Standardization of Three Familiar Sound Recognition Tests in Hearing and Deaf Adult Populations

*Clémence Debiève, †Flora Rosenzweig, and †Justine Wathour

*UCLouvain, Louvain-La-Neuve, Belgium; and †Cliniques universitaires Saint-Luc, Brussels, Belgium

Objective: Recognition of familiar noises is crucial for understanding and reacting appropriately to our auditory environment. Its improvement is one of the benefits expected after cochlear implantation. The aim of this study was to standardize three environmental sounds noise recognition tests and to illustrate their application to a population of deaf adults with cochlear implants.
Method: Norms were established on a sample of 126 normal-hearing adults divided into 6 age groups. Three familiar sound recognition tests were used: 1) the Blue Mouse "First Familiar Sounds" (BM), 2) the UCL-IRSA test (TI), and 3) the Bernadette Piérart Familiar Sounds Test (TBF). These tests were also administered to 61 implanted deaf ears.
Results: We observed a significant effect of age on the accuracy scores of the TI and TBF tests for the hearing group and on the time scores of the TI and BM tests. Overall, the performance of the deaf participants was poorer and more variable than that of the hearing participants.
Conclusion: We have three tests that can be used in practice to measure the performance of deaf people (with cochlear implants) at different stages of their pre- and post-implant rehabilitation.
Key Words: Cochlear implant—Familiar sound—Test.

Otol Neurotol **00:**00–00, 2024.

INTRODUCTION

Although improving speech perception is often described as the main goal of cochlear implantation, improved perception of environmental sounds, particularly stimuli associated with safety (cars, alarms, etc.), is also reported as an important factor of quality of life by implant users (1–3). To date, few studies have focused specifically on the perception of environmental sounds in cochlear implanted deaf individuals (1,3). However, the perception and recognition of these environmental sounds are crucial for understanding and navigating our surroundings (2,3).

Currently, complete restoration of both speech and environmental sound perception remains limited by technical constraints of the implant itself. Indeed, despite the considerable advances offered by the cochlear implant (CI) in restoring peripheral hearing, its limited frequency resolution may be responsible for distortion of incoming sensory input. Consequently, users may experience both verbal and nonverbal hearing difficulties (4,5), and have been shown to demonstrate poor to average performance in environmental stimuli recognition tasks (5). Interestingly, CI users do not always complain about environmental sounds. This may be due in part to their limited awareness of the richness of their auditory environment or because priority is often given to speech perception. The absence of complaints does not therefore necessarily predict the level of ability of CI users (2,3,5). Recently, Shafiro et al. (3) carried out a review of the literature on this subject. Nineteen studies were identified, the majority of which focused on postlingually deaf adults (one study on prelingually deaf adults and three on children). In most studies, patient identification performances were assessed through the presentation of isolated environmental sounds in quiet followed by designation among a suggested written word list. The reported outcomes were mostly mediocre and variable. Moreover, several of these studies highlight a link with other assessments (word recognition, sentence recognition, etc.) (5–8). It is not always easy to compare the various studies, given the methodological differences (number of stimuli, number of response options, different types of sound used, etc.) (3).

In general, there are few standardized speech and language tests for deaf people. However, norms can be used to compare a patient's performance with that of a reference population (9). It should be noted that familiar sound recognition tests are of interest not only in the context of hearing pathologies but also in the field of neuropsychological pathologies and language delay (10).

In Belgium, at the Centre d'Audiophonologie des Cliniques Universitaires Saint-Luc in Brussels, three tests are used in

Address correspondence and reprint requests to Justine Wathour, Ph.D., Cliniques Universitaires Saint-Luc, Avenue Hippocrate 100, 1200 Brussels, Belgium; E-mail: justine.wathour@uclouvain.be
The authors have no funding sources to report.
The authors disclose no conflicts of interest.
Ethics approval: This study was approved by the local ethical review board (Hospital St-Luc—B4032021000101).
Research involving human participants: Prospective study.
Informed consent: Patients signed a consent form to participate in the study.
Code availability: Not applicable.
DOI: 10.1097/MAO.0000000000004215

practice to assess these abilities: the Blue Mouse (BM) "the first familiar sounds," the Uclouvain—Institut Royal pour Sourds et Aveugles (UCL-IRSA) test (Royal Institute for the Deaf and Blind), and Bernadette Piérart's "Test des Bruits Familiers" (TBF).

This study reviews the standardization of these three tests on a hearing adult population and illustrates their application to a population of deaf adults with cochlear implants.

MATERIAL AND METHOD

Subjects

One hundred twenty-six hearing participants (53 men and 73 women) were tested between February and June 2022 (average age: 47.2 yr). All patients were 18 years or older and had a good command of the French language. Written informed consent was obtained from all participants patients before inclusion. Participants with hearing loss (tonal thresholds over 30 dB), a neurological disorder, or an established disorder preventing them from taking the tests were excluded. Participants were divided into six age groups (Table 1).

Fifty deaf patients, representing 61 cochlear implanted ears (Table 2) with a bilateral hearing loss, were tested between February and June 2022 (mean age: 53.02 yr). All patients were 18 years or older, had had at least one cochlear implant for at least 1 year, and had a good command of the French language. Patients with a neurological disorder or a known disorder that prevented them from taking the tests were excluded. Each patient had been implanted and evaluated in our academic hospital.

There was no monetary incentive to participate in the study.

Procedure

The data were prospectively collected over a period of 4 months.

Tests Used

Pure-tone audiometry

For hearing participants, auditory thresholds were assessed at 125, 250, 500, 1000, 2000, 4000, and 8000 Hz using the "Mimi Hearing Test" application (Mimi Hearing Technologies GmbH, https://apps.apple.com/us/app/mimi-hearing-test/id932496645).

For deaf participants, audiometric thresholds for warble tones were obtained in free-field condition at 250, 500, 1000, 2000, 4000, 6000, and 8000 Hz. The Bureau International d'Audiophonologie (BIAP) score was calculated from the average loss in dB (HL) at frequencies 500, 1000, 2000, and 4000 Hz. This corresponds to the average acoustic level representing the average total hearing loss (BIAP, n.d.).

TABLE 1. *Distribution of participants by age (years) and gender*

	18–29	30–39	40–49	50–59	60–69	70–85	Total
Male	14	5	10	10	7	7	53
Female	15	12	11	15	11	9	73
Total	29	17	21	25	18	16	126

TABLE 2. *Distribution of participants by ear tested*

Ear Tested	Left	Right	Two Ears	Total
Total	25	31	5	61

Three sound recognition tests are as follows:

1. The familiar sounds panel of the BM test: The BM is a software used for auditory stimulation and to assess children's auditory recognition skills. However, it is also offered to deaf adults. It is used by means of a computer. It was created over 20 years ago by Alain Vinet with the help of a team of speech therapists. It consists of 11 different boards made up of sounds and associated images. The board of familiar sounds presents 10 very familiar items perceived from an early age. It includes the following sounds and their images: a car, a dog, a telephone, a mother singing a lullaby, a cat, a vacuum cleaner, a doorbell, a father reading a story, and a television. The board also contains an additional image, the hat, which acts as an intruder. Patients are required to associate the presented sound with the corresponding image. (A. Vinet, personal communication, June 3, 2022).
2. Bernadette Pierart's TBF: The TBF has been available since May 2020 on the Ortho Édition website. It was created by Bernadette Piérart. It is a computerized test comprising 20 familiar sounds. Each sound is presented to the patient at the same time as 4 black and white images on the screen, arranged in quadrants. One image corresponds to the sound presented, whereas the other three have a distractor function (an acoustically close distractor, a conceptually close distractor and an unrelated distractor). The patient must click on the image corresponding to the auditory stimulus. Each noise lasts 30 seconds, appears only once, and is worth 1 point if successful. There are three categories of noises: cries (11 sounds), environmental noises (7 sounds related to nature or the home), and community noises (2 sounds). They come from the sound bank of the Audiovisual Centre of the Université Catholique de Louvain. The entire test lasts 15 minutes. This test has been calibrated for children as well as for adolescents and adults. For adolescents and adults, norms are available for three adult/adolescent age groups (12–39, 40–59, and >60 yr).
3. The UCL-IRSA familiar sound recognition test (TI): The UCL-IRSA familiar sounds recognition test contains 10 sounds and is free of rights downloadable from the Internet (F. Vander Linden, personal communication, July 27, 2022). No other information is available about the sounds or images used. The following 10 sound extracts are supplied, with the 10 images corresponding to the sounds: a telephone, a crying baby, a horse, a police car, applause, a dog, a dripping tap, a cow, a car, and a piano.

Experimental Design

Testing was performed in a quiet environment. Sounds were presented at a comfortable intensity (50–60 dB). For the deaf group, in the event of residual hearing, the contralateral ear was masked. The order in which the tests were presented was determined at random. The instructions and evaluation protocols were identical for all participants. The instructions given were: "You will hear sounds and you will have to identify the image corresponding to the sound you hear. The three tests will be timed, but don't worry about the stopwatch."

The items were presented in the same order. The TBF test does this automatically, whereas for the BM and IRSA tests, the experimenter must trigger the sounds in the correct order. Bilaterally implanted deaf people had to take the tests three times: once with their left speech processor, once with their right speech processor, and once with both speech processors. The three tests were carried out on the same day. To avoid a learning effect and to ensure that the same ear would always start, the tests alternate between the three conditions.

The BM and TI tests were scored in the same way:

- Mark "1" if the item was correctly designated and "0" if the item was not correctly designated. Immediate self-correction was accepted.
- Each sound was only played once per trial.
- The total time was recorded in seconds.
- The total score out of 10 was then converted to a percentage of correct responses (in practice, speech and language tests for the deaf are expressed as percentages).

The TBF was corrected automatically by the computer (each correct answer was worth 1 point). However, the temporal score had to be added manually to the total. This total then had to be converted into a percentage.

Statistical Analysis

Statistical analyses were carried out using IBM SPSS (Chicago, IL) Statistics Version 27 software.

The normal distribution of the data for the different variables was tested using the Kolmogorov–Smirnov and Shapiro–Wilk tests. These tests showed, with a 5% risk of error, that all the variables except total TBF time did not follow a normal distribution. The following nonparametric tests were therefore applied: Kruskal–Wallis, Mann–Whitney, Wilcoxon, Friedman, and Spearman's coefficient.

An alpha risk of 5% ($p < 0.05$) for rejecting the null hypothesis was chosen for all statistical analyses. For some, this p value had to be adjusted for the number of comparisons carried out on the sample.

A correlation is considered strong enough not to be due to chance if it is significant and ≥0.6 or ≤−0.6.

Ethics Commission

This paper was approved by the Comité d'Éthique Hospitalo-Facultaire on January 11, 2022 (Belgian Registration No.: B4032021000101).

RESULTS

Hearing Population

Standardization of the Three Tests

The three standardized tests (accuracy scores and time scores) are the BM, the UCL-IRSA test (TI), and the Family Noise Test (TBF).

Age Group Comparison

Precision score

The precision scores of the three tests were analyzed.

A significant difference between age groups was found for the TI test ($\chi^2(5) = 17.882$, $p < 0.05$) and the TBF test ($\chi^2(5) = 19.684$, $p < 0.05$) but not for the BM test ($\chi^2(5) = 7.551$, $p = 0.183$).

The Mann–Whitney test for two independent samples was used to determine precisely which age groups differ from each other in terms of precision scores. The p value was adjusted with the Bonferroni correction according to the number of 2-to-2 comparisons made.

The 70–85 age group differs significantly from the 18–29 age group ($U = 174$, $p < 0.0167$) and the 50–59 age group ($U = 150$, $p < 0.0167$).

For the TBF test, the 70–85 age group differed significantly from the 18–29 age group ($U = 65.5$, $p < 0.00625$) and the 50–59 age group ($U = 98$, $p < 0.00625$).

Time score

The Kruskal–Wallis test was not significant for the TBF test ($\chi^2(5) = 2.465$, $p = 0.782$), but it was significant for the TI test ($\chi^2(5) = 41.924$, $p < 0.05$) and for the BM test ($\chi^2(5) = 24.320$, $p < 0.05$).

For BM, the age groups that differed significantly were 70–85, 18–29 ($U = 71$, $p < 0.0071$), 30–39 ($U = 46$, $p < 0.0071$), 40–49 ($U = 43$, $p < 0.0071$), 50–59 ($U = 79.5$, $p < 0.0071$), and 60–69 ($U = 60$, $p < 0.0071$).

For TI, the age groups that differed significantly were the 70–85 age group and the 18–29 ($U = 23$, $p < 0.0056$), 30–39 ($U = 12$, $p < 0.0056$), 40–49 ($U = 13$, $p < 0.0056$), 50–59 ($U = 10$, $p < 0.0056$), and 60–69 ($U = 39.5$, $p < 0.0056$) age groups.

There was no significant and sufficiently strong correlation between the three tests (Spearman coefficient >0.6 or <−0.6).

Correlation between Precision and Time Scores for Each Test

Spearman's nonparametric coefficient shows a significant and sufficiently strong correlation (−0.649) between the accuracy and time scores for the BM test.

Deaf Population

Comparison With Results for Hearing People

Qualitatively, the deaf patients had poorer scores (precision and time) than the hearing participants (Figs. 1, 2, and 3). The deaf patients' scores also showed greater variability, as evidenced by the confidence interval error bars

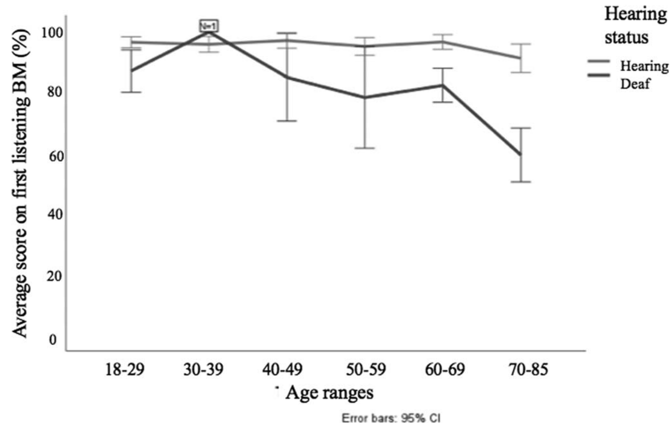

FIG. 1. Comparison of first listening accuracy scores in the BM test of deaf participants with those of hearing participants.

within which the scores of 95% of the participants fall. Finally, the ceiling effect (accuracy score) observed in the hearing group for the TI and BM tests was not found in the deaf participants.

Precision Scores

The precision scores of the three tests differed significantly ($\chi^2(2) = 64.599$, $p < 0.05$). The differences between the three tests were analyzed in detail using Wilcoxon tests, adapting the p value (0.0167). The TBF total score differed significantly from the TI first listening score ($W = 6.177$, $p < 0.0167$) and the BM first listening score ($W = -5.808$, $p < 0.0167$).

Correlation Between Precision and Speed Scores

The correlations (Spearman's coefficient) between the precision and time scores for BM and TI are significant ($p < 0.05$) and are −0.834 and −0.821, respectively. The correlation for TBF is significant but not strong enough (<+0.6 and >−0.6).

Comparison Between Test Precision Scores and Pure-Tone Audiometry

The results of the three tests (precision and time) of the participants with a poorer BIAP score were compared with the results of the participants with a better BIAP score. The two groups were created based on their median (26.25 dB):

- The 27 participants with worse tonal thresholds had a BIAP score >26.25.
- The 29 participants with better tonal thresholds had a BIAP score </= 26.25.

The Mann–Whitney statistical test showed that the BM accuracy, and TI and BM time scores of participants with a higher BIAP were significantly better than those with a lower BIAP ($p < 0.05$).

DISCUSSION

Regarding *precision*, in hearing people, the Family Noise Test and the UCL-Irsa test (IT) show significant differences according to age. Participants aged between 70 and 85 years obtained significantly lower accuracy scores than participants aged between 18 and 29 years and between 50 and 59 years for the Family Noise Test and the UCL-IRSA (TI) test. These results are not surprising given that the peripheral auditory system deteriorates with age. This phenomenon was controlled by selecting only participants with sufficient tonal thresholds. The central auditory system and

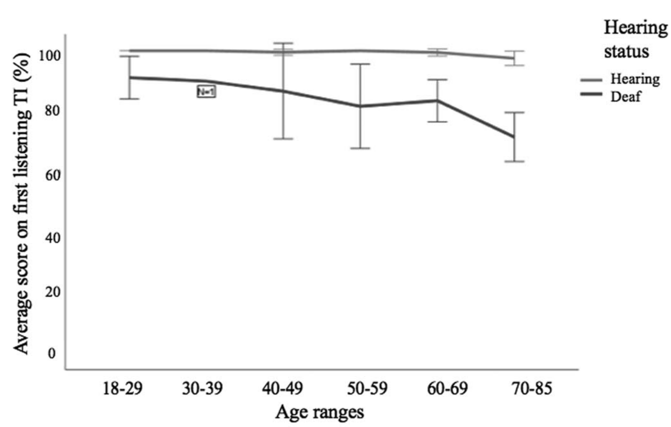

FIG. 2. Comparison of first listening accuracy scores in the TI test of deaf participants with those of hearing participants.

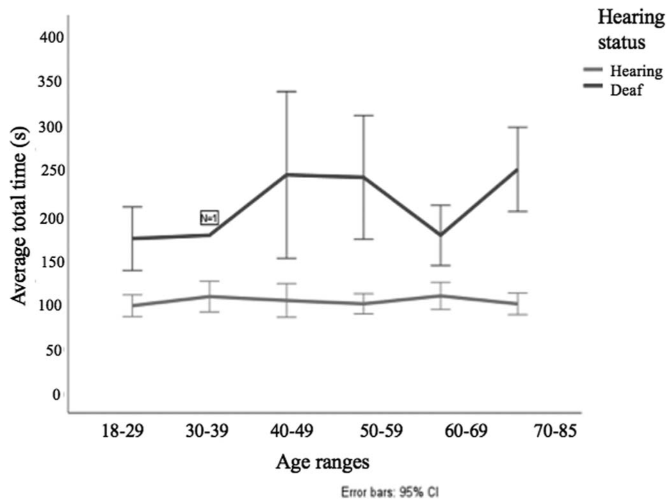

FIG. 3. Comparison of the TBF test accuracy scores of deaf participants with those of hearing participants.

the cognitive system also deteriorate with age, which would also explain the poorer results of the older participants (11). For practical reasons, we were unable to check these two parameters (central hearing and cognitive abilities).

Regarding *time*, hearing participants between the ages of 70 and 85 years also took longer than younger participants (<70) to complete the BM and UCL-IRSA tests. These results are consistent with the fact that processing speed is slower in the elderly. It begins to decline in the third decade, and this decline continues throughout life (12). On the other hand, for the TBF, there is no significant difference in time scores, probably because the sounds are more complex and some are only recognizable thanks to an acoustic event arriving later (e.g., the sound of cutlery for item 20). For this TBF, processing speed would therefore have little influence.

Regarding *the link between speed and accuracy*, for the hearing group, we found a single mean negative monotonic correlation between speed and accuracy scores for The BM test. In contrast, for deaf participants, the UCL-IRSA test and the BM test showed a strong negative monotonic correlation between accuracy and speed scores. The higher the accuracy score, the faster the participants responded. We did not find the same correlation in the hearing population, probably due to the ceiling effect of these two tests. On the other hand, no link between speed and accuracy was found for the Familiar Noise Test, again perhaps due to the more complex noises and the sometimes late appearance of a stimulus triggering auditory recognition.

Regarding *deaf and hearing population comparison*, we observed that the performance of the deaf participants in the study is highly variable and is poorer than that of the hearing participants, which corroborates the current scientific literature (6). The hearing participants obtained a mean score of between 91.25 and 97.14% for BM, between 97.5 and 100% for TI, and between 76.88 and 90.34% for TBF, which is close to the results of McMahon et al. (1), who found 90% of correct answers.

Deaf participants obtained an identification mean of 78.52% for BM, 83.11% for TI, and 62.21% for TBF. Some studies show an average identification of familiar environmental sounds close to our results for the deaf, whereas others present rather different averages. For example, participants in Reed and Delhorne (8) averaged 79.2%, which is close to the results of the BM test and the UCL-IRSA test, and further away from the Familiar Sounds Test. In contrast, the study by Shafiro et al. (5) found an average of 45.3%, which is lower. These variations may be explained by methodological differences in the task:

- Most studies use a designation of written words and not images, as in our study.
- The nature of the link to be made is not the same. The sound–verbal label relationship is more arbitrary, whereas the sound–image relationship is more iconic and causal (13,14). It is possible that the brain processes involved in these types of tasks are not exactly the same, especially as there are many links between hearing and vision that influence recognition abilities. Direct connections modulate neural activity between auditory and visual cortex, as well as with associative cortex (multisensory integration) (15,16).
- Response modalities (free-response format, closed-response format with many or few options)
- Number of sounds
- Distractor characteristics
- Sample characteristics (age, type of deafness, implant use, etc.) sound quality of noises
- Reading abilities of tested subjects

It should be noted that few studies exist in the scientific literature on the recognition of familiar sounds in deaf and hearing people, so it is difficult for us to draw firm conclusions (3).

We found no studies that examined *the time performance* of hearing and deaf participants. Yet, they provide important information in relation to auditory recognition. It is necessary to be able to react to our environment adequately, but also quickly. In our study, hearing participants performed

BM between 34.58 and 63.08 seconds, TI between 31.27 and 48.46 seconds, and TBF between 97.59 seconds (1 min 38 s) and 108.56 seconds (1 min 48.56 s). The deaf are slower, completing BM in 111.84 seconds (1 min 52 s), TI in 118.96 seconds (1 min 58.96 s), and TBF in 205 seconds (3 min 25 s).

For both hearing and deaf participants, the UCL-IRSA is on average the most successful test, followed by the BM and finally the Familiar Noise Test. The UCL-IRSA and the BM have a ceiling effect in hearing participants, but not in the deaf. The Familiar Noise Test shows no ceiling effect for the hearing group; it seems more difficult, perhaps because it mobilizes more high-level cerebral processes and integration capacities. For example, for sounds in the "community noise" category, the ability to segregate auditory stimuli is important. To recognize item 20 (the canteen), you need to identify the sound of cutlery among all the sounds. This test also requires greater semantic precision to link the sound to its referent and thus recognize the auditory object, because the distractors are specific to each sound, sometimes leading to causal uncertainty. The nature of the distractors influences the level of semantics required to recognize an object. Multisensory processing is influenced not only by the low-level properties of stimuli but also by other factors such as the congruence of perceptual and semantic features of sensory inputs (17). Cross-modal semantic incongruence interferes with audiovisual object recognition (18).

As a general observation, participants with lower tonal thresholds perform significantly better on the BM and UCL-IRSA tests than those with higher tonal thresholds. This effect is qualitatively the same for the Familiar Noise Test. Low-level hearing ability seems to play a role in the recognition of familiar noises, whatever the complexity of the task.

To summarize, in practical terms, we have three tests capable of assessing the recognition of familiar noises, standardized on adult hearing subjects, and verified on deaf subjects. The aim is to enrich the practice of clinical audiometry, in particular by evaluating this recognition in implant patients before and after implantation, monitoring their identification progress, and offering them customized training.

Given the information this study provides, we suggest using BM and UCL-IRSA at the start of programming, and when the patient has achieved a good score, moving on to TBF to monitor progress.

REFERENCES

1. McMahon KR, Moberly AC, Shafiro V, Harris MS. Environmental sound awareness in experienced cochlear implant users and cochlear implant candidates. *Otol Neurotol* 2018;39:e964–71.
2. Harris MS, Moberly AC, Hamel BL, et al. A longitudinal comparison of environmental sound recognition in adults with hearing aids before and after cochlear implantation. *J Speech Lang Hear Res* 2021;64:1040–52.
3. Shafiro V, Luzum N, Moberly AC, Harris MS. Perception of environmental sounds in cochlear implant users: a systematic review. *Front Neurosci* 2022;15:788899.
4. Pralus A, Hermann R, Cholvy F, et al. Rapid assessment of non-verbal auditory perception in normal-hearing participants and cochlear implant users. *J Clin Med* 2021;10:2093.
5. Shafiro V, Gygi B, Cheng MY, Vachhani J, Mulvey M. Perception of environmental sounds by experienced cochlear implant patients. *Ear Hear* 2011;32:511–23.
6. Harris MS, Boyce L, Pisoni DB, Shafiro V, Moberly AC. The relationship between environmental sound awareness and speech recognition skills in experienced cochlear implant users. *Otol Neurotol* 2017;38:e308–14.
7. Inverso Y, Limb CJ. Cochlear implant–mediated perception of nonlinguistic sounds. *Ear Hear* 2010;31:505–14.
8. Reed CM, Delhorne LA. Reception of environmental sounds through cochlear implants. *Ear Hear* 2005;26:48–61.
9. Bernaud J-L. Méthodes de tests et questionnaires en psychologie. Paris: Dunod; 2005;26:48–61..
10. Piérart B, Dehant A, Chapelle P. Une épreuve d'identification de bruits familiers—intérêt et limites. *Glossa* 1996;53:26–35 https://www.glossa.fr/index.php/glossa/article/view/250.
11. Profant O, Jilek M, Bures Z, et al. Functional age-related changes within the human auditory system studied by audiometric examination. *Front Aging Neurosci* 2019;11:26.
12. Harada CN, Natelson Love MC, Triebel KL. Normal cognitive aging. *Clin Geriatr Med* 2013;29:737–52.
13. Saygin AP, Dick F, Bates E. An on-line task for contrasting auditory processing in the verbal and nonverbal domains and norms for younger and older adults. *Behav Res Methods* 2005;37:99–110.
14. Sirri L, Guerra E, Linnert S, et al. Infants' conceptual representations of meaningful verbal and nonverbal sounds. *PLoS One* 2020;15:e0233968.
15. Bizley JK, Maddox RK, Lee AKC. Defining auditory-visual objects: Behavioral tests and physiological mechanisms. *Trends Neurosci* 2016;39:74–85.
16. Opoku-Baah C, Schoenhaut AM, Vassall SG, et al. Visual influences on auditory behavioral, neural, and perceptual processes: A review. *J Assoc Res Otolaryngol* 2021;22:365–86.
17. Thomas RL, Nardini M, Mareschal D. The impact of semantically congruent and incongruent visual information on auditory object recognition across development. *J Exp Child Psychol* 2017;162:72–88.
18. Viggiano MP, Giovannelli F, Giganti F, et al. Age-related differences in audiovisual interactions of semantically different stimuli. *Dev Psychol* 2017;53:138–48.