. 4(b)). This indicates the same tendency as that for the /120/240/ temporal pattern in that the spatial disparity presented with shorter time intervals was underestimated, and/or conversely, the spatial disparity presented with longer time intervals was overestimated.

Note that the slopes for psychometric functions under S₁–S₂–S₃ conditions were clearly gradual rather than those under the other conditions (Fig. 3). This indicates that the accuracy of sound localization decreased in the complex sound sequence.

The extreme mislocalization of S₂ under S₁–S₂–S₃ conditions was not solely attributed to the presentation of either a preceding (S₁) or following (S₃) sound, because the results under S₁–S₂–S₃ and S₂–S₃ conditions indicated that the effect of presentation of either S₁ or S₃ was to shift S₂ in the direction opposite to these sounds, and the values for this effect were less than 3.4°. S₂ was drastically mislocated only when three sounds were successively presented, i.e., only when a temporal pattern consisting of two or more time intervals was perceived. The perceptual location of S₂ was strongly influenced by the temporal pattern, and the auditory Tau effect demonstrated in Experiment 1 might have been caused by the mislocalization of S₂.

CONCLUSIONS

The results presented in this report demonstrate that the auditory Tau effect occurs when a simple sequence consisting of three sounds (which was more commonly used in earlier studies on other modalities) was used. The perceptual, not physical, temporal pattern affects the experience of auditory ambient space. This illusory phenomenon might be caused by the mislocalization of S₂. To further understand the interactive interference between time and space in auditory modality, the results of the present study need to be analyzed further to determine what effect spatial variations have on the perceptual organization of temporal patterns.

We believe that the dramatic effect of a temporal pattern on spatio-temporal interference rather than that of other modalities can be attributed to the superiority of auditory modality in the time domain. The total IOI of the sound sequence used...
in the present experiment was 360 ms. Under shorter temporal conditions, the perceptual connection of these sounds in the temporal domain is indivisible. However, there is no evidence that shows how the spatial information on successive sounds is integrated into a sequential spatial pattern or a moving pattern. To discuss the asymmetrical effect on the interactive interference between time and space, we need to study what effect variations in spatial patterns have on the experience of the temporal pattern.

(a) /120/240/
Spatial pattern: \( \bigcirc \) \( S_1 \) \( \bigcirc \) \( \bigcirc \) \( S_2 \) \( \bigcirc \) \( S_3 \)
Perceptual location of \( S_2 \)
Temporal pattern: 120 ms 240 ms
\( S_1 \) \( S_2 \) \( S_3 \)

(b) /240/120/\nSpatial pattern: \( \bigcirc \) \( \bigcirc \) \( S_1 \) \( S_2 \) \( \bigcirc \) \( S_3 \)
Perceptual location of \( S_2 \)
Temporal pattern: 240 ms 120 ms
\( S_1 \) \( S_2 \) \( S_3 \)

(c) /180/180/\nSpatial pattern: \( \bigcirc \) \( \bigcirc \) \( S_1 \) \( S_2 \) \( \bigcirc \) \( S_3 \)
Perceptual location of \( S_2 \)
Temporal pattern: 180 ms 180 ms
\( S_1 \) \( S_2 \) \( S_3 \)

Figure 4. Diagram of results under \( S_1 \)-\( S_2 \)-\( S_3 \) conditions. Dotted circle indicates physical location of \( S_2 \). Solid circles indicate perceptual locations of \( S_2 \).

ACKNOWLEDGEMENTS

This research was supported by a Grant-in-Aid for Young Scientists (B) (21730584) and by a Grant-in-Aid for Specially Promoted Research (to Y. Suzuki, 19001004) from the Japan Society for the Promotion of Science (JSPS).

REFERENCES