ACUSTICA · acta acustica
Vol. 86 (2000) 1034 – 1037

# Perception of a Secondary Auditory Image with Three Sound Sources

B. T. G. Tan, S. H. Tang, Gongqiang Yu
Department of Physics, National University of Singapore 10 Kent Ridge Crescent, Singapore 119260

**Summary**
In a three-source sound system, one of which is undelayed and the other two delayed, a listener facing these sources perceives an unexpected secondary auditory image in addition to the expected primary auditory image. The two delayed sources have time delays relative to the undelayed source within the range in which the precedence effect operates. The secondary auditory image is believed to be due to summing localization taking place between the two sound sources which have been delayed, when the relative time delay between them is within the range for summing localization to occur.

PACS no. 43.66.Qp

## 1. Introduction

### 1.1. The precedence effect and summing localization

When the time delay between two laterally separated sound sources, such as loudspeakers, in front and in the same horizontal plane of a listener is below a certain threshold value, which lies between 630 $\mu$s and 1 ms [1, 2], summing localization [1, 3] operates, and the position of the resultant auditory image is perceived to be between the two loudspeakers. The exact position is determined by a combination of time delay and level differences between the two [4]. As the time delay is increased beyond this threshold value, which we call the precedence threshold, the auditory image appears to be from the direction of the earlier source, i.e. the leading loudspeaker. This perceptual phenomenon is known as the precedence effect, or the Haas effect, or the "law of the first wavefront" [5, 6, 7]. If the time delay is increased further, an echo threshold is reached beyond which the delayed sound is perceived separately as an echo. Quantitative estimates of echo thresholds vary tremendously (2–50 ms), depending on a number of variables [2]. Previous work [1, 8, 9] on the time delay effects between two sources have classified them into three ranges, i.e. summing localization, precedence effect and echo effect in order of increasing relative time delay. The precedence threshold marks the onset of the precedence effect while the echo threshold marks the onset of the echo effect.

### 1.2. Three-source system

The possible combinations of inter-source time delay with three sources are more complex and have not been as fully explored as the two-source case. However, multi-source sound systems are now widely used, such as in sound-reinforcement systems as well as in home theatre systems. In both sound reinforcement and home theatre systems with multiple sources, the interaction between the sources due to time delays between them is more complex than for two source systems.

In this paper, we report on the resultant auditory images due to a system with three sound sources with inter-source time delays. In particular, we report on the detection of an unexpected and previously unreported secondary auditory image under certain conditions of inter-source time delay. This secondary auditory image was perceived when one source was undelayed, and the other two sources had time delay values within the precedence effect range, but had a time delay relative to each other within the summing localization range.

Received 1 October 1999,
accepted 2 October 2000.

## 2. Method

### 2.1. Subjects

The three subjects were one of the authors (30-year-old male postgraduate student), a 29-year-old female biology postgraduate student and a 22-year-old male physics undergraduate student. All subjects had normal hearing (<15 dB hearing loss between 500 Hz and 8.0 kHz, measured using a B&K Audiometer Type 1800). Each subject received extensive training with the stimuli and procedures used in this experiment before the actual experimental measurements were taken. The subjects were trained in localizing the speakers used in the experiment, and to distinguish the positions of the auditory images formed by the precedence effect.

### 2.2. Experimental layout and apparatus

The three sound sources were three loudspeakers A, B and M in the same horizontal plane as the listener as shown in Figure 1. A and B were placed in front of and at equal distances of 2.5 m from the listener with A at 45° to the left of the forward position and B at 45° to the right. This is a standard configuration for two stereophonic loudspeakers. The third loudspeaker (M) was placed directly in front of and also at 2.5 m from the listener i.e. at 0°. The signals

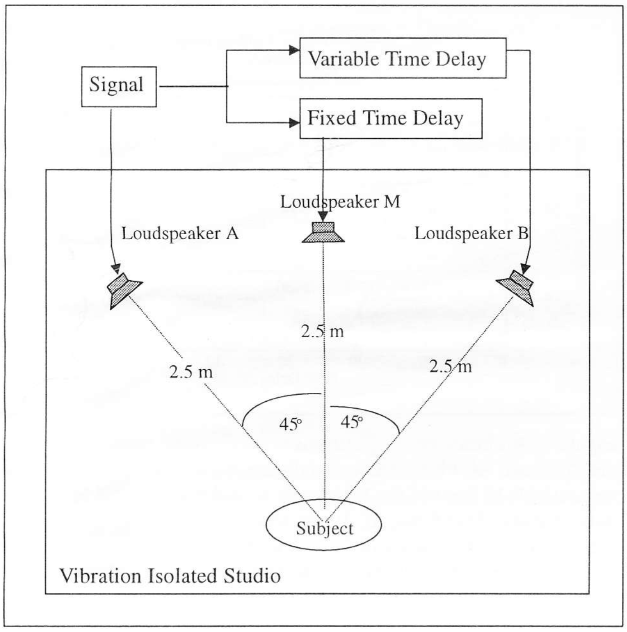


Figure 1. Experimental layout and block diagram of apparatus (pointer loudspeakers omitted for clarity).

used were male speech recordings (approximately 5 syllables per second), and the sound intensity of each loudspeaker was set to give the same level of 60 dB SPL(A) at the listener's position.

The loudspeakers used were Yamaha 4" full-range graphite moving coil units mounted in small totally sealed and enclosed cabinets measuring 0.13 m wide by 0.18 m high by 0.13 m deep. The frequency response of these loudspeakers was within 3 dB over the working range of 250 Hz to 4 kHz. All distances and angular positions of the loudspeakers were defined relative to the midpoint of the subject's head. The major axis of each loudspeaker was always pointed to the subject. The experiments took place in a vibration-isolated studio. The studio's surfaces were covered with sound absorbent material so that it had a very short reverberation time (0.3 s).

In order to help the subjects to locate the auditory image positions, 8 additional loudspeakers (The same model as mentioned above) were placed at equal distances 2.5 m away from and facing the subject. Four of these loudspeakers were positioned between loudspeakers A and M, and 4 between loudspeakers M and B, in the same horizontal plane as A, M and B. The three loudspeakers A, M and B and these 8 additional loudspeakers thus formed an array of 11 loudspeakers equally spaced in front of the subject with an angle between each adjoining loudspeaker of 9°. In this experiment, the 11 loudspeakers were used as "pointer loudspeakers" to enable the subjects to confirm auditory image positions between A and B. In order to use the loudspeakers A, M and B either as pointer loudspeakers or as the experimental sound sources, a two-way switch was used to select their required function.

The chair in which each subject was seated was adjusted so that the subject was approximately in the same horizontal plane as the loudspeakers' centres. The subject was seated facing loudspeaker M, i.e. at 0° and with his or her head position kept constant as far as possible within an adjustable head clamp.

## 2.3. Procedure

The test signal was sent to loudspeakers A, M and B. The signal to loudspeaker A was not subject to a time delay, but the signal to loudspeaker M had a fixed time delay which was beyond the precedence threshold but below the echo threshold. Thus the signal to M had a fixed time delay at values of (a) 6 ms (b) 10 ms and (c) 20 ms. These values were selected according to other researchers' data [1, 2], and also according to our previous studies. For each fixed value of M's time delay, the listener was able to vary the time delay of B between zero and a value well into the echo effect range. This could be affected by the listener via a time delay controller (YAMAHA D2040) with variation of time delay from zero to 999 ms in steps of 21 $\mu$s. Using this controller, the listener could vary the time delay upward and downward repeatedly as desired by him/her in order to fix a time delay value, which corresponded to a particular auditory image. During each experiment, the test signal was sent continuously to the three loudspeakers A, M and B, and the settings of the time delay between loudspeaker B and the pair of loudspeakers A and M could be varied by the subject.

When the listener varied the time delays of B from zero upwards, she/he was asked to note the positions of any auditory images. For auditory images in between the loudspeakers A and B, the 11 pointer loudspeakers (only one of which could be active at any one time) could be used to determine the positions of the auditory images. The speech signal sent to the pointer loudspeakers was different from that sent to loudspeakers A, M and B. The subject could use the control panel to choose which pointer loudspeaker was active. By using the control panel, when one pointer loudspeaker was chosen, the signal to this pointer loudspeaker was switched on, and the test signal to loudspeakers A, M and B was simultaneously switched off. The subject could then switch off the pointer loudspeaker to reactivate loudspeakers A, M and B with the test signal to restore the auditory image. In this way, the subject could freely choose one pointer loudspeaker and compare the auditory image position with the position of the chosen pointer loudspeaker. By adjusting the time delay value of B, the subject could shift the auditory image position until it matched the position of one of these 11 pointer loudspeakers. Typically each listener would match a position several times with a variability of time delay reading within $\pm 10\%$. From our experimental experience, we found that the auditory image position could thus be fixed more accurately than if the subject had to estimate its position unaided by the pointer loudspeakers.

In this experiment, the subjects were asked to mainly pay attention to the positions of the auditory image. The variations of overall spatial extent noticed by subjects were not recorded. For a certain auditory image position,

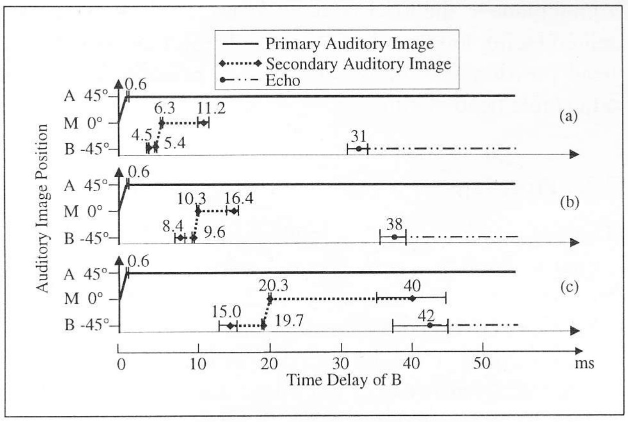


Figure 2. Auditory image positions relative to the three loudspeakers A, B and M which are at the directions of +45°, -45° and 0° respectively in front of the listener. A is undelayed and M is at fixed delay values of (a) 6 ms, (b) 10 ms and (c) 20 ms. The numbers on the graphs are the values of B's time delay for the corresponding events. The bars indicate the variability of the data points among the three subjects.

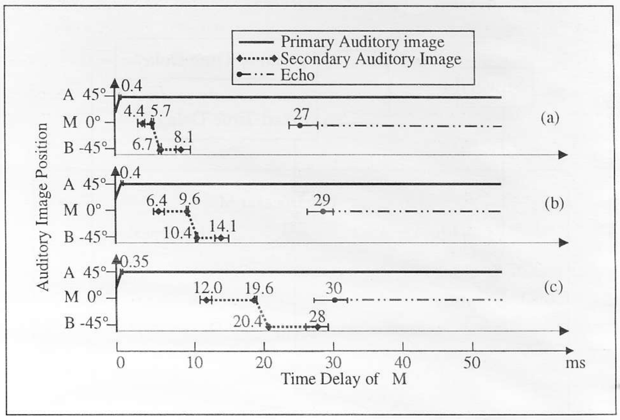


Figure 4. Auditory image positions relative to the three loudspeakers A, B and M which are at the directions of +45°, -45° and 0° respectively in front of the listener. A is undelayed and B is at fixed delay values of (a) 6 ms, (b) 10 ms and (c) 20 ms. The numbers on the graphs are the values of M's time delay for the corresponding events. The bars indicate the variability of the data points among the three subjects.

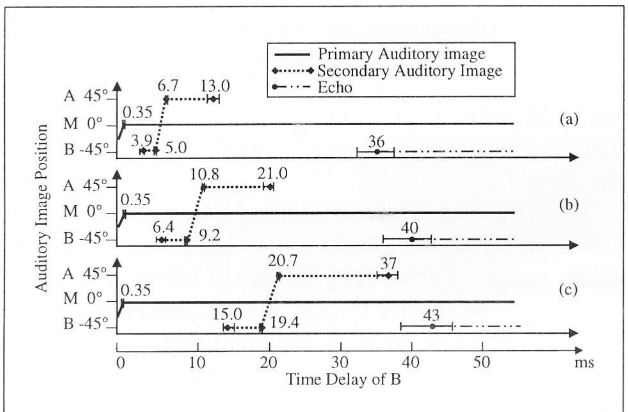


Figure 3. Auditory image positions relative to the three loudspeakers A, B and M which are at the directions of +45°, -45° and 0° respectively in front of the listener. M is undelayed and A is at fixed delay values of (a) 6 ms, (b) 10 ms and (c) 20 ms. The numbers on the graphs are the values of B's time delay for the corresponding events. The bars indicate the variability of the data points among the three subjects.

Table I. Time delay of B (to A) and auditory image position in the task of fixed time delay of 6 ms between A (leading) and M.

| Subjects | YG | YJ | PK | Average |
|---|---|---|---|---|
| Auditory image position | Time delay of loudspeaker B | | | |
| PAI at A | 0.60 | 0.64 | 0.60 | 0.6 |
| SAI appeared | 3.8 | 5.3 | 4.5 | 4.5 |
| SAI at B | 5.4 | 5.3 | 5.4 | 5.4 |
| SAI at M | 6.4 | 6.3 | 6.3 | 6.3 |
| SAI disappeared | 12.2 | 10.4 | 11 | 11.2 |
| Echo at B | 28 | 33 | 32 | 31.0 |

Table II. Time delay of B (to A) and auditory image position in the task of fixed time delay of 10 ms between A (leading) and M.

| Subjects | YG | YJ | PK | Average |
|---|---|---|---|---|
| Auditory image position | Time delay of loudspeaker B | | | |
| PAI at A | 0.55 | 0.64 | 0.60 | 0.6 |
| SAI appeared | 7.8 | 9.0 | 8.3 | 8.4 |
| SAI at B | 9.5 | 9.6 | 9.6 | 9.6 |
| SAI at M | 10.3 | 10.4 | 10.3 | 10.3 |
| SAI disappeared | 17.8 | 15.0 | 16.4 | 16.4 |
| Echo at B | 35.0 | 41.0 | 38.0 | 38.0 |

the subject could repetitively adjust the time delay upwards and downwards until she/he reached a conclusion. This time delay was thus recorded as a result of that auditory image position.

## 3. Results

The results are shown in Tables I to III. Using the data from the tables, we have plotted Figure 2. In Figure 2, the Y-axis denotes the positions of the auditory images in the horizontal plane of the three speakers with respect to the listener. The corresponding values of the time delay of B are shown on the X-axis. Each data point in the figures represents the average of observations by the 3 subjects. The curves were plotted from these average data points. The variability among the three subjects was also shown in the figures. Because two types of auditory image were perceived by the subjects, the main or primary auditory image is indicated by solid lines while the secondary auditory image is indicated by dotted lines. The listeners were required to record the time delays for all positions of the auditory images observed, and they were able to do this for any of the image positions at A, M, B and the 8 pointer loudspeakers in between. In Tables I, II and III and Figures 2, 3 and 4, only the time delay values for auditory images at A, M and B are shown, as these are the key po-

Table III. Time delay of B (to A) and auditory image position in the task of fixed time delay of 20 ms between A (leading) and M.

| Subjects | YG | YJ | PK | Average |
|---|---|---|---|---|
| Auditory image position | Time delay of loudspeaker B | | | |
| PAI at A | 0.60 | 0.64 | 0.60 | 0.6 |
| SAI appeared | 14.0 | 15.8 | 15.2 | 15.0 |
| SAI at B | 19.7 | 19.6 | 19.7 | 19.7 |
| SAI at M | 20.3 | 20.3 | 20.3 | 20.3 |
| SAI disappeared | 34.0 | 45.0 | 42.0 | 40.3 |
| Echo at B | 36.0 | 45.0 | 44.0 | 41.7 |

sitions along the path of the shifting auditory images. For clarity, other values of time delay obtained by the listeners for positions at the other than A, M or B are not explicitly shown as data points on the curves.

We describe the behaviour of the primary auditory image first. For all three fixed values of M's delay, as shown by (a), (b) and (c) in Figure 2, the primary auditory image behaves similarly as the time delay of B is increased from zero. When the time delay of B is zero, the primary auditory image is at M, the midpoint of A and B. As B's time delay is increased, the primary auditory image moves towards A and stays at A when the time delay goes beyond the precedence threshold, as indicated by the solid lines. These results are similar to the two-loudspeaker case i.e. if only A and B were present [1, 6, 7].

This could be easily explained by considering A and M as forming a resultant auditory image at A due to the precedence effect. Summing localization between B and this resultant at A gives an auditory image that starts at the midpoint of A and B, i.e. M, and moves towards A as the time delay of B is increased from zero towards and beyond the precedence threshold with respect to A. Hence the behaviour of the primary auditory image is not surprising, acting as though only A and B were present.

However, a weaker secondary auditory image was perceived by all three subjects when the time delay of B was increased further, as shown by the dotted lines in (a), (b) and (c) in Figure 2. For all three cases, this secondary auditory image was initially perceived at B and then it moved from B towards M as the time delay of B was increased. After it had reached M, it stayed there as the time delay of B was further increased, and then disappeared.

Further increases in the time delay of B beyond the fusion threshold resulted in the perception of B separately from A as an echo. The dashed-dotted lines indicate this.

## 4. Discussion and summary

We believe that the secondary auditory image is due to a summing localization interaction between B and M, because the secondary auditory image was perceived only when the time delay of B relative to M came within the summing localization range. Furthermore, the movement of the secondary auditory image from B towards M is consistent with summing localization behaviour between B and M, for an increase from zero of B's time delay relative to M. This secondary auditory image was unlikely to be due to an echo effect because the time delay values of B relative to both A and M were not in the echo effect range when the secondary auditory image was perceived.

What is surprising is that M is capable of interacting in this way with B, even though M is already combining with A in a precedence effect interaction. Hence M can simultaneously form a precedence effect auditory image with A and a summing localization image with B, since the primary auditory image formed by the precedence effect could be heard simultaneously with the weaker secondary auditory image formed by summing localization. Hence the "suppressed" delayed-source in a precedence effect situation is not really suppressed at all. In our experiment, it behaves like any other sound source in the environment.

We also repeated the above experiment but with the roles of A, M and B interchanged. The results at each time were in accordance with the above findings, but with the corresponding roles of each loudspeaker A, M and B interchanged. These results are shown in Figures 3 and 4. These other experiments therefore confirmed the validity of our results with the original configuration.

The fact that summing localization could operate between the two lagging sources in our experiments is reinforced by Blauert's assertion [1] that when there are more than two sound sources, regardless of where they are placed, summing localization can also occur. Litovsky et al [2] also stated that the precedence effect does not represent complete suppression of the lag, but rather a strong dominance by the lead. This discovery of a secondary auditory image, which is part of a more comprehensive study of three-source systems, may hopefully provide a better understanding of auditory perception with multiple sound sources, such as occurs in domestic theatre systems as well as in sound reinforcement systems.

### References

[1] J. Blauert: Spatial hearing–the psychophysics of human sound localization. The MIT Press, Cambridge, Massachusetts, 1997, Revised edition.

[2] R. Y. Litovsky, H. S. Colburn, W. A. Yost, S. J. Guzman: The precedence effect. J. Acoust. Soc. Am. **106** (1999) 1633–1654.

[3] H. Warncke: Die grundlagen der raumbezuglichen stereophonischen uberstrgung im tonfilm. Akust **Z. 6** (1941) 174–188.

[4] D. M. Leakey: Further effects produced by two loudspeakers in echo-free conditions. J. Acoust. Soc. Am. **29** (1957) 966.

[5] H. Wallach, E. B. Newman, M. R. Rosenzweig: The precedence effect in sound localization. J. Psychology **13** (1949) 315–336.

[6] H. Haas: The influence of a single echo on the audibility of speech. J. Audio Eng. Soc. **20** (1972) 146–159.

[7] P. M. Zurek: The precedence effect and its possible role in the avoidance of interaural ambiguities. J. Acoust. Soc. Am. **67** (1980) 952–964.

[8] W. A. Yost: Sound source determination and segregation. J. Acoust. Soc. Am. **102** (1997) 3097.

[9] W. M. Hartmann: Listening in a room and the precedence effect. – In: Binaural and spatial hearing in real and virtual environments. R. Gilkey, T. Anderson (eds.). Lawrence Erlbaum Associates, Mahwah, New Jersey, 1997, 191–210.