Available online at www.sciencedirect.com
ScienceDirect
Hearing Research 238 (2008) 118–123
Hearing Research
www.elsevier.com/locate/heares

Research paper

# Spectral integration time of the auditory localisation system

Russell L. Martin *, Ken I. McAnally [1]

Air Operations Division, Defence Science and Technology Organisation, P.O. Box 4331, Melbourne 3001, Australia

Received 15 June 2007; received in revised form 27 August 2007; accepted 30 August 2007
Available online 5 September 2007

**Abstract**

For the elevation and front-versus-back hemifield of a sound source to be accurately determined, the sound must contain a broad range of frequencies.

Experiment 1 of this study examined the spectral integration time of the auditory localisation system by measuring the accuracy with which frequency-modulated (FM) tones of modulation periods ranging from 0.5 to 200 ms can be localised. For each of the four participants, judgements of sound–source elevation and front–back hemifield were most accurate for a modulation period of 5 ms. Accuracy levels for the 5 ms modulation period approached those for a pink-noise stimulus. This suggests that the spectral integration time of the auditory localisation system is around 5 ms.

Supporting evidence for this conclusion was sought in experiment 2, in which two participants localised noise stimuli that had magnitude spectra identical to those of 5 ms equivalent-rectangular-duration samples of the FM tones from experiment 1. For both participants, functions relating localisation error measures (i.e., elevation error and frequency of front–back confusion) to modulation period for spectrally matched noises were similar to those for FM tones.


*Keywords:* Sound localisation; Time window; Temporal window

## 1. Introduction

It is well-established that a sound's energy must be distributed across a broad range of frequencies for the elevation and front-versus-back hemifield of its source to be accurately determined (e.g., Hebrank and Wright, 1974; King and Oldfield, 1997; Roffler and Butler, 1968). However, the effect on sound localisation of varying the time interval within which the required range of frequencies is presented has not been widely studied. To the present authors' knowledge, the only study to have examined this issue is that by Hofman and Van Opstal (1998), in which sound–source locations were restricted to the oculomotor range (i.e., were within 30° of azimuth and elevation of directly ahead). Stimuli in that study were linear frequency sweeps (sweeping downwards from 16 to 0 kHz) that ranged in duration from 1.28 to 40.96 ms and were presented consecutively for 500 ms. Hofman and Van Opstal found that the slope of the line-of-best-fit relating the true and perceived elevation of the source (which they called the elevation gain) decreased as sweep period increased above 2.56 ms. For sweep periods of 2.56 ms or less, the elevation gain approached that for a broadband noise stimulus. As the period of the next slower sweep in Hofman and Van Opstal's study was 5.12 ms, these results suggest that, for the purpose of estimating sound–source elevation, spectral information is integrated over an interval of less than 5.12 ms.

The spectral integration time of the auditory localisation system is an important issue that warrants further investigation. Experiment 1 of the present study examined the

*Abbreviations:* FM, frequency-modulated; HRIR, head-related impulse response

* Corresponding author. Tel.: +61 3 96267115.
*E-mail addresses:* russell.martin@dsto.defence.gov.au (R.L. Martin), ken.mcanally@dsto.defence.gov.au (K.I. McAnally).
[1] Tel.: +61 3 92627251.

0378-5955/$ - see front matter 
doi:10.1016/j.heares.2007.08.006

accuracy with which frequency-modulated (FM) tones of modulation periods ranging from 0.5 to 200 ms can be localised when source locations are not restricted to the frontal hemifield. The inclusion of source locations in the front and back hemifields allowed the assessment of front–back hemifield, as well as elevation, discrimination. As will be described later, the results of that experiment suggest that the localisation system's spectral integration time is around 5 ms. Supporting evidence for this conclusion was sought in Experiment 2, in which participants localised noise stimuli that had magnitude spectra identical to those of 5 ms equivalent-rectangular-duration samples of the FM tones from experiment 1. For both experiments, stimuli were presented in virtual acoustic space rather than from a free-field source. This allowed a more rapid presentation rate of localisation trials and precluded the possibility that cues associated with head movements could be used to localise sources.

## 2. Methods

### 2.1. Participants

One female and three male employees of the Defence Science and Technology Organisation participated in experiment 1. Two of these employees also participated in experiment 2. The ages of the four ranged from 30 to 46 years.

Each participant's absolute thresholds for 1, 2, 4, 8, 10, 12, 14 and 16 kHz pure tones were measured using procedures described by Watson et al. (2000). For all participants, all thresholds were lower than the relevant age-specific norm plus one standard deviation (as reported by Corso, 1963, and Stelmachowicz et al., 1989).

All participants had considerable previous experience in experiments assessing sound localisation. One, P2 in Figs. 1 and 2, was a co-author of this manuscript. Informed consent was received from all.

### 2.2. Head-related impulse response measurement

Head-related impulse responses (HRIRs) were measured for each participant using a blocked-ear-canal technique as described in detail by Martin et al. (2001). Briefly, miniature microphones (Sennheiser, KE4-211-2) were inserted in the participant's left and right ear canals. The participant was seated in a 3 m × 3 m, sound-attenuated, anechoic chamber at the centre of a 1 m radius hoop on which a loudspeaker (Bose, FreeSpace tweeter) was mounted. The participant placed his or her chin on a rest that helped position the head at the centre of rotation of the hoop. Head position and orientation were monitored using a magnetic tracking device (Polhemus, 3Space Fastrak). HRIR measurements were not made unless the participant's head was stationary, no more than 3 mm from the hoop centre in the $x$, $y$, and $z$ directions, and oriented within 1° of straight and level.

HRIRs were measured for 448 locations distributed more-or-less evenly around the participant and extending from 0 to 350° in azimuth and from −50 to 80° in elevation. For each location, two 8192-point Golay codes were generated at a sampling rate of 50 kHz (Tucker-Davis-Technologies, System II) and presented through the hoop-mounted loudspeaker. Impulse responses were derived from samples of the signals at the two in-ear microphones (Zhou et al., 1992). Immediately following HRIR measurement, the impulse responses of the microphones were measured together with those of the headphones (Sennheiser, HD520 II) through which stimuli would be presented during localisation trials. This was achieved by playing Golay codes through the headphones, which had been carefully placed on the participant's head, and sampling the responses of the microphones. The impulse response of the hoop-mounted loudspeaker, which had been measured previously, and those of the microphones and headphones were deconvolved from the HRIRs by division in the complex frequency domain.

### 2.3. Stimulus generation

FM tones (400 ms in duration including 50 ms cosine-squared rise and fall time) with exponential sine frequency contours (sweeping across the frequency range from 0.5 to 19.5 kHz) were generated at a sampling rate of 50 kHz using the Matlab (Mathworks) function *modulate*. The use of continuously FM tones precluded the possibility that stimuli would contain transients associated with frequency discontinuities. The rate of frequency modulation was varied across conditions such that the modulation period was 0.5, 1, 2, 5, 10, 20, 50, 100, or 200 ms.

Stationary noise stimuli (328 ms in duration including 50 ms cosine-squared rise and fall times) were generated at a sampling rate of 50 kHz from 10 ms Hanning-windowed (i.e., 5 ms equivalent-rectangular-duration) samples of the above-described FM tones. Each noise stimulus had a magnitude spectrum identical to that of the sample from which it was derived. For tones with modulation periods of 0.5, 1, 2, 5, 10, or 20 ms, samples were centred at random points anywhere on the modulation's cycle. For tones with modulation periods of 50, 100, or 200 ms, samples were centred at random points on the rising or falling phase of the modulation's cycle where the instantaneous frequency of the tone was less than or equal to 13.65 kHz. This ensured that all samples contained frequencies in the range within which all participants had reasonable auditory sensitivity. In all cases, no sample was taken from the first or the last 50 ms of the tone.

As the frequency of the FM tones was modulated by an exponential sinusoid, the energy in each sweep was distributed approximately equally across critical bands. Localisation of FM tones (experiment 1) and noises with magnitude spectra that matched those of samples of the FM tones (experiment 2) was therefore compared with localisation of pink noise. Bursts of pink noise (400 or 328 ms in