Speech-in-Noise Audiometry in Adults: A Review of the Available Tests for French Speakers

Pierre Reynard^{a, b, c} Josée Lagacé^d Charles-Alexandre Joly^{a, c} Léon Dodelé^e
Evelyne Veuillet^{a, b, c} Hung Thai-Van^{a, b, c}

^aHospices Civils de Lyon, Hôpital Edouard Herriot, Service d'Audiologie & Explorations Otoneurologiques, Lyon, France; ^bUniversité Lyon 1, Villeurbanne, France; ^cInstitut de l'Audition, Institut Pasteur, Paris, France; ^dAudiology and Speech-Language Pathology Program, Faculty of Health Sciences, University of Ottawa, Ottawa, ON, Canada; ^eCentre Auditif Dodelé, Braine-l'Alleud, Belgium

Keywords

Speech understanding in noise · Speech-in-noise tests ·
Speech recognition · Francophone

Abstract

Background: Difficulty understanding speech in background noise is the reason of consultation for most people who seek help for their hearing. With the increased use of speech-in-noise (SpIN) testing, audiologists and otologists are expected to evidence disabilities in a greater number of patients with sensorineural hearing loss. The purpose of this study is to list validated available SpIN tests for the French-speaking population. **Summary:** A review was conducted according to Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines. PubMed and Scopus databases were searched. Search strategies used a combination of 4 keywords: speech, audiometry, noise, and French. There were 10 validated SpIN tests dedicated to the Francophone adult population at the time of the review. Some tests use digits triplets as speech stimuli and were originally designed for hearing screening. The others were given a broader range of indications covering diagnostic or research purposes, determination of functional capacities and fitness for

duty, as well as assessment of hearing amplification benefit.

Key Messages: As there is a SpIN test for almost any type of clinical or rehabilitation needs, both the accuracy and duration should be considered for choosing one or the other. In an effort to meet the needs of a rapidly aging population, fast adaptive procedures can be favored to screen large groups in order to limit the risk of ignoring the early signs of forthcoming presbycusis and to provide appropriate audiological counseling.

© 2021 S. Karger AG, Basel

Introduction

Understanding speech in the presence of background noise is often difficult, even for listeners with normal hearing sensitivity. Although it is particularly more challenging to recognize SpIN for most individuals with hearing loss, some clinical populations with normal hearing thresholds have greater difficulties in such acoustic conditions [Smoorenburg, 1992]. Speech understanding difficulties in noise are reported as the most common reason for consultations in audiology or hearing health services [McArdle et al., 2005]. However, most clinical hearing

test batteries typically only include pure-tone audiometry and speech audiometry in quiet condition. The limits of this assessment strategy have been discussed for decades as they generally underestimate speech difficulties in noise [Carhart and Tillman, 1970; Plomp and Mimpen, 1979; Smoorenburg, 1992].

On the other hand, hearing health professionals have access to an ever-increasing number of speech-in-noise (SpIN) tests for diagnostic, functional evaluation, and/or screening purposes. The speech material varies widely, from digits to non-sense sentences, as well as the type of masking involved (from energetic to informational masking) and test procedures (e.g., adaptive procedure or fixed level) [Theunissen et al., 2009].

The use of SpIN tests allows many advantages for hearing-aid adjustment (device selection and fitting [Taylor, 2003]). Furthermore, the interest of the tests has been shown in the early evaluation of presbycusis [British Society of Audiology, 2019] and also for the diagnosis of hidden hearing loss or auditory neuropathy spectrum disorder (ANSO) [Moore et al., 2014; Iliadou et al., 2017].

On top of time constraints often reported as a reason for not evaluating speech recognition abilities in noise [Wilson and McArdle, 2005], hearing health professionals may not always have a clear idea of which SpIN test to use. The objective of this study was to provide a review of the available and validated SpIN tests for the adult French speaking population. As there are about 300 million speakers of French in the world, which represents about 4% of the total planet population [Organisation Internationale de la Francophonie, 2018], the relevance of such work appears incontestable. Having a precise idea of the available SpIN tests for the French speaking population may provide some guidance for the clinical practices and the future efforts in test development for that specific population.

Main characteristics of the SpIN tests are listed below. Having a better understanding of the impact of these characteristics on the performance may help the clinicians and researchers to select the appropriate test for the need or goal and permits a more accurate interpretation of these test results.

Materials and Methods

Relevant current publications in peer-reviewed journals were sourced through a computerized literature search. The review was conducted according to Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines (PRISMA, see Moher et al. [2009]).

Searches were undertaken in PubMed (NLM NIH) and Scopus (Elsevier) databases. The indicators “speech,” “audiometry,” and “noise” were entered as search terms for all years up until 2020 in both databases. Reports from 1957 to 2020 were sourced. Non-French or non-English reports, letters, and conference papers were excluded. Pediatric reports were excluded. Full-text articles were then retrieved for all remaining studies. Reports were only included if they detailed the SpIN test, its target population, and its procedure, and if there was indeed an adult version adapted to the French population.

Test Characteristics

When looking at the available SpIN tests, there are different characteristics that need to be considered in order to select the appropriate one for the intended application. The following paragraphs provide a description of these SpIN test characteristics, including their advantages and limitations.

Testing Procedure

Two test strategies are employed for the measures of speech perception in noise abilities. Fixed-level presentation procedures consist of keeping constant the intensity level of both the speech items and the competing noise during the test. In general, the final score obtained is the number of correctly identified speech items, often expressed as a percentage. As speech items stimuli are presented at supra-threshold intensity level, the listening conditions of SpIN tests using fixed-level procedures can be more representative of the ones found in real life, when appropriate levels and signal-to-noise ratios are selected.

For adaptive procedures, the goal is to measure a detection or recognition threshold over the shortest amount of time, without sacrificing accuracy [Leek, 2001]. Compared to the fixed-level procedure, the adaptive procedure does not require as many test items for obtaining the final score and helps avoiding ceiling effects [Theunissen et al., 2009].

Some SpIN tests are being presented bilaterally, while others are being presented unilaterally. The results obtained with this latter test condition represent “ear-specific” functions, which might present some advantages over the bilateral testing conditions for diagnostic purposes. On the other hand, results obtained with tests using bilateral presentation may be more representative of the functional listening abilities.

Speech Stimuli

Compared to monosyllabic words, the score obtained with more realistic speech stimuli such as sentences may be more representative of the listener’s ability in real-life situations; scores obtained with SpIN tests using sentences as speech stimuli may be influenced by non-auditory abilities such as linguistic and cognitive functions; sentences are also less sensitive to the hearing deficit because of the redundancy of the testing material [McArdle et al., 2005; Wilson and McArdle, 2005]. While some sentence-based tests offer sentences with the same syntactic structure and generated with a limited number of items, other tests offer unrelated sentences of varying length and syntactic structure.

The familiarity of the corpus items also has to be examined when considering any SpIN tests. In comparable acoustic conditions for example, speech intelligibility is higher for familiar than for unfamiliar words [Epstein et al., 1968]. Close-set corpus (reducing the impact of the obsolescence or familiarity of the speech

stimuli on the SpIN test scores) has been used for a long time but is not representative of daily-life speech and phoneme distribution [Clopper et al., 2006]. It is well established that using pre-recorded material is preferable over live voice for the presentation of the speech items when administering SpIN tests [Stach et al., 1995; British Society of Audiology, 2019] and most are available on CD or via internet.

Types of Masking Noise

Energetic and informational masking are essentially the 2 ways that a background noise can affect speech intelligibility. Informational masking occurs when the signal and masker are both audible, but the listener is unable to disentangle the elements of the target signal from a similar-sounding distracter [Pollack, 1975; Brungart, 2001]. One single talker babble or multitalker babble noises are used to create informational masking. The latter include 2 or more talkers and are produced by electronically mixing individual recording of talkers while reading [Kalikow et al., 1977], for example. The number of talkers included in a babble noise influence the magnitude of the masking effect [Freyman et al., 2004; Simpson and Cooke, 2005], but this effect does not follow a monotonical trend. Rather, tests including single or multitalker babble noise are presumed to involve informational masking and are believed to provide insight on the higher order auditory processes [Brungart, 2001].

Energetic masking occurs when both the signal and the background noise contain energy in the same critical frequency bands at the same time, and portions of the speech signals are difficult, or impossible, to recognize [Brungart, 2001]. SpIN tests including speech spectrum or white noise as a background, are assumed to involve energetic masking.

Masking noises can be categorized as stationary or fluctuant. Stationary noises have a constant intensity over time, while the intensity of fluctuant noises changes over time. Stationary noises may not be representative of the general daily life background noises [Wagener and Brand, 2005]. However, for reasons of ease of calibration and test processing, the majority of the available SpIN tests includes a stationary noise.

Performance Variables

Another important aspect to consider when selecting a SpIN test relates to its performance variables, which are influenced by its reliability, validity, sensitivity, and specificity [Theunissen et al., 2009]. Most tests have been validated for a specific clinical population and for a specific clinical purpose, for example, for hearing screening [Smits et al., 2004] or for the determination of speech perception problems in noise [Killion et al., 2004]. When administering a SpIN test, it is important to be sure that it was designed for the population for whom it is being used with.

Availability of normative data is also an important aspect to consider for the selection of a SpIN tests. Normative data allows comparison of an individual test results to that of the normal hearing population. Subject variables are important in the identification of factors that could explain results that deviate from the documented norms [Theunissen et al., 2009].

By having a better understanding of the impact of the test characteristics on the performance, it is possible to customize the hearing evaluation with respect to the needs (e.g., diagnostic and fitness for duty assessment) and personal factors (e.g., attention span and linguistic abilities) of the listener.

Results

The systematic literature search produced 1,013 unique articles. Screening by title and abstract eliminated 796 articles (Non-French or non-English reports, letters, and conference papers, pediatric reports), leaving 37 studies based on the aim of the study with SpIN tests for full-text review (Fig. 1). Ten different SpIN tests were identified, all of which provided a French version and were suitable for adult patients.

Test Validation Procedure

Validation of any SpIN tests requires different phases. First, the speech signal's items and the competitive noise must be created before being selected at the development phase. Then, during the optimization phase, the acoustic energy must be equalized across items that should be sorted in lists with the same difficulty (phonological difficulty or level of difficulty per se). Finally, during the validation phase, the normative data are defined for different testing procedures, and/or different listening conditions and/or types of different hearing loss statuses. Both training effect and test-retest variability are evaluated during this last phase.

Normative data can be either the scores obtained by normal hearing (NH) subjects for a fixed signal-to-noise ratio (SNR), or the characteristics of the psychometric function representing changes in speech recognition with the SNR. According to the ISO 7029:2017 norm (<https://www.iso.org/standard/42916.html>), a normal hearing person is a subject free from all signs or symptoms of ear disease and from obstructing wax in the ear canal and who has no history of undue exposure to noise, exposure to potentially ototoxic substances or familial hearing loss. In the tests reviewed here, normative values were obtained for NH subjects aged from 16 to 57 years for adults, from 5 to 10 years for the French Simplified Matrix Test (Frasimat) children's group. While classification of hearing loss may slightly differ between countries, the pure tone average between 0.25 and 8 kHz of NH subjects never exceeded 25 dB HL in the selected studies. The target characteristics of the psychometric function are the SNR level that allows a subject to correctly identify 50% of the signal (50% speech reception threshold = SRT50), and the slope of the function.

SpIN Tests Available in French

The following paragraphs provide a description of 10 SpIN tests that can be used with French speaking adult population. It was felt important to only provide detailed

information about the SpIN tests that are available to hearing health professionals and researchers (free of charge, or by purchasing) and in a recorded version. For many obvious reasons, the use of recorded stimuli has been recommended for sometimes, and it was felt important to encourage best practice by providing the details of recorded tests only. Availability of scientific report or literature about the development and the performance of the test was another condition for the inclusion in the following description. SpIN test normative values (with fluctuation interval, when available) in French NH speakers, both for tests with adaptive and fixed-level procedure are reported in Figure 2.

Tests Using Digits

The Digit Triplet Test (DTT) was originally developed for hearing screening and in the Dutch language [Smits et al., 2004]. During this test, digit triplets (e.g., 3-2-5) are presented simultaneously with a competing noise. The listener has to enter the digit sequence heard on a numeric keypad or touch screen. The DTT uses an adaptive scoring procedure to compute a SRT which corresponds

to the SNR level at which the triplets are recognized 50% of the time (SRT50). A strong relationship between digit triplet scores and pure-tone thresholds has been reported [Smits et al., 2004; Watson et al., 2012]. The DTT has been adapted in several languages [Zokoll et al., 2012] within the European Union framework of the HearCom project (<http://www.hearcom.org>), including a European French version. The International Collegium of Rehabilitative Audiology (ICRA) [Akeroyd et al., 2015] serves as a blueprint for the construction and implementation of the test across languages [Zokoll et al., 2012]. The ICRA's blueprint promotes a high level of consistency in test properties (e.g., efficiency, reproducibility, and steepness of psychometric function) that optimizes comparability across different versions (Table 1).

French Digit Triplets Test

The French Digit Triplets Test (FrDigit3) [Jansen et al., 2010] was developed following the model of the original version of the DTT. It includes 10 lists of 27 triplets, uttered by a female speaker of European French. In each list, each number is presented 3 times per position in the

Fig. 2. SpIN test normative data in French-speaking adults. Normative values (with fluctuation interval, when available) are the SRT50 values for adaptive procedures (**a**) and intelligibility scores (%) for fixed-level procedures (**b**). **b** FrBio. Yellow (bottom) multitalker noise. Blue (top) speech noise. **c** SUN test: scoring procedure. SpIN, speech-in-noise; SUN, speech understanding in noise; HINT, hearing in noise test.

Table 1. Available words/digits in noise tests

European context (France)

FrDigit3
[Jansen et al., 2010]
Test procedure: adaptive. Closed set
Monaural condition: phone version
Binaural condition: internet version
Stimuli: digit triplets. Female adult speaker of European French (Paris)
Noise: LTAS of the test digits material
Access to online version: <http://testauditif.be>

North American context (Canada)

CDTT
[Giguère et al., 2020]
Test procedure: adaptive, binaural presentation. Closed set
Stimuli: digit triplets. Male/female adult speaker of Canadian French and English
Noise: speech-shaped noise matching the LTAS of each digit material (French female, French male, English female, English male)
For availability: cgiguere@uottawa.ca

LTAS, long-term average spectrum; CDTT, Canadian version of the Digit Triplet Test; FrDigit3, French Digit Triplets Test; LTAS, long-term average spectrum.

triplet. The normative values were developed from normo-hearing subjects ($n = 6$) between the ages of 16 and 49. The quasi-stationary noise that was created for the FrDigit3 includes the same frequency spectrum as the

long-term average of the digits material. Each triplet is preceded by a carrier phrase “les numéros” (meaning: the numbers) and is pronounced with a natural intonation and without co-articulation between each digit.

Table 2. Psychometric data

Test	Population	Subjects, <i>n</i>	Training effect	Test-retest variability	Normative data									
FrDigit3	Belgian and French age range = (16–49) PTA 250-8k < 20 dB HL	Development and optimization: <i>n</i> = 20 NH Normative data: <i>n</i> = 30 NH Test-retest/learning: <i>n</i> = 24 NH + HI	From 1st to 2nd evaluation: 0.9 dB SNR (NH + HI)	0.7 dB SNR (NH + HI)	SRT50 = −5.4 dB SNR; SD = 0.7 (adaptive procedure)									
CDTT	Canada (French and English speakers) age range = (18–30) PTA 250-8k < 15 dB HL	Development and optimization: <i>n</i> = 48 NH (24 French) List equivalence, normative data and test-retest variability: <i>n</i> = 64 NH (32 French) Test-retest/learning: <i>n</i> = 64 NH (32 French)	Not evaluated	0.51 dB SNR (French male voice) 0.62 dB (French female voice)	SRT50 = −12.66 dB SNR; SD = 0.62 (French male voice, adaptive procedure) SRT50 = −12.45 dB SNR; SD = 0.62 (French female voice, adaptive procedure)									
Framatrix	Belgian age range = (20–54) PTA 250-8k ≤ 20 dB HL	Optimization: <i>n</i> = 27 NH Validation: <i>n</i> = 30 NH	From 1st to 2nd list (word scoring): 1.2 dB SNR	0.4 dB SNR (word scoring)	SRT50 = −6 dB SNR; SD = 0.6 dB SNR Slope = 14%/dB SNR; SD = 1.6 (word scoring)									
Frasimat	Unknown nationality adults: age range = (19–57) Children: age range = (5–10) PTA 125-8k < 15 dB HL	Validation: <i>n</i> = 80 NH = 20 adults + 20 children (5–6 yr) + 20 children (7–8 yr) + 20 children (9–10 yr)	Not evaluated	Not evaluated	5–6 yr: SRT50 = −4.2 dB SNR; SD = 1.3 7–8 yr: SRT50 = −4.6 dB SNR; SD = 1.2 dB 9–10 yr: SRT50 = −5.5 dB SNR; SD = 1.5 dB (adaptive procedure)									
Canadian French HINT	French-Canadian age range = (18–45) PTA 250-8k < 15 dB HL	Development: <i>n</i> = 9 NH Optimization: <i>n</i> = 50 NH Validation: <i>n</i> = 36 NH	From 1st to 2nd trial Silence = 1.2 dB SNR Noise = 1.1 dB SNR Noise and composite = 0 dB SNR	Silence SRT50 = 16.4 dB A; SD = 3.8 Front noise SRT50 = −3.0 dB SNR; SD = 1.1 Side noise SRT50 = −11.4 dB SNR; SD = 1.3 Composite score SRT50 = −7.2 dB SNR; SD = 0.8										
FIST	Belgian and French age unknown PTA 125-8k < 25 dB HL (NH)	Development: <i>n</i> = 433 NH + HI (French) Optimization: <i>n</i> = 12 NH (Belgian) Validation: <i>n</i> = 20 NH (10 Belgian + 10 French)	Not evaluated	Inter-list variability: Belgian = 1.1 dB SNR French = 1.0 dB SNR	SRT50 = −7.4 dB SNR; SD = 0.7/slope = 20.2%/dB; SE = 2.6 (Belgian: SRT50 = −7.1 dB SNR; SD = 0.7/slope = 19.2%/dB; SE = 2.7) (French: SRT50 = −7.8 dB SNR; SD = 0.6/slope = 21.2%/dB; SE = 2.5) (Adaptive procedure)									
FrBIO	French-Canadian age range = (18–30) (for the validation group) PTA unknown	Development: <i>n</i> = 20 (4 talkers + 16 NH for sentence intelligibility evaluation) Optimization: <i>n</i> = 30 NH + 25 HI Validation: <i>n</i> = 30 NH	Not evaluated	One-month interval (<i>n</i> = 15) <1.35%, not significant	<table><tr><th>Silence</th><th>Speech noise, dB SNR</th><th>Multitalker noise, dB SNR</th></tr><tr><td>+10</td><td>+5</td><td>0</td></tr><tr><td>+10</td><td>+5</td><td>0</td></tr></table> Mean, % SD SRT50 = 0.08 dB SNR; SD = 0.6 Slope = 14%/dB; SD = 1.6	Silence	Speech noise, dB SNR	Multitalker noise, dB SNR	+10	+5	0	+10	+5	0
Silence	Speech noise, dB SNR	Multitalker noise, dB SNR												
+10	+5	0												
+10	+5	0												
VRB	French NH age range = (19–39) PTA 250-8k ≤ 20 dB HL HI age range = (40–90) 20 dB HL < PTA 250-8k < 80 dB HL	Development: <i>n</i> = 10 NH Optimization: <i>n</i> = 12 NH Validation: <i>n</i> = 6 NH + 29 HI	Not evaluated	Not evaluated										
PTA, pure-tone average; NH, normal hearing; SNR, signal-to-noise ratio; CDTT, Canadian version of the digit triplet test; FrDigit3, French digit triplets test; Frasimat, French Simplified Matrix Test; HINT, Hearing-in-Noise Test; FIST, French Intelligibility Sentence Test; FrBIO, French version of the AzBio; VRB, Vocale Rapide dans le Bruit.														

PTA, pure-tone average; NH, normal hearing; SNR, signal-to-noise ratio; CDTT, Canadian version of the digit triplet test; FrDigit3, French digit triplets test; Frasimat, French Simplified Matrix Test; HINT, Hearing-in-Noise test; FIST, French Intelligibility Sentence Test; FrBIO, French version of the AzBio; VRB, Vocale Rapide dans le Bruit.

The FrDigit3 was implemented as an automatic self-screening test on home telephone. Forty ears from 24 subjects were tested, of which 32 were tested with retest. A test-retest variability of 0.7 dB was found and the correlation between SRT and pure-tone average (0.5, 1, 2, and 4 kHz) was 0.77 [Jansen et al., 2010]. The procedure is now available online, and on digital tablets which can be used with broadband headphones. Details about how to get access to the FrDigit3 are included in Table 1. For this test, SRT50 is -5.4 dB SNR (SD = 0.7) (Table 2; Fig. 2).

Canadian Version of the DTT

The Canadian version of the DTT (CDTT) was recorded in both official languages of the country (English and French), with 2 fluently bilingual adult talkers (1 male, 1 female) [Giguère et al., 2020]. As for the other language versions of the DTT, the recordings were processed and optimized, according to the ICRA recommendations. Normative data were obtained from 64 adult patients. A speech-shaped noise was developed for each language-talker combination. There are 4 lists of 24 triplets for each language (French and English) and each talker (female and male), and data were obtained with normal hearing adults for each version.

The administration of the test requires limited hardware, that is, a pair of earphones, an external keypad, and external sound card. The CDTT can be presented with a computer or a laptop. At the present time, the CDTT is only available for research purposes and the details about how to obtain a copy of the test are included in Table 1. For this test, SRT50 is -12.66 dB SNR (SD = 0.62) with a French male voice, and -12.45 dB SNR (SD = 0.62) with a French female voice (Table 2; Fig. 2).

Tests Using Words

Test de Mots dans le Bruit. The Test de Mots dans le Bruit (TMB) is a word in noise test using a fixed-level procedure available for speakers of Canadian French [Lagacé, 2010; Lagacé et al., 2013]. The TMB includes 4 lists of 35 monosyllabic words which are presented simultaneously with a speech babble of speakers of French (4 females and 4 males). The familiarity of the words has been verified and the recordings were produced by a female speaker of Canadian French. The TMB is available on split track compact disk; the words stimuli and the speech babble are on 2 different channels; they are presented ipsilaterally at a fixed level of SNR of +5 dB. The total number of correct responses for each ear is then converted in percentage and compared to the age-appropriate normative data (normal hearing adult speakers of Ca-

nadian French) [Lagacé et al., 2010]. The TMB is available for clinical and research purposes.

Tests Using Sentences

The idea of using a closed-set paradigm for the design of sentences in noise tests was first proposed by Hagerman [1982]. Performances measured with sentences in noise tests are then presumably less influenced by the listener's linguistic knowledge and ability.

The original version of the matrix sentence tests consists of syntactically fixed but semantically unpredictable sentences [Hagerman, 1982] and was developed in the Swedish language. It was generated from words randomly taken from an inventory of 50 items (10 alternatives per word group, i.e., name, verb, numeral, adjective, and object). The fact that the matrix sentences tests are based on limited word material implies that the whole list of items may be learnt. However, the possible permutations lead to a large corpus of quite unpredictable sentences, which gives to this test the major advantage of allowing various trials in the same subject with no repetition [Jansen et al., 2012].

Although it is possible to administer the sentences using a fixed-level procedure, the most popular version of the matrix sentence tests uses an adaptive procedure and word scoring [Hagerman, 1982; Zokoll et al., 2013] to determine SRT in noise. One SRT measurement (i.e., 1 list of 20 sentences) takes approximately 2 min [Zokoll et al., 2013]. The matrix sentences test has been adapted for use in almost 20 languages (each one with its own normative values), including a European French version (see below). All versions were developed according to the HearCom minimum quality standards for speech intelligibility tests [HearCom, 2006] (Table 3).

French Matrix Test

The FrMatrix [Jansen et al., 2012] is the French version of the matrix sentences test. It includes 28 lists of 10 sentences of 5 words generated from 50 words (10 nouns + 10 verbs + 10 numbers + 10 objects + 10 colors), recorded with a female speaker of European French (Paris). The normative values were collected on normal hearing patients ($n = 20$) aged 20–29 years. The FrMatrix includes a stationary long-term average speech spectrum noise, which was generated by superimposing the 280 sentences several times [Jansen et al., 2012]. The FrMatrix can be presented via headphones or loudspeakers. The procedure is adaptive with the noise being set at 65 dB SPL. Initially, the signal (speech) level is set at 65 dB SPL (SNR = 0 dB). The SNR is adjusted after each sen-

Table 3. Available sentences in noise tests

<i>European context (France)</i>	
FrMatrix [Jansen et al., 2012] Frasimat [Prang et al., 2020]	Adaptive procedure. Closed set Stimuli: syntactically fixed (name, verb, numeral, adjective, object) unpredictable sentences. Female French (Paris) spk Noise: stationary LTAS noise generated by the superimposition of the sentence corpus Availability: www.auritec.de/fr/logiciels-et-audiometrie-vocale/framatrix.html
FIST [Luts et al., 2008]	Adaptive procedure. Open set Stimuli: 14 lists of 10 sentences. Recorded with a female French (Paris) speaker Stationary LTAS noise generated by the superimposition of the sentence corpus Female French (Paris) speaker
VRB [Leclercq et al., 2018]	Adaptive procedure. Open set Stimuli: 120 sentences, 3 keywords each. Recorded with female spk of European French Noise: 4-talker babble (superimposition of 2 dyadic conversations: English and French) Availability: biotone.fr
<i>International French (Canada-France)</i>	
International FrBio test [Bergeron et al., 2019]	Fixed-level procedure. Open set Stimuli: 30 lists of 20 sentences. Recorded by 4 spk of French, composite' accent Two types of noise: Universal LTASS Four-talker babble noise (superimposition of 8 sentences from the corpus, 2 sentences per spk) Available on CD format. See details at: https://audiotests.wixsite.com/audiotests
<i>North American context (Canada)</i>	
HINT Canadian version [Vaillancourt et al., 2008]	Adaptive procedure. Open set Stimuli: male adult speaker of Canadian French Noise: speech-shaped noise matching the LTAS of the sentence material At the moment of writing: commercial access is restricted
<p>LTASS, long-term average speech spectrum; Spk, speaker; FIST, French Intelligibility Sentence Test; HINT, Hearing-in-Noise Test; FrBio, French version of the AzBio test; VRB, Vocale Rapide dans le bruit; FrMatrix, French version of the Matrix test; Frasimat, French Simplified Matrix Test.</p>	

tence by taking into account the number of words correctly repeated and the number of inversions. The software automatically determines the SRT50 at the end of the list. The realization of 2 training lists is necessary before the test to limit the risk of bias (training effect). Re-test showed a rather low inter-subject variability (0.4 dB) from the 3rd list. When administered unilaterally in headphones, the FrMatrix can be used to diagnose a loss of intelligibility in noise when a subject's SRT50 is not within the normal range. It can also be used to measure prosthetic gain or to measure the contribution of binaural hearing. Details about how to get the test material are provided in Table 3. For this test, SRT50 is -6 dB SNR (SD = 0.6) (Table 2; Fig. 2).

Recently, a Fra-simat version containing 3 words sentences (numeral-adjective-object) was created in order to reduce the patient's cognitive load during measurements

[Prang et al., 2020]. This simplified version is also used to cope with the specific constraints of testing either the pediatric population or cognitively impaired individuals. For this test, SRT50 are provided for children: SRT50 is -4.2 dB SNR (SD = 1.3), -4.6 dB SNR (SD = 1.2), and -5.5 dB SNR (SD = 1.5) for children aged 5–6 years, 7–8 years, and 9–10 years, respectively.

The remaining tests that are available in the French language are in the category of what is called open set, and an example of such test is the Hearing-in-Noise Test (HINT) [Nilsson et al., 1994]. The HINT was originally developed in the American English language to measure sentence recognition thresholds in quiet and in background noise, and to determine the advantage of binaural directional hearing. The original version of the HINT includes 12 equivalent lists of 20 sentences that have been normed for naturalness, difficulty, and reliability. The

sentence lists were intended for adaptive measurement of SRT in quiet and in spectrally matched noise.

The standard HINT protocol uses 4-dB step sizes for the 1st 4 sentences and 2-dB step sizes for the remaining 16 sentences [Soli and Wong, 2008]. The level of the noise is kept at 65 dBA for the 3 noise conditions (noise front, noise right, and left). The listener is asked to repeat each sentence that is heard. The SRT is defined as the average presentation level of sentences 5 through 20 plus the presentation level for a 21st sentence determined from the response to the 20th sentence. Since its development, the HINT has been used in many studies, such as to measure the benefit of amplification [Pumford et al., 2000; Ricketts and Henry, 2002; Nordrum et al., 2006], cochlear implantation [Dorman et al., 2002; Eisenberg et al., 2004], and the use of personal FM systems [Johnston et al., 2009].

Canadian Version of the HINT

The HINT has been adapted in several languages [Soli and Wong, 2008] including Canadian French [Vaillancourt et al., 2008]. The Canadian adult version of the HINT includes 12 equivalent lists of 20 sentences that were recorded at the House Ear Institute by a male professional actor and speaker of Canadian French [Vaillancourt et al., 2005]. The long-term average spectrum (LTAS) of the sentences was calculated and used to specify the frequency response for a linear phase finite impulse response filter. Then, a white noise was synthesized, filtered with the created finite impulse response filter, and scaled to the same A-weighted root mean square average amplitude of the sentence set [Vaillancourt et al., 2005]. The resulting speech-spectrum noise, matched to the actual speaker, ensures that the SNR is kept approximately the same (on average) at all frequencies; the same procedure as in the original version of the HINT [Soli and Wong, 2008] was used. The normative values were distributed over normo-hearing Quebec subjects ($n = 36$) between the ages of 18 and 45 years. The procedure is adaptive for which the noise is set at 65 dBA. Initially the signal level is set at 60 dBA. The speech level decreases when the subject repeats the sentence correctly or increases when the sentence is not repeated correctly (4 dB steps). At the end, the SRT50 is calculated over 16 measures. The test can be carried out in binaural diotic, dichotic condition with headphones or in open field (2 HP required). For this test, SRT50 is -3.0 dB SNR ($SD = 1.1$) for the front noise condition and SRT50 is -11.4 dB SNR ($SD = 1.3$) for the side noise conditions (Table 2; Fig. 2).

The HINT is widely used for research purposes and included in test batteries in centers or clinics where avail-

able. However, at the time of writing, the original version and the many adapted language versions of the HINT test are unfortunately hardly available commercially.

French Intelligibility Sentence Test

The French Intelligibility Sentence Test (FIST) [Luts et al., 2008] was developed, using the recommended procedure for the development of the HINT [Soli and Wong, 2008]. Unlike other versions of the HINT, the sentences were equalized without modifying the sentence level [Luts et al., 2008]. The FIST includes 14 test lists of 10 sentences that are equivalent when presented in noise with regard to SRT and the slope of the performance-intensity function. The sentence material has been evaluated with 2 group of 10 normo-hearing adults (native French from France and Belgium), using the same adaptive test procedure and the stationary speech weighted noise that has the same LTAS as the sentences at a fixed level of 65 dB SPL. A small, but significant difference of 0.7 dB between the SRT measured with the group of French and Belgian listeners has been reported [Luts et al., 2008]. The authors explained this small difference in favor of the French listeners by the fact that the recording of the FIST sentences was made with a female speaker from Paris, who may have a different accent than the one found in Belgium. For the FIST, $SRT50 = -7.4$ dB SNR ($SD = 0.7$) (Table 2; Fig. 2).

The AzBio sentence test [Spahr et al., 2012] is another example of an open set SpIN test that has been adapted in French language. The original version was developed in American-English language for the assessment of improvement in auditory perception in adults, following any type of therapeutic interventions, but initially cochlear implantation. The AzBio test includes 15 lists of 20 sentences with similar levels of difficulty, which allows within-subject comparisons. The recordings were made with 4 different speakers, 2 females and 2 males, and they can be presented in quiet or in a 10-talker babble noise at fixed SNRs. Because of its large corpus (15 lists of 20 sentences), the AzBio can be used to evaluate changes in performance over time or across conditions, and even with listeners with extensive exposure to sentence test materials.

French Version of the AzBio (FrBio)

An international French version of the AzBio (FrBio) was developed in a collective effort between collaborators from France and Canada [Bergeron et al., 2019] and the unique characteristic of the test lies on its expected international use. A consensus on the familiarity of the lexicon

and expressions from the collaborators of both countries was obtained for the final corpus of 660 sentences, which were recorded by 4 speakers of French, with a “composite accent.” These speakers were either natives from France and lived in Canada for many years, or native Canadians educated in a European French environment. The equivalency of the 30 lists of 20 sentences of the final FrBio was confirmed with 30 adults with normal hearing from Canada and France, and 25 with adults with hearing loss [Bergeron et al., 2019]. Normative data were developed with an additional sample of 30 adults with normal hearing from France and Canada, for 1 quiet condition and 2 noise conditions (speech spectrum noise and multi-talker noise) at 3 levels of SNR (+10, +5 and 0 dB SNR). Normalization data showed an average performance between 96% and 99% with low standard deviation. A retest was performed at 1 month and did not show any significant difference. There are 2 competing noise included in the FrBio: a 4-talker babble noise and a stationary speech spectrum noise. The babble noise was created by superimposing 8 sentences from the corpus, 2 sentence recordings per speaker. The stationary speech noise was generated by using the same spectrum as the LTAS of these 8 sentences.

The FrBio test is useful for evaluating differences in the speech recognition performance between experimental conditions or changes in performance over time by referring to the binomial Thornton and Raffin table for 50 items corpus [Thornton and Raffin, 1978]. The test material is available on a CD format (Table 3).

One more sentence in noise test that falls into the open set category is the Quick Speech In Noise (QuickSIN) test [Killion et al., 2004]. It was originally developed in American English Language for fast evaluation of speech-in-noise abilities of listeners with normal hearing or with hearing loss. According to British Society of Audiology [2019], the QuickSIN test is the most widely quoted speech-in-noise test in the literature when referring to clinical practice. It includes 12 equivalent lists of 6 sentences spoken by a female speaker, comprising 5 keywords each. While the 4-talker babble remains at a constant intensity level, the test uses an adaptive procedure in which each sentence of a list is presented at different intensity levels, from +25 dB to 0 dB SNR [Killion et al., 2004]. The listener has to repeat the sentences presented, and each correctly repeated word is awarded 1 point for a total possible score of 30 points per list [Duncan and Aarts, 2006]. The result of the test is expressed as an “SNR loss,” which is essentially the difference between the SNR required by a listener to achieve 50% correct

sentence identification and the normative value of SNR from a group without any hearing problem [Killion et al., 2004].

Le Test Vocale Rapide dans le Bruit

The Vocale Rapide dans le Bruit (VRB) test [Leclercq et al., 2018] is a French adaptation of the QuickSIN. The speech material consisted of 120 sentences of equal difficulty and spoken by a female talker of European French, with SNR-50s clustering within a range of 3.34 dB (standard deviation = 0.55 dB). Each sentence includes 3 keywords and is presented with a sample (8 s) of the “onde vocale globale” (global vocal wave: OVG) noise [Dodele and Dodelé, 2000]. It consists of several simultaneous speech signals (4 voices mixed: 1 couple speaking in French, while the 2nd is speaking English at the same time). A particular interest of this noise is its steadiness in intensity. The VRB test uses an automated descending procedure, similar to the QuickSIN, with 9 SNR per list, from +24 dB to 0 dB, by 3 dB steps. The score obtained with the VRB test is also expressed in SNR loss. Normative values were collected on normo-hearing subjects ($n = 12$) between 19 and 39 years of age. The major interest of the test is its short duration (about 5 min with a training list and a minimum of 4 lists as recommended by the authors). The VRB test is available via the HubSound software (Biotone, Villepinte, France) which also includes different audiometric tests to be used with the listeners who are speaker of European French (Table 3). For this test, the SRT50 is 0.08 dB SNR (SD = 0.6) (Table 2; Fig. 2).

Tests Using Non-Words

The speech understanding in noise (SUN) test [Paglialonga et al., 2011], initially developed in Italian, has been adapted into several languages including French. Vocal material includes list of 12 balanced intervocalic consonants (VCV) and a speech weighted noise, unilaterally presented via headphones. The test was optimized on 150 subjects with different levels of hearing divided into 3 classes. The procedure is described as fast and reliable (<1 min per ear) [Paglialonga et al., 2013]. A fixed SNR training (+8 dB) is recommended before the evaluation, the training and test together represents <2 min. This is a fixed SNR procedure where the speech level is fixed (it is advisable to test at the speech level of 60 dB SPL) [Paglialonga et al., 2014]. The initial presentation noise level is 58 dB SPL (+2 dB SNR) and increases by 2 dB with each presentation up to 68 dB SPL (−8 dB SNR) (see Fig. 2c). At each presentation, the subject must choose between 3

Table 4. Available non-words in noise tests

<i>European context (France)</i>	
SUN test	Adaptive procedure. Closed set
[Paglialonga et al., 2011, 2013]	Stimuli: intervocalic consonants (VCV) Noise: speech-weighted noise Availability: unknown
SUN, speech understanding in noise.	

answers via a touch screen (forced choice). The number of correct answers is counted at the end of the list. Nine or more correct answers: no listening difficulty; 7 or 8 correct answers: hearing control desirable; 6 correct answers or less: hearing test recommended. It is a screening test for sensorineural hearing loss allowing a rapid assessment of hearing in order to refer the patient to a specialist when the number of correct answers is <7. By using VCVs and an automated forced choice response paradigm, the influence of auditory-cognitive interactions and the role of non-auditory functions are limited compared to an open task and would therefore allow testing listeners with poor French language skills (Table 4).

Discussion

The purpose of the present article was to provide information on SpIN tests in order to facilitate the undertaking of speech-in-noise measures into routine clinical assessments. After having defined characteristics of the SpIN tests (for a better understanding of their impact on the test outcomes), a review of the available SpIN tests for the French speaking population was provided.

Even if the utility of testing speech perception in noise abilities is not questioned anymore, it was not incorporated in European Francophone regular practice until the recent review and update of procedures for hearing aids prescription and reimbursement in Belgium and France. They now clearly recommend the use of SpIN tests for assessing patients with slight to mild hearing loss [Belgium National Institute for Health and Disability Insurance, 2015]. In 2018, updated policies were released in France. They state that any hearing aid prescribed with well-documented abnormal SpIN testing, in the presence of normal PTA as well as Speech Reception-in-Quiet (SIQ) testing, is eligible for partial or total reimbursement. Also, both ANSD and auditory processing disorder, 2 conditions

frequently associated with poor SpIN test results (despite normal PTA and SIQ thresholds), are now listed as newly eligible for hearing aid reimbursement (<https://www.legifrance.gouv.fr/jorf/id/JORF-TEXT000037615111>). With the growing use of SpIN tests, clinicians are expected to diagnose a greater number of patients with any sensorineural deficit, better identify those with ANSD or auditory processing disorder, and conduct more robust and reliable clinical trials in the field of deafness.

At the time of the writing, there were essentially 10 validated SpIN tests available for use with the French speaking population in international literature. Two tests are using digits triplets as speech stimuli and were primarily created for self-administered hearing screenings using phone calls, internet, or smartphone applications. The other tests are targeting a wider range of indications that cover diagnostic issues, assessment of the effectiveness of various therapeutic interventions, and also research protocols: 2 using closed-set test using sentences and 4 open-set tests using sentences as speech material. Only 1 test using non-words was found (the SUN test), for which normative data are available. However, the way in which the material can be obtained is not known. It is to note that one of the most widely used non-words test in France and Belgium – but not published internationally – is the Audiométrie Verbo-fréquentielle (AVfB) test. It was created to complete the auditory assessment in silent condition [Dodelé and Dodelé, 2000]. The test is based on the use of logatomes so as to avoid lipreading and any context effects that could lead to an unwanted effect of top-down processes on the measurement outcomes in the presence of noise. For the 2 speaker “automatic” version of the test, SNR gradually decreases from +9 to –9 dB SNR. For the “manual” version, 2–3 speakers are recommended and SNR can be adjusted.

We found only 1 test using words in noise, adapted for Canadian Francophones. There are currently no recorded words in noise tests available for European French speakers. Historical, but widely used, words-in-noise tests used in France used live voices during testing procedure; however, no article presenting validated standards has been made available in the international literature for both tests. Some in-house procedures were developed over the years, without any validation verifications. Lafon [1964] proposed a simple technique to assess hearing in noise – via a classic 2-channel audiometer – the phonetic test for French speakers. It is composed of 2 lists of 51 monosyllables with 3 phonemes, one

of these is “characteristic”; incorrect identification of the latter leads to the highest number of incorrect word reporting. An old and easy way to perform audiometry in noise was to use the HIRSH test [Hirsh et al., 1952]. At 1st, the percentage of words (mainly LAFON lists) correctly repeated by a subject is determined with a white noise (fixed at 65 dB). The signal and the noise are emitted from the same loudspeaker (0° azimuth), and then with the competing noise coming from a 2nd loudspeaker (90°).

This is not to say there are no other tests being used by the hearing health professionals serving francophone populations, or conducting research with French speakers, however. As one can understand from this updated state of the art, there is no such thing as a perfect test. It is hoped that by having a better understanding of the SpIN test characteristics and their impact on the outcomes, the hearing health professionals will be more confident in their selection of the test, based on the needs (e.g., diagnostic and fitness for duty assessment) and personal factors (e.g., attention span and linguistic abilities) of their client or patient.

When choosing a SpIN test, several criteria should be considered depending on clinical purpose: the characteristics of the speech material, the type of noise, the listening conditions, the type of procedure and its duration, the target scores, and the availability of normative values. Choosing a test that has been validated for the specific purpose and target population is paramount for obtaining reliable and accurate results. Although not detailed in the present paper, the use of recorded SpIN tests instead of monitored live voice presentation is highly recommended, as individual differences between speakers affects the reliability of the results [Hughes and Scott, 1967; Theunissen et al., 2009], and this recommendation is part of most practice guidelines [British Society of Audiology, 2019]. Then, it should be ensured that the speech material can be fully understood by the intended audience. While all the tests listed in this review were developed for French-speaking subjects, the speech material intelligibility may be affected either by the talker accent or dialect characteristics. Some tests have been developed in France and Belgium, and as such, the accent of the speech stimuli may correspond to what is called the European French. Reciprocally, the recordings of the SpIN tests that have been developed in Canada present some accent and intonation features that are more typical of the Canadian French. For example, among the 12 sentence lists provided by the Canadian French version of the HINT,

only 5 can be effectively used in France. As the effect of accent may reduce listener performance [Marriage et al., 2001; Sockalingham et al., 2004], it is essential to use a country specific SpIN test, even though all of them have been developed for the French speaking population. The exception to this general recommendation may be for the FrBio test, for which special precautions were taken during the recordings to ensure that the accent could not be identified as being Canadian or European French.

Second, some SpIN tests are dedicated to specific indications. In particular, the digit triplet-based tests were designed for screening purpose. Most of them are freely available and can be auto administered (<https://www.who.int/health-topics/hearing-loss/hearwho>).

For diagnostic purpose and determination of hearing deficit, adaptive procedures aiming at SRT50 measurement can be recommended. Compared to fixed-level procedures, they are less sensitive to potential ceiling effects and, as such, more appropriate for standardized audiometric scores calculation [Theunissen et al., 2009]. In addition, renewed legislative guidelines for hearing aid services in Belgium and France now refer to SpIN test results expressed in dB SNR (<https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000037615111>).

Assessment of binaural hearing and hearing aids benefits should rely on SpIN tests allowing multiple trials with non-redundant speech materials delivered in free field. While a single loudspeaker may be sufficient to compare different hearing aids settings, a minimum of 2 are needed to fully evaluate binaural hearing (i.e., summation, and head shadow and squelch effects), and at least 4 are required to reproduce spatial impression of diffuse sound field [Hiyama et al., 2002].

A few SpIN tests have been validated in other languages than French (namely, FrBio, HINT, and Matrix), and thus are of particular interest for cross linguistic studies. Their utilization ensures identical procedures and testing conditions in case of multicenter studies [Kollmeier et al., 2015].

If the aim is to evaluate SpIN in the most realistic conditions, sentence recognition materials should be preferred to digits or words and fluctuating to stationary noise [Wilson and McArdle 2005; Theunissen et al., 2009]. Notably, hearing impaired subjects may be more comfortable with procedures using SNRs that allow >50% of speech recognition [Jansen et al., 2012; Bergeron et al., 2019]. The FrMatrix test, for instance, enables the determination of SRT70 or SRT80 [Jansen et al., 2012]. Similarly, in the most difficult test condition of the FrBIO, the speech recognition never drops below 80% in NH sub-

jects [Bergeron et al., 2019]. This can help avoid discouragement of patients over the time course of their hearing rehabilitation.

Conclusion

As there are about 300 million French speakers in the world, it appeared pertinent to have a precise idea of the available SpIN tests to guide the clinical practices and future efforts in test development for the Francophone population. The characteristics of the SpIN tests were also described, and as these characteristics are not unique to French language, the analyses and advice presented may be applied to tests selection in any language. If the purpose of conducting such measures is only for diagnostic purposes, it seems that there is no ideal test paradigm, and it may simply be the accuracy and duration of the test that should be considered. SpIN tests that use adaptive procedures may then be the type of test to consider limiting the risk of ignoring the early signs of hearing loss as well as for providing appropriate therapeutic and rehabilitative services.

There are many types of SpIN tests available in the French language – some use adaptive procedure, others use fixed-level strategies, some include sentences as speech items, others include single words, etc. There is a SpIN test for almost any type of clinical or rehabilitation needs. In an effort to meet the needs of a rapidly growing aged population, fast adaptive procedures can be favored to screen large populations in order to limit the risk of ignoring the early signs of forthcoming presbycusis and to provide appropriate audiological counseling.

References

- Akeroyd MA, Arlinger S, Bentler RA, Boothroyd A, Dillier N, Dreschler WA, et al. International Collegium of Rehabilitative Audiology (ICRA) recommendations for the construction of multilingual speech tests. ICRA working group on multilingual speech tests. *Int J Audiol*. 2015;54(Suppl 2): 17–22.
- Belgium National Institute for Health and Disability Insurance (Institut national d'assurance maladie-invalidité, INAMI). 2015 Jan. Available from: <https://www.inami.fgov.be/fr/professionnels/sante/audiciens/Pages/circulaires-audicien.aspx>.
- Bergeron F, Berland A, Fitzpatrick EM, Vincent C, Giasson A, Leung Kam K, et al. Development and validation of the FrBio, an international French adaptation of the AzBio sentence lists. *Int J Audiol*. 2019;58(8): 510–5.
- British Society of Audiology. *Practice guidance assessment of speech understanding in noise in adults with hearing difficulties*. 2019.
- Brungart DS, Simpson BD, Ericson MA, Scott KR. Informational and energetic masking effects in the perception of multiple simultaneous talkers. *J Acoust Soc Am*. 2001; 110(3):2527–38.
- Carhart R, Tillman TW. Interaction of competing speech signals with hearing losses. *Arch Otolaryngol*. 1970;91(3):273–9.
- Clopper CG, Pisoni DB, Tierney AT. Effects of open-set and closed-set task demands on spoken word recognition. *J Am Acad Audiol*. 2006;17(5):331–49.
- Dodele L, Dodele D. L'audiométrie vocale en présence de bruit et filetest AVfB. *Cahiers de L' Audition*. 2000;13(6):15–22.
- Dorman MF, Loizou PC, Spahr AJ, Maloff E. A comparison of the speech understanding provided by acoustic models of fixed-channel and channel-picking signal processors for cochlear implants. *J Speech Lang Hear Res*. 2002;45(4):783–8.

Acknowledgements

The authors are grateful to Prof. Birger Kollmeier for his helpful comments on a previous version of the manuscript, including the paragraph dedicated to the SpIN test he and his team have developed and to Mrs. Verena Landel for proofreading.

Statement of Ethics

An ethics statement was not required for this study type, no human or animal subjects were used.

Conflict of Interest Statement

The authors report no conflicts of interest.

Funding Sources

The authors acknowledge the support to Institut de l'Audition from Fondation pour l'Audition (FPA IDA09).

Author Contributions

P.R., C.-A.J. and J.L. took the lead in drafting then finalizing the article. E.V. and L.D. were involved in identifying the major points to be considered and framing the questions raised by the study. H.T.-V. supervised both the review and the writing, and also helped in drawing up the key findings. All authors contributed in shaping the manuscript and provided critical revision.

Data Availability Statement

All data generated or analyzed during this study are included in this article. Further inquiries can be directed to the corresponding author.

- Duncan K, Aarts N. A comparison of the HINT and quick SIN tests. *J Speech Lang Pathol Audiol*. 2006;30.
- Eisenberg LS, Kirk KI, Martinez AS, Ying EA, Miyamoto RT. Communication abilities of children with aided residual hearing: comparison with cochlear implant users. *Arch Otolaryngol Head Neck Surg*. 2004;130(5):563–9.
- Epstein A, Giolas TG, Owens E. Familiarity and intelligibility of monosyllabic word lists. *J Speech Hear Res*. 1968;11(2):435–8.
- Freyman RL, Balakrishnan U, Helfer KS. Effect of number of masking talkers and auditory priming on informational masking in speech recognition. *J Acoust Soc Am*. 2004;115(5):2246–56.
- Giguère C, Lagacé J, Ellaham NN, Pichora-Fuller MK, Goy H, Bégin C, et al. Development of the Canadian digit triplet test in English and French. *J Acoust Soc Am*. 2020;147(3):EL252.
- Hagerman B. Sentences for testing speech intelligibility in noise. *Scand Audiol*. 1982;11(2):79–87.
- HearCom. D-1-3: protocol for implementation of communication tests in different languages. 2006.
- Hirsh IJ, Davis H, Silverman SR, Reynolds EG, Eldert E, Benson RW. Development of materials for speech audiometry. *J Speech Hear Disord*. 1952;17(3):321–37.
- Hiyama K, Komiya S, Hamasaka K. The minimum number of loudspeakers and its arrangement for reproducing the spatial impression of diffuse sound field. *J Audio Engineer Soc*. 2002;50(10):864.
- Hughes RL, Scott RJ. Speech audiometry. tape versus disc recording. *Arch Otolaryngol*. 1967;86(2):175–8.
- Iliadou VV, Ptok M, Grech H, Pedersen ER, Brechmann A, Deggouj N, et al. A European perspective on auditory processing disorder-current knowledge and future research focus. *Front Neurol*. 2017;8:622.
- Jansen S, Luts H, Wagener KC, Frachet B, Wouters J. The French digit triplet test: a hearing screening tool for speech intelligibility in noise. *Int J Audiol*. 2010;49(5):378–87.
- Jansen S, Luts H, Wagener KC, Kollmeier B, Del Rio M, Dauman R, et al. Comparison of three types of French speech-in-noise tests: a multi-center study. *Int J Audiol*. 2012;51(3):164–73.
- Johnston KN, John AB, Kreisman NV, Hall JW III, Crandell CC, Johnston KN, et al. Multiple benefits of personal FM system use by children with auditory processing disorder (APD). *Int J Audiol*. 2009;48(6):371–83.
- Kalikow DN, Stevens KN, Elliott LL. Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *J Acoust Soc Am*. 1977;61(5):1337–51.
- Killion MC, Niquette PA, Gudmundsen GI, Revit LJ, Banerjee S. Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. *J Acoust Soc Am*. 2004;116(4):2395–405.
- Kollmeier B, Warzybok A, Hochmuth S, Zokoll MA, Usilar V, Brand T, et al. The multilingual matrix test: principles, applications, and comparison across languages: a review. *Int J Audiol*. 2015;54(Suppl 2):3–16.
- Lafon JC. [Audiometry with the phonetic test]. *Acta Oto-Rhino-Laryngol Belgica*. 1964;18:619–33.
- Lagacé J. Développement du test de mots dans le bruit: mesure de l'équivalence des listes et données préliminaires sur l'effet d'âge. *Can Acoust*. 2010;38(2):19–30.
- Lagacé J, LeBlanc L, Boisvert V, Arseneau M, Breau-Godwin S. Mise à jour sur le développement du test de mots dans le bruit. *Can Acoust*. 2013;41(2):65–72.
- Leclercq F, Renard C, Vincent C. Speech audiometry in noise: development of the French-language VRB (vocale rapide dans le bruit) test. *Eur Ann Otorhinolaryngol Head Neck Dis*. 2018;135(5):315–9.
- Leek MR. Adaptive procedures in psychophysical research. *Percept Psychophys*. 2001;63(8):1279–92.
- Luts H, Boon E, Wable J, Wouters J. FIST: a French sentence test for speech intelligibility in noise. *Int J Audiol*. 2008;47(6):373–4.
- Marriage J, King J, Briggs J, Lutman ME. The reliability of the SCAN test: results from a primary school population in the UK. *Br J Audiol*. 2001;35(3):199–208.
- McArdle RA, Wilson RH, Burks CA. Speech recognition in multitalker babble using digits, words, and sentences. *J Am Acad Audiol*. 2005;16(9):726–4.
- Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ*. 2009;339:b2535.
- Moore DR, Edmondson-Jones M, Dawes P, Fortnum H, McCormack A, Pierzycki RH, et al. Relation between speech-in-noise threshold, hearing loss and cognition from 40–69 years of age. *PLoS One*. 2014;9(9):e107720.
- 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(2):1085–99.
- Nordrum S, Erler S, Garstecki D, Dhar S. Comparison of performance on the hearing in noise test using directional microphones and digital noise reduction algorithms. *Am J Audiol*. 2006;15(1):81–91.
- Organisation Internationale de la Francophonie. *La langue française dans le monde, synthèse*: Gallimard; 2018. Available from: <http://observatoire.francophonie.org/2018/synthese.pdf>.
- Paglalonga A, Tognola G, Grandori F. SUN-test (speech understanding in noise): a method for hearing disability screening. *Audiol Res*. 2011;1(1):e13.
- Paglalonga A, Grandori F, Tognola G. Using the speech understanding in noise (SUN) test for adult hearing screening1. *Am J Audiol*. 2013;22(1):171–4.
- Paglalonga A, Tognola G, Grandori F. A user-operated test of suprathreshold acuity in noise for adult hearing screening: the SUN (speech understanding in noise) test. *Comput Biol Med*. 2014;52:66–72.
- Plomp R, Mimpfen AM. Speech-reception threshold for sentences as a function of age and noise level. *J Acoust Soc Am*. 1979;66(5):1333–42.
- Pollack I. Auditory informational masking. *J Acoust Soc Am*. 1975;57(S1):S5.
- Prang I, Parodi M, Coudert C, Legoff S, Exter M, Buschermöle M, et al. Le matrix simplifié français : un outil d'évaluation de l'intelligibilité de la parole dans le bruit. *Eur Ann otorhinolaryngol Head Neck dis*. 2021 Sep;138(4):253–6.
- Pumford JM, Seewald RC, Scollie SD, Jenstad LM. Speech recognition with in-the-ear and behind-the-ear dual-microphone hearing instruments. *J Am Acad Audiol*. 2000;11(1):23–35.
- Ricketts T, Henry P. Evaluation of an adaptive, directional-microphone hearing aid. *Int J Audiol*. 2002;41(2):100–12.
- Simpson SA, Cooke M. Consonant identification in N-talker babble is a nonmonotonic function of N. *J Acoust Soc Am*. 2005;118(5):2775–8.
- Smits C, Kapteyn TS, Houtgast T. Development and validation of an automatic speech-in-noise screening test by telephone. *Int J Audiol*. 2004;43(1):15–28.
- Smoorenburg GF. Speech reception in quiet and in noisy conditions by individuals with noise-induced hearing loss in relation to their tone audiogram. *J Acoust Soc Am*. 1992;91(1):421–37.
- Sockalingam R, Murrison R, Cole D, James S, Morin S, Potter S. Comparing SCAN-A scores between speakers of standard Australian English and American English: a preliminary study. *Aust New Zeal J Audiol*. 2004;26(2):110.
- Soli SD, Wong LL. Assessment of speech intelligibility in noise with the hearing in noise test. *Int J Audiol*. 2008;47(6):356–61.
- Spahr AJ, Dorman MF, Litvak LM, Van Wie S, Gifford RH, Loizou PC, et al. Development and validation of the AzBio sentence lists. *Ear Hear*. 2012;33(1):112.
- Stach BA, Davis-Thaxton ML, Jerger J. Improving the efficiency of speech audiometry: computer-based approach. *J Am Acad Audiol*. 1995;6:330–3.
- Taylor B. Speech-in-noise tests: how and why to include them in your basic test battery. *Hear J*. 2003;56(1):40–3.

- Theunissen M, Swanepoel de W, Hanekom J. Sentence recognition in noise: variables in compilation and interpretation of tests. *Int J Audiol*. 2009;48(11):743–57.
- Thornton AR, Raffin MJ. Speech-discrimination scores modeled as a binomial variable. *J Speech Hear Res*. 1978;21(3):507–18.
- Vaillancourt V, Laroche C, Mayer C, Basque C, Nali M, Eriks-Brophy A, et al. The Canadian French hearing in noise test. *Int J Audiol*. 2008;47(6):383–5.
- Vaillancourt V, Laroche C, Mayer C, Basque C, Nali M, Eriks-Brophy A, et al. Adaptation of the hint (hearing in noise test) for adult canadian francophone populations. *Int J Audiol*. 2005;44(6):358–69.
- 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(3):144–56.
- Watson CS, Kidd GR, Miller JD, Smits C, Humes LE. Telephone screening tests for functionally impaired hearing: current use in seven countries and development of a US version. *J Am Acad Audiol*. 2012;23(10):757–67.
- Wilson RH, McArdle R. Speech signals used to evaluate functional status of the auditory system. *J Rehabil Res Dev*. 2005;42(4 Suppl 2):79–94.
- Zokoll MA, Hochmuth S, Warzybok A, Wagener KC, Buschermöhle M, Kollmeier B. Speech-in-noise tests for multilingual hearing screening and diagnostics1. *Am J Audiol*. 2013;22(1):175–8.
- Zokoll MA, Wagener KC, Brand T, Buschermöhle M, Kollmeier B. Internationally comparable screening tests for listening in noise in several European languages: the German digit triplet test as an optimization prototype. *Int J Audiol*. 2012;51(9):697–707.