OPEN ACCESS GUIDE TO AUDIOLOGY AND HEARING AIDS FOR OTOLARYNGOLOGISTS

SPEECH AUDIOMETRY

Marianne van Zyl

In order to properly assess the degree and nature of a patient's hearing loss, and to plan the most appropriate intervention, the following audiological tests are recommended following a detailed history and clinical examination: pure tone audiometry; tympanometry; and speech audiometry. This chapter will focus on the theory and practice of speech audiometry.

Speech audiometry is used to measure the ability of a patient to perceive speech signals. Speech materials (pre-recorded or read by examiner) are presented; the patient repeats the speech materials to determine how well it was perceived.

Why speech audiometry?

Speech audiometry assesses a patient's auditory ability using words, ***which are much more representative of everyday listening experience than pure tones***. Measuring the ability to perceive speech gives the clinician a clearer picture of the patient's ***functional hearing ability*** and is extremely valuable to predict a patient's ***success with hearing aids.***

Equipment / Setup requirements

Speech audiometry, especially when using live voice (speech materials read out by examiner) requires a test set-up where the patient is seated in a sound-treated room with the examiner in an adjacent room. A plexiglass partition between patient and examiner is preferred so as to permit mutual visibility *(Figure 1)*. The audiometer should have a microphone attached to a Volume Unit (VU) meter, and a talk-back facility whereby the patient's voice is audible to the examiner via a microphone installed in the sound booth.

Figure 1: Example of test setup and supra-aural earphones

If pre-recorded materials are used, the audiometer should have an auxiliary input to which a CD-player can be connected. As with pure tone audiometry, speech audiometry can be done using earphones (inserts or supra-aural) or sound field speakers as transducers. Proper calibration of equipment is essential.

Types of speech tests

The speech audiometry tests generally used are ***threshold*** and ***supra-threshold tests***.

Threshold tests are employed to confirm pure tone thresholds and to determine a suitable presentation level for supra-threshold tests.

Supra-threshold tests are used to determine the percentage of speech recognition a patient can obtain. It also provides an impression of how accurately speech is recognized by the listener at a typical conversational speech level. Supra-threshold tests are also used to determine the benefit a patient obtains from a hearing aid.

A third type of speech test that is sometimes used in clinical practice is s***peech-in-noise tests.*** These provide an impression of a patient's functioning in everyday communication situations which mostly involve some degree of background noise; tests are available in a number of languages [1].

Many types of speech materials can be used to conduct speech audiometry. These commonly include single words (mono- or bisyllabic words), sentences or phonemes (vowels and consonants). It is preferable to assess speech perception in a patient's ***native language***, or at least in a language in which he/she is highly capable. Therefore unless such standardised speech materials are available, new materials should be developed for every language in which speech audiometry is to be conducted.

Presentation of test materials

Speech audiometry tests are presented either as ***"monitored live-voice"*** or by using ***pre-recorded test materials***.

Monitored live-voice testing: The test administrator presents speech materials through a microphone that is connected to a VU meter which measures the intensity level of the voice *(Figure 2)*.

The average output level at the transducer is controlled via the audiometer, but fluctuations in signal intensity are a function of fluctuations in the live-voice intensity. The ***advantages*** of monitored live-voice testing are that it is easy to administer, and allows flexibility of the test procedure.

Figure 2: Typical audiometer with microphone and VU meter

The ***disadvantages*** are the lack of precise control over the signal intensity, along with variations that occur across test administrators - different clinicians testing the same patient with the same test materials might yield large differences in results due to individual differences in the testers' voice frequencies and pronunciations.

Pre-recorded materials: Pre-recorded speech materials are recorded in an ideal acoustic environment, digitally edited, and presented via a CD-player or personal computer through the auxiliary input of the audiometer. ***Advantages*** include test-retest reliability of the materials, as well as a greater degree of control of signal intensity. ***Disadvantages*** include the effort required to produce such materials and the lack of flexibility in terms of the time intervals between presentations of individual test items (words).

The following sections describe two speech tests that are commonly used in clinical practice, i.e. the speech reception or recognition threshold (SRT), and the speech discrimination score; the author also describes the test procedures as well as the interpretation of the results.

Speech recognition threshold (SRT)

The SRT is the most frequently used speech threshold test. It is a measure of the intensity level at which the listener is able to ***correctly repeat 50% of words presented.*** This level should correspond roughly to the average of the pure tone audiometry thresholds at 500, 1000 and 2000 Hz [2]. When the pure tone audiogram is steeply sloping (high frequency hearing loss), the SRT corresponds better with either the best two of three pure tone average (PTA) frequencies (500, 1000 & 2000 Hz) or the best one of these three frequencies [3]. The SRT is generally con-ducted using ***spondee words*** *i.e.* single words that comprise two syllables with equal emphasis placed on each syllable.

SRT test procedure

Before testing, the listener is ***made familiar with the test materials*** to ensure that all the words are known to him/her. This is done by providing a written list of the words to the patient to read aloud, or by presenting the words at a comfortable intensity (e.g. 30-40 dB above the PTA of 500, 1000 and 2000 Hz) and asking the patient to repeat the words (if the patient finds this intensity too low, words can also be presented at a slightly higher intensity for the purposes of familiarisation). Any words that the listener cannot repeat correctly or that is unfamiliar are excluded from the test. Before starting the test, instruct the patient to repeat the words he/she hears, even if the words become very soft, and he/she is allowed to guess if uncertain.

Chaiklin-Ventry SRT test procedure

The SRT may be measured in many ways. A simple technique is that described by *Chaiklin and Ventry* [4]; it is a descending method, meaning that the test commences at an intensity level clearly audible to the patient, and is then lowered in predetermined step sizes to a level at which 50% recognition is attained.

1. Examine the pure tone thresholds obtained for 500, 1000 and 2000 Hz in the test ear
2. Use the lowest two of the three values to calculate a two-frequency PTA
3. Start testing at an intensity level 25dB above the two-frequency PTA by presenting one spondee word. If using monitored live-voice, take care to present the two syllables of the word with equal intensity by carefully watching the VU-meter and ensuring that the meter indicates the same level for both syllables
4. If the listener is unable to repeat the word at the initial presentation level, increase the intensity by 10dB and present another word
5. Continue this process until the listener is able to correctly repeat the presented word
6. Once the listener correctly repeats a word, lower the intensity by 5dB and present another word
7. Continue lowering the intensity in steps of 5dB until the listener makes an error
8. Once this occurs, the main part of the test is commenced
9. Start the main threshold search 10dB above the level where the listener made the first error. Present three words at this level
10. If all three words are repeated correctly, lower the intensity by 5dB and present three more words
11. Once the listener makes an error with any of the three words, present three more words at the same level to get a score out of six
12. Continue to descend in 5dB steps, each time presenting six words, until a level is reached where the patient

correctly repeats three of six (50%) of the words

13. This level is then recorded on the test form as the SRT. If a patient repeats more than three of six words correctly at one level (e.g. 4/6 at 25dB) and less than three of six at the next presentation level (e.g. 2/6 at 20 dB), record the SRT as 20-25dB on the test form
14. Conduct the test in the opposite ear, so as to determine the SRT in each ear independently

Masking during SRT testing

Both insert and supra-aural earphones can cause cross-hearing (a response from the non-test ear) due to the transmission of the presented sound via bone conduction to the other ear. The presented stimulus, however, reaches the non-test ear at a lower intensity (an attenuated level) than the level at which it was presented at the test ear. The attenuation that the sound undergoes as it travels from the test ear to the non-test ear is called interaural attenuation, and the amount of attenuation depends on the type of transducer (inserts or supra-aural earphones) used. Masking is used to prevent the sounds being heard in the non-test ear due to cross-hearing.

Insert earphones (Figure 3)
The interaural attenuation for insert earphones is about 60dB [5]. Therefore masking is required when the SRT of the test ear minus the best pure tone bone conduction threshold in the non-test ear is >60dB; or when the SRT of the test ear minus the SRT of the non-test ear is >60dB For example if the SRT of the test ear is 50dB, and the best bone conduction threshold in the non-test ear is 5dB, the calculation is as follows: 50dB - 5dB = 45dB. As 45dB is less than 60dB, no masking is needed.

Figure 3: Insert earphones

Supra-aural earphones (Figure 1)

The interaural attenuation for supra-aural earphones is about 40dB [5]. Therefore masking is required when the SRT of the test ear minus the best pure tone bone conduction threshold in the non-test ear is >40dB, or when the SRT of the test ear minus the SRT of the non-test ear is >40dB.

Measuring a masked SRT

1. The masked SRT is measured after the unmasked SRT has been determined.
2. Instruct the patient that he/she will hear a noise in one of the ears, but that he/she should ignore this noise and keep repeating the words heard in the other ear to the best of his/her ability.
3. Introduce the masking noise (preferably the noise marked “speech noise” on the audiometer, or if this is unavailable, “white noise”) to the non-test ear. The noise is presented at 10dB above the SRT of the non-test ear.
4. Present 6 spondee words at the same level at which the threshold was measured before, with the masking noise presented to the non-test ear.
5. If the patient still attains 50% recognition at this level, increase the masking by 5dB and present six more words.
6. Repeat this process, until the masking level has been increased by 15dB

(three steps of 5dB) from the original masking level.

7. If the patient still attains 50%, then the masked threshold is the same as the unmasked threshold. Record this on the audiogram, along with the masking level that was used (e.g. 20-35dB).
8. If the patient is unable to attain 50% recognition at the previously measured SRT level once masking is introduced, keep the masking level constant and increase the speech level in the test ear by 5dB, presenting six words at that level.
9. Continue this process until the patient attains 50% recognition.
10. Increase the masking level by 5dB and present six more words, repeating this process until the masking level is 15dB higher than the initial level.

Speech discrimination score

The speech discrimination score (also called word recognition score) is the most commonly used test of ***supra-threshold speech perception***.

This is an important test in the audiological test battery, as it ***indicates the patient's ability to hear and understand speech at typical conversational levels***. It also indicates how well a patient can perceive speech if the presentation level is increased; this helps the clinician to predict the potential benefits from amplification. Some patients, especially those with sensorineural hearing loss and hearing losses that do not affect all frequencies equally, are unable to achieve 100% speech recognition, regardless of how much the presentation level is increased. These patients will benefit from strategies to improve the signal-to-noise ratio towards greater clarity of speech signals (e.g. advanced digital hearing aids, FM systems etc.). Speech audiometry results can also be used to indicate the ***site-of-lesion*** if pathology is present. Patients with cochlear or retrocochlear hearing loss or pathology are generally unable to attain 100% speech recognition.

Speech discrimination can be measured using a variety of test materials. Monosyllabic single words preceded by a carrier phrase are most frequently used. Because measuring speech discrimination involves testing at three or more intensity levels, test materials should include a number of lists so that a complete list of words can be used at each presentation level. A number of commercially available tests in English are available *(*see *Lawson and Peterson* for a list of tests [5]). Two frequently used English word lists are the *Northwestern Auditory Test No. 6 (NU-6)* [6] and the *CID Auditory Test W-22* [7]. Lists that are phonetically balanced (contain the same number of each particular phoneme in each list) are particularly popular as they ensure that listeners with a steeply sloping audiogram who have difficulty with high frequency phonemes will not find some of the lists easier than others. Ideally patients should be tested in their native language (mother tongue). If test materials are not available in the patient's first language and they are proficient in English, it should be possible to conduct the test using English materials, provided that the patient is familiar with all the words used. If the patient has limited English vocabulary it might be necessary to use a ***closed set test*** where the listener chooses a response from a limited number of options e.g. the *Modified Rhyme Hearing Test* [9].

Standard word discrimination/recognition score measurement

1. Instruct the patient to listen carefully to the presented words and to repeat them aloud.

2. Presentation of each test item may be preceded with a "carrier phrase" e.g. "say the word"; although results of research into the differences between this method and where no carrier phrase is used are inconclusive [2]. What is most important is that, whichever method is selected (*i.e.* with/ without carrier phrase), it should be used consistently throughout the test and across lists.
3. ***If audiometric results (air and bone conduction pure tone audiometry) indicate normal hearing or a conductive loss***, start testing at 30dB above the measured SRT, or if the SRT is not available, 30dB above the average pure tone thresholds measured at 500, 1000 and 2000 Hz. Most patients with normal hearing should attain 100% speech recognition at 25-40dB above their SRT [10]. Patients with conductive loss are also usually able to attain 100% recognition at intensity levels 30-40 dB above their own SRT.
4. ***If a patient's audiogram indicates sensorineural hearing loss (both air and bone conduction thresholds elevated)***, start testing at 40dB above the patient's SRT [8]. Patients with cochlear or retro-cochlear hearing loss or pathology are generally unable to attain 100% speech recognition, regardless of presentation level. The objective with such patients is to determine what the maximum percentage of speech recognition is, and at what intensity level this occurs.
5. Start the test by presenting a complete list of words (typical word lists usually comprise 50 words) at the selected starting intensity level.
6. Record how many of the words the patient repeats correctly. It is useful to mark words in the list that the patient finds difficult or repeats incorrectly; when a patient mistakes a particular phoneme for another, make a note of this. Calculate the percentage of words repeated correctly by doubling the score out of 50.
7. If the patient attains 100% recognition at the initial presentation level, present a 2nd list at a level 20dB below this level, or at a conversational level (*i.e.* 60dB HL) if the patient's thresholds were elevated, to obtain an impression of how well the patient would hear at normal conversational levels.
8. If the patient is unable to attain 100% recognition (or very close to 100%) at the initial presentation level, present another list at an intensity level 20dB above the initial level. If time allows, a 3rd list is presented at a level 20dB higher than that used for the 2nd list, or 10dB higher if the presentation level is uncomfortable for the patient.
9. With ***cochlear hearing loss***, speech recognition is expected to increase up to a certain point with increasing intensity (maximum recognition point), and then to stabilise.
10. With ***retrocochlear lesions*** speech recognition improves as intensity increases up to a point (maximum recognition point) after which recognition deteriorates at higher intensities. This phenomenon is called "roll-over" and is considered to be an indication of retrocochlear pathology. However, the degree of deterioration in speech recognition at increased intensities that can be considered a clinically significant roll-over may differ according to the speech materials used [3] and should be interpreted with caution and in conjunction with other site-of-lesion tests such as acoustic reflexes, reflex decay test or Auditory Brainstem Response measurements. The formula used to calculate the degree of roll-over is as follows: (maximum score – minimum score)/maximum score; the maximum score indicates the maximum % of recognition the patient

attained, and the minimum score indicates the minimum % recognition attained at an intensity level higher than the level where the maximum score was obtained.

Masking during word recognition testing

Unlike with SRTs, it is not necessary to first measure unmasked thresholds; the decision whether or not to mask is based on the presentation level which can be determined before the test is started. Masking for supra-threshold speech recognition testing is recommended under the following conditions [5]

1. The presentation level at the test ear minus the best bone conduction threshold in the non-test ear exceeds the interaural attenuation (40 dB for supra-aural earphones, 60 dB for inserts).
2. The presentation level at the test ear minus the SRT of the non-test ear exceeds interaural attenuation.

The ***level of masking*** depends on a number of factors:

1. ***Air-bone gap*** exceeding 15dB at any frequency. Examine the air and bone conduction thresholds in the non-test ear, and note the biggest significant air-bone gap in the non-test ear.
2. ***Interaural attenuation*** depends on the transducer used, e.g. 40dB for supra-aural, and 60 dB for insert earphones). Calculate the level of masking required by using the following formula: (Presentation level in test ear) minus (interaural attenuation) plus (the largest significant air-bone gap in the non-test ear) plus (safety factor of 20dB). Example: Word recognition is to be tested at 70dB in the test ear; supra-aural earphones are to be used; and the largest air-bone gap in the non-test is 20dB. The masking level is as follows: 70dB (presentation level) – 40dB (interaural attenuation) + 15dB (air-bone gap) + 20dB (safety factor) = 65dB of masking required in the non-test ear.

Reporting speech audiometry results

Speech reception threshold is recorded on the audiometric test form for the two ears separately e.g. the SRT of the right ear is reported as 15dB, and the SRT of the left ear as 10dB. A typical version of SRT results in an audiology report may be worded as follows: "The patient's speech recognition threshold (the loudness level where approximately 50% of speech is understood) was 15dB in the right ear and 10dB in the left ear. These findings confirm the pure tone thresholds". In cases where the SRT does not correspond closely to pure tone audiometry results, an explanation should be given if possible. In the case of a steeply sloping hearing loss, for instance, the SRT might correspond only to the best of the pure tone thresholds at 500, 1000 and 2000 Hz. If the patient's responses during pure tone audiometry were inconsistent and seemed unreliable, this may also explain a difference between pure tone thresholds and SRT. In cases like these, additional results (such as acoustic reflex thresholds measured with pure tones and broadband noise stimuli) may need to be considered in order to estimate accurate thresholds. If the SRT was measured at an intensity that was much higher than expected according to the pure tone results, the examiner should refer to the word recognition results to determine the possibility of a specific deficit with regards to the processing of speech sounds.

Word recognition test results are usually reported in a graph *(Figure 4)*. Each point on the graph illustrates the result of one completed list of a word recognition test as a percentage of the words that the listener repeated correctly at the corresponding

intensity indicated on the x-axis of the graph.

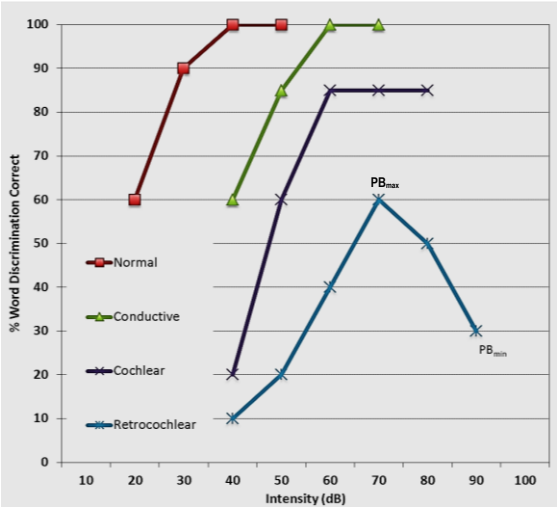

Figure 4: Reporting of word recognition tests

Conductive hearing losses generally follow the same pattern of results as normal hearing, but at a higher intensity *(Figure 4)*.

Cochlear pathology tends to demonstrate a "plateau effect", reaching a ceiling of performance at less than 100% and no improvement in the score despite a rise in intensity *(Figure 4)*.

"PB_{max}" indicates the maximum % speech recognition a listener is able to attain with phonetically balanced word lists; with retrocochlear pathology this value may be less than expected according to the pure tone results [3]. The last value for the retrocochlear loss is labelled the "PB_{min}", and refers to the lowest recognition score obtained at an intensity above the intensity at which the maximum score was measured *(Figure 4)*. As shown in *Figure 4*, patients with retrocochlear pathology may show a decrease in speech recognition as intensity increases, which is called a "rollover". However, the degree to which recognition decreases may depend on the word list that was used; the presence of rollover should therefore be considered in conjunction with other audiological findings to determine the likelihood of a retrocochlear lesion being present.

Table 1 provides examples of how to report word recognition results in an audiology report.

Site of lesion	Examples of reported results
No lesion (normal hearing)	100% speech recognition at presentation level 30dB above PTA
Conductive	100% correct speech recognition, but only at increased presentation level. This corresponds with the pure tone results and indicates conductive hearing loss
Cochlear	Unable to reach 100% speech recognition, regardless of increase in presentation level. This corresponds with presence of sensory hearing loss due to cochlear lesion
Retrocochlear	Unable to reach 100% speech recognition, and speech recognition abilities deteriorate with increase in presentation level. This corresponds with presence of sensorineural hearing loss, and may be indicative of retrocochlear pathology

Table 1: Reporting word recognition

References

1. Soli SD, Wong LLN. Assessment of speech intelligibility in noise with the Hearing in Noise Test. *International Journal of Audiology*, 2008; **47**(6): 356-61
2. Martin FN. *Introduction to Audiology*, (1997) 6th ed, Allyn & Bacon, Boston
3. Gelfand SA. *Essentials of Audiology*, (2009) 3rd ed, Thieme Medical Publishers, Inc., New York.
4. Chaiklin JB, Ventry IM. Spondee threshold measurement: A comparison of 2- and 5-dB steps. *Journal of Speech and Hearing Disorders*. 1964; **29**: 47-59

5. Lawson G. Peterson M. *Speech Audiometry*, (2011) Plural Publishing, Inc., San Diego
6. Tillman TW, Carhart R. An expanded test for speech discrimination utilizing CNC monosyllabic words, *Northwestern University Auditory Test No. 6 Tech Report SAM-TR-66-55*, (1966) USAF School of Aerospace Medicine, Brooks Air Force Base, Texas.
7. Hirsh IJ, Davis H, Silverman SR, Reynolds EG, Eldert E, Benson, RW. Development of materials for speech audiometry. *Journal of Speech and Hearing Disorders.* 1952; **17**(3): 321-37
8. Kramer SJ. *Audiology: science to practice*, (2008) Plural Publishing, Inc., San Diego
9. Kruel EJ, Nixon JC, Kryter KD, Bell DW, Lang JS, Schubert ED. A proposed clinical test of speech discrimination, *Journal of Speech and Hearing Research.* 1968; **11**(3): 536-52
10. Schoepflin JR (2012). Back to basics: speech audiometry.
11. http://www.audiologyonline.com/articles/back-to-basics-speech-audiometry-6828. Last accessed on 15 December 2012

Author

Marianne van Zyl, M.Comm.Path
Bioengineering Group
Department of Electrical, Electronic and Computer Engineering
University of Pretoria
Pretoria, South Africa
marianne.vanzyl@up.ac.za

Editors

De Wet Swanepoel PhD
Associate Professor
Department of Communication Pathology
University of Pretoria
Pretoria, South Africa
dewet.swanepoel@up.ac.za

Claude Laurent, MD, PhD
Professor in ENT
ENT Unit
Department of Clinical Science
University of Umeå
Umeå, Sweden
claude.laurent@ent.umu.se

Johan Fagan MBChB, FCORL, MMed
Professor and Chairman
Division of Otolaryngology
University of Cape Town
Cape Town
South Africa
johannes.fagan@uct.ac.za

OPEN ACCESS GUIDE TO AUDIOLOGY & HEARING AIDS FOR OTOLARYNGOLOGISTS
http://www.entdev.uct.ac.za