A CASE OF SUPERIOR AUDITORY SPATIAL ATTENTION

Robert Quelch[1], Gurjit Singh[1,2] , Kathy Pichora-Fuller[1,2]
[1]Dept. of Psychology, University of Toronto, 3359 Mississauga Rd North, Mississauga, Ontario, Canada L5L 1C6
[2]Toronto Rehabilitation Institute, 550 University Ave, Ontario, Canada M5G 2A2

1. INTRODUCTION

A growing body of research investigating the behavioural and neurophysiological responses of professional musicians to simple auditory stimuli has revealed enhanced responding to a number of auditory cues, including spatial location, pitch, and timing [1]. With regard to auditory spatial attention, the research suggests that musical experience involving the simultaneous deployment of attention toward both the entire musical 'scene' and its constituent components (i.e., orchestral conducting) can enhance mechanisms associated with both the acoustic (bottom-up) and cognitive (top-down) processing of sound [2].

Using more complex stimuli, other researchers have been able to show how the allocation of attention to spatial locations can significantly improve speech recognition scores in listening situations involving multiple talkers and uncertainty about the spatial location of the target [3]. Further, it appears that the top-down cognitive processing of sounds is influenced by a listener's ability to take advantage of the binaural auditory cues associated with spatial separation [4]. In the present experiments, we investigate the hypothesis that individuals with prolonged musical experience are better able to report on speech in multi-talker listening situations.

2. EXPERIMENT 1

2.1 Participant

RQ is a native English speaking 40-year old musician with 7 years of formal piano training. For 20 years, he played music professionally in a number of bands. RQ has a moderate bilateral hearing loss above 4 kHz and chronic bilateral tinnitus (Table 1).

2.2 Stimuli

The stimuli were sentences from the Coordinate Response Measure (CRM) corpus spoken by four male talkers [5]. Sentences have the format: "Ready [callsign] go to [colour] [number] now". For each talker, the CRM corpus contains sentences with all possible combinations of eight callsigns, four colours, and eight numbers.

2.3 Equipment and Presentation Conditions

All data were collected in a 3.3m^2 single-walled sound–attenuating booth. The stimuli were presented via Grason-Stadler Inc. loudspeakers: two when there was simulated spatial separation or three when there was real spatial separation. All stimuli were routed from a computer to a Tucker-Davis Technologies (TDT) System III to a Harmon/Kardon (HK3380) amplifier. Sentences were converted from digital files (sampling rate 24.414 kHz) to analogue using two TDT System III RP2.1s. Visual cueing and feedback was displayed on a Compar 17" NEC touch screen monitor positioned on a table located in front of the listener, below shoulder level.

2.4 Design

The dependent variable was accuracy of word identification. Three independent variables were systematically manipulated: callsign cue, location certainty, and presentation method (real or simulated spatial separation). Each block included 30 trials and the starting callsign cue condition for a block of trials was randomly determined. Within each session, for each callsign condition, the probability specifications were selected randomly without replacement. Data was collected on a total of 240 trials for each of the 16 conditions.

2.5 Procedure

On each trial, three sentences were selected randomly from the CRM corpus and presented simultaneously to the listener. All sentences were presented at 60 dBA. Participants were instructed to face forward (0° azimuth). One of the three sentences in each trial was randomly designated as the target sentence by providing the listener with its callsign and the other two sentences were considered maskers. The task was to select the colour and number in the target sentence using the touch screen. Both the correct colour and number were required for a correct response. The callsign cue was provided either 1 s before ('before' condition) or immediately after ('after' condition) each trial in a block. Four probability specifications were provided in advance to indicate the proportion of trials that the target sentence would be presented from the left, center and right spatial locations (*0-100-0, 10-80-10, 20-60-20, 33-33-33*). On any given trial, the location of the target sentence was randomly selected from among the three loudspeakers with the limitation that the probability cue was accurate across a block of 30 trials.

2.6 Results

In the callsign 'after' conditions, RQ performed similarly to the normal hearing younger adults tested in prior studies [3,4] with the exception of the certain location condition (0-100-0) in which he performed better than the 99% confidence interval (CI) of the younger adult group tested previously in identical conditions [4]. In the callsign 'before' conditions, RQ significantly outperformed the other listeners in all but one of the probability conditions (0-80-0). When the stimuli were presented using the precedence effect to simulate spatial separation, the reduced auditory cues associated with real spatial separation had an overall negative impact on RQ's performance, although, he significantly outperformed the young group in all of the location probability conditions (Fig. 1).

Table 1. Left (L) and right (R) hearing thresholds (dB HL) for RQ.

	Frequency (kHz)									
	0.25	0.5	0.75	1	1.5	2	3	4	6	8
L	0	15	20	20	25	10	10	20	35	45
R	0	5	10	5	20	10	10	15	40	50

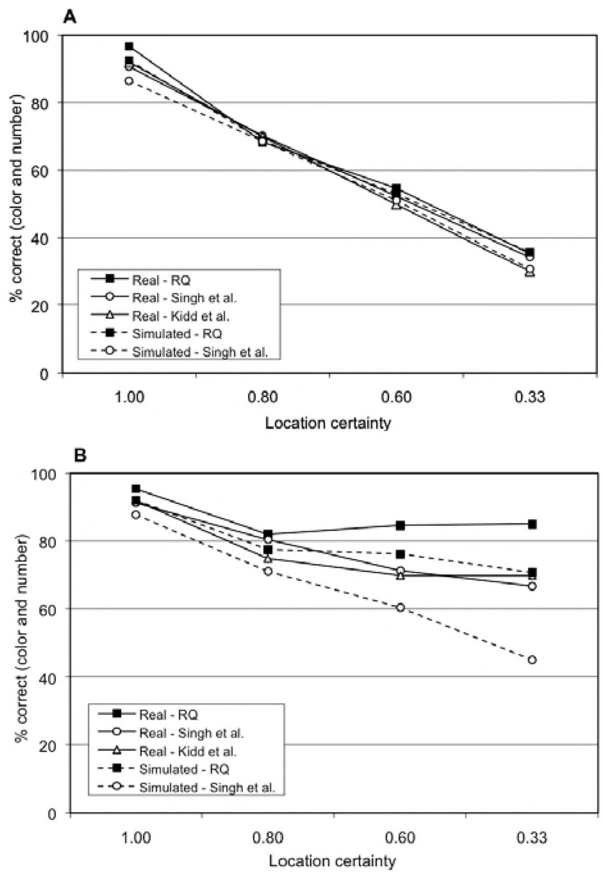

Fig 1. Mean proportion correct identification scores for RQ, Kidd *et al.*, Singh *et al.* plotted as a function of increasing target location uncertainty when the target identity was (A) unknown prior to the presentation of the stimuli and (B) known prior to the presentation of the stimuli.

3. EXPERIMENT 2

3.1 Participants

The listeners were 6 normal hearing musicians aged 17-25 years, each with a minimum of 5 years of music experience (mean = 6.5 yrs, *SD* = 3.2 yrs). All but one of the participants had formal music training (mean = 10 yrs, *SD* = 3.4 yrs) and all reported playing music at least weekly.

3.2 Procedure

The stimuli and procedures were similar to Experiment 1 with the exception that the callsign cue after and simulated spatial separation conditions were not presented. Data were collected on a total of 120 trials for each of the 4 conditions for each participant.

3.3 Results

When the location of the target sentence was certain, 4 of the 6 musicians significantly outperformed the non-musicians (P1, P3, P4, P5). When the spatial location of the target was uncertain, two of the musicians significantly outperformed the nonmusicians (P3, P4) (Fig. 2).

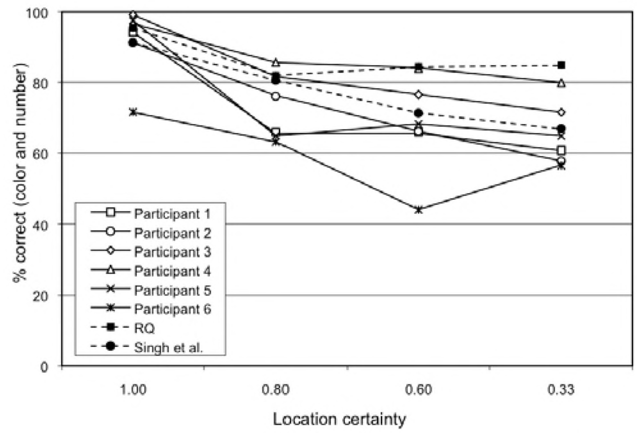

Fig 2. Mean proportion correct identification scores for the musician group, RQ, and the previously tested young group [4] plotted as a function of increasing target location uncertainty when the target identity was known prior to stimulus presentation.

4. DISCUSSION

Although the results of these experiments are not sufficient to define the precise mechanisms supporting the high levels of performance reported for RQ and some of the musicians, overall, they suggest that prolonged musical experience may, in some cases, enhance a listener's ability to comprehend speech in complex listening environments. This may be due to an improved ability to process the various auditory cues due to spatial separation resulting in an enhanced ability to selectively attend to and/or divide attention among the competing talkers. The results obtained for RQ in Experiment 1, in which the precedence effect was used to reduce the auditory cues associated with real spatial separation, underscore the importance of top-down cognitive processing of auditory information. Further, it appears that this ability may help offset the detrimental effects that high-frequency hearing loss and chronic tinnitus might have on hearing, listening and comprehending in complicated listening environments.

REFERENCES

[1] Münte, T.F., Nager, W., Beiss, T., Schroeder, C., & Altenmüller, E. (2003). Specialization of the specialized: Electrophysiological investigations in professional musicians. *Ann. N.Y. Acad. Sci., 999,* 131-139.

[2] Nager, W., Kohlmetz, C., Altenmüller, E., Rodriguez-Fornels, A., Münte, T.F. (2003). The fate of sounds in conductor's brains. *Cognitive Brain Research*, *17*, 83-93.

[3] Kidd, G., Arbogast, T.L., Mason, C.R., & Gallun, F.J. (2005). The advantage of knowing where to listen. *J. Acout. Soc. Am.*, *118*, 3804-3815.

[4] Singh, G., Pichora-Fuller, M.K. & Schneider, B. (in press). The effect of age on auditory spatial attention in conditions of real and simulated spatial separation. *J. Acout. Soc. Am.*

[5] Bolia, R.S., Nelson, W.T., Ericson, M.A., & Simpson, B.D. (2000). A speech corpus for multitalker communications research. *JASA, 107*, 1065-1066.

This work was funded by the Canadian Institutes of Health Research and the Natural Sciences and Engineering Research Council of Canada.