Research Article

Contribution of Spectral and Temporal Cues in Timbre Recognition among Normal-Hearing and Hearing-Impaired People

Mohammad Mohammadi, Mohammad Maarefvand*, Akram Pourbakht

Department of Audiology, School of Rehabilitation Sciences, Iran University of Medical Sciences, Tehran, Iran

Use your device to scan and read the article online

Citation: Mohammadi M, Maarefvand M, Pourbakht A. Contribution of Spectral and Temporal Cues in Timbre Recognition among Normal-Hearing and Hearing-Impaired People. Aud Vestib Res. 2024;33(4):322-9.

https://doi.org/10.18502/avr.v33i4.16650

Highlights

- Temporal and spectral cues for timbre recognition are both affected by hearing loss
- Temporal cues may be more affected by hearing loss in timbre recognition
- Sensorineural hearing loss can reduce timbre recognition

Article info:
Received: 11 Jan 2024
Revised: 12 Feb 2024
Accepted: 14 Feb 2024

* **Corresponding Author:**
Department of Audiology, School of Rehabilitation Sciences, Iran University of Medical Sciences, Tehran, Iran.
maarefvandm@gmail.com

ABSTRACT

Background and Aim: Timbre perception is a multidimensional feature of music perception and hearing loss may deteriorate timbre and instrument recognition. Deficits in timbre recognition may originate from poor temporal or spectral coding of musical notes subsequent to hearing loss However, it is not well understood which of these features are more dominant in instrument recognition with normal hearing and hearing loss. This study aimed to evaluate the relative importance of temporal and spectral cues instrument recognition in people with and without nearing loss.

Methods: Two groups of adults with normal-hearing and hearing loss were tested with the timbre subtest of the Clinical Assessment of Music Perception (CAMP). Originally, a series of notes is played with eight musical instruments and participants were asked to identify the played instrument among them. In two other conditions, either the rise time or spectral cues of the same notes were masked using a special masking technique. Among eight instruments.

Results: Instrument recognition scores were not significantly different between the original and spectral-manipulated conditions but were lower at the temporal-manipulated condition. The difference between the two groups was significant ($p<0.05$), where the normal-hearing group recognized the musical instruments significantly better than participants with hearing loss group ($p<0.05$).

Conclusion: Temporal cues may have greater importance on timbre recognition regardless of hearing status. This suggests that the interventions and assistive devices which are designed to improve timbre recognition for people with hearing loss should prioritize improvement in temporal coding.

Keywords: Timbre recognition; hearing loss; music; musical instrument; temporal; spectral

Introduction

Music has an important role in human life, contributing to good feeling and overall well-being, which can even be used as a therapy. Like other sounds, musical notes have various physical features such as frequency, intensity, and duration. These physical features are related to the perception of one-dimensional aspects of pitch and loudness. However, musical notes have also a multi-dimensional aspect called timbre or tone color, determined simultaneously by duration, intensity, and frequency. Timbre enables listeners to differentiate between two notes with the same pitch and loudness [1], allowing for the recognition of different musical instruments. Unlike pitch and loudness, timbre lacks a reliable individual scale for measurement and is often described using words such as pleasant, harsh, sharp, and muffled [2], which none of them provides a complete and quantitative description. Previous studies have indicated that temporal (duration), spectral (frequency), and intensity cues of musical notes may play roles in timbre perception [3, 4]. Temporal cues encompass the rise and fall times of notes and temporal changes in amplitude during the steady-state region of notes (i.e. plateau). Changes in temporal cues can result in alterations in the temporal envelope of notes. Although many notes naturally exhibit some degree of amplitude fluctuations in their plateau, these fluctuations are less variable and perceptually distinctive compared to amplitude changes in the rise and fall times. In a study, it was shown that the most effective temporal cue in musical instrument recognition was the rise time [5]. On the other hand, the fall time can be influenced by the reverberation of the note being played and may become unreliable across different environments. In addition to temporal cues, musical instruments exhibit varying spectral cues in different harmonics even with the same fundamental frequency (F0). This implies that, depending on the family of musical instruments, harmonics can have different spectral power distributions, which may contribute to timbre and instrument recognition. The distribution of spectral power across different harmonics and the number of resolved harmonics can be quantified by spectral centroid, which is an indicator of the center of gravity among different harmonics.

Hearing loss may impair the timing or spectral coding of musical instruments, making instrument recognition challenging. Hearing loss can disrupt the precise timing of neural discharge and impair the coding of the rise time which depends on highly synchronized neural responses [6]. Furthermore, hearing loss can widen auditory filters, resulting in decreased spectral resolution and resolvability, as two or more harmonics may fall within a widened auditory filter. Consequently, timbre recognition seems to be more challenging for the auditory brain in adults with Sensorineural Hearing Loss (SNHL) [6].

While the importance of temporal and spectral cues for musical instrument recognition is acknowledged, the relative importance of these cues needs to be investigated, because there may be situations where one or both cues are inaccessible. For example, in noisy places such as concert halls, one of these cues may be more important than the other one, potentially impairing instrument and timbre recognition. Both noise and hearing loss can impair timbre recognition, reduce music perception and appreciation. Therefore, it is pertinent to explore which of these cues is more reliable for individuals with normal hearing when only one of them is available. Additionally, it is important to ascertain whether there are differences in the use of these cues for instrument recognition between individuals with normal hearing and those with SNHL. Thus, this study aimed to investigate the relative importance of temporal and spectral cues for musical instrument recognition in people with normal hearing and those with SNHL.

Methods

Participants

Participants were adults aged 18–65 years with normal hearing (n=29) and with moderate to severe SNHL (n=28). The two groups were matched for age and gender. Since the task in this experiment was musical instrument recognition, only people who could successfully complete the training were included in the study. Participants in the SNHL group had symmetrical descending hearing loss. None of the groups had pure tone average more than 70 dB HL. As another inclusion criterion, the participants should receive scores above 90% in the dichotic digits test (which was used for central auditory processing assessment). The participants had no history of using hearing aids and tumors involving the auditory nerve, and no middle ear problems such as infection, bone adhesion or TM perforation. They were Persian speakers.

Measure

The timbre subtest of the Clinical Assessment of Music Perception (CAMP) test was utilized for instrument recognition in this study. The CAMP is a valid and reliable musical test [7]. In the timbre subtest, a five-note sequence from middle frequency range (C4-A4-F4-G4-C5) was played with eight musical instruments from four instrument families. These notes were presented at the most comfortable level for each participant using a loudspeaker placed at a distance of 1 meter from the person's ear level.

Experimental conditions

There were three experimental conditions. In the first condition, the musical notes of the CAMP played with different instruments were presented to the participants (original condition). In the second condition, the same musical notes were presented to the participants while the rise-time and fall-time cues of the notes were masked and inaccessible (temporal-masked condition). In the third condition, all harmonics in the plateau part were masked, and the rise-time cue was available to the participants (spectrum-manipulated condition). Before the start of the main test, a trial was conducted to ensure that the participants understand and have consistent performance. For all participants, the experiment started with the first condition followed by either second or third condition. The second and third conditions were randomly used.

To test whether binaural hearing can help with timbre recognition, temporal-manipulated notes were presented to one ear and spectral-manipulated notes were presented to the other ear through headphones (TDH-39, Telephonics, USA). The participants were asked to identify the instrument being played. This condition was considered as the integrated condition. The score of the integrated condition was compared with those of the three above mentioned conditions to investigate whether the integration of the two manipulated conditions can change the instrument recognition performance.

Experimental stimuli

The rise and fall times and plateau of each note were calculated using Praat software v.6.4. The frequency range in the spectrogram of this software was 0–5000 Hz, over a 70 dB dynamic range, which included all fundamental frequencies and harmonics. To add masking, the notes were set in one mono-channel of Audacity Software Version 3.3 and noise segments were set in the other mono-channel. The noise segments were aligned in the second channel precisely to mask only the rise and fall times of the notes in the first channel, and the plateau remained unmasked. Then, the two mono-channels were mixed and presented to the participants. The same procedure was followed for the third condition, except that the noises were aligned in the second mono-channel to mask plateau and fall time regions of the notes in the first mono-channel. The intensity level of masking was the same for all notes and instruments, determined to mask temporal and spectral cues with the highest intensity.

Statistical analysis

The percentage of correct instrument recognition was calculated for both groups quantitatively for statistical analysis. A mixed ANOVA was used to measure between-group and within-group differences. Data analysis was done in SPSS v.17 software. The significance level was set at $p<0.05$.

Results

Descriptive statistics for both groups in the three conditions are presented in Table 1. The normal-hearing group had 29 participants (15 females, 14 males) and the SNHL group had 28 participants (14 females, 14 males). The scores represent the correct percentage of instrument recognition. The participants had age 34.4±12.2 years in the normal-hearing group and 39.1±11.2 years in the SNHL group.

As shown in Figure 1, the scores in the original and spectral-manipulated conditions were similar and not significantly different ($p>0.05$). However, the score in the temporal-manipulated condition was significantly lower than in other conditions ($p<0.05$).

The ANOVA results for three conditions showed

Table 1. Mean and standard deviation of timbre recognition scores in different conditions for two groups with normal hearing and sensorineural hearing loss (values are in percentage)

	Mean(SD)	
	Normal	SNHL
Original signal	56.1(17.8)	23.5(14.2)
Spectral manipulated signal	57.9(17.7)	21.1(11.9)
Temporal manipulated signal	42.9(18.2)	14.4(11.8)

SNHL; sensorineural hearing loss

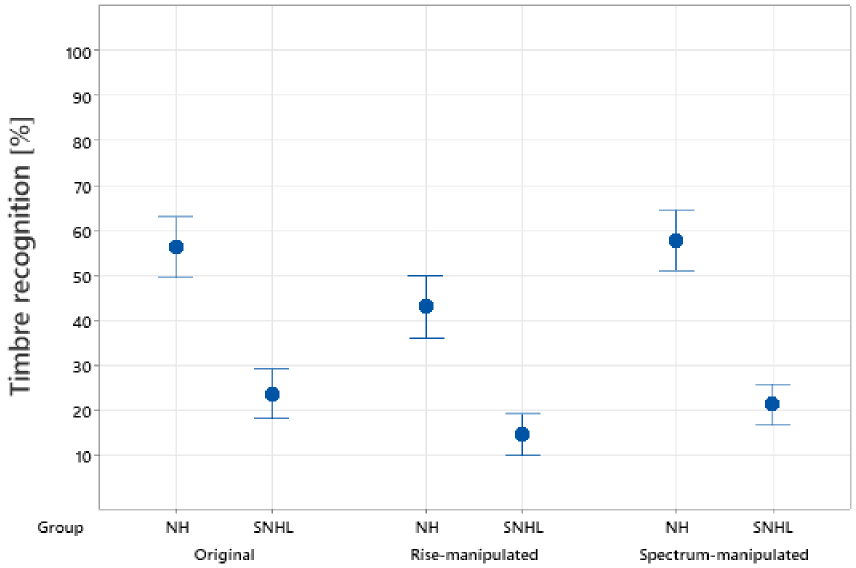

Figure 1. Mean timbre recognition score in different conditions for normal and sensorineural hearing loss participants. NH; normal hearing, SNHL; sensorineural hearing loss

that both effects of condition and group was significant. The interaction effect of group and condition was also significant. A post-hoc test (Bonferroni test) was used to investigate this effect further. Figure 2 illustrates that the scores of the normal-hearing group were significantly higher than those of the SNHL group. The temporal manipulation significantly reduced the percentage of correct instrument recognition, while spectral manipulation had no significant effect on timbre (instrument) recognition. The difference in the recorded scores for different conditions can reveal the relative importance of temporal and spectral cues in both groups.

Twenty normal-hearing and SNHL people completed the integrated condition. The mean score in the normal-hearing group were 55.3±13.3. In the SNHL group, the mean score was 15.2±5.4. Figure 3 demonstrates that the difference between the two groups was significant, where the normal-hearing group recognized the musical instruments significantly better than the SNHL group ($p<0.05$).

The comparison of the integrated condition with the three conditions (original, temporal-manipulated, and spectral-manipulated) showed that the normal-hearing group had significantly higher scores in the integrated condition than in the temporal-manipulated condition. There were no significant differences between the integrated condition and the two original and spectral-manipulated conditions. The SNHL group had significantly higher scores in the original condition compared to the integrated condition. In the spectral-manipulated condition, they had significantly higher scores compared to the integrated condition, but there was no significant difference between the temporal-manipulated and integrated conditions.

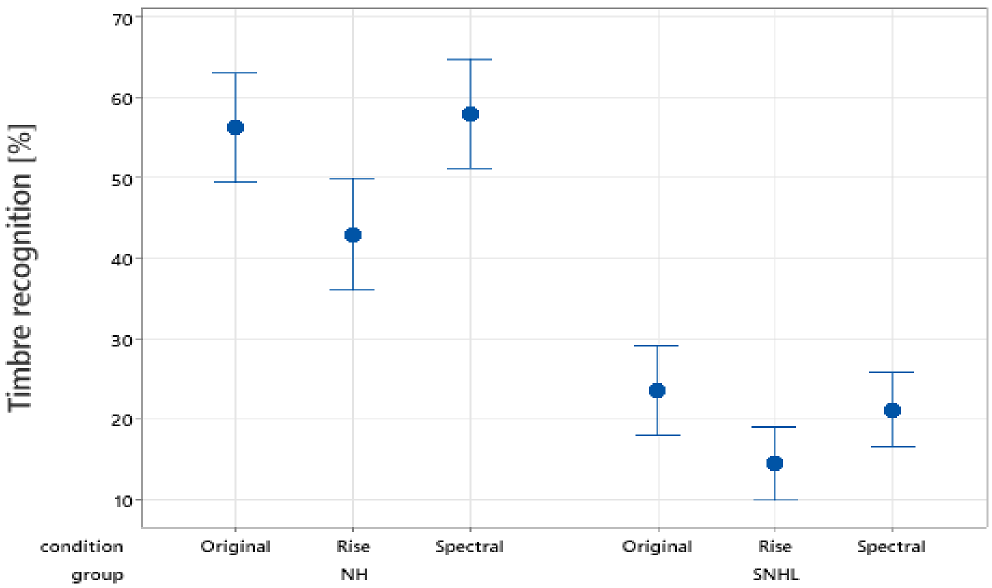

Figure 2. The percentage of timbre recognition in two normal and sensorineural hearing loss groups in different conditions. NH; normal hearing, SNHL; sensorineural hearing loss

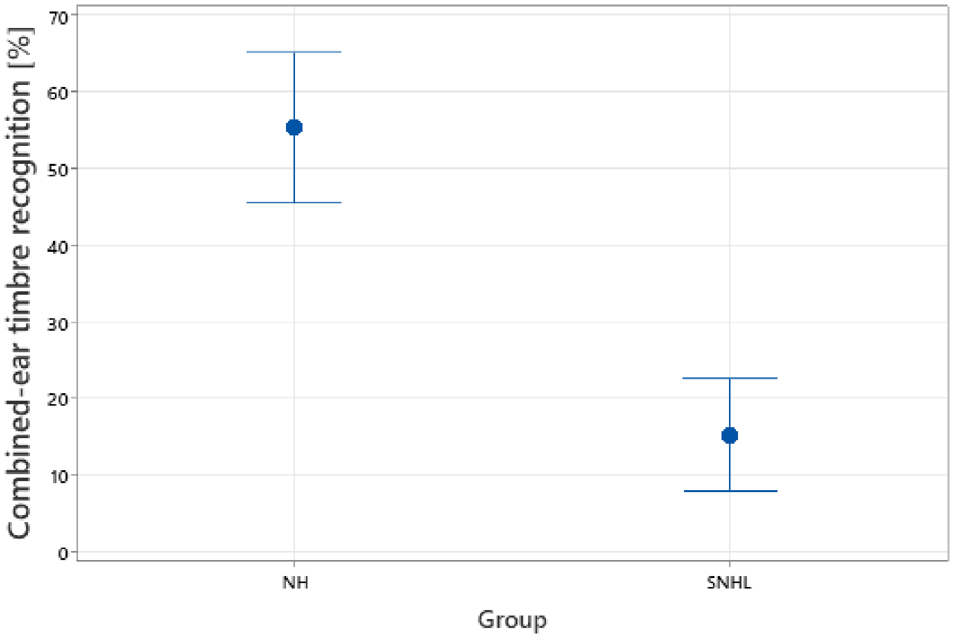

Figure 3. The mean and confidence interval of the timbre recognition score in integration condition for two normal and sensorineural hearing loss groups. NH; normal hearing, SNHL; sensorineural hearing loss

Discussion

In this study, the relative importance of spectral and temporal cues in the instrument (timbre) recognition among people with normal hearing and with SNHL was investigated. It has long been known that, among the temporal cues of musical notes (rise time, fall time and plateau), the most important cue is the rise time [5]. In addition, it is known that the strength of the resolved harmonics also has an important role in the recognition of timbre. However, the degree of contribution and relative importance of spectral and temporal cues in comparison with each other have not been investigated even in people with normal hearing, to the best of our knowledge. Therefore, the participants in this study experienced three conditions (original, temporal-manipulated and spectral-manipulated) to reveal which cues were more reliable in instrument recognition. The study revealed that participants without access to temporal cues had poorer musical instrument recognition compared to when spectral cues were unavailable.

Timbre recognition plays a crucial role in helping to differentiate between familiar and unfamiliar voices, locate a person's voice in a crowded place, and identify different musical instruments [5]. Despite its importance, this multi-dimensional and complex conception is poorly understood. In the plateau phase, the reliance on temporal cues for timbre recognition may shift to spectral cues [8, 9]. The reliability of the fall time can be affected by the reverberation of the note, leading to inconsistency across different environments [2]. Therefore, the most reliable temporal cue in various listening conditions is the rise time. Our study revealed that, when participants had no access to temporal cues, their ability to recognize musical instruments was poorer compared to when spectral cues were unavailable. In one study, the notes were played in reverse temporal order, resulting in a severe reduction in discrimination ability of people with normal hearing, although the notes had the same spectral content [10]. A study suggested that it is easier to understand the timbre of musical instruments with a fast rise time such as piano and guitar, highlighting the importance of temporal cues [11]. A study demonstrated the significance of temporal cues by combining temporal information from one instrument with spectral information from another instrument (referred to as chimeras). The participants, including both normal-hearing people and cochlear implant users, relied on temporal information to recognize the musical instrument [12].

The perception of timbre is also influenced by the spectral-intensity component, particularly during the plateau phase. This component allows for a detailed analysis of the frequency components of a note, determining the presence of various harmonics. The distributed energy among different harmonics is a key factor for distinction between musical instruments. For instance, the piano primarily maintains stable energy at the fundamental frequency, while the violin and accordion distribute energy across multiple harmonics. The flute exhibits the most energy in the first five harmonics, while the saxophone notes mainly contain energy in the first and second harmonics. Some instruments, such as the clarinet and flute, lack energy in higher harmonics, while others such as the trumpet, violin, and saxophone, have energy in higher harmonics [13]. To quantify the importance of different harmonics, a measure known as the spectral centroid is used, representing the frequency at which there is a balanced energy on both sides of the spectrum. However, it requires averaging several stable cycles of harmonics, which is practically challenging during the transient periods such as the rise and fall times. Previous studies have indicated that individuals with normal hearing rely more on temporal cues and spectral centroid than on the Temporal Fine Structure (TFS) of notes [6, 14]. As long as they can extract necessary information from these two factors, they do not rely much on the TFS. The rise time of notes can affect their temporal envelope, while the spectral centroid and TFS are linked to the spectral characteristics of notes. Considering the spectral manipulation in the current study, in line with previous studies, it highlights the significance of temporal cues in timbre recognition. These findings have broader implications for cochlear implant or hearing aid users.

Since hearing loss not only affects speech perception but also the perception and enjoyment of music. Despite the similarities between speech and music, they have important differences. Thus, there was a need for their further investigation in individuals with normal hearing and those with SNHL. The current study revealed that temporal cues were important for music instrument in both normal-hearing and SNHL people. When temporal cues were masked, the mean timbre recognition decreased from 56.1% to 42.9% in the normal-hearing group and decreased from 23.5% to 14.4% in the SNHL group. Another finding of the current study was the similar effect of SNHL on temporal and spectral cues. There was no significant interaction effect of group and condition, indicating that SNHL affected the conditions with a similar pattern. The scores of individuals with SNHL decreased under original, spectral-manipulated and temporal-manipulated conditions with a similar pattern. While spectral cues in different forms of TFS and spectral centroid may be available to the individuals, temporal cues are available from temporal envelope cues. According to the results, hearing loss significantly reduced the perception of timbre. However, it should be noted that people in this study had moderate to severe hearing loss. With an increase in the degree of hearing loss, it is expected that the width of the frequency tuning of neurons in the auditory cortex also increase.

In this study, the scores at the original condition and two manipulated conditions equally decreased. Since the normal-hearing group had the lowest score at the temporal manipulated condition, the same trend can be seen in people with SNHL. However, the rate of timbre recognition dropped dramatically possibly due to two destructive factors of temporal manipulation and hearing loss. The rate of correct timbre recognition was 56.1% in the normal-hearing people and 23.5% in the SNHL people. In one study, the rate of correct timbre recognition in the SNHL people was in a range of 50.3–73.9%, while this rate was 95.2% in the normal-hearing people [11]. In another study, the timbre recognition was assessed using the Korean version of the CAMP test, and the score in the SNHL people was 33.07% [6], which is close to the rate reported in our study. It seems that the participants in our study were less familiar with the instruments since all instrument were for the western music culture.

In this study, at the integrated condition, normal-hearing group obtained a score of 55.3%, which was not significantly different from the score at the original condition (56.1%). In the SNHL people, no significant difference was reported between these two conditions. Two possible reasons can be suggested to explain this finding. First, the brain does not combine information, but uses a cue that has a higher efficiency for instrument recognition, which was the temporal cue in this study. This weighting of the cues has already been mentioned in other cases [11]. Another possible reason is that the brain engages in intelligent integration of temporal information from one ear with useful spectral information from the other ear. Since the score at the integrated condition was not better that that at the spectral-manipulated and original conditions, it is not possible to support or reject these possible reasons. Furthermore, non-significant difference between the scores of the integrated and original conditions in the SNHL group is an intriguing finding that need further exploration.

Due to the absence of a standardized test for measuring the timbre recognition performance in the Iranian population, the lower scores of timbre recognition at the original condition in individuals may be attributed to their lack of cultural familiarity with Western instruments. As a result, a huge effort was made to train individuals to mitigate this potential impact on the results. It is recommended to conduct similar studies on hearing aid users and evaluate the development of targeted interventions with a focus on enhancing temporal processing of auditory signals to improve timbre recognition in individuals with hearing impairment. The results of our study can be useful for music and sound processing technologies to better fit them to the individuals with hearing loss. It is also crucial to have access to valid and reliable music tests using the Iranian culture-based instruments. The results can also help in improving the design and effectiveness of hearing aids for Iranian people with hearing loss.

Conclusion

When people do not have access to temporal cues, they had poorer musical instrument recognition than when spectral cues are not available to them. This indicates that temporal cues have greater importance for musical instrument recognition in both normal-hearing and SNHL people. Moreover, hearing loss has similar effect on temporal and spectral cues.

Ethical Considerations

Compliance with ethical guidelines

This study with an ethics code of IR.IUMS.REC.1400.630 was conducted in the School of Rehabilitation, Iran University of Medical Sciences.

Funding

This research did not receive any grant from funding agencies in the public, commercial, or non-profit sectors.

Authors’ contributions

MM: Study design, acquisition of data, drafting the manuscript; MM: Study design, interpretation of the results, statistical analysis; AP: Study design, interpretation of the results.

Conflict of interest

No conflicts of interest were disclosed by the authors.

References

1. Houtsma AJM. Pitch and timbre: Definition, meaning and use. J New Music Res. 1997;26(2):104-15 [DOI:10.1080/09298219708570720]

2. Howard DM, Angus JAS. Acoustics and Psychoacoustics. 3rd ed. Abingdon UK: Routledge; 2012.

3. Plomp R. Timbre as a multi-dimensional attribute of complex tones. Frequency analysis and periodicity detection in hearing. 1970;397-414.

4. Grey JM. Multidimensional perceptual scaling of musical timbres. J Acoust Soc Am. 1977;61(5):1270-7. [DOI:10.1121/1.381428]

5. McAdams S, Winsberg S, Donnadieu S, De Soete G, Krimphoff J. Perceptual scaling of synthesized musical timbres: common dimensions, specificities, and latent subject classes. Psychol Res. 1995;58(3):177-92. [DOI:10.1007/BF00419633]

6. Kim HJ, Lee JH, Shim HJ. Effect of Digital Noise Reduction of Hearing Aids on Music and Speech Perception. J Audiol Otol. 2020;24(4):180-90. [DOI:10.7874/jao.2020.00031]

7. Kang R, Nimmons GL, Drennan W, Longnion J, Ruffin C, Nie K, et al. Development and validation of the University of Washington Clinical Assessment of Music Perception test. Ear Hear. 2009;30(4):411-8. [DOI:10.1097/AUD.0b013e3181a61bc0]

8. Lakatos S. A common perceptual space for harmonic and percussive timbres. Percept Psychophys. 2000;62(7):1426-39. [DOI:10.3758/bf03212144]

9. Marozeau J, Lamping W. Timbre Perception with Cochlear Implants. In: Siedenburg K, Saitis C, McAdams S, Popper AN, Fay RR, editors. Timbre: Acoustics, perception, and cognition. 1st ed. New York: Springer; 2019. p. 273-93.

10. Wei Y, Gan L, Huang X. A Review of Research on the Neurocognition for Timbre Perception. Front Psychol. 2022;13:869475. [DOI:10.3389/fpsyg.2022.869475]

11. Prentiss SM, Friedland DR, Nash JJ, Runge CL. Differences in Perception of Musical Stimuli among Acoustic, Electric, and Combined Modality Listeners. J Am Acad Audiol. 2015;26(5):494-501. [DOI:10.3766/jaaa.14098]

12. Heng J, Cantarero G, Elhilali M, Limb CJ. Impaired perception of temporal fine structure and musical timbre in cochlear implant users. Hear Res. 2011;280(1-2):192-200. [DOI:10.1016/j.heares.2011.05.017]

13. Marozeau J, de Cheveigné A, McAdams S, Winsberg S. The dependency of timbre on fundamental frequency. J Acoust Soc Am. 2003;114(5):2946-57. [DOI:10.1121/1.1618239]

14. Caclin A, McAdams S, Smith BK, Winsberg S. Acoustic correlates of timbre space dimensions: a confirmatory study using synthetic tones. J Acoust Soc Am. 2005;118(1):471-82. [DOI:10.1121/1.1929229]