European Annals of Otorhinolaryngology, Head and Neck diseases 133 (2016) 101–106

Available online at
ScienceDirect
www.sciencedirect.com

Elsevier Masson France
EM|consulte
www.em-consulte.com/en

Original article

Validation of the French-language version of the OTOSPEECH automated scoring software package for speech audiometry

F. Venail [a,b,*], E. Legris [b], B. Vaerenberg [c], J.-L. Puel [b], P.J. Govaerts [c], J.C. Ceccato [c]

[a] *Service ORL, université Montpellier, CHU Gui-de-Chauliac, 80, avenue Augustin-Fliche, 34295 Montpellier cedex, France*
[b] *Plateforme d'audiologie I-PAudioM, institut des neurosciences de Montpellier, Inserm U1051, 34091 Montpellier cedex, France*
[c] *The Eargroup, Herentalsebaan 75, 2100 Antwerp-Deurne, Belgium*

ARTICLE INFO

Keywords:
Speech audiometry
Automated scoring
Validation study

ABSTRACT

Objectives: To validate a novel speech audiometry method using customized self-voice recorded word lists with automated scoring.
Patients and methods: The self-voice effect was investigated by comparing results with prerecorded or self-recorded CVC (consonant-vowel-consonant) word lists. Then customized lists of 3-phoneme words were drawn up using the OTOSPEECH software package, and their scores were compared to those for reference lists. Finally, the customized list scores were compared on automated (Dynamic Time Warping [DTW]) versus manual scoring.
Results: Self-voice did not change scores for perception of CVC words at 10, 20 and 30 dB (ANOVA > 0.05). Scores obtained with pre-recorded and self-recorded lists correlated ($n = 10$, $R^2 = 0.76$, $P < 0.01$). Customized list scores correlated strongly with the reference cochlear lists of Lafon in normal-hearing ($n = 77$, $R^2 = 0.83$, $P < 0.001$) and hearing-impaired populations ($n = 13$, $R^2 = 0.89$, $P < 0.001$). Results on the automated and manual scoring methods correlated in both populations ($n = 77$, $R^2 = 0.71$, $P < 0.01$; and $n = 13$, $R^2 = 0.76$, $P < 0.01$, respectively), with DTW scores ranging from 24.17 to 53.24.
Conclusions: Automated scoring of customized self-voice recorded lists for speech audiometry displayed results similar to conventional audiometric techniques.
© 2016 Elsevier Masson SAS. All rights reserved.

1. Introduction

Speech audiometry is an everyday practice in audiology, not only to diagnose hearing loss but also to assess auditory rehabilitation (conventional or implanted hearing aid, cochlear implant or speech therapy).

Implementation involves the subject hearing then repeating phonemes presented at varying intensity. The examiner judges whether the phonemes have been repeated correctly, and scores the subject's performance (usually as a percentage).

Speech audiometry performance depends not only on auditory acuity and presentation intensity but also on the type of vocal material presented [1]. The stronger the semantic content of the material (e.g., real words rather than nonsense words or logatoms), the more understanding calls upon the subject's lexical field [2].

Thus speech audiometry assessment runs up against insurmountable obstacles if subjects are not being tested in their native language, if their lexical field is limited or if a strong regional accent makes it difficult for the examiner to understand the repeated phonemes [3]. Assessment may further be biased in the context of rehabilitation programs, where iterative presentation of the same word-lists induces learning bias over successive sessions, artifactually boosting the subject's scores.

The present study assessed the OTOSPEECH software package (Eargroup, Antwerp, Belgium), which uses customized word-lists drawn from the individual subject's own lexical field, with automated scoring. Briefly, word-lists are taken from the subject's everyday lexical field according to the occurrence of phonemes in his or her native language (including dialects) [4], then recorded with the subject's own voice, thereby allowing for individual acoustic and articulatory parameters and accent. They are then presented at varying intensity, just as in conventional speech audiometry. The words as repeated by the subject are then compared against the recorded words, using an algorithm to determine automatically the number of correctly repeated phonemes.

This semi-automatic speech audiometry procedure was assessed by first comparing subjects' performance when repeating logatoms recorded in their own voice or by another speaker.

* Corresponding author. Tel.: +33 4 67 33 68 00; fax: +33 4 67 33 67 28.
E-mail address: frederic.venail@inserm.fr (F. Venail).

http://dx.doi.org/10.1016/j.anorl.2016.01.001
1879-7296/© 2016 Elsevier Masson SAS. All rights reserved.

Performance was next compared between customized word-lists and calibrated lists, both self-recorded. Finally, automatic scoring was compared to against the examiner's manual scoring.

2. Materials and methods

2.1. Subjects

The study population comprised native French-speaking adults able to read easily from a computer screen placed at a comfortable distance. Educational levels were at least middle-school, and speech intelligibility (Speech Intelligibility Rating [SIR]) was excellent (SIR 1).

Pure-tone audiometric hearing thresholds were checked after ruling out external or middle ear pathology on otoscopic examination. Subjects with conduction or perceptual hearing loss were excluded. All measurements concerned acoustic stimuli in the right ear.

Two subgroups were studied: normal-hearing and hearing-impaired according to ISO 7029.

The study was considered as comprising routine care, and did not require institutional review board approval.

2.2. Material

The study was conducted at the I-PAudioM audiology platform (INSERM U1051), Montpellier (France). Testing was performed in soundproof booths, using a calibrated Affinity system (Interacoustics, Denmark). Audiometric thresholds were studied for each ear separately, using TDH-39 headphones, and conduction hearing loss was ruled out by bone conduction measurements using a B-71 vibrator.

The computer linked to the Affinity system was connected up to a USB sound card (Aureon 7.1) and microphone to record the subject's voice. Visualization equipment comprised two screens: one in the booth, visible to the subject, and one outside, visible to the examiner.

The OTOSPEECH software is part of the A§E suite (Otoconsult, Antwerp, Belgium) and was installed on the computer connected up to the measurement system; the output intensity of the word-lists emitted by the program was calibrated using a sonometer.

2.3. Voice recording and results analysis on OTOSPEECH software

The OTOSPEECH software enabled use of prerecorded word-lists pronounced by a speaker or self-recorded lists read by the subject from a text. Customized lists of 3-phoneme CVC (consonant-vowel-consonant) words were taken by the software from a text of more than 300 words chosen by the subject in a field familiar to him or her. The words were then recorded by the subject, after the software had checked the audio quality of the recording set-up. Intensity of each word in the lists was calibrated automatically by the software. The words were then saved for future use. In this way, subjects created their own speech audiometry word-list, consisting of familiar vocabulary and recorded by their own voice.

In the test phase, OTOSPEECH compared each word as repeated by the subject to the source-word recorded by the subject, on acoustic comparison using a dynamic time warping (DTW) algorithm [5]. Unlike other speech-recognition methods, DTW requires no phoneme data-base for comparison. The algorithm extracts indices (Mel-Frequency Cepstral Coefficients [MFCC] [6,7]) and, after normalization, compares the Euclidean distance between them, in the form of vectors calculated for the target signal (recording) and test signal (repetition): the shorter the distance, the closer the target and test signals: i.e., the more faithful the repetition. This type of analysis limits intra-subject variability and requires no learning phase. Implementation of DTW distance was validated in Dutch in a previous study [8]; the DTW score, after polynomial transformation to express it as a value between 0 and 100%, showed the same psychometric properties as a speech audiometry curve using the Bruges CVC logatom test (see [8] for details).

In the light of these results, we did not perform the transformation but, in subsequent analyses, compared DTW scores and correct repetition scores after manual validation.

In parallel, the self-voice effect was studied using CVC logatoms in a subgroup of normal-hearing subjects. Word-lists were recorded using Adobe Audition CS5 software (Adobe Corp., San Jose, CA, USA) and intensity was normalized using Matlab software (Mathworks, Natick, MA, USA).

2.4. Assessment of the self-voice effect in speech audiometry

The impact of using the subject's own voice in speech audiometry was assessed in 2 sessions, in part of the normal-hearing population:

- in the first session, 4 prerecorded (neutral male voice) lists of 17 CVC Dodelé logatoms each [9] were delivered at 5, 10, 20 and 30 dB HL and correct repetition was scored manually;
- at end of session, the 4 lists were recorded by the subject and saved for session 2;
- in the second session, 2 weeks after self-voice recording, the lists were delivered at the same intensities as in session 1 (5, 10, 20 and 30 dB HL) and scored manually.

2.5. OTOSPEECH validation

The OTOSPEECH software was assessed in 2 sessions.

In the first session, 2 Lafon cochlear lists of 17 3-phoneme words each [10] were recorded by the subject and saved to the software. Also, 2 customized lists of 24 3-phoneme words each were put together from a text familiar to the subject.

The second session was held at least 2 weeks later, to limit memory bias. Speech audiometry was performed using the 2 Lafon lists and 2 customized lists, all self-recorded. Lists were presented at 2 intensity levels: one judged week and the other comfortable by the subject; levels thus varied between subjects. Instead of reporting the speech perception scores for fixed intensities, we rather reported the results for two subjective intensities (low and high), chosen by the test subject according to his or her hearing threshold, to be able to show the variations in scores between subjects. The number of correctly repeated phonemes was scored manually for both types of list, and DTW scores were calculated automatically for the customized lists only.

2.6. Statistical analysis

For logatoms, raw values were compared on 2-factor ANOVA ($P < 0.05$ significance threshold) to discern effects of intensity (5, 10, 20 and 30 dB HL) and list type (pre- versus self-recorded).

Correlations between manually and automatically-scored Lafon and customized lists (percentage correctly repeated phonemes) were assessed on Pearson correlation test ($P < 0.05$ significance threshold).

Finally, the correlation between manual scoring (percentage correctly repeated phonemes) and automated scoring of customized lists (DTW score) was assessed on Pearson correlation test ($P < 0.05$ significance threshold).

Download English Version:

https://daneshyari.com/en/article/4109890

Download Persian Version:

https://daneshyari.com/article/4109890

Daneshyari.com