

# Musical training enhances neural processing of comodulation masking release in the auditory brainstem

Soheila Rostami,<sup>1</sup> Abdollah Moossavi<sup>2</sup>

<sup>1</sup>Department of Audiology, Faculty of Rehabilitation Sciences; and <sup>2</sup>Department of Otolaryngology, Head and Neck Surgery, Faculty of Medicine, Iran University of Medical Sciences, Tehran, Iran

## Abstract

Musical training strengthens segregation the target signal from background noise. Musicians have enhanced stream segregation, which can be considered a process similar to comodulation masking release. In the current study, we surveyed psychoacoustical comodulation masking release in musicians and non-musicians. We then recorded the brainstem responses to complex stimuli in comodulated and unmodulated maskers to investigate the effect of musical training on the neural representation of comodulation masking release for the first time. The musicians showed significantly greater amplitudes and earlier brainstem response timing for stimulus in the presence of comodulated maskers than non-musicians. In agreement with the results of psychoacoustical experiment, musicians showed greater comodulation masking release than non-musicians. These results reveal a physiological explanation for behavioral enhancement of comodulation masking release and stream segregation in musicians.

## Introduction

Musical performance, like speech, is one of the most complicated tasks performed by people. Skilled musicians have spent

thousands of hours rehearsing by the time they are 21 years of age,<sup>1</sup> which is assumed to be the main reason for their increased auditory perceptual abilities as well as the structural and functional modifications noticed at the cortical and subcortical levels for music and speech.<sup>2-4</sup>

Weak signals embedded in modulated noise can be detected more efficiently than those embedded in unmodulated noise. It is known that an increase in noise bandwidth, a psychoacoustic phenomenon called comodulation masking release (CMR), facilitates signal detection in modulated noise.<sup>5</sup> The detection threshold of a sinusoidal signal in an on-frequency masker can be improved by presenting further off-frequency maskers having the same envelope fluctuations across frequency bands.<sup>6</sup> CMR can manifest itself through within-channel and across-channel mechanisms.<sup>7</sup> The within-channel process is based on changes in the temporal envelope within a filter band, while the across-channel mechanisms are mediated by the temporal correlation between on-frequency and off-frequency filter bands. CMR can be calculated as the difference between the threshold of a sinusoidal signal with an unmodulated masker (UM) and its threshold with a comodulated masker (CM) at the same bandwidth.<sup>5</sup> CMR for a masker bandwidth larger than the critical band signal will be greater than CMR for a bandwidth equal to the critical band. This difference is known as a true or across-channel CMR.<sup>6,8</sup>

The brainstem response to complex stimuli (cABR) can be used to evaluate the integrity of auditory functioning by showing evidence of the activity of the subcortical nuclei.<sup>9,10</sup> The cABR shows the neural encoding of acoustical features of a stimulus with significant reliability.<sup>11</sup> Musicians have displayed enhanced subcortical encoding of stimulus features by giving faster responses than non-musicians.<sup>12</sup> Moreover, the auditory brainstem responses are influenced by linguistic experience and are flexible.<sup>13</sup> CMR is a complex task requiring auditory stream segregation. Formation of an auditory stream requires the fundamental ability to appropriately represent, group and store auditory units. It is considered to be an essential aspect of CMR. Researches have shown that listeners with greater attention and working memory will be least affected by masker and perform better on speech perception in noise skills.<sup>14,15</sup> Because musicians can separate out the sounds of instruments in an orchestral performance (stream segregation), it has been hypothesized that a musician's life experience of musical stream segregation may results in improved comodulation masking release. The current study surveyed psychoacoustical comodulation masking release in musicians and non-musicians, then recorded cABR in quiet and in comodulated and unmodulated maskers to investigate the effect of musical training on the neural representation of comodulation masking release. It was hypothesized that musicians have less-degraded brainstem responses in the presence of a comodulated masker than non-musicians.

Correspondence: Abdollah Moossavi, Department of Otolaryngology, Head and Neck Surgery, Faculty of Medicine, Iran University of Medical Sciences, Hemmat Highway, Tehran, Iran.  
Tel.: +98.21.86701 - Fax: +98.21.88602217.  
E-mail: moossavi.a@iums.ac.ir

Key words: Musical training; auditory brainstem response; comodulation masking release.

Contributions: SR, designed the study, performed experiments, analyzed data, and wrote the manuscript; AM, designed the study, performed data analysis and edited the manuscript.

Conflict of interest: the authors have no conflict of interest to declare.

Received for publication: 18 May 2017.

Accepted for publication: 19 July 2017.

This work is licensed under a Creative Commons Attribution NonCommercial 4.0 License (CC BY-NC 4.0).

©Copyright S. Rostami and A. Moossavi, 2017  
Licensee PAGEPress, Italy  
Audiology Research 2017;7:185  
doi:10.4081/audiores.2017.185

## Materials and Methods

### Subjects

The participants were 36 right-handed normal-hearing adults 18 to 30 years of age with no history of audiological, otological or neurological disorders. Their audiometric thresholds were 15 dB HL or better at octave intervals of 250 to 8000 Hz. The subject classified as musicians (N = 19) had 10 years or more of experience and started training before the age of seven. They had practiced at least four times weekly over the previous three years before registering for the study. Non-musicians (N = 17) were classified by the inability to meet the musician criteria.

### Stimulus

Separate psychoacoustical and electrophysiological experiments were performed. All stimuli were processed in MATLAB R2014a. In the psychoacoustical experiment, the signal was a pure tone at 700 Hz, 300 ms in duration. The masker consisted of seven noise bands, one centered at a frequency of 700 Hz (on-frequency band) and the flanking bands were tones at 300, 400, 500, 900, 1000 and 1100 Hz, each with a bandwidth of 24 Hz and a duration of 1000 ms. Both on-frequency and flanking bands were 100% amplitude-modulated (AM) tones. In the electrophysiological experiment, the signal was the 40 ms speech syllable /da/ containing an initial stop consonant burst followed by a consonant-to-vowel transition synthesized using a Klatt synthesizer.<sup>16</sup> For the masked conditions, a speech-shaped noise was presented in two modes with and without comodulation (CM, UM). In the comodulated condition, the maskers were modulated with a 100% amplitude modulation depth. The stimuli were converted from digital to analogue (44,100 Hz sampling rate; 16 bit) and equalized for RMS power. They were amplified using a programmable attenuator (TDT PA5) and a headphone buffer (TDT HB7). In both experiments, the signal and masking noise were presented to the right ear while the left ear remained silent.

### Psychoacoustical experiment

The psychoacoustical experiment took place in a double-walled soundproof booth. The signal and masker were presented to the right ear of listeners through a TDH-39 headphone. The masker was presented at an intensity of 60 dB SPL for a duration of 1000 ms. The signal was presented in the last third of the masker. All tests were performed using a three-alternative forced choice procedure with adaptive signal-level adjustment. Each trial contained three intervals separated by gaps of 500 ms. The signal was presented to the subject in a randomly-chosen interval and the task was to indicate which interval contained the signal. The signal initiated at an intensity of 70 dB SPL. The signal intensity was adjusted according to the two-down, one-up procedure to estimate 70.7% of the psychometric point.<sup>17</sup> The first step size was 4 dB, which was halved at every second reversal until the step size reached 1 dB. The process then continued for six reversals. In the next stage, the mean intensity level of the final six reversals was calculated and considered to be the estimated threshold. The final threshold was calculated using the mean of the three estimated thresholds.

The masked signal thresholds were obtained under three masking conditions. In the first condition, the masker consisted of two independent narrow bands of noise (on-frequency and flanking bands) without modulation and was denoted as the UM condition. In the CM condition, both the on-frequency band and flanking bands had the same envelope of square-wave modulation. In the reference (RF) condition, the threshold was measured for the mod-

ulated on-frequency masker alone. CMR was calculated by subtraction of CM from UM. The true or across-frequency CMR (AF CMR) was quantified by subtracting the threshold for the RF condition from the threshold obtained for the CM condition.

### Electrophysiological experiment

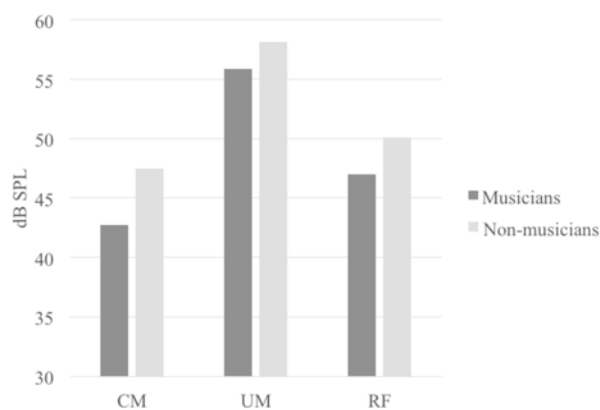
The cABR was recorded using a Biologic Navigator Pro system using BioMARK software (Natus Medical; USA) and all experiments were carried out in a double-walled, electrically and acoustically sealed sound booth. The subjects were seated in a comfortable chair. During the experiment, listeners could watch their favorite subtitled movies while keeping quiet and motionless.

Electrophysiological responses were recorded using a vertical array of three Ag-AgCl electrodes (vertex Cz active, high forehead ground, ipsilateral earlobe reference). Throughout data recording, electrode impedance was less than 5 k $\Omega$  and the inter-electrode impedance difference was below 3 k $\Omega$ . An online bandpass filter was employed at 100 to 2000 Hz. Online artifact rejection was used by a criterion of  $\pm 23$   $\mu$ v. The time window was 75 ms, including a 15 ms pre-stimulus period. Two blocks of 3000 artifact-free sweeps were collected and averaged. The recording sessions lasted for about 2 h. The electrophysiological responses to the syllable /da/ were recorded in quiet and masking conditions (CM, UM). It was presented to the right ear of listeners through an insert earphone (ER-3; Etymotic Research) at 80 dB SPL at a rate of 10.9 Hz with alternating polarity. The stimulus was introduced at signal-to-noise ratios of +10 dB. The signal-to-noise ratio was selected based on pilot tests. Both the stimulus and maskers were presented through an earphone inserted into the right ear while the left ear was kept in silence. The experiment and data collection were completed without interruption in one session. The latency and amplitude of the onset peak: V; onset trough: A; transient peak: C; FFR peaks: D, E, and F; offset peak: O of the cABR were analyzed off-line.

## Results

### Psychoacoustical experiment

Statistical analysis was conducted in SPSS v. 18.0 software. The Kolmogorov-Smirnov test indicated that all data followed a normal distribution ( $P < 0.05$ ). No statistically significant differences were found between groups for age and pure-tone audiometry. Figure 1 shows the average psychoacoustic thresholds of the signal for the three conditions (CM, UM and RF) in both groups.



**Figure 1.** Comparison of average psychoacoustic thresholds of signal for the CM, UM and RF conditions between groups.

The musicians demonstrated lower thresholds than non-musicians for all conditions. An independent-samples t-test was conducted to compare the thresholds of the signals in musicians and non-musicians for the three conditions. There are significant differences in the masked thresholds of signals in musicians (CM condition: M = 42.72, SD = 1.49; UM condition: M = 55.89, SD = 1.16; RF condition: M = 46.96, SD = 2.02) and non-musicians (CM condition: M = 47.50, SD = 2.32; UM condition: M = 58.14, SD = 2.32; RF condition: M = 50.10, SD = 2.24); CM condition:  $t(34) = 7.41$  ( $P < 0.05$ ); UM condition:  $t(34) = 3.72$  ( $P < 0.05$ ); RF condition:  $t(34) = 4.40$  ( $P < 0.05$ ). These results demonstrate the positive effects of musical training on detection of signal in noise.

The average CMR and AF-CMR of both groups are shown in Figure 2. The musicians demonstrated higher CMRs than non-musicians (musicians: M = 13.16, SD = 1.85; non-musicians: M = 10.49, SD = 3.25);  $t(34) = 3.06$  ( $P < 0.05$ ). The average AF-CMR was significantly greater for the musicians group (M = 4.23, SD = 2.25) than the non-musicians group (M = 2.50, SD = 1.03);  $t(34) = 2.89$  ( $P < 0.05$ ). The musicians demonstrated greater comodulated release from masking than the non-musicians.

### Electrophysiological experiment

Figure 3 is a time-domain representation of grand average brainstem responses to the syllable /da/ from -15 to 60 ms in quiet and in two comodulated and un-modulated masking conditions (CM, UM) for a signal-to-noise ratio of +10 dB. The effect of the masking conditions on the brainstem responses was examined using a one-way analysis of variance (ANOVA). For all peaks of both groups, the ANOVA revealed a significant reduction in amplitude and a significant increase in latency for two masking conditions ( $P < 0.05$ ), but these changes were significantly greater for the UM than CM condition. The peak latencies and amplitudes of responses for the musician and non-musician groups were compared for response peaks V, A, C, D, E, and O. For all peaks, the independent-samples t-tests revealed that the two groups had nearly the same average peak latencies and amplitudes in quiet and UM masking condition ( $P > 0.05$ ). The musicians showed significantly greater comodulated release from masking and earlier response timing than non-musicians for the syllable /da/ in the presence of comodulated masker (Table 1). Musicians showed greater comodulated release from masking than non-musicians in agreement with results of the psychoacoustical experiment.

### Discussion

In accordance with the hypothesis, the performance of the musicians was better on CMR and the brainstem correlates of CMR. In previous studies, better speech perception in the presence

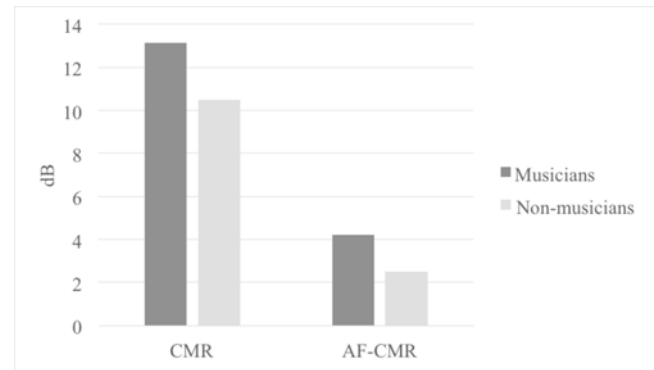


Figure 2. Comparison of average CMR and AF-CMR between groups.

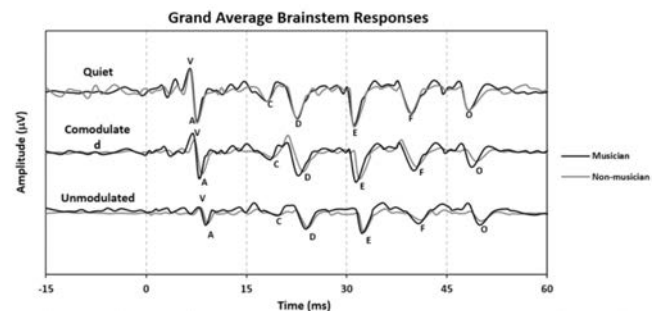


Figure 3. Comparison of grand average auditory brainstem responses in quiet and comodulated and unmodulated maskers for musicians (denoted in black) and non-musicians (denoted in gray).

Table 1. Group comparisons of brainstem measures.

Latency (ms)	Quiet						Comodulated masker						Un-modulated masker						
	Musicians		Non-musicians		t(34)	P	Musicians		Non-musicians		t(34)	P	Musicians		Non-musicians		t(34)	P	
Mean	SD	Mean	SD	Mean			SD	Mean	SD	Mean			SD	Mean	SD	Mean			SD
V	6.633	0.226	6.607	0.226	0.345	0.733	7.07	0.223	7.26	0.303	2.153	0.039*	7.96	0.558	8.192	0.691	1.114	0.273	
A	7.503	0.269	7.515	0.244	0.135	0.894	7.89	0.245	8.27	0.592	2.467	0.022*	8.893	0.701	9.144	0.508	1.213	0.234	
C	18.19	0.446	18.164	0.772	0.121	0.905	18.613	0.48	18.99	0.54	2.217	0.033*	19.735	0.64	19.978	0.759	1.039	0.306	
D	22.495	0.222	22.49	0.293	0.054	0.957	22.836	0.187	23.346	0.335	5.55	0.001*	23.776	0.679	24.107	0.461	1.687	0.101	
E	31.064	0.184	31.096	0.048	0.723	0.478	31.348	0.287	31.87	0.44	4.147	0.001*	32.277	0.501	32.558	0.529	1.63	0.112	
F	39.569	0.202	39.544	0.247	0.33	0.743	39.967	0.255	40.36	0.588	2.548	0.019*	40.789	0.62	41.1	0.46	1.687	0.101	
O	48.237	0.225	48.248	0.261	0.134	0.894	48.725	0.411	49.414	0.473	4.665	0.001*	49.835	0.777	50.108	0.624	1.15	0.258	
Amplitude (µV)																			
V	0.151	0.022	0.141	0.029	1.138	0.263	0.124	0.016	0.08	0.033	4.886	0.001*	0.039	0.018	0.037	0.015	0.322	0.749	
A	-0.191	0.016	-0.202	0.025	1.666	0.105	-0.17	0.029	-0.139	0.013	3.02	0.005*	-0.077	0.023	-0.065	0.024	1.509	0.141	
C	-0.06	0.017	-0.062	0.026	0.387	0.702	-0.046	0.017	-0.032	0.013	2.613	0.013*	-0.014	0.005	-0.01	0.008	1.803	0.08	
D	-0.175	0.032	-0.173	0.036	0.152	0.88	-0.149	0.013	-0.125	0.018	4.576	0.001*	-0.101	0.038	-0.092	0.044	0.629	0.533	
E	-0.218	0.029	-0.231	0.042	1.048	0.302	-0.194	0.015	-0.174	0.035	2.211	0.038*	-0.13	0.028	-0.123	0.032	0.684	0.498	
F	-0.14	0.029	-0.142	0.035	0.222	0.826	-0.118	0.024	-0.092	0.034	2.51	0.018*	-0.067	0.035	-0.043	0.022	1.021	0.315	
O	-0.121	.031	-0.122	0.029	0.132	0.896	-0.096	0.024	-0.058	0.041	3.331	0.003*	-0.075	0.034	-0.05	0.028	0.067	0.947	



of noise was observed in musicians for prosody,<sup>18</sup> melody,<sup>19</sup> pitch,<sup>20</sup> temporal component and speech discrimination.<sup>21,22</sup> These results suggest that musical training enhanced the neural processing of speech. Moreover, musicians showed enhanced attention and working memory.<sup>23,24</sup> The present data show that a musician's life experience of musical stream segregation results in improved comodulation-masking release. Because perceptual cues are significant for segregate the target signal from background noise, those who listen with enhanced auditory perceptual skills can identify fine acoustical signals and show improved ability to auditory stream segregation.

The Hebbian principal states that the relations between neurons are simultaneously active and do not weaken over time.<sup>25</sup> This can be one explanation for the increase in the abilities of musicians. It is possible that extended music practice will improve neural connections. The results indicate that the Hebbian principle can apply for learning at lower stages (bottom-up) and also be required at higher stages (top-down). The nervous system extracts the relevant signal and suppresses irrelevant noise in bottom-up and top-down processing.<sup>26</sup> Bottom-up and top-down interactive processing results in subcortical plasticity with musical training.

The outcomes of the present study indicate that musicians have the potential to benefit from brief temporal minima in modulated background noise to catch signal cues (also known as *listening-in-the-dips*). The masking release decreases for maskers with flat temporal and spectrally steady-state noise and is likely to mask the weaker portions of the signal.<sup>27</sup> Spectral and temporal resolutions are required for listening-in-the-dips. The results indicate that lifelong experience with stream segregation improves neural signal encoding and enhances representation of the speech signal in comodulated noise. It can be concluded that musical experience is an advantage for the dip-listening mechanism.

The results indicate that the average true or AF-CMR is significantly greater for the musicians group than the non-musicians group. Across-frequency processing can also explain enhanced CMR in the musician group. CMR is a complex task requiring across frequency processing.<sup>7</sup> Across-frequency modulation can collaborate with other activities and result in auditory grouping, stream segregation and auditory object formation. The present data show that a grouping mechanism and masking release are associated with one another. These results indicate that musicians performed better on across-frequency modulation processing, auditory grouping and stream segregation.

## Conclusions

The data from the current study indicates that musical experience is an advantage for comodulation masking release. Musicians had enhanced subcortical representations of the syllable /da/ in comodulated maskers. Thus, musicians demonstrate improved neural synchrony and less-degraded brainstem responses for comodulated masker than non-musicians. In agreement with the results of the psychoacoustical experiment, musicians showed greater comodulated release from masking than non-musicians. It can be concluded that musical experience is an advantage for the dip-listening mechanism and across-frequency processing. These results suggest a physiological explanation for psychoacoustical enhancement in musicians for comodulation masking release.

## References

- Ericsson KA, Krampe RT, Tesch-romer C, et al. The role of deliberate practice in the acquisition of expert performance. *BMC Med Educ* 1993;100:363-406.
- Micheyl C, Delhommeau K, Perrot X, Oxenham AJ. Influence of musical and psychoacoustical training on pitch discrimination. *Proc Meet Acoust* 2006;219:36-47.
- Strait D, Skoe E, Kraus N, Ashley R. Musical experience influences subcortical processing of emotionally-salient vocal sounds. *Eur J Neurosci* 2009;29:661-8.
- Peretz I, Zatorre RJ. Brain organization for music processing. *Annu Rev Psychol* 2005;56:89-114.
- Hall JW, Haggard MP, Fernandes MA. Detection in noise by spectro-temporal pattern analysis. *J Acoust Soc Am* 1984;76:50-56.
- Verhey JL, Pressnitzer D, Winter IM. The psychophysics and physiology of comodulation masking release. *Exp Brain Res* 2003;153:405-17.
- Schooneveldt GP, Moore BCJ. Comodulation masking release (CMR) as a function of masker bandwidth, modulator bandwidth, and signal duration. *J Acoust Soc Am* 1989;85:273-81.
- Carlyon RP, Buus S, Florentine M. Comodulation masking release for three types of modulator as a function of modulation rate. *J Acoust Soc Am* 1989;85:S36.
- Jewett D, Romano M, Williston J. Human auditory evoked potentials: possible brain stem components detected on the scalp. *Science* 1970 ;167:1517-8.
- Chandrasekaran B, Kraus N. The scalp-recorded brainstem response to speech: Neural origins and plasticity. *Psychophysiology* 2010;47:236-46.
- Song J, Nicol T, Kraus N. Test-retest reliability of the speech-evoked auditory brainstem response. *Clin Neurophysiol* 2011;122:346-55.
- Kraus N, Skoe E, Parbery-Clark A, Ashley R. Experience-induced malleability in neural encoding of pitch, timbre, and timing. *Ann N Y Acad Sci* 2009;1169:543-57.
- Krishnan A, Xu Y, Gandour J, Cariani P. Encoding of pitch in the human brainstem is sensitive to language experience. *Cogn Brain Res* 2005 ;25:161-8.
- Heinrich A, Schneider BA, Craik FIM. Investigating the influence of continuous babble on auditory short-term memory performance. *Q J Exp Psychol* 2008;61:735-51.
- Strait DL, Kraus N. Can you hear me now? musical training shapes functional brain networks for selective auditory attention and hearing speech in noise. *Front Psychol* 2011;2.
- Klatt DH. Software for a cascade/parallel formant synthesizer. *J Acoust Soc Am* 1980;67:971-95.
- Levitt H. Transformed up-down methods in psychoacoustics. *J Acoust Soc Am* 1971;49:467-77.
- Magne C, Schön D, Besson M. Musician children detect pitch violations in both music and language better than nonmusician children: behavioral and electrophysiological approaches. *J Cogn Neurosci* 2006;18:199-211.
- Fujioka T, Trainor LJ, Ross B, et al. Automatic encoding of polyphonic melodies in musicians and nonmusicians. *J Cogn Neurosci* 2005;17:1578-92.
- Bidelman GM, Gandour JT, Krishnan A. Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. *J Cogn Neurosci* 2011;23:425-34.
- Kühnis J, Elmer S, Meyer M, Jäncke L. The encoding of vowels and temporal speech cues in the auditory cortex of professional musicians: an EEG study. *Neuropsychologia*

- 2013;51:1608-18.
22. Strait D, O'connell S, Parbery-Clark A. Musicians' enhanced neural differentiation of speech sounds arises early in life: developmental evidence from ages 3 to 30. *Cereb Cortex* 2014;24:2512-21.
  23. Baumann S, Meyer M, Jäncke L. Enhancement of auditory-evoked potentials in musicians reflects an influence of expertise but not selective attention. *J Cogn Neurosci* 2008;20:2238-49.
  24. George E, Coch D. Music training and working memory: an ERP study. *Neuropsychologia* 2011;49:1083-94
  25. Hebb DO. *The organization of behavior: a neuropsychological theory*. Abingdon: Psychology Press Ltd.; 2005.
  26. Luo F, Wang Q, Kashani A, Yan J. Corticofugal modulation of initial sound processing in the brain. *J Neurosci* 2008;28:11615-21.