Audiology

Development of a novel Italian speech-in-noise test using a roving-level adaptive method: adult population-based normative data

Sviluppo di un innovativo test audiometrico vocale nel rumore in italiano che utilizza un metodo “roving-level” adattivo: dati normativi basati sulla popolazione adulta

P. CANZI1, M. MANFRIN1, G. LOCATELLI1, P. NOPPI1, M. PEROTTI2, M. BENAZZO1

1 Department of Otorhinolaryngology, IRCCS Policlinico San Matteo Foundation and University of Pavia, Italy; 2 MED-EL GmbH, Innsbruck, Austria; 3 SS Otoproteesica, Implanting Center Regione Piemonte, Ospedale Civile “SS Antonio and Biagio e C. Arrigo”, Alessandria, Italy

SUMMARY

In recent years the increasing development of hearing devices has led to a critical analysis of the standard methods employed to evaluate hearing function. Being too far from reality, conventional investigation of hearing loss based on pure-tone threshold audiometry and on mono/disyllabic word lists, presented in quiet conditions, has been shown to be inadequate. A speech-in-noise test using a roving-level adaptive method employs target and competing signals varying in level in order to reproduce everyday life speaking conditions and explore a more complete sound range. Up to now, only few roving-level adaptive tests have been published in the literature. We conducted a roving-level adaptive test in healthy Italian adults to produce new normative data on a language of Latin origin.

KEY WORDS: Speech-in-noise test • Roving-level adaptive test

RIASSUNTO

Negli ultimi anni, il crescente sviluppo di dispositivi acustici ha condotto a un’analisi critica dei metodi standard che sono stati impiegati per valutare la funzione uditiva. Gli esami audiologici tradizionali, basati sulla soglia audiometrica tonale e sulle liste di parole mono/bisillabiche nel silenzio, si sono nel tempo dimostrati inadeguati perché troppo distanti dalla realtà. Un test audiometrico vocale nel rumore, che utilizza un metodo “roving-level” adattivo, adopera segnali target e segnali competitivi modificabili con lo scopo di riprodurre le condizioni di eloquio della vita quotidiana, quindi esplorare un più ampio range uditivo. A oggi, solamente pochi test “roving-level” adattivi sono disponibili in letteratura. Gli autori hanno condotto un test “roving-level” adattivo in adulti italiani sani, al fine di ottenere nuovi dati normativi in una lingua di origine latina.

PAROLE CHIAVE: Audiometria vocale nel rumore • Test “roving-level” adattivo

Acta Otorhinolaryngol Ital 2016;36:506-512

Introduction

In recent years the increasing development of hearing devices has led to a critical analysis of the standard methods employed to evaluate the hearing function. It is renowned that two main factors influence the outcomes of hearing aids: type and severity of hearing impairment and environmental acoustic characteristics. Clinical practice outlines that aid patients who are apparently homogeneous for anatomical and audiological features may report different hearing perception under the same fitting conditions. Moreover, speech intelligibility involves complex perceptual processes that emphasise the need for audiological tests reflecting reality as much as possible. Normal-hearing listeners show a natural and exceptional ability to select the target voice in a background of competing voices. Being too far from reality, conventional investigation of hearing loss based on pure-tone thresholds and on mono/disyllabic word lists, presented in quiet conditions, has been shown to be inadequate. In fact, the international diffusion of sentence-type speech-in-noise tests has witnessed the emerging necessity of a more accurate audiological assessment. These tests have been conceived to explore speech communication in noisy settings at a fixed or varied signal-to-noise ratio (SNR) as an expression of two basic philosophies. Fixed SNR tests are easily adopted and are administered to assess performances at a specific SNR level established before the examination. The listener’s abilities are evaluated in terms of percent intelligibility, thus showing meaningful results.
not only for the clinician, but for the tested patient as well. Nevertheless, outcomes are influenced by the presentation levels of the target speech and the noise of competition. When clinicians select SNR parameters, the measured results not only depend on subjective hearing, but also on the fixed SNR. For example, a really challenging or an easy SNR may lead, respectively, to under- or overestimate the status of the patient. The introduction of adaptive tests has overcome the described limitations by reversing the investigation perspective. No longer the tester, but the performance of the patient himself determines the SNR conditions. In particular, either the target signal level (e.g. speech) or the competing signal level (e.g. noise) increases/decreases by a set amount, according to the correct answers given by the patient. The test starts with an easy SNR level, and subsequently the SNR conditions vary depending on the accuracy of the preceding response until the subject correctly repeats 50% of speech signals. The SNR presentation level, which is necessary for a listener to correctly recognise the speech messages 50% of the time, defines the speech reception threshold (SRT). In comparison to the fixed SNR tests, outcomes provided by the adaptive ones are expressed in decibels instead of percentage intelligibility and show a reduced risk of bias arising from the limits of the method used. Many adaptive tests have been developed and all are characterised by a varied SNR, but typically either the target signal or the competing signal is fixed. It is self-evident that real life is much different: rarely do the speech signals or noise signals change while the other one remains unmodified, but all vary independently. Furthermore, each hearing device performs differently at different sound pressure levels in relation with its own specific input dynamic range. The effects of environmental noise and technical features of hearing aids on speech intelligibility should be detected as much as possible by hearing tests. A roving-level adaptive test employs target and competing signals varying in level in order to reproduce everyday life speaking conditions and to explore a more complete sound range. In particular, speech material is presented at a roving level and the noise is adapted depending on the number of correct words answered by the tester to obtain patients’ SRT. Up to now, only few roving-level adaptive tests have been published. The aim of our study was to introduce a novel Italian roving-level adaptive test to yield normative data in healthy Italian adults.

Materials and methods

We evaluated 50 native normal hearing Italian individuals (25 men and 25 women), aged between 23 and 50 years, showing a pure tone average (PTA) lower than 20 dB HL at the frequencies of 500, 1000, 2000 and 4000 Hz. To carry out the test, we employed the sentences developed by Cutugno, Prosser and Turrini, selecting only those made up of 5 words, establishing 6 lists piloted for equal difficulty and composed of 20 sentences each. Basically, some specific criteria were adopted for the construction of the sentence: choice of the lexical items, morphosyntactic structures and phrasal constituents order. The aim was to develop simple sentences that are easily understood by persons of different age groups, regional origins and cultural level. Therefore, enigmatic and abstract terms (e.g. solitude, height) or uncommon words (e.g. plausible) were avoided. With regard to the verb tenses, their usage frequency was taken into account, as well as their morphophonological structure in relation to perceptibility criteria. Considering the decreasing frequency, the most commonly used indicative verb tenses were: simple present, present perfect, imperfect tense, simple future and simple past. Sentences were short and understandable, made up of a number of syllables between 9 and 13 (Table I). Lists 1, 2 and 3 were used for the test (overall 60 sentences), while lists 4, 5 and 6 for the re-test (overall 60 sentences). The noise of competition applied was a cocktail party background. Sentences were separated by a 10-second pause while the noise of competition remained uninterrupted. The exam was performed in a soundproof booth; each sound was sent through a loudspeaker set in front of the subject, at eye-level and 1.2-meter distance. Each subject could reply by means of a microphone without any comments related to the correct answer. The procedure consisted of 60 sentences presented with a noise of competition. It lasted 14 minutes and was administered twice with a 30-minute break, in order to obtain an average result. Overall, each subject was submitted to a test and a re-test, and was instructed to repeat aloud completely or partially all the sentences heard or understood. During the test, speech material was randomly roved: sentences were casually organised in 20 groups of three sentences (overall 60 sentences) and each of the three sentences composing a triplet was casually presented at 55 dB HL, 65 dB HL, or 75 dB HL. Within each triplet all three levels were used once, the test started at +10 dB SNR with an easy competing signal and the noise level (“cocktail party” background) changed according to Table II and Table III. In particular, SNR varied across presentation levels and the SNR level was reduced progressively depending on the accuracy of the preceding response, until a conversion point that allowed calculation of the SRT value at which the subject was able to understand 50% of the sentences. The noise variation index ranged ± 5 dB HL from the 1st to the 5th triplet and ranged ± 2 dB HL from the 6th to the 20th triplet. The three presentation levels of the signal (speech) and the SNR conditions were selected to represent the range of speech levels that may typically be encountered in everyday life. In particular, 55 dB HL corresponds to a comfortable conversational speech level, while 75 dB HL represents the limit within which a proper hearing aid fitting can result beneficial in terms of auditory perception (hearing loss > 75 dB represents the selection criteria for cochlear implantation).
Table I. Lists of sentences used in the adaptive speech in noise test.

| List 1* | List 2* | List 3* |
|---------|---------|---------|
| È rimasto solo al mondo | È scappato ieri dalla prigione | La nave scivola sulle onde |
| Ieri hai comprato poco latte | Milano ha un clima freddo | L’aquilonia vola alto nel cielo |
| Francesca è incinta di nuovo | Gli ospiti sono arrivati già | Certamente ha viaggiato in treno |
| L’acqua bolle a cento gradi | Il turista passeggiava nel museo | Il ristorante ha cambiato gestione |
| Il giornale contiene cattive notizie | Il detersivo rovina la lana | Gli operai lavorano nel cantiere |
| La macchina funzionava con difficoltà | Le stelle brillano in cielo | Il fulmine ha colpito l’albero |
| Il tassista guida con prudenza | Oggi hanno camminato per ore | Tremava ancora per la paura |
| Stamattina sono andato in banca | Domani balleremo fino a tardi | I genitori sono molto apprensivi |
| Il recinto separa i giardini | Laura diceva bugie a tutti | Prepariamo la colazione alla mamma |
| L’auto abanda sempre in curva | La signora leggeva un libro | Il vino invecchia nelle botti |
| Il cameriere porterà le pizze | Le scale sono molto faticose | Il direttore convoca la riunione |
| Il mare era molto agitato | La partita iniziava in anticipo | Laura insegnava inglese ai bambini |
| Il compito è molto difficile | Le collezioni maturano a giugno | Ha amato una sola donna |
| Le cigliere maturano a giugno | Il meccanico aggiustò il motore | Il core cantava una canzone |
| Il reparto è arrivato alla metà | La ragazza aveva lunghi capelli | Spesso mangiamo con troppa fretta |
| Il nonno dormiva due ore | Il cameriere apparecchia il tavolo | I cittadini pagano le tasse |
| I pompieri spensero il fuoco | Le ragazze seguono la moda | Il fritto misto è pesante |

* Sentences used during the test

# Sentences used during the re-test

Table II. Noise variation index from the 1st to the 5th triplet of sentences.

| NUMBER OF CORRECT WORDS ANSWERED | NOISE VARIATION INDEX* |
|----------------------------------|------------------------|
| 0                                | -5 dB HL               |
| 1                                | -3 dB HL               |
| 2                                | -1 dB HL               |
| 3                                | +1 dB HL               |
| 4                                | +3 dB HL               |
| 5                                | +5 dB HL               |

*Noise variation index changed depending on the number of correct words answered (all employed sentences were composed of 5 words). For example, if the number of correct words answered was 0, then the noise presentation level of the next triplet was decreased of 5 dB HL.

Table III. Noise variation index from the 6th to the 20th triplet of sentences.

| NUMBER OF CORRECT WORDS ANSWERED | NOISE VARIATION INDEX* |
|----------------------------------|------------------------|
| 0                                | -2 dB HL               |
| 1                                | -1 dB HL               |
| 2                                | 0 dB HL                |
| 3                                | 0 dB HL                |
| 4                                | +1 dB HL               |
| 5                                | +2 dB HL               |

*Noise variation index changed depending on the number of correct words answered (all employed sentences were made up of 5 words). For example, if the number of correct words answered was 0, then the noise presentation level of the next triplet was decreased of 2 dB HL.
The technological setting was composed of two loudspeakers, a clinical audiometer Biomedica-Amplifon Interacoustics AC 40® and a power amplifier Interacoustics AP70®. To facilitate administration of the test, an automated computer controlled procedure was developed using an Excel calculation sheet (Fig. 1).

Statistical analysis
To determine the repeatability and consistency of SRTs measured over time, test-retest reliability was calculated, using Spearman’s Rho-correlation. To verify whether age has an influence on the overall measured mean SRT value, univariate analysis of variance (ANOVA) was performed.

Results
The mean results obtained through the test and re-test were calculated considering each employed intensity; subsequently, overall test + re-test outcomes were analysed (Table IV). The overall mean SRT value was -13.3 dB HL (standard deviation: 0.64; minimum value -14.6 dB HL; maximum value -12.1 dB HL) and was calculated at the end of both test and re-test, with respect to all SRT values for each intensity examined. The low standard deviation measured (close to 0) suggested that the mean SRT value was highly indicative of each listener’s performance. Statistical analysis showed a significant relationship between the overall mean SRT of the test and the overall mean SRT of the re-test (r = 0.496; p < 0.001), confirming the repeatability and consistency of SRTs measured over time. However, when comparing the different levels between the test and the re-test, a significant relationship between the test and the re-test was found at 65 dB HL (r = 0.366; p = 0.009) and at 75 dB HL (r = 0.362; p = 0.010), but no significant relationship was found at 55 dB HL between the test and re-test (r = 0.176; p = 0.223). The results of univariate ANOVA showed that age significantly affected overall mean SRT (F = 4.500; p = 0.039) (Fig. 2).
Discussion

Over the past decades, advances in audiological research have improved to such an extent that they have considerably and truly changed patients’ lives. Furthermore, the evolution of traditional hearing aids and the birth of total or semi-implantable hearing devices have led the clinician to re-evaluate the approach to clinical practice. The necessity to understand whether a certain patient could benefit from the current hearing technology, as well as quantitatively determining the functional improvement levels, has promoted the integration of better hearing diagnostic methods for the purpose of further exploration of the complexity of the auditory system functions. Being too far from reality, conventional investigations of hearing loss based on pure-tone threshold audiometry and on mono/di-syllabic word lists presented in quiet conditions have been shown to be inadequate. The lack of information derived from traditional test battery is typically relevant if we consider the “cocktail-party” problem, hence the loss of SNR score in patients with hearing disorders. In particular, the population repeated SRT changes in relation to different speech source positions, with a background of diffuse noise. Outcomes highlighted a lack of efficacy of bone-anchored hearing aid simulators in discriminating speech sources located in different acoustic settings. Beltrame et al. showed the auditory results of the Vibrant Soundbridge coupled with the round windows in patients with mixed hearing loss. The measurements of aided SRT in background noise of 55 and 70 dB SPL documented only a slight increase compared with normal reference data.

Among adaptive sentence-type speech-in-noise tests cited in literature, the introduction of roving-level tests may represent an additional and useful tool in speech intelligibility investigation. It is notorious that each hearing device performs differently at different sound pressure levels in relation with its own specific input dynamic range. Furthermore, in everyday speaking conditions speech and noise signals change independently of one another. A roving-level adaptive test employs target and competing noise signals change independently of one another. A roving-level adaptive test employs target and competing noise signals varying in level to mimic everyday life speaking conditions and to explore a more complete sound range. Haumann et al. recently analysed a relative homogeneous sample of 55 cochlear implant (CI) users with different CI systems submitted to fixed and roving-level adaptive level tests. When a fixed-level method was applied, the groups using different devices reached very similar outcomes. On the other hand, remarkable changes were recorded employing roving-level adaptive tests that are able to detect the effects of CI processors on everyday speech perception. In our personal experience, the roving condition concerned three different speech material intensities (55 dB HL, 65 dB HL, 75 dB HL) and as a consequence three different mean SRT values were analysed. Speech messages were randomly roved while noise level varied depending on the accuracy of the preceding response until the subject repeated 50% of the speech signal correctly. The final mean value was -13.3 dB HL (standard deviation: 0.64; minimum value -14.6 dB HL; maximum value -12.1 dB HL) and was calculated at the end of both the test and re-test, with respect to all SRT values for each intensity examined. In particular, the population repeated 50% of the presented speech material correctly, when the confounding signal was overcoming the speech message of 13.3 dB HL. It is plausible that a normal hearing Italian adult submitted to the test obtains a result ranging from -14.6 dB HL to -12.1 dB HL, although Italian listeners with hearing impairment will probably show a SRT score worse than -12.1 dB HL. As far as the literature is concerned, each of the 50 subjects evaluated achieved better results during the re-test probably due to the cognitive
gain achieved in the preceding test. Statistical analysis showed a significant relationship between the overall mean SRT of the test and the overall mean SRT of the re-test \( r = 0.496; p < 0.001 \), confirming the repeatability and consistency of SRTs measured over time. However, when comparing the different levels examined between the test and re-test, no significant relationship was found at 55 dB HL between the test and re-test \( r = 0.176; p = 0.223 \). This result may be explained by a strong learning effect (much better performances during the re-test in comparison with the test) during the presentation level of 55 dB HL. The results of the univariate ANOVA showed that age significantly affected the overall mean SRT, despite the fact that this study did not include elderly subjects (age ranged between 23 and 50 years). Younger subjects achieved meaningfully better SRTs values demonstrating age as a relevant influencing factor on the hearing processing under conditions of noise (Fig. 2). Current literature describes the development of many adaptive SNR tests and their comparison is actually challenging because of several procedural aspects that may influence outcomes:

- different language adoption with respect to a large number of speech audiological parameters: number of syllables/words, word frequency, sentence redundancy, understandable or meaningless sentences, list length, pronunciation, breaks between sentences;
- different noise presentation modes: noise masking type, continuous/ fluctuating/synchronous noise delivering;
- different environment parameters: booth dimensions, type of soundproofing, use of headphones/speakers, number of speakers, distance between patient and speaker;
- different adaptive SNR test features: starting SNR presentation level, fixed element (speech or noise?), intensity of the fixed signal, adaptive method (signal change criteria according to tester answers), re-test execution, use of visual aids;
- different sample enrollment; number of listeners, nationality, age, education level, comorbidities.

In fact, despite the rather innovative aspects of adaptive speech intelligibility tests, they still have some limits:

- the extreme procedural variability does not allow an international comparison of SRT outcomes;
- paediatric age, education level and native language may highly influence correct performance of the test;
- the need for a relatively long execution time may condition the feasibility of adaptive tests during daily clinical practice.

Conclusions

We introduced a novel Italian roving-level adaptive test to obtain normative data in healthy Italian adults. The adoption of roving-level tests may represent an additional and useful tool in speech intelligibility investigation to quantify the actual benefits related to hearing rehabilitation devices (which hearing device? when to prescribe it?) in acoustic conditions simulating everyday life.

References

1. Prosser S, Pulma M, Mancuso A, et al. Speech perception with hearing aids: effects of noise reduction and directional microphone systems on amplified signals. Audiol Med 2009;7:106-11.
2. Taylor B. Speech-in-noise tests: how and why to include them in your basic test battery. Hear J 2003;56:40-3.
3. Nilsson M, Soli SD, Sullivan JA. Development of the hearing in noise test for the measurement of speech reception thresholds in quiet and in noise. J Acoust Soc Am 1994;95:1085-99.
4. Cherry E. Some experiments on the recognition of speech with one and with two ears. J Acoust Soc Am 1953;25:975-9.
5. Walden BE, Schwartz DM, Williams DL, et al. Tests of the assumption underlying comparative hearing aid evaluations. J Sp Hear Dis 1983;48:264-73.
6. Nielsen JB, Dau T. Development of a Danish speech intelligibility test. Int J Audiol 2009;48:729-41.
7. Luts H, Boon E, Wable J, et al. FIST: a French sentence test for speech intelligibility in noise. Int J Audiol 2008;47:373-4.
8. Vainio M, Suni A, Jarveläinen H, et al. Developing a speech intelligibility test based on measuring speech reception thresholds in noise for English and Finnish. J Acoust Soc Am 2005;118:1742-50.
9. Wong LL, Soli SD. Development of the Cantonese hearing in noise test (CHINT). Ear Hear 2005;26:276-89.
10. Soli SD, Vermiglio A, Wen K, et al. Development of the Hearing In Noise Test (HINT) in new languages. In: Short course presented at the 14th Convention of the American Academy of Audiology. Philadelphia, PA 2002.
11. Hagerman B. Clinical measurements of speech reception threshold in noise. Scand Audiol 1984;13:57-63.
12. Laurence RF, Moore BC, Glasberg BR. A comparison of behind-the-ear high-fidelity linear hearing aids and two-channel compression aids in the laboratory and in everyday life. Br J Audiol 1983;17:31-48.
13. Hagerman B. Sentences for testing speech intelligibility in noise. Scand Audiol 1982;11:79-87.
14. Plomp R, Mimpren AM. Improving the reliability of testing the speech reception threshold for sentences. Audiology 1979;18:43-52.
15. Bovo R, Prosser S, Oratore RP, et al. Speech recognition with BAHA simulator in subjects with acquired unilateral sensorineural hearing loss. Acta Otolaryngol 2011;131:633-9.
16. Dietz A, Buschermöhl M, Aarnisalo AA, et al. The development and evaluation of the Finnish Matrix Sentence Test for speech intelligibility assessment. Acta Otolaryngol 2014;134:728-37.
17. Brand T, Kollmeier B. Efficient adaptive procedures for threshold and concurrent slope estimates for psycho- physics and speech intelligibility tests. J Acoust Soc Am 2002;111:2801-10.
Noordhoek IM, Houtgast T, Festen JM. Measuring the threshold for speech reception by adaptive variation of the signal bandwidth. I. Normal-hearing listeners. J Acoust Soc Am 1999;105:2895-902.

Ciorba A, Zattara S, Loroni G, et al. Quantitative enhancement of speech in noise through a wireless equipped hearing aid. Acta Otorhinolaryngol Ital 2014;34:50-3.

Haumann S, Lenarz T, Büchner A. Speech perception with cochlear implants as measured using a roving-level adaptive test method. ORL J Otorhinolaryngol Relat Spec 2010;72:312-8.

Cutugno F, Prosser S, Turrini M. Audiometria vocale – vol. III. First edition. Padova: GN. Resound 2005.

Killion MC. SNR loss: I can hear what people say, but I can’t understand them. Hear Rev 1997;4:8-14.

Moore BCJ. Perceptual consequences of cochlear hearing loss and their implication for the design of hearing aids. Ear Hear 1996;17:133-60.

Beltrame AM, Martini A, Prosser S, et al. Coupling the Vibrant Soundbridge to cochlea round window: auditory results in patients with mixed hearing loss. Otol Neurotol 2009;30:194-201.

Wagener KC, Brand T. Sentence intelligibility in noise for listeners with normal hearing and hearing impairment: Influence of measurement procedure and masking parameters. Int J Audiol 2005;44:144-56.

Received: March 4, 2016 - Accepted: July 28, 2016