



Auditory attentional filter in the absence of masking noise

Elan Selvi Anandan¹ · Ruby Husain¹ · Kumar Seluakumaran¹

Accepted: 13 November 2020 / Published online: 3 January 2021
© The Psychonomic Society, Inc. 2021

Abstract

Signals containing attended frequencies are facilitated while those with unexpected frequencies are suppressed by an auditory filtering process. The neurocognitive mechanism underlying the auditory attentional filter is, however, poorly understood. The olivocochlear bundle (OCB), a brainstem neural circuit that is part of the efferent system, has been suggested to be partly responsible for the filtering via its noise-dependent antimasking effect. The current study examined the role of the OCB in attentional filtering, particularly the validity of the antimasking hypothesis, by comparing attentional filters measured in quiet and in the presence of background noise in a group of normal-hearing listeners. Filters obtained in both conditions were comparable, suggesting that the presence of background noise is not crucial for attentional filter generation. In addition, comparison of frequency-specific changes of the cue-evoked enhancement component of filters in quiet and noise also did not reveal any major contribution of background noise to the cue effect. These findings argue against the involvement of an antimasking effect in the attentional process. Instead of the antimasking effect mediated via medial olivocochlear fibers, results from current and earlier studies can be explained by frequency-specific modulation of afferent spontaneous activity by lateral olivocochlear fibers. It is proposed that the activity of these lateral fibers could be driven by top-down cortical control via a noise-independent mechanism.

Significance

The neural basis for auditory attentional filter remains a fundamental but poorly understood area in auditory neuroscience. The efferent olivocochlear pathway that projects from the brainstem back to the cochlea has been suggested to mediate the attentional effect via its noise-dependent antimasking effect. The current study demonstrates that the filter generation is mostly independent of the background noise, and therefore is unlikely to be mediated by the olivocochlear brainstem reflex. It is proposed that the entire cortico-olivocochlear system might instead be used to alter the hearing sensitivity during focus attention via frequency-specific modulation of afferent spontaneous activity.

Keywords Attention: Neural mechanisms · Attention: Selective · Hearing · Olivocochlear · Auditory efferent

Introduction

Selectively focusing on a specific signal frequency improves hearing sensitivity towards the attended frequency. This phenomenon has been traditionally studied using the probe-signal method, which measures detection of near-threshold tonal signals in the presence of background noise (Botte, 1995; Dai et al., 1991; Greenberg & Larkin, 1968; Scharf et al., 1987; Tan et al., 2008). Subjects are typically led to focus on a certain signal frequency (targets) either by presenting them on most trials or by adding preceding cue tones that match

the target frequency. In a small proportion of trials, instead of targets, equally detectable tones of distant frequencies (probes) are presented. When detection rates are plotted as a function of signal frequencies, sharply tuned frequency-response curves centered at the target frequencies (i.e., auditory attentional filters) are obtained (Botte, 1995; Dai et al., 1991; Greenberg & Larkin, 1968; Scharf et al., 1987; Tan et al., 2008). The width of attentional filters is roughly one critical band (CB) centered at the attended frequency while its depth amounts to about 6 dB in equivalent signal level (Dai et al., 1991; Tan et al., 2008).

The neural basis underlying the generation of auditory attentional filters is poorly understood. Early studies argued for the presence of a theoretical filter that could differentially process relevant and irrelevant auditory inputs either at an early or later stages of information processing (Broadbent, 1958; Norman, 1968). Since then, there has been a long-

✉ Kumar Seluakumaran
kumarselvakumaran@um.edu.my

¹ Department of Physiology, University of Malaya Faculty of Medicine, Kuala Lumpur, Malaysia

standing debate regarding the level at which the attentional filter is generated (central vs. peripheral) and whether such filtering is achieved by facilitating the attended stimuli, suppressing unattended stimuli, or both (reviewed in Giard et al., 2000).

When listeners direct their attention to a particular signal frequency, there are rapid changes in the neuronal response pattern at appropriate frequency areas in the auditory cortex (Da Costa et al., 2013; Grady et al., 1997; Riecke et al., 2018). These changes are likely mediated via short-term neuronal modulations that involve both excitatory and inhibitory inputs (Jääskeläinen et al., 2011; Jääskeläinen & Ahveninen, 2014). Interestingly, focused listening can also lead to frequency-specific changes in auditory midbrain neurons (Riecke et al., 2018; Slee & David, 2015) and cochlear responses (Maison et al., 2001), suggesting that attention-mediated changes in the cortex may be relayed back to lower auditory processing centers including the peripheral organ via a top-down control mechanism (Giard et al., 2000; Yakunina et al., 2019). The auditory cortex is linked to the subcortical auditory centers and cochlea via a complex neural network called the auditory efferent system. It includes multiple corticofugal pathways that project from the auditory cortex to the thalamus, midbrain, and other lower brainstem auditory centers (Robertson, 2009; Terreros & Delano, 2015). One of the efferent components, the olivocochlear bundle (OCB), projects from the brainstem back to the cochlea (Rasmussen, 1946). There is a direct anatomical and functional link between the corticofugal fibers and the OCB neurons (Mulders & Robertson, 2000; Yakunina et al., 2019) that allows the auditory cortex to directly influence the incoming auditory signals at the earliest stage of information processing.

Several studies have implicated the cortico-olivocochlear pathway in top-down modulation of cochlear responses during selective attention (Lukas, 1980; Maison et al., 2001; Puel et al., 1988). However, the strongest evidence supporting the involvement of OCB in the generation of auditory attentional filters comes from studies conducted in a group of patients with Ménière's disease who underwent vestibular neurectomy, a surgical procedure that also severed the OCB (Scharf et al., 1994; Scharf et al., 1997). The underlying disease itself did not appear to affect the normal functioning of the OCB in these patients before the surgery (Scharf et al., 1997). However, when attentional filters were measured pre- and post-surgery, patients who initially detected unexpected probe tones poorly compared to the expected targets, could clearly detect the probes after the operation (Scharf et al., 1997). Based on the post-operative changes, Scharf et al. (1997) postulated that in a normal individual, the OCB could selectively suppress non-attended signal frequencies during focused attention on the target frequency. The OCB consists of two separate divisions: the medial olivocochlear (MOC) and the lateral olivocochlear (LOC) fibers. Both neuronal

groups receive sound-driven inputs from the cochlear nucleus, but the MOC brainstem reflex pathway (involving auditory nerve fibers, cochlear nucleus neurons, and MOC fibers) and its cochlear effects have been better characterized (Liberman & Brown, 1986; Warren & Liberman, 1989). Besides acoustic input, MOC and LOC fibers can also be influenced by the basal activity of the descending cortical neurons (Dragicevic et al., 2015; Jäger & Kössl, 2016; León et al., 2012; Suga et al., 2002). It is unclear whether the action of the OCB during selective frequency listening is achieved via its brainstem-level acoustic reflex circuitry or by the influence of corticofugal projections.

Tan et al. (2008) have argued that the attentional filter, at least in part, is generated by the MOC brainstem acoustic reflex via its antimasking effect in the cochlea (Kawase et al., 1993; Kawase & Takasaka, 1995; Winslow & Sachs, 1988). In quiet, MOC stimulation leads to a suppression of the auditory afferent sound-evoked activity. This effect is mediated via the MOC inhibitory action on the cochlear outer hair cell micromechanical response (Gifford & Guinan, 1987; Warren & Liberman, 1989). However, when continuous background masking noise is present, MOC activation leads to an increase in the afferent response to transient tonal signals by suppressing the afferent activity to the noise and releasing the neurons from adaptation (Kawase et al., 1993; Kawase & Takasaka, 1995; Winslow & Sachs, 1988). According to Tan et al. (2008), a similar antimasking effect could take place when listeners focus on target frequencies during the probe-signal task. MOC neurons, when primed by the cue tones, could suppress the background noise-driven afferent activity. This effect would reduce the amount of masking on target tones, leading to an improvement in their detection. On the other hand, without MOC activation at distant frequencies, probe tones would relatively receive more masking from the background noise and its detection would be suppressed (Tan et al., 2008). This mechanism is possible as the OCB innervation to the cochlea is frequency-specific (Liberman & Brown, 1986; Robertson et al., 1987).

Tan et al.'s (2008) antimasking hypothesis, however, is at odds with findings reported by Scharf et al. (1997). Without an MOC antimasking effect in the operated ears of vestibular neurectomy patients, their detection of target tones during the probe-signal task should have been poorer after the surgery, but this was not the case. Target detection levels elicited pre- and post-surgery were largely unchanged. Tan et al. (2008) argued that this disagreement could arise partly due to a non-optimal activation of the MOC brainstem reflex by the monoaural stimuli used by Scharf et al. (1997). Due to the crossed and uncrossed patterns of the MOC fibers, binaural input is much more effective than monoaural signals in driving the MOC brainstem reflex (Liberman, 1988). However, attentional filter depths measured by Scharf et al. (1997) in some of their patients were similar to those found in normal

subjects using binaural stimulation (Dai et al., 1991; Tan et al., 2008), suggesting that the attentional effect with monaural condition can be as strong as the binaural.

An alternative method to study the role of OCB in selective frequency listening, and in particular the validity of the MOC antimasking hypothesis, is by comparing auditory attentional filters generated with and without the presence of background noise in normal-hearing subjects. During a probe-signal task, the presence of background noise could influence the functioning of MOC brainstem reflex in at least two ways. Firstly, continuous binaural broadband noise, as used in almost all earlier studies (Botte, 1995; Dai et al., 1991; Greenberg & Larkin, 1968; Scharf et al., 1987; Tan et al., 2008), is a potent stimulus for direct activation of MOC neurons. Secondly, the MOC-mediated antimasking effect is only effective if background noise is available to mask the tonal signals (Kawase et al., 1993; Kawase & Takasaka, 1995). If Tan et al.'s (2008) suggestion is correct, when tested without the presence of binaural background noise (i.e., in quiet), subjects' target detection should be impaired during the probe-signal task, leading to a reduced attentional filter depth.

A previous study investigated the effect of varying background noise levels (overall levels between 20 and 70 dB) on attentional filters (Botte, 1995). Attentional filters measured at the lowest noise level were almost similar to those obtained at higher noise levels (Botte, 1995), implying that the MOC-mediated antimasking effect could not have played a major role in the filter generation (low levels of noise would have produced only a small degree of masking). However, even at low levels, continuous binaural background noise used by Botte (1995) could have still directly activated the MOC brainstem reflex (Lieberman, 1988). Also, in some afferents with high spontaneous activity, MOC activation could still enhance (or unmask) the detection of weakly masked tonal signals by suppressing the spontaneous neural activity (Kawase et al., 1993). To confidently exclude the contribution of background noise in the generation of attentional filters, it is important to extend Botte's (1995) study to include measurement of attentional filters in quiet.

In the current study, the involvement of OCB during focus listening was addressed in three separate experiments that involved measurement of attentional effects in quiet and in the presence of continuous binaural masking noise in a group of normal-hearing listeners.

Experiment 1

The main aim of the first experiment was to compare auditory attentional filters elicited using signals presented with and without background noise in the same group of subjects. This was done to evaluate the contribution of background noise, if any, to the attentional filter generation.

Methods

Subjects

Five adults (one male and four females) aged 20–38 years with normal otoscopic and tympanometric findings served as subjects. Their audiometric thresholds were < 20 dB HL at octave test frequencies between 0.5 and 8 kHz. Two of the subjects were authors of this paper who were unaware of the study outcome when the task was performed, while the remaining subjects were naïve paid volunteers. Previous studies of similar nature have shown that neither training nor awareness of the experimental design affects a subject's performance (Scharf et al., 1987; Dai et al., 1991). Based on the large effect size ($d_z = 15.15$) of the attentional filter depth estimated from a similar study (Tan et al., 2008), a sample size of five subjects should be adequate to achieve a statistical power of 0.80. Informed consent was obtained from volunteers and the study protocol was approved by the university's medical ethics committee (MEC 201310-0360).

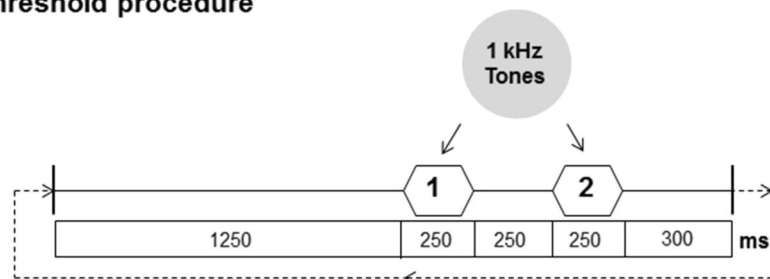
Sound stimuli

Tones and white noise were generated digitally at a sampling rate of 44.1 kHz (16-bit precision) via a computer fitted with a Creative Sound Blaster X-Fi sound card and running LabVIEW 13.0 software (National Instruments, Austin, TX, USA). Stimuli were presented diotically via Sennheiser HD 201 (Wedemark, Germany) headphones and calibrated using an ear simulator (KEMAR 43AG, G.R.A.S. Sound and Vibration, Holte, Denmark) connected to a sound-level meter (Norsonic Nor140, Tranby, Norway). For the noise condition, broadband noise (0.2–18 kHz) was set at an overall level of 65 dB SPL. Spectrum levels of the noise recorded from the headphones were 20 ± 1 dB/Hz dB between 0.6 and 1.6 kHz.

Threshold measurements

Before the probe-signal task, threshold levels for 1-kHz tones (both in quiet or noise) were first determined for each subject using two-interval forced-choice trials (2IFCTs). The temporal structure of a single threshold tracking trial is shown in Fig. 1a. Cue tones were not present, and the to-be-detected signals were always 1-kHz tones. Signals appeared at equal probability during one of the two 250-ms observation intervals. The other interval contained no stimulus. The two observation intervals were separated by a 250-ms blank interval. They were marked with the numerals "1" and "2" and displayed visually on a computer screen inside a sound-treated booth. A "respond now" message appeared after the second interval, and subjects indicated which of the two temporal intervals they thought contained the signal by clicking either the left or the right mouse button. The subject's response initiated the

a Threshold procedure



b Experiment 1

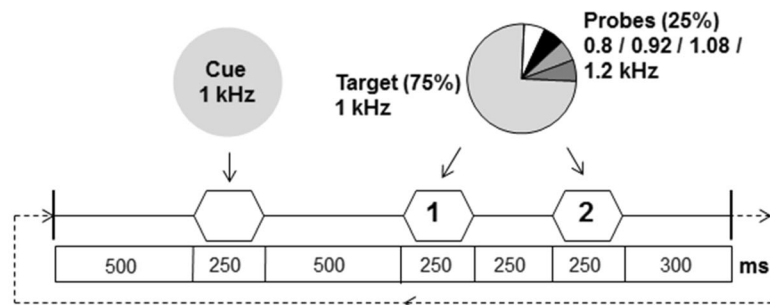


Fig. 1 Temporal structure of a single two-interval forced-choice trial used for the threshold tracking procedure (a) and probe-signal task of Experiment 1 (b). Only one of the two intervals (1 or 2) contained the signal. See text for further explanation

next trial sequence and visual feedback was provided for their responses after every trial.

Signal levels were varied adaptively using a “one-up, three-down” staircase procedure that tracks the 79.4% correct detection point on the psychometric function (Levitt, 1971). The initial step size of 5 dB was reduced to 1 dB after the first incorrect response, and the average of the last five reversals was taken as their threshold. Each threshold run consisted of a total of 80 trials and took about 10 min to complete. Prior to the first threshold measurement, a short practice session lasting about 20 trials using clearly audible 1-kHz tones was provided to the naïve subjects.

Experimental procedure

The probe-signal method used for the measurement of attentional filters involves a signal detection task with frequently presented target tones and occasionally presented probe tones. Subjects’ detection rates were measured using 2IFCTs similar to the threshold trials except that they always began with a suprathreshold 1-kHz cue tone set at 14 dB above the threshold level. The cue was followed, after a 500-ms delay, by the two observation intervals, one of which was equally likely to contain either a target tone (1 kHz) or one of the four possible probe tones (Fig. 1b). In each block of trials, the target frequency (1 kHz) was presented during 75% of the trials while the four probe tones appeared at random with equal probability on remaining 25% of the trials.

Targets and probes were always presented at the threshold level determined from the threshold runs. Based on the

estimated CB of about 160 Hz at 1 kHz (Scharf, 1970), two of the probes were set at frequencies one-half CB away from the target (0.92 and 1.08 kHz) while the remaining two were set at slightly more than one CB away (0.8 and 1.2 kHz). Initially, each subject performed a total of three testing sessions in quiet: one session per day for three consecutive days. Daily sessions consisted of one threshold run and three blocks of probe-signal trials (128 trials in each block), which lasted approximately 40 min. Hence, after the 3-day session, each subject completed a total of nine blocks of trials with 864 observations at the target frequency and 72 observations at each of the four probe frequencies. Subsequently, subjects underwent a similar 3-day testing in noise condition. The background noise was present throughout the 2IFCTs. Subjects’ performance was calculated from percent correct detection of the signals from the trial blocks.

Results and discussion

Subjects’ threshold levels, as determined from the adaptive threshold-tracking procedure, were 9.0 ± 2.2 (3.3–11.6) [$M \pm SD$ (range)] dB SPL in quiet and 37.7 ± 0.8 (36.4–38.5) dB SPL when the background noise was present. Their detection rates as a function of signal frequencies in both conditions are shown in Fig. 2 and were qualitatively similar. Mean performance at the target frequency (1 kHz) was close to the expected 79.4% detection ($77.8\% \pm 2.7$) in noise but was slightly higher in quiet ($84.4\% \pm 4.7$). Detection rates declined close to chance level (50%) for the most deviant probe frequencies in both conditions, indicating that the subjects

mostly did not hear them. Results obtained in quiet and noise conditions were compared using two-way repeated-measures ANOVA (background noise and signal frequencies as factors) after verifying data normality using the Kolmogorov-Smirnov test. There was a highly significant effect on the detection of signal frequencies, $F(4,16) = 35.75$, $p < 0.0001$, $\eta^2 = 0.721$, but not for the presence of background noise, $F(1,4) = 0.926$, $p = 0.390$, $\eta^2 = 0.013$, and noise by frequency interaction, $F(4,16) = 0.453$, $p = 0.769$, $\eta^2 = 0.012$. Tukey's post hoc multiple-comparison test revealed significantly higher detection levels for 1-kHz target tones compared to all four probe tones ($p < 0.01$) for both quiet and noise conditions. However, detection rates across different probe frequencies in the two conditions did not differ significantly ($p > 0.05$).

In order to compare attentional filters obtained in quiet and noise, the main parameter of interest was the filter depth. Attentional filter depths can be obtained by transforming percent correct detections into attenuation in dB using d' values (Botte, 1995; Buus et al., 1986; Dai et al., 1991). However, this procedure was not applied to the current data as only a limited number of signal frequencies were available. Furthermore, the derived attenuation values from d' can be overestimated when detection levels are close to chance level (this was the case for the two most distant probes) (Botte, 1995; Dai et al., 1991). Instead, the filter depths (in percent detection) were calculated from the difference between the target detection level and the average detection rates of the two most distant probes. Comparison of filter depths in the two conditions (Quiet: $31.5\% \pm 5.6$; Noise: $27.0\% \pm 6.5$) showed no significant difference (paired t-test, $p = 0.388$, $d_z = 0.429$). This simplified procedure is justified as the attentional filters are symmetrical on both sides of the 1-kHz center frequency when they are measured using low-to-moderate noise levels (Botte, 1995). The shape of listeners' psychometric functions for detection of target and probe tones in both

noise as well as quiet is similar, with an average slope of about 5%/dB (Buus et al., 1986; Dai et al., 1991; Watson et al., 1972; Wright, 1973). Hence, for every 1-dB change in tone intensity, subjects' percent correct detection should change by about 5% in both conditions. Based on this assumption, the depth of the attentional filters should be approximately 6.3 dB in quiet and 5.4 dB in noise. These values are close to the 6-dB depth reported by Tan et al. (2008) when a similar testing paradigm and 60 dB of background noise were used to measure the attentional filters. The current results do not support the involvement of a noise-dependent MOC antimasking effect in the generation of attentional filters. As discussed in the *Introduction*, if the antimasking effect was responsible for the attentional filter, listeners' target enhancement during the probe-signal attentional task should have been impaired when the noise was absent. This, in turn, should have produced a shallower attentional filter in quiet compared to those measured in noise. Neither of these was found in the current study.

Besides the filter depth, we can also surmise from the current data that the width of attentional filters in quiet and noise should be roughly similar. When measured using tones in noise, attentional filter widths are typically about one CB, indicating that attentional focus is able to improve the hearing sensitivity (or signal-to-noise ratio) within one CB centered at the attended target frequency (Buus et al., 1986; Dai et al., 1991; Tan et al., 2008). Based on the CB value of 160 Hz at 1 kHz (Scharf, 1970), the upper and lower cut-off frequencies for the CB are 0.92 and 1.08 kHz. From the current data, when 0.92 and 1.08 kHz probes were used, detection rates fell to about 60% in both quiet and noise conditions. This resembles the results obtained in earlier studies (Dai et al., 1991; Greenberg & Larkin, 1968; Scharf et al., 1987; Tan et al., 2008) and indicates that attentional filter widths in both quiet and noise measured in the current study should be within one CB.

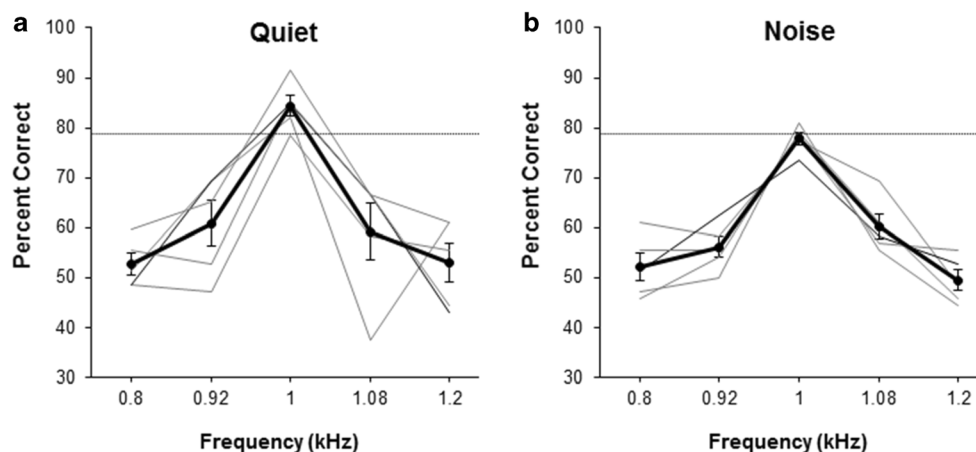


Fig. 2 Percentage correct detection (Mean \pm SEM) of 1-kHz targets and each of the four probes (0.8, 0.92, 1.08 and 1.2 kHz) for quiet and noise conditions. Grey lines represent results of individual subjects while horizontal dotted lines represent the expected detection level based on the

threshold setting procedure (79.4%). Data points are based on 4320 trials (864 trials per subject) for targets and 360 trials (72 trials per subject) for each of the probes ($N = 5$)

A more sophisticated way to compare attentional filter bandwidth in quiet and noise is by calculating the equivalent rectangular bandwidth (ERB) value derived from a curve-fitting procedure involving the attentional filter and the equivalent rounded exponential (ROEX) equation for peripheral auditory filters obtained from separate notch-noise experiments (Glasberg & Moore, 1990). This was not done in the present study as the limited number of probe frequencies measured here would give rise to large errors in filter-shape estimation and the corresponding ERB calculation. A relevant study (Botte, 1995) that obtained ERB values for attentional filters measured using a larger number of signal frequencies (but within a similar frequency range to the current study) revealed similar ERB values for filters in low (spectrum level: -2 dB SPL/Hz) and moderate noise levels (18 dB SPL/Hz) (see Table II in Botte, 1995). Their moderate level of background noise is close to that used in the current study (20 dB SPL/Hz). However, slight widening of the attentional filters was noted at much higher noise levels (>28 dB SPL/Hz; Botte, 1995), but this could simply reflect the broadening of the underlying peripheral auditory filters (Botte, 1995; Glasberg & Moore, 1990). Taken together with the present results, Botte's (1995) finding that varying background noise levels did not have a major effect on the shape of attentional filters again suggests that the underlying mechanism for the generation of these filters is likely to be independent of the background noise.

Besides the role of MOC antimasking effect in target enhancement, Tan et al. (2008) also proposed that the suppression of probes during attentional filter measurements could be mediated by a long-term expectation of the target signals, possibly via an experience-dependent cortical mechanism. In order to investigate if there is any evidence for a gradual increase in the probe suppression after the beginning of the task, data from the 3-day sessions in both quiet and noise conditions were also analyzed according to individual trial blocks (Fig. 3). Subjects' target detection remained consistently higher than the probes in both conditions. The poorer detection of probe tones was apparent from the first block of testing and remained close to chance level throughout the 3-day testing period. It generally takes only < 10 min for subjects to complete each trial block. Hence, the underlying mechanism of probe suppression should have been engaged within the order of minutes and no further systematic effect occurred for the duration of testing. While there are reports available on the lack of any systematic trend on target detection during prolonged probe-signal task-testing periods (Greenberg & Larkin, 1968), the absence of any clear long-term trend in the probe suppression suggests that its underlying mechanism is also likely to be independent of any long-term learning effect, at least within the 3 days of testing.

Experiment 2

In the first experiment, subjects were led to focus on the target frequency in two ways. Firstly, cue tones that matched the target frequency were added to all trials, and, secondly, targets were presented on most trials (75%) to induce an expectation of the target frequency. Probe signals, which were presented infrequently (25% of the trials), were detected at significantly lower levels compared to the targets. However, before the task, only thresholds for 1-kHz tones were measured in individual subjects, and probes with deviant frequencies were presented at equivalent intensities with the assumption that all signals would be equally detectable without any focused listening. Although this assumption is acceptable for tones presented in background noise (Green et al., 1959), detection thresholds in quiet may vary at these frequencies. Large differences in absolute thresholds have been reported when test frequencies only differed by 100 Hz (Long, 1984) and this could be a source of error in the estimation of filter depth in the quiet condition. In order to investigate this issue, the same five subjects were subjected to additional fixed-frequency trials in quiet in which along with 1-kHz tones, the two most distant frequencies used in the first experiment (0.8 and 1.2 kHz) were also set as targets, appearing in 100% of the trials in each block. If the three signal frequencies have equivalent thresholds, when presented at equivalent intensities, all of them should be detected close to the expected level of 79.4% (Levitt, 1971).

A separate issue related to the first experiment is that targets in quiet were detected at a slightly higher level (84.4%) compared to the targets in noise (77.8%). This difference was statistically significant (paired t-test, $p < 0.05$; $d_z = 1.485$). During the threshold setting procedure, cue tones were absent, but they were introduced during the testing. One possibility is that the presence of cue during the trials somehow further improved the detection of targets in quiet but not in noise. Although the benefit offered by the cue appears to be small, it is possible that the presence of probe tones during the remaining 25% of the Experiment 1 trials introduced some amount of uncertainty in the detection of the signals and this effect countered the actual benefit offered by the cue. To investigate whether the presence of cue tone would further enhance detection of signals that are already expected to occur in every trial, cued trial blocks were also added in the current experiment.

Methods

Thresholds in quiet were determined using methods similar to those described in the first experiment and using only 1-kHz tones. Experimental trials were identical to the first study except that the frequency of the to-be-detected signals was now fixed throughout each block (without any probes) at one of the

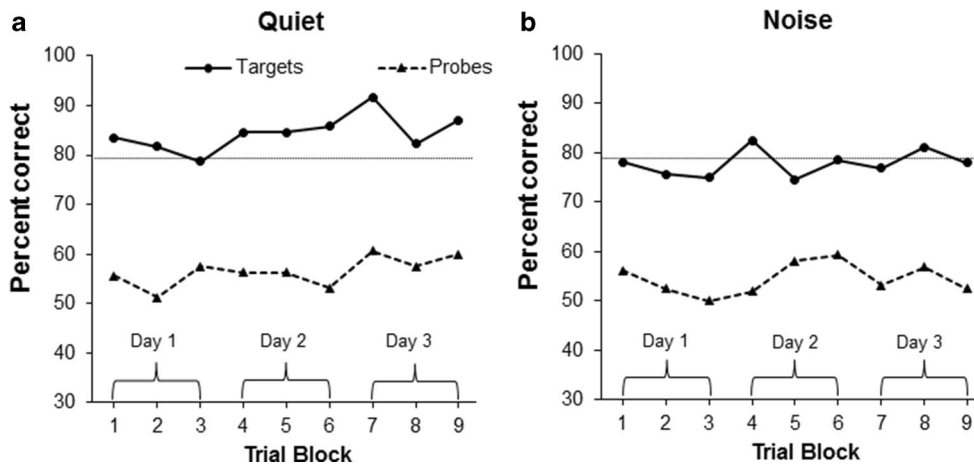


Fig. 3 Mean percentage correct detection of 1-kHz targets and probes for individual trial blocks (block 1-9) during the 3-day testing sessions for quiet and noise conditions. Data for the probes were pooled from four signal frequencies (0.8, .92, 1.08 and 1.2 kHz). Data points are based on a

total of 480 trials (96 trials per subject) for targets and 160 trials (32 trials per subject) for probes ($N = 5$). Horizontal dotted lines represent the expected detection level based on the threshold setting procedure (79.4%)

following target frequencies: 0.8, 1, or 1.2 kHz (Fig. 4a). There were two conditions for each target frequency – cued and uncued. Subjects completed a total of three uncued and

three cued blocks (100 trials per block) in a single testing session that lasted about 1 h. Order of presentation for the frequency and cue conditions were randomized.

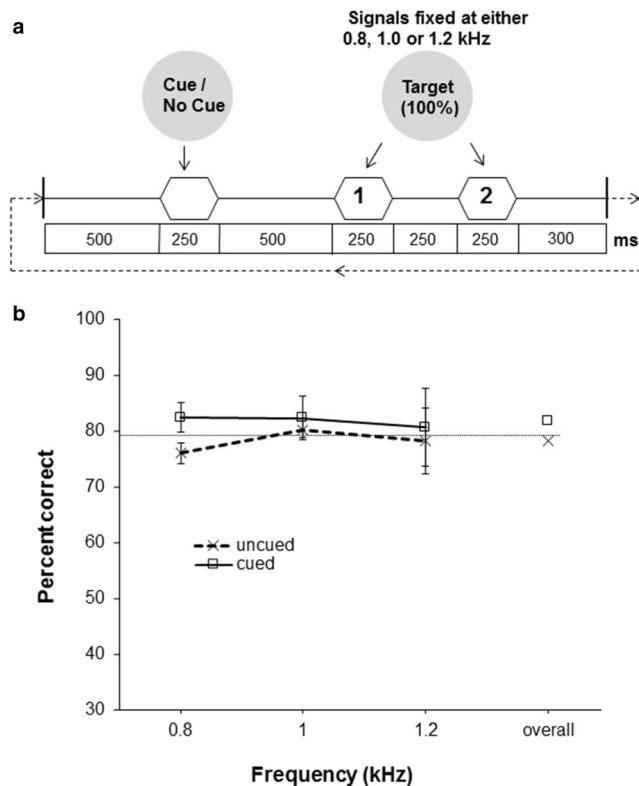


Fig. 4 a Trial structure for Experiment 2. Target frequency was fixed throughout a single block at either 0.8, 1 or 1.2 kHz. Cue, when present, always matched the target frequency. See text for further details. b Mean percentage correct detection (\pm SEM) of uncued and cued signals at all three signal frequencies. Each datum point is based on a total of 500 trials (100 trials per subject, $N = 5$). Overall column represents the average detection levels across all three frequencies for the uncued (\times) and cued (\square) conditions. Horizontal dotted lines represent the expected detection level based on the threshold setting procedure (79.4%)

Results and discussion

Subjects’ threshold in quiet was 10.6 ± 2.1 (8.1–13.3) dB SPL. When presented as the only signal during a given block of trials, the detection of all three signal frequencies (0.8, 1, and 1.2 kHz) were now at equal levels (Fig. 4b). Two-way repeated-measures ANOVA with cue and signal frequencies as factors revealed no significant main effects of cue, $F(1,4) = 2.008$, $p = 0.229$, $\eta^2 = 0.041$, frequency, $F(2,8) = 0.093$, $p = 0.912$, $\eta^2 = 0.011$, or a cue by frequency interaction, $F(2,8) = 0.406$, $p = 0.679$, $\eta^2 = 0.012$. The equivalent detection levels of these signals indicate that subjects’ hearing sensitivity did not differ across the tested frequencies. This validates the interpretation in Experiment 1 that the detection of probes to near chance level in quiet was due to the auditory attentional filter effect.

The average overall detection rates for cued and uncued trials were $81.77\% \pm 8.01$ and $78.16\% \pm 2.73$, respectively, and this difference was also not statistically significant (paired t-test, $p = 0.229$; $d_z = 0.633$). Post hoc power analysis with G*Power program (Faul et al., 2007) with $\alpha = 0.05$ showed power ($1-\beta$) of 0.196. Hence, the non-significant cue-evoked effect can be attributed to lower statistical power, mainly due to the small sample size used. Nevertheless, the cue-evoked enhancement obtained is relatively small (average improvement of $\sim 3.6\%$ across the three frequencies, as shown in the overall column of Fig. 4b).

Trials introduced during the uncued condition were similar to those used for threshold estimation, except for the variation in signal levels that is inherent to the adaptive tracking procedure. Hence, it is not surprising that subjects’ performance

during the uncued blocks is close to the expected level of 79.4%. However, when cues were introduced, only a small improvement in their performance were noted, and this agrees with the finding in Experiment 1. As individual trial blocks in the current experiment consisted only of target signals, the lack of a clear cue-evoked enhancement effect cannot be explained by any uncertainty introduced by the occasionally presented probes (as in the first experiment). It is also unlikely that the cue effect was limited by a ceiling effect at the upper portion of subjects' psychometric function as the psychometric function slope for detection of tones in quiet remains steep at the 79% point (Watson et al., 1972; Wright, 1973).

Tan et al. (2008) found that when subjects performed the probe-signal task with cued and frequently occurring targets, detection of targets was enhanced while probes were detected at chance level. This resulted in 6-dB depth of the attentional filter. However, when both cues and to-be-detected signals were selected randomly from a range of available signals, the presence of a cue only enhanced the detection of the frequency-matched targets (about 3 dB) but failed to suppress the detection of wrongly cued probes to chance level. In other words, the usual 6 dB depth of the attention filter was halved to 3 dB. Cue signals used by Tan et al. (2008) in the latter randomized testing paradigm were 0.84, 0.92, 1.0, 1.08, and 1.16 kHz, while the to-be-detected signals were five possible multiples (0.84, 0.92, 1, 1.08, and 1.16) of each cue frequency. This gave rise to a total of 25 possible frequency combinations during the trials. Since the frequency of the cues and to-be-detected signals was randomized from trial-to-trial, the trial design should not have induced any sustained expectation towards a particular signal frequency. Analysis of the data according to individual subjects and cue frequencies (available in Tan's PhD thesis (2008)) revealed that Tan et al.'s (2008) finding did not result from averaging the data across individual observers or trials and there was no systematic bias of attention towards a particular cue frequency. Tan et al. (2008) argued that the finding provided evidence that active probe suppression occurred as a result of the long-term expectation of target signals and was not a consequence of the preceding cue signal. Hence, the authors suggested that the cue-evoked enhancement effect and the suppression of probes were mediated via separate mechanisms, the former being a rapid trial-to-trial effect while the latter is a more enduring effect that is sustained throughout the trial block. If the cue-evoked target enhancement is a transient process and is independent of the long-term expectation of target frequency, why did the cueing in the current experiment fail to further improve target detection when the signals were presented repeatedly at the same frequency?

The lack of cue-evoked effect can be explained by considering that the target enhancement effect described by Tan et al. (2008) was, in fact, sustained from one trial to another. Utilizing the single-interval procedure and varying the delay

between onset of brief cue tones and the observation interval, Scharf et al. (2007) noted that frequency-matched cues could improve the hearing sensitivity to target signals presented as early as 52 ms after the onset of the cue. Once evoked, the cue-induced effect could completely overcome any frequency uncertainty at a cue-signal delay of 352 ms. This rapidly acting cue-effect can be sustained for at least several seconds after its onset without any significant decay (Green & McKeown, 2001). In the uncued trials (and the threshold-tracking trials), the to-be-detected signals were of the same frequency. Subjects could have used the repeatedly presented targets as cues to sustain their attention towards the signal frequency from one trial to another (in the current task, each trial lasted for 2.3 s). As there was only a small improvement of target detection when cues were added later during the cued blocks, the near-threshold "cross-trial cueing" during the uncued trials appears to be nearly as effective as the suprathreshold cues presented during the cued trials.

Experiment 3

The first experiment showed that the depth of attentional filters measured in quiet and noise was approximately 6 dB. If the argument provided by Tan et al. (2008) is correct, the filter depth, which represents the difference in listeners' hearing sensitivity to targets and the two most deviant probes, would have been contributed by both target enhancement (3 dB) as well as active probe suppression (3 dB). It is the cue-evoked target-enhancement component that has been suggested to be mediated by the MOC antimasking effect (Tan et al., 2008). However, during the probe-signal task in the first experiment, the cue-evoked effect was not measured independent of probe suppression. The third experiment was designed to isolate the cue-evoked enhancement component of the attentional filter using modified probe-signal trials, in which signal frequencies were varied randomly trial-to-trial. By varying the signal frequencies, the sustained inhibitory effect due to the long-term expectation of the signal frequencies can be excluded (Tan et al., 2008).

Methods

The same five subjects who participated in the first two experiments also took part in this study. The threshold tracking procedure and temporal structure of 2IFCTs used in the current experiment (Fig. 5a) were similar to those used earlier. However, trials contained signals at one of the following five frequencies (0.8, 0.92, 1.0, 1.08, or 1.2 kHz) selected at random, with each signal frequency appearing in 20 trials per block (probability of 0.20 on a given trial). Subjects completed a total of three uncued and three cued blocks (100 trials per block), initially in the quiet condition. Then, similar testing

was done in the presence of background noise with levels that were comparable to Experiment 1. In cued blocks, cue tones that matched the to-be-detected signal frequencies were added. Testing sessions alternated between uncued and cued blocks, and the cue-evoked effect was determined by comparing subjects' performance between the two. The six blocks of trials in each condition (quiet and noise) were completed in a single testing session that lasted about 1 h.

Results and discussion

The mean signal level used for the quiet and noise conditions were 10.5 ± 2.5 (7.4–13.7) dB SPL and 38.1 ± 0.38 (37.8–38.7) dB SPL, respectively. These levels were comparable with those used in the first two experiments. In the absence of any cues, the detection of randomly selected signals deteriorated in both quiet and noise conditions (Fig. 5b). The presence of a cue, however, clearly improved the detection of these tones. Three-way repeated-measures ANOVA with cue, background noise, and signal frequencies as factors revealed a highly significant effect of cue on the detection of the signals, $F(1,4) = 75.92$, $p < 0.001$, $\eta^2 = 0.263$, and a smaller but significant effect for the frequency condition, $F(4,16) = 3.694$, $p < 0.05$, $\eta^2 = 0.050$, indicating that some signal frequencies were better detected compared to others. However, the presence of background noise did not have any significant effect, $F(1,4) = 1.494$, $p = 0.288$, $\eta^2 = 0.046$. In the quiet

condition, detection of higher signal frequencies (1.08 and 1.2 kHz) appeared to be slightly better than the lower frequencies (0.8, 0.92, and 1.0 kHz) for both cued and uncued conditions (the cued graph is essentially an upward shift of the uncued). In addition, percent correct responses of the low frequencies (0.8, 0.92, and 1.0 kHz) in noise were better than those seen in quiet. However, Tukey's multiple comparison test failed to show any significant effect for the above comparisons.

Frequency-specific changes in detection rates (differences in percent detection in cued and uncued conditions) in both quiet and noise conditions are shown in Fig. 6a. The cue-evoked effects in both conditions were very similar for all test frequencies except for 0.8 kHz, where the effect was higher in noise compared to quiet. However, two-way repeated-measures ANOVA with presence of noise and signal frequencies as factors revealed no significant effects of background noise, $F(1,4) = 0.133$, $p = 0.733$, $\eta^2 = 0.004$, frequency, $F(4,16) = 1.380$, $p = 0.285$, $\eta^2 = 0.093$, as well as noise-by-frequency interaction, $F(4,16) = 0.368$, $p = 0.828$, $\eta^2 = 0.032$. Figure 6b depicts the change in detection rates in individual subjects when detection rates from all five signal frequencies were combined (overall) and for the 0.8-kHz signal alone. There is a large variation in the cue-evoked effects across individual subjects. Comparison of the overall cue-mediated improvements across both conditions (Quiet: $11.64\% \pm 4.78$; Noise: $12.90\% \pm 4.65$) revealed no significant difference (paired t-

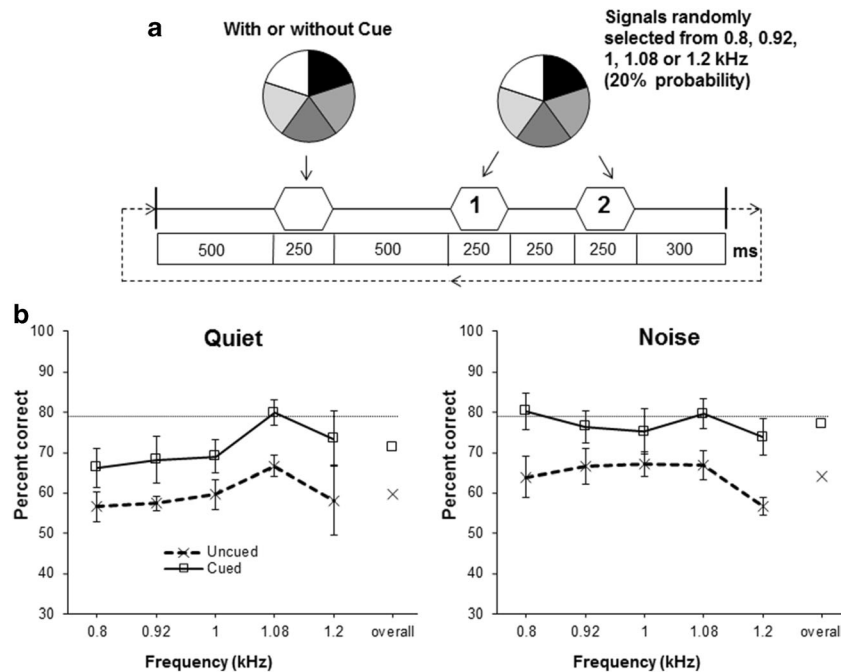


Fig. 5 a Trial structure for Experiment 3. To-be-detected signals were selected randomly from a set of five different frequencies (0.8, 0.92, 1.0, 1.08 or 1.2 kHz). Signals in the cued condition were always preceded by frequency-matched cues. **b** Mean percentage correct detection (\pm SEM) of uncued and cued signals in quiet and noise conditions. Each datum point is

based on 360 trials (60 trials per subject) ($N = 5$). Overall column represents the average detection levels across all five frequencies for the uncued (\times) and cued (\square) conditions. Horizontal dotted lines represent the expected detection level based on the threshold setting procedure (79.4%)

test, $p = 0.706$; $d_z = 0.181$). Post hoc power analysis with the G*Power program showed a power ($1-\beta$) of 0.061 ($\alpha = 0.05$). With a small effect size, it is unlikely that the difference would reach significance if the sample size was increased further. A very large sample size of 242 subjects will be required if any statistical significance was to be achieved with a power of 0.80. With the slope of psychometric functions taken as 5%/dB (see earlier description), the overall increase in detection levels ($\sim 12\%$) will amount to about 2.5 dB in equivalent sound level. This is only slightly smaller than the overall 3 dB cue-evoked effect reported by a separate study that used a larger range and a higher number of randomly selected signal frequencies (Scharf et al., 2007). A similar analysis for individual data for the 0.8-kHz signal frequency alone (Quiet: $9.60\% \pm 5.50$, Noise: $16.26\% \pm 9.82$) also showed no significant difference between the cue-evoked effects in both conditions but the effect size was larger (paired t-test, $p = 0.138$; $d_z = 0.827$, power = 0.296).

Evaluating the frequency-specific effects of the cue is important because MOC effects in the cochlea appear to be strongest at the lower frequencies (0.5–1 kHz) (Zhao & Dhar, 2012). Based on qualitative observation of data in both Figs. 5 and 6, it is tempting to suggest that the background noise has an influence on the performance of subjects at lower frequencies, particularly at 0.8 kHz. Firstly, the detection of lower signal frequencies (0.8, 0.92, and 1.0 kHz) in the noise condition was better than those in quiet (Fig. 5b). Secondly, the cue-evoked effect at 0.8 kHz was much larger when the background noise was present. However, as pointed out earlier, these comparisons were not statistically significant. Even if statistical significance is reached with a larger sample size, the role of background noise in the cue-evoked effects, if any, is likely to be minor. For example, at 0.8 kHz, about 10% improvement in detection levels (or ~ 2 -dB effect) can still be achieved by the cue despite the absence of background noise. Furthermore, the cue-evoked effect is not necessarily the strongest at 0.8 kHz as an equivalent effect is also seen at 1.2 kHz (Fig. 6a). Besides the MOC system, a separate sound-evoked brainstem efferent pathway that could elicit antimasking effects in the cochlea, particularly at the lower frequencies, is the middle ear muscle (MEM) reflex (Liberman & Guinan, 1998; Kawase et al., 1993). Unlike the MOC action on cochlear outer hair cells, effects elicited by MEM reflex are achieved by contraction of middle ear muscle and stiffening of the ossicular chain (Liberman & Guinan, 1998). However, current results are unlikely to be influenced by this reflex as the stimuli levels used here are far below the MEM reflex threshold (Gelfand, 1984).

The equivalent cue-evoked effects in quiet and noise conditions at most signal frequencies again suggests that the mechanism underlying the generation of the auditory attentional filters, particularly the cue-evoked enhancement, is largely independent of the background noise and a noise-

dependent process such as the MOC-mediated antimasking effect is unlikely to play any role in the attentional filtering.

General discussion

The neural basis underlying selective frequency listening remains one of the most fundamental but poorly understood areas in auditory neuroscience. There has been a suggestion that the MOC system, via its noise-dependent antimasking effect, could be responsible for the generation of auditory attentional filters (Scharf et al., 1994; Scharf et al., 1997; Tan et al., 2008). Since auditory selective attention is generally thought to aid the detection of anticipated signals in a noisy environment, previous probe-signal studies have always measured attentional filters in the presence of background noise (Botte, 1995; Dai et al., 1991; Greenberg & Larkin, 1968; Scharf et al., 1987; Tan et al., 2008). It is unknown

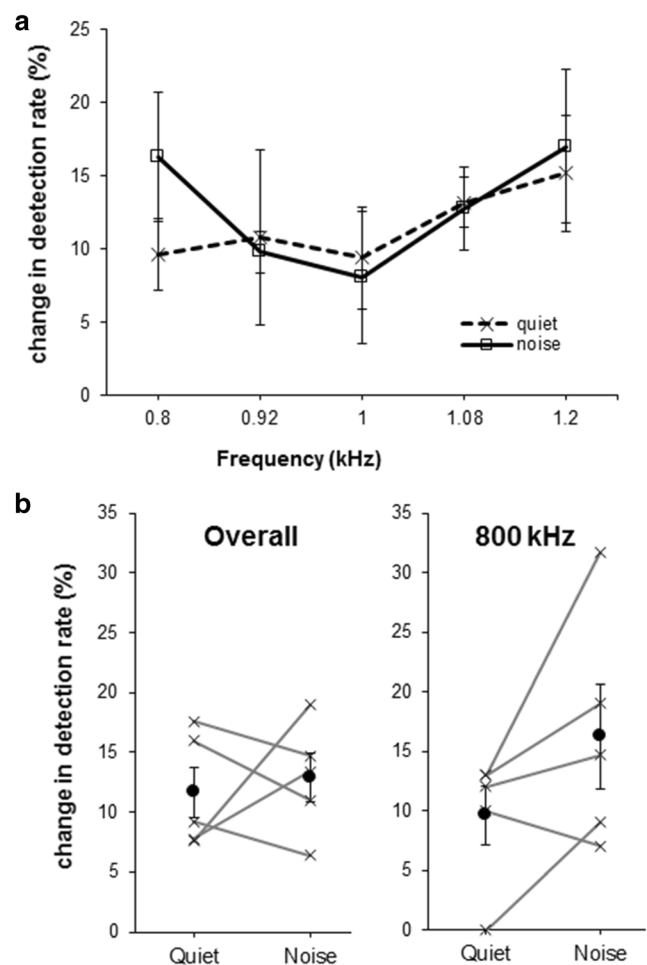


Fig. 6 **a** Mean (\pm SEM) change in detection rates (cued-uncued) at individual signal frequencies in quiet and noise conditions. **b** Change in detection rates in individual subjects for overall detection levels (all five signal frequencies combined) and 0.8 kHz alone. Filled circle represents the average (\pm SEM) change in all subjects ($N = 5$). Connecting lines across data points represent data obtained from the same subject

whether the background noise itself is critical for the generation of attentional filters. In the current study, the role of MOC was evaluated by comparing attentional filters in quiet and noise. As filters in both conditions had comparable depths, by inference, it is unlikely that OCB functioning during focused listening is mediated via a noise-dependent process such as the MOC antimasking effect.

Focused attention to auditory signals can evoke rapid and dynamic changes in the response pattern of tonotopically organized neurons at the primary and non-primary auditory cortices (Da Costa et al., 2013; Grady et al., 1997; Riecke et al., 2018). These changes are likely to be mediated via a short-term modulation in center-excitatory and surround-inhibitory neuronal inputs (Jääskeläinen et al., 2011; Jääskeläinen & Ahveninen, 2014), and could be driven by the feedback from the frontal cortex (Fritz et al., 2010). The quick and dynamic reshaping of the neuronal receptive field could readily explain the rapid attention-mediated changes (on the time scale of minutes or less) seen in the current study (Experiment 1) and the ability of the subjects to shift their focus from one frequency to another when these frequencies were presented randomly but cued on a trial-to-trial basis (Experiment 3). More importantly, in agreement with the current finding, such changes are likely independent of background masking noise (Fritz et al., 2007) and, therefore, would be equally effective in both quiet as well as noisy backgrounds. However, findings from patients with severed OCB (Scharf et al., 1994; Scharf et al., 1997) indicate that the changes in higher auditory centers alone cannot account for the attentional filtering. The reduced depth of the attentional filters (due to an improvement of probe detection) after the OCB was severed was noted only in the operated and not in the unoperated ear of the same patient (Scharf et al., 1997), implying that the post-operative deficit must involve a failure in OCB functioning. A likely explanation is that when the OCB was cut, cortical changes could not be relayed back to the cochlea (via the descending cortico-olivocochlear pathway) to elicit frequency-specific changes in hearing sensitivity during focused attention (Giard et al., 2000; Robertson, 2009). In support of this, results from a recent human imaging study showed that the superior olivary complex (the site of origin of the OCB) is the active target of cortical top-down modulation during attentional focus (Yakunina et al., 2019).

Attentional filters measured before and after the surgery in Scharf et al.'s (1997) patients showed that target-detection thresholds were unchanged at the expected frequency. However, detection of unattended probes improved markedly after the operation. In normal individuals, this argues for a strong role of the OCB, suppressing tones at unexpected frequencies, and, interestingly, no role at all for enhancing tones at the expected frequency. In other words, when someone listens for a tonal signal, the OCB acts to suppress hearing sensitivity over almost the entire frequency range, except for

the critical band containing the attended tone, where it has no effect. Based on the physiology of the OCB, the most straightforward mechanism to explain probe suppression in quiet is via the MOC inhibitory actions on auditory nerve responses at the unattended frequencies (Gifford & Guinan, 1987; Warren & Liberman, 1989). Instead of acoustic input, descending corticofugal fibers could also provide frequency-specific activation of the MOC fibers, and this is likely to be a noise-independent process. However, the antimasking effect would still apply to the actions of MOC fibers in the cochlea for the detection of masked tones. Hence, the mechanism cannot explain probe suppression when substantial background noise is present.

The interpretation that attentional filters are generated by probe suppression alone, however, assumes that the reference level for basal OCB activity is during the target detection where subjects know which signal frequency to attend to (frequency certainty), and probe detection is 6 dB below this sensitivity level. However, it is more intuitive to regard the basal reference as the level at which listeners do not know which signal frequency to focus on (frequency uncertainty). The filter depth can then be attributed by two processes, the 3-dB target enhancement (detection of targets above the uncertainty level) and 3-dB active-probe suppression (detection of probes below the uncertainty level) (Tan et al., 2008). In the ensuing discussion, the latter interpretation was used to explain an alternative OCB mechanism that reconciles the current finding (being noise-independent) with the two separate mechanisms proposed by Tan et al. (2008) and post-operative changes seen in vestibular neurectomy patients (Scharf et al., 1997).

Detection of weak signals is dependent not only on the presence of external noise but also the internal noise, which is added to the stimulus-evoked neuronal activity (Green, 1964). One main source of internal noise is the neural noise provided by a high level of spontaneous activity in auditory neurons (Eggermont, 2015). It is possible that changes in hearing sensitivity to near-threshold signals observed during the probe-signal task are caused by a change in the internal noise, particularly the afferent spontaneous activity. Instead of the MOC neurons, this role is likely mediated via the lateral division of the OCB (i.e., LOC fibers). There is evidence from animal experiments that LOC fibers could provide a tonic, excitatory effect that raises the spontaneous activity of primary auditory neurons (Liberman, 1990; Zheng et al., 1999). This effect is likely achieved via the release of neuroactive substances from the LOC terminals that directly contact the dendrites of primary afferents (Le Prell et al., 2014). The activity of LOC fibers can, in turn, be regulated by the basal activity of the cortex via its corticofugal projections (León et al., 2012), and this is consistent with the earlier suggestion that the changes during the attentional process are likely

driven by the whole cortico-olivocochlear system via a mechanism that is independent of the external noise.

The proposed mechanism for the LOC-mediated effects during focus attention is illustrated in Fig. 7. In normal individuals, tonic excitatory effects of LOC fibers could render the auditory nerve fibers to be spontaneously active, leading to a substantial level of internal neural noise in the afferents (Fig. 7a). When a listener focuses on the target frequency, LOC fibers innervating the attended frequency area could be inhibited by the corticofugal fibers (Fig. 7b). This could lead to a decrease in afferent spontaneous activity at the target frequency, giving rise to lower neural noise and improvement in hearing sensitivity. The results obtained from Experiment 3 (cued and uncued trials) can be explained by this mechanism. It represents the difference in listener's performance when he/she knows what signal frequency to listen to (frequency certainty) compared to when he/she does not know (frequency uncertainty) (Creelman, 1960; Green, 1961; Scharf et al.,

2007). This is one of the two mechanisms suggested by Tan et al. (2008), and was referred to as target enhancement.

During the probe-signal task, excitatory effects of LOC fibers could increase in CBs other than the band containing the attended signal, causing a further elevation of spontaneous activity (and neural noise) at the unattended frequencies (Fig. 7b). Such an effect would lead to a deterioration in sensitivity to the probes below the levels expected for no cue (or frequency uncertainty) condition. This was termed “active probe suppression” by Tan et al. (2008). It represents the performance difference when the listener does not know which frequency to listen for (frequency uncertainty) compared to those measured for unattended signals when his/her attention is focused to a different frequency (wrongly cued signals). Findings related to signal enhancement caused by frequency-matched cues and active suppression of wrongly cued signals are also reported elsewhere (Johnson & Hafter, 1980). However, the obtained effects were much smaller (in the order of 1–1.5 dB) compared to the 3 dB suggested by Tan et al. (2008), possibly due to differences in the trial structure and psychophysical task employed (see Table 3 in Johnson & Hafter, 1980).

When the OCB is cut (as in vestibular neurectomy), the loss in the excitatory LOC input to the auditory nerve fibers can lead to a drop in afferent spontaneous activity (Fig. 7c). Support for this comes from animal models that showed reduced afferent spontaneous activity either when the entire OCB was cut (Lieberman, 1990; Zheng et al., 1999) or when the LOC neurons were selectively disrupted at the brainstem (Le Prell et al., 2014). During the attentional task, the reduced afferent spontaneous activity is expected to remain in patients with severed OCB at both target and probe frequencies (Fig. 7d). The lower basal activity at the target frequency is not expected to produce any change in the sensitivity to targets as the spontaneous activity would be comparable to before the OCB was cut (see target frequency area during focused attention in Fig. 7b). This concurs with the finding that absolute and masked thresholds (as well as target detection) remained the same before and after the surgery in Scharf et al.'s (1997) patients as all these tasks would have been performed under focused attention. However, after the OCB was cut, a drop in spontaneous activity in the afferents innervating non-attended frequency areas would have led to better detection of the probes (Scharf et al., 1997).

It is important to consider how much of the overall attentional filter depth is contributed by OCB effects in the cochlea. Tan et al. (2008) suggested that only half of the 6-dB filter depth is mediated by the actions of the OCB, while the remaining 3 dB is established intrinsically within the cortex. This argument was supported by the analysis of filter depth changes in all of Scharf et al.'s (1997) patients (Tan, 2008). On average, the reduced depth of the attentional filter (due to better probe detection) was about 15% difference in detection rates, which corresponds to about 3 dB in effective stimulus

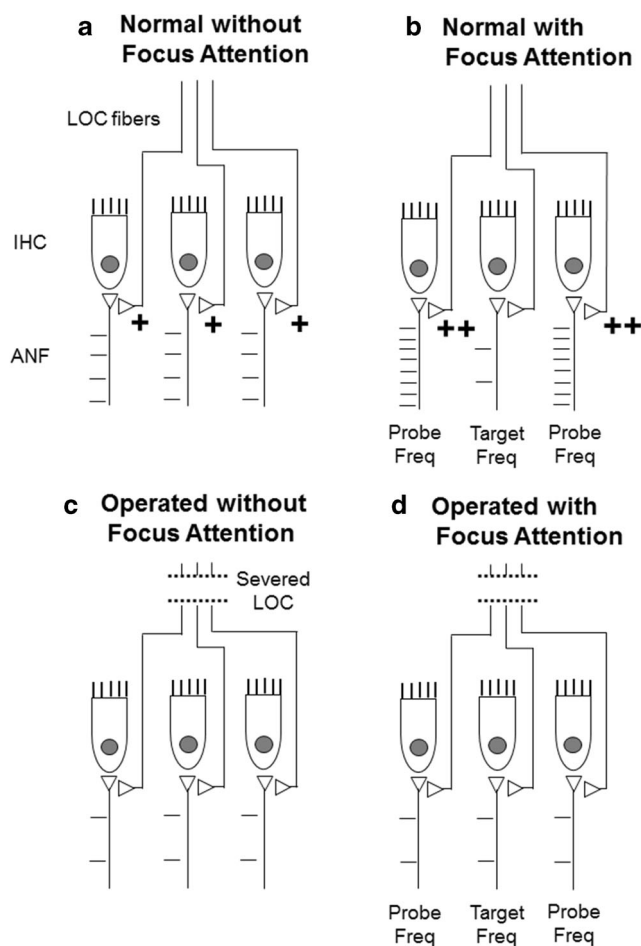


Fig. 7 Schematic diagram illustrating the possible role of LOC fibers during focus attention in normal (intact LOC; **a**, **b**) and operated (severed LOC; **c**, **d**) individuals. Dashed lines along the ANF represents the level of spontaneous activity (more lines indicate higher spontaneous activity). ‘+’ and ‘++’ indicate mild and strong excitatory effects. ANF: auditory nerve fiber; LOC: lateral olivocochlear; IHC: inner hair cell; Freq: Frequency. Refer to text for further details

level (Tan, 2008). Hence, the remaining 3 dB was attributed to the cortical mechanism. However, changes in filter depth varied substantially from patient to patient (see Fig. 6 in Scharf et al., 1997). In some cases, the loss of probe suppression was complete (probes were detected as well as targets). In these patients, the OCB-mediated mechanism should account for the entire 6-dB depth of the filter. Others showed only a small loss of filter depth, indicating that some aspects of the attentional filter can be retained despite severing the OCB. The variability in the loss of probe suppression among the patients might be contributed by individual differences in the basal OCB activity. In individuals with a weaker OCB activity, both the cortical process and OCB could act synergistically to produce a stronger attentional effect in the cochlea (Robertson, 2009), and severing the OCB would only partially reduce the attentional filter depth. Support for this also comes from a more recent study that measured attentional filters in a group of cochlear implant subjects in whom the OCB targets within the cochlea are likely disrupted (Bester et al., 2016). While some of these subjects showed a complete absence of attentional filtering, others retained a varying degree of ability to produce it, suggesting that in certain individuals, central mechanisms could also independently contribute towards the attentional filter generation.

The present study has several limitations. Firstly, comparison of attentional filter parameters in quiet and noise was limited by the number of probe frequencies tested on either side of the target frequency. This did not allow statistical comparison of filter widths in quiet and noise. Secondly, unlike the filter depth measured in the first experiment, the cue-evoked effects measured in Experiments 2 and 3 were small (around 1 dB). While this supports our conclusion that the presence of background noise did not play a major role in the attentional filter generation, it is possible that any minor but significant effect was missed as a result of low statistical power due to the small sample size used. This is supported by additional statistical analysis using the two one-sided tests (TOSTs) for equivalence (Lakens et al., 2018; Equivalence_Tests_TOSTER.xlsx, Version: 7, available at <https://osf.io/q253c/>). When the smallest effect size of interest was taken as the just noticeable difference of 1 dB (5%), the TOST procedure indicated that the observed overall cue-evoked effects were not significantly within the equivalent bounds of $\pm 5\%$ detection rates in both Experiment 2 ($t(4) = -0.55$, $p = 0.306$; based on data from Fig. 4b) and Experiment 3 ($t(4) = -1.23$, $p = 0.142$; data from Fig. 6b). This indicates that these cue effects, while small, cannot be rejected as non-meaningful. Thirdly, the stimuli used here were limited to 250-ms tones presented in continuous background noise. Attentional filtering could be influenced not only by signal frequency, but also its duration. For example, when brief tone bursts (5 ms) were presented in continuous noise, the measured attentional filters were much wider (Wright & Dai, 1994a). Also, the sustained nature of the background noise

could be filtered separately from gated signals with transient onsets and offsets (Wright & Dai, 1994b). In future, it might be useful to compare the present results with effects obtained using brief signals and tones and noise that are gated simultaneously.

In conclusion, the current study shows that the neural mechanism underlying auditory attentional filter generation is mostly independent of the presence of external noise, and this does not support the involvement of the noise-dependent MOC brainstem reflex. Instead, it is proposed that the changes in hearing sensitivity during focus attention are established by the descending cortico-olivocochlear pathway, possibly via the modulatory effects of the LOC fibers on the spontaneous activity of auditory neurons. The proposed LOC mechanism is, however, speculative at this stage, and will have to await further experimental evidence.

Acknowledgements The authors thank Prof. Donald Robertson and Prof. Geoff Hammond (The University of Western Australia) for useful discussions on the earlier version of the paper. Prof. Geoff Hammond also provided the Labview codes for stimulus generation. The study was funded by the University of Malaya Postgraduate Research Grant (PG143-2015A).

Open Practice Statement The data of the experiments will be made available after acceptance, and none of the experiments was preregistered.

References

- Bester, C., Jayakody, D.M., Hammond, G. & Robertson, D. (2016). Selective attentional processes in cochlear implant recipients: Measurements of the attentional filter. *Journal of the Acoustical Society of America*, 140(6):4091.
- Botte, M. C. (1995). Auditory attentional bandwidth: effect of level and frequency range. *The Journal of the Acoustical Society of America*, 98(5 Pt 1), 2475-2485.
- Broadbent, D. E. (1958). *Perception & Communication*. London: Pergamon Press.
- Buus, S., Schorer, E., Florentine, M., & Zwicker, E. (1986). Decision rules in detection of simple and complex tones. *The Journal of the Acoustical Society of America*, 80(6), 1646-1657.
- Creelman, C. D. (1960). Detection of signals of uncertain frequency. *The Journal of the Acoustical Society of America*, 32(7), 805-810.
- Da Costa, S., van der Zwaag, W., Miller, L. M., Clarke, S., & Saenz, M. (2013). Tuning in to sound: frequency-selective attentional filter in human primary auditory cortex. *The Journal of Neuroscience*, 33(5), 1858-1863.
- Dai, H. P., Scharf, B., & Buus, S. (1991). Effective attenuation of signals in noise under focused attention. *The Journal of the Acoustical Society of America*, 89(6), 2837-2842.
- Dragicevic, C. D., Aedo, C., Leon, A., Bowen, M., Jara, N., Terreros, G., ... Delano, P. H. (2015). The Olivocochlear Reflex Strength and Cochlear Sensitivity are Independently Modulated by Auditory Cortex Microstimulation. *Journal of the Association for Research in Otolaryngology*, 16(2), 223-240.
- Eggermont, J. J. (2015). Animal models of spontaneous activity in the healthy and impaired auditory system. *Frontiers in neural circuits*, 9, 19.
- Faul, F., Erdfelder, E., Lang, A., & Buchner, A. (2007). A flexible statistical power analysis program for the social, behavioral and biomedical sciences. *Behavior Research Methods*, 39, 175-191.

- Fritz, J. B., David, S. V., Radtke-Schuller, S., Yin, P., & Shamma, S. A. (2010). Adaptive, behaviorally gated, persistent encoding of task-relevant auditory information in ferret frontal cortex. *Nature Neuroscience*, *13*(8), 1011.
- Fritz, J. B., Elhilali, M., David, S. V., & Shamma, S. A. (2007). Does attention play a role in dynamic receptive field adaptation to changing acoustic salience in A1? *Hearing Research*, *229*(1-2), 186-203.
- Gelfand, S.A. (1984). The contralateral acoustic-reflex threshold. In: S. Silman (ed.), *The Acoustic Reflex* (pp. 137-186). New York : Academic Press.
- Giard, M. H., Fort, A., Mouchetant-Rostaing, Y., & Pernier, J. (2000). Neurophysiological mechanisms of auditory selective attention in humans. *Frontiers in Bioscience*, *5*, d84-94.
- Gifford, M. L., & Guinan, J. J., Jr. (1987). Effects of electrical stimulation of medial olivocochlear neurons on ipsilateral and contralateral cochlear responses. *Hearing Research*, *29*(2-3), 179-194.
- Glasberg, B. R., & Moore, B. C. (1990). Derivation of auditory filter shapes from notched-noise data. *Hearing Research*, *47*(1-2), 103-138.
- Grady, C. L., Van Meter, J. W., Maisog, J. M., Pietrini, P., Krasuski, J., & Rauschecker, J. P. (1997). Attention-related modulation of activity in primary and secondary auditory cortex. *Neuroreport*, *8*(11), 2511-2516.
- Green, D. M. (1961). Detection of auditory sinusoids of uncertain frequency. *The Journal of the Acoustical Society of America*, *33*(7), 897-903.
- Green, D. M. (1964). Consistency of auditory detection judgments. *Psychological Review*, *71*(5), 392.
- Green, D. M., McKey, M. J., & Licklider, J. (1959). Detection of a pulsed sinusoid in noise as a function of frequency. *The Journal of the Acoustical Society of America*, *31*(11), 1446-1452.
- Green, T. J., & McKeown, J. D. (2001). Capture of attention in selective frequency listening. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(5), 1197.
- Greenberg, G. Z., & Larkin, W. D. (1968). Frequency-response characteristic of auditory observers detecting signals of a single frequency in noise: the probe-signal method. *The Journal of the Acoustical Society of America*, *44*(6), 1513-1523.
- Jääskeläinen, I. P., & Ahveninen, J. (2014). Auditory-cortex short-term plasticity induced by selective attention. *Neural plasticity*, *2014*, 216731.
- Jääskeläinen, I. P., Ahveninen, J., Andermann, M. L., Belliveau, J. W., Raji, T., & Sams, M. (2011). Short-term plasticity as a neural mechanism supporting memory and attentional functions. *Brain Research*, *1422*, 66-81.
- Jäger, K., & Kössl, M. (2016). Corticofugal modulation of DPOAEs in gerbils. *Hearing research*, *332*, 61-72.
- Johnson, D. M., & Hafter, E. R. (1980). Uncertain-frequency detection: Cuing and condition of observation. *Perception & Psychophysics*, *28*(2), 143-149.
- Kawase, T., Delgutte, B., & Liberman, M. C. (1993). Antimasking effects of the olivocochlear reflex. II. Enhancement of auditory-nerve response to masked tones. *Journal of Neurophysiology*, *70*(6), 2533-2549.
- Kawase, T., & Takasaka, T. (1995). The effect of contralateral noise on masked compound action potential in humans. *Hearing Research*, *91*(1), 1-6.
- Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence testing for psychological research: A tutorial. *Advances in Methods and Practices in Psychological Science*, *1*(2), 259-269.
- Le Prell, C. G., Dolan, D. F., Hughes, L. F., Altschuler, R. A., Shore, S. E., & Bledsoe Jr, S. C. (2014). Disruption of lateral olivocochlear neurons with a dopaminergic neurotoxin depresses spontaneous auditory nerve activity. *Neuroscience letters*, *582*, 54-58.
- León, A., Elgueta, D., Silva, M. A., Hamamé, C. M., & Delano, P. H. (2012). Auditory cortex basal activity modulates cochlear responses in chinchillas. *PLoS ONE*, *7*(4), e36203.
- Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America*, *49*(2B), 467-477.
- Liberman, M. C. (1988). Response properties of cochlear efferent neurons: monaural vs. binaural stimulation and the effects of noise. *Journal of Neurophysiology*, *60*(5), 1779-1798.
- Liberman, M. C. (1990). Effects of chronic cochlear de-efferentation on auditory-nerve response. *Hearing Research*, *49*(1-3), 209-223.
- Liberman, M. C., & Brown, M. C. (1986). Physiology and anatomy of single olivocochlear neurons in the cat. *Hearing Research*, *24*(1), 17-36.
- Liberman, M. C., & Guinan, J. J. (1998). Feedback control of the auditory periphery: antimasking effects of middle ear muscles vs. olivocochlear efferents. *Journal of Communication Disorders*, *31*(6), 471-82.
- Long, G. R. (1984). The microstructure of quiet and masked thresholds. *Hearing Research*, *15*(1), 73-87.
- Lukas, J. H. (1980). Human auditory attention: the olivocochlear bundle may function as a peripheral filter. *Psychophysiology*, *17*(5), 444-452.
- Maison, S., Micheyl, C., & Collet, L. (2001). Influence of focused auditory attention on cochlear activity in humans. *Psychophysiology*, *38*(1), 35-40.
- Mulders, W., & Robertson, D. (2000). Evidence for direct cortical innervation of medial olivocochlear neurones in rats. *Hearing Research*, *144*(1-2), 65-72.
- Norman, D. A. (1968). Toward a theory of memory and attention. *Psychological Review*, *75*(6), 522.
- Puel, J.-L., Bonfils, P., & Pujol, R. (1988). Selective attention modifies the active micromechanical properties of the cochlea. *Brain Research*, *447*(2), 380-383.
- Rasmussen, G. L. (1946). The olivary peduncle and other fiber projections of the superior olivary complex. *Journal of Comparative Neurology*, *84*, 141-219.
- Riecke, L., Peters, J. C., Valente, G., Poser, B. A., Kemper, V. G., Formisano, E., & Sorger, B. (2018). Frequency-specific attentional modulation in human primary auditory cortex and midbrain. *Neuroimage*, *174*, 274-287.
- Robertson, D. (2009). Centrifugal control in mammalian hearing. *Clinical and Experimental Pharmacology and Physiology*, *36*(7), 603-611.
- Robertson, D., Anderson, C.-J., & Cole, K. S. (1987). Segregation of efferent projections to different turns of the guinea pig cochlea. *Hearing Research*, *25*(1), 69-76.
- Scharf, B. (1970). Critical Bands. In J. V. E. Tobias (Ed.), *Foundations of Modern Auditory Theory* (Vol.1, pp. 159-202). New York: Academic Press.
- Scharf, B., Magnan, J., & Chays, A. (1997). On the role of the olivocochlear bundle in hearing: 16 case studies. *Hearing Research*, *103*(1-2), 101-122.
- Scharf, B., Magnan, J., Collet, L., Ulmer, E., & Chays, A. (1994). On the role of the olivocochlear bundle in hearing: a case study. *Hearing Research*, *75*(1-2), 11-26.
- Scharf, B., Quigley, S., Aoki, C., Peachey, N., & Reeves, A. (1987). Focused auditory attention and frequency selectivity. *Perception & Psychophysics*, *42*(3), 215-223.
- Scharf, B., Reeves, A., & Suci, J. (2007). The time required to focus on a cued signal frequency. *The Journal of the Acoustical Society of America*, *121*(4), 2149-2157.
- Slee, S. J., & David, S. V. (2015). Rapid task-related plasticity of spectrotemporal receptive fields in the auditory midbrain. *Journal of Neuroscience*, *35*(38), 13090-13102.
- Suga, N., Xiao, Z., Ma, X., & Ji, W. (2002). Plasticity and corticofugal modulation for hearing in adult animals. *Neuron*, *36*(1), 9-18.

- Tan, M. N. (2008). Selective listening processes in humans (PhD Thesis), University of Western Australia, Perth, WA.
- Tan, M. N., Robertson, D., & Hammond, G. R. (2008). Separate contributions of enhanced and suppressed sensitivity to the auditory attentional filter. *Hearing Research*, 241(1-2), 18-25.
- Terreros, G., & Delano, P. H. (2015). Corticofugal modulation of peripheral auditory responses. *Frontiers in systems neuroscience*, 9, 134.
- Warren, E. H., & Liberman, M. C. (1989). Effects of contralateral sound on auditory-nerve responses. I. Contributions of cochlear efferents. *Hearing Research*, 37(2), 89-104.
- Watson, C. S., Franks, J. R., & Hood, D. C. (1972). Detection of tones in the absence of external masking noise. I. Effects of signal intensity and signal frequency. *The Journal of the Acoustical Society of America*, 52(2B), 633-643.
- Winslow, R. L., & Sachs, M. B. (1988). Single-tone intensity discrimination based on auditory-nerve rate responses in backgrounds of quiet, noise, and with stimulation of the crossed olivocochlear bundle. *Hearing Research*, 35(2), 165-189.
- Wright, B. A., & Dai, H. (1994a). Detection of unexpected tones with short and long durations. *The Journal of the Acoustical Society of America*, 95, 931-938.
- Wright, B. A., & Dai, H. (1994b). Detection of unexpected tones in gated and continuous maskers. *The Journal of the Acoustical Society of America*, 95(2), 939-948.
- Wright, H. N. (1973). Psychometric Functions in Quiet for Long-and Short-Duration Tones at 1000 Hz. *The Journal of the Acoustical Society of America*, 54(1), 337-337.
- Yakunina, N., Tae, W. S., Kim, S. S., & Nam, E. C. (2019). Functional MRI evidence of the cortico-olivary efferent pathway during active auditory target processing in humans. *Hearing Research*, 379, 1-11.
- Zhao, W., & Dhar, S. (2012). Frequency tuning of the contralateral medial olivocochlear reflex in humans. *Journal of Neurophysiology*, 108(1), 25-30.
- Zheng, X., Henderson, D., McFadden, S. L., Ding, D. L., & Salvi, R. J. (1999). Auditory nerve fiber responses following chronic cochlear de-efferentation. *Journal of Comparative Neurology*, 406(1), 72-86.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.