

SPEECH AUDIOMETRY

Authored by
Mohammed looti

November 14, 2025

RECOMMENDED CITATION

Mohammed looti (2025). *SPEECH AUDIOMETRY*. Encyclopedia of psychology. Retrieved from <https://encyclopedia.arabpsychology.com/?p=17712>

Introduction and Definition of Speech Audiometry

Speech audiometry constitutes a fundamental and indispensable component of the comprehensive audiological evaluation, providing critical insights into an individual's ability to hear, process, and understand spoken language. It is defined precisely as the measurement of how accurately and efficiently a person perceives speech stimuli across a range of acoustic intensities. While pure-tone audiometry establishes the softest sounds an individual can detect--the thresholds--speech audiometry moves beyond simple detection to assess the functional capacity for communication, which relies heavily on complex spectral and temporal processing. This technique utilizes standardized speech material, such as single words, sentences, or spondee words, presented at various loudness levels to evaluate the integrity of the auditory system from the outer ear through the central auditory pathways. The results obtained are paramount for differential diagnosis, assisting clinicians in distinguishing between conductive hearing losses, sensorineural deficits, and retrocochlear pathologies, thereby guiding appropriate intervention strategies, whether they involve medical referral, hearing aid fitting, or auditory rehabilitation.

The core distinction between speech audiometry and pure-tone testing lies in the complexity of the stimulus. A **pure tone** is a single frequency, whereas speech is a dynamic, broadband signal comprising numerous frequencies, varying intensities, and rapid temporal changes. Because speech is highly redundant and carries meaningful linguistic information, testing with speech stimuli offers a more ecologically valid measure of everyday hearing performance than simple tonal thresholds alone. Effective speech perception requires not only the detection of acoustic energy but also sufficient resolution and processing speed within the central nervous system to decode phonemes into meaningful words and sentences. Therefore, a poor performance on speech audiometry, even when pure-tone thresholds are only mildly impaired, often indicates significant underlying difficulties, such as distortion or clarity issues commonly associated with cochlear damage or neural processing disorders.

The implementation of speech audiometry is systematic, involving precise calibration of equipment and controlled acoustic environments, typically utilizing a sound-treated booth to minimize ambient noise interference. The stimuli can be delivered either via monitored live voice (MLV) or, preferably, through recorded materials (CDs or digital files). Recorded materials are highly favored in clinical practice because they ensure consistent intensity, presentation rate, and speaker characteristics, thus maximizing test reliability and allowing for direct comparison of results across different clinical settings. The various tests within the speech audiometry battery are designed to probe different aspects of auditory function, including the softest level at which speech is detected, the softest level at which speech is understood, the ability to understand speech at comfortable listening levels, and the intensity level that causes discomfort, providing a comprehensive profile of the patient's communication abilities and limitations.

Historical Context and Evolution

The origins of formalized speech testing can be traced back to the middle of the 20th century, particularly driven by the need to effectively rehabilitate soldiers returning from World War II with noise-induced hearing loss. Early audiological practice often relied solely on subjective assessments or rudimentary pure-tone measures, which proved inadequate for predicting real-world communication success. Pioneering work conducted at institutions such as the Harvard Psycho-Acoustic Laboratory led to the standardization of speech materials, including the development of phonetically balanced word lists (PB lists) and **spondee words**, which are two-syllable words with equal stress on both syllables (e.g., 'hotdog,' 'baseball'). This standardization was crucial, as it allowed clinicians to quantify speech understanding in a reliable manner for the first time, moving the field of audiology toward objective, measurable parameters of hearing function.

Initially, speech audiometry was frequently administered using the **monitored live voice (MLV)** technique, where the audiologist speaks the test words to the patient while monitoring their voice level on a calibrated meter. While convenient, the MLV method introduced variability due to differences in the audiologist's speaking pattern, accent, and consistency of intensity delivery. Recognizing the inherent limitations in reliability and validity imposed by MLV, the field progressively shifted toward using high-quality recorded materials. This transition marked a significant evolutionary step, ensuring that the acoustic characteristics of the stimuli remained constant, thereby enhancing the precision of diagnostic findings and improving the ability to track changes in a patient's hearing over time or compare outcomes from different clinical trials or rehabilitative interventions.

Contemporary speech audiometry has further evolved beyond simple word recognition to incorporate more complex, real-world stimuli, such as sentence recognition tests often presented in the presence of competing noise (speech-in-noise testing). These advanced protocols, like the QuickSIN or HINT tests, are designed to mimic challenging listening environments, offering a far superior prediction of hearing aid benefit and everyday communication difficulties than quiet-based word tests alone. This evolution reflects a broader trend in audiology to move from threshold-based assessment to functional assessment, providing clinicians with detailed metrics on how hearing loss impacts daily life and guiding the development of personalized amplification and assistive listening technologies. The continual refinement of speech materials and presentation techniques underscores the commitment to providing accurate and clinically relevant measures of auditory performance.

Core Components: Speech Recognition Threshold (SRT) and Speech

Detection Threshold (SDT)

Two foundational measurements within the speech audiometry battery are the **Speech Detection Threshold (SDT)**, sometimes referred to as the Speech Awareness Threshold (SAT), and the **Speech Recognition Threshold (SRT)**. The SDT is defined as the lowest intensity level, measured in decibels hearing level (dB HL), at which a patient can just barely detect the presence of speech sound, recognizing that a sound is present without necessarily understanding the content. This threshold is primarily used when a patient cannot perform the SRT task, often due to severe cognitive impairment, linguistic barriers, or profound hearing loss where no speech understanding is possible. The stimulus used for SDT is usually continuous speech or words presented slowly until the patient indicates awareness of the sound, and the resulting threshold should closely align with the best pure-tone threshold in the 500 Hz, 1000 Hz, or 2000 Hz range.

The SRT, conversely, requires the patient to not only detect the presence of speech but also to correctly identify and repeat the spoken material, thereby involving a cognitive component of recognition and understanding. For this test, **spondee words** are the standard stimuli due to their high redundancy and predictable acoustic shape, which facilitates easy recognition at low intensity levels. The procedure for determining the SRT follows a bracketing technique similar to pure-tone audiometry, seeking the softest level at which the patient can correctly repeat 50% of the presented spondee words. The SRT serves a crucial dual function: first, it provides a functional measure of threshold for complex stimuli; and second, it acts as a fundamental validation check for the pure-tone audiogram. Clinically, the SRT is expected to agree closely (within ± 6 dB) with the **Pure-Tone Average (PTA)** of the thresholds at 500 Hz, 1000 Hz, and 2000 Hz. Significant discrepancies between the SRT and PTA often suggest inconsistencies in patient response, potential malingering, or technical errors in testing.

The relationship between SDT and SRT highlights the distinction between simple acoustic awareness and linguistic processing. In most cases involving mild to moderate hearing loss, the SDT is typically 5 to 10 dB lower (softer) than the SRT, reflecting the fact that detection requires less auditory effort than recognition. However, in cases of severe or profound hearing loss, especially those with significant slope in the high frequencies, the gap between SDT and SRT may widen, or the thresholds may converge if the patient relies only on low-frequency acoustic cues for both detection and recognition. Accurate determination of both thresholds is vital for establishing the dynamic range of hearing, which is the difference between the softest sound heard and the loudest sound tolerated, informing the settings for hearing aid compression and output limits.

Evaluating Suprathreshold Abilities: Word Recognition Score (WRS)

The **Word Recognition Score (WRS)**, also frequently referred to as the Speech Discrimination Score (SDS), is arguably the most diagnostically powerful measure within the speech audiometry

battery. Unlike the SRT, which measures threshold, the WRS assesses the clarity of hearing at suprathreshold levels--that is, at intensity levels well above the patient's threshold of audibility. This test uses lists of phonetically balanced (PB) monosyllabic words, designed to contain the approximate proportion of speech sounds found in the English language. The words are typically presented at an intensity level optimized to maximize the patient's potential for understanding, known as the **Most Comfortable Loudness (MCL)** level or, alternatively, at a standard high level (e.g., 40 dB SL, sensation level above the SRT) to ensure audibility.

The WRS is expressed as a percentage of correctly identified words out of a fixed set (usually 25 or 50 words). A high WRS (90% to 100%) indicates excellent clarity and minimal distortion, even if the pure-tone thresholds are elevated, suggesting a primarily conductive component or a very mild sensorineural loss. Conversely, a poor WRS (e.g., below 70%) is often highly indicative of a sensorineural hearing loss, particularly one involving damage to the cochlea (e.g., significant hair cell loss) or, more critically, damage to the auditory nerve or central processing pathways. The reduction in the ability to discriminate speech sounds, even when presented loudly enough to be clearly audible, signifies a loss of clarity, a phenomenon known as recruitment or distortion, which significantly impairs communication.

A particularly important clinical finding derived from WRS testing is the phenomenon of **rollover**. Rollover occurs when the word recognition score actually decreases as the intensity of the presentation stimulus increases beyond a certain point. While a slight decrease at extremely high levels is normal, significant rollover (defined by specific formulas comparing maximum performance to performance at higher levels) is a hallmark sign highly suggestive of a retrocochlear lesion, such as an acoustic neuroma or other tumor affecting the VIIIth cranial nerve. In contrast, typical sensorineural hearing loss resulting from cochlear damage usually exhibits a plateau in performance once the optimal presentation level is reached, meaning the score does not improve with added volume but also does not significantly decline. The presence or absence of rollover is therefore crucial for determining the necessity of advanced neurological imaging, such as an MRI, to rule out serious underlying pathology.

Clinical Application of Loudness Measures: MCL and UCL

Beyond threshold and recognition abilities, speech audiometry provides essential metrics related to the patient's perception of loudness, specifically the **Most Comfortable Loudness (MCL)** and the **Uncomfortable Loudness Level (UCL)**, also known as the Loudness Discomfort Level (LDL). These measures define the usable dynamic range of hearing, which is critical for the appropriate fitting and programming of hearing aids. The MCL is the intensity level at which the patient reports that speech is most pleasant and easy to listen to, often determined using continuous discourse or standard speech lists, with the patient adjusting the volume or responding to volume changes directed by the clinician. Establishing the MCL is essential because it dictates the primary output

level target for speech input provided by amplification devices.

The UCL, or LDL, defines the upper limit of the patient's comfortable listening range--the intensity level at which speech becomes uncomfortably loud or painful, though importantly, not necessarily damaging. This measurement is vital for setting the maximum output limits (MPO) of a hearing aid. If the hearing aid output exceeds the patient's UCL, the user will experience discomfort and likely reject the device, regardless of the clarity provided. For individuals with sensorineural hearing loss, particularly those suffering from cochlear damage, the phenomenon of **recruitment** often leads to a dramatically reduced dynamic range. Recruitment causes sound loudness to grow abnormally fast; soft sounds may be inaudible, while moderately loud sounds quickly become painfully loud. In such cases, the difference between the SRT and the UCL can be very narrow, demanding sophisticated compression circuitry in hearing aids to manage the limited dynamic range effectively.

The relationship between MCL and UCL provides a clear picture of the patient's tolerance for loud sound. A healthy ear typically has a dynamic range of 90 to 100 dB. In contrast, a patient with significant sensorineural loss might have a dynamic range compressed to only 30 or 40 dB. By precisely measuring the MCL and UCL, the audiologist can tailor the hearing aid to compress incoming speech signals, ensuring that soft sounds are amplified to the MCL and loud sounds are limited to just below the UCL, maximizing audibility without causing discomfort. These measures transform the theoretical thresholds derived from pure-tone testing into practical, functional parameters necessary for successful auditory rehabilitation.

Instrumentation and Testing Procedures

The execution of accurate speech audiometry relies upon specialized and precisely calibrated instrumentation, primarily the **audiometer**, coupled with sound-treated environments. The audiometer must be capable of generating and controlling the intensity of the speech signal, routing it to the appropriate transducer (headphones, inserts, or loudspeakers), and monitoring the overall level of the output. Crucially, the audiometer must have a calibrated speech circuit that allows the clinician to monitor the level of the input signal, ensuring that the presentation level to the patient is accurate, particularly when using monitored live voice or when setting the level for recorded materials.

Standard testing protocols adhere to a specific sequence to ensure validity and reliability. The procedure typically begins with the measurement of the SRT using spondee words, utilizing a descending-ascending technique to accurately bracket the threshold. Once the SRT is established, it is used as a reference point for setting the presentation level for the subsequent WRS test. The selection of the presentation level for WRS is crucial; it must be high enough to overcome the hearing loss but not so high as to cause distortion or discomfort. For most clinical purposes, the

presentation level is set at 30 or 40 dB above the SRT (known as 30 or 40 dB SL, Sensation Level), maximizing the likelihood of achieving the patient's best possible score.

The method of response for speech audiometry is also standardized. For SRT and WRS, the patient is typically instructed to repeat the word they hear. If the patient has articulation or communication difficulties, alternative response modes, such as pointing to pictures or written words, may be employed, though these methods introduce potential cognitive or visual confounds. For MCL and UCL measurements, the patient usually provides subjective feedback, often using a standardized loudness scale (e.g., "too soft," "just right," "too loud"). Strict adherence to these standardized methods, including the use of recorded materials in a quiet environment, minimizes procedural variance, ensuring that observed changes in performance are attributable to changes in the patient's auditory function rather than testing inconsistencies.

Interpretation and Correlation with Pure-Tone Results

The true diagnostic power of speech audiometry emerges when its results are analyzed in conjunction with the pure-tone audiogram. This correlation provides critical cross-validation and aids in localizing the site of lesion. As previously noted, the most fundamental correlation is the expected agreement between the **SRT and the Pure-Tone Average (PTA)**. A strong correlation confirms that the patient's subjective responses are reliable and consistent across simple (tones) and complex (speech) stimuli. A significant divergence, such as an SRT that is substantially better than the PTA, often suggests non-organic hearing loss (malingering), requiring further objective testing or counseling.

Furthermore, the relationship between the degree of pure-tone loss and the WRS is highly informative. For conductive hearing losses, where sound transmission is impaired but the inner ear is healthy, the WRS is typically excellent (90% to 100%) once the volume is raised sufficiently to overcome the conductive barrier. This indicates that the clarity of hearing is preserved. Conversely, for sensorineural hearing losses, particularly those involving cochlear damage, the WRS often decreases disproportionately to the degree of pure-tone loss. For example, a patient with a moderate pure-tone loss might only achieve a WRS of 60%, reflecting underlying distortion and loss of frequency resolution caused by inner ear damage.

A key diagnostic pattern involves analyzing the shape of the **Performance-Intensity (PI) function**, which plots the WRS percentage against increasing presentation intensity. In normal hearing, the PI function rises steeply to 100% and plateaus. In conductive loss, the function is shifted to the right (requires higher intensity) but still reaches 100%. In cochlear loss, the function reaches a lower maximum score and plateaus. In retrocochlear loss, the function exhibits the characteristic rollover, where performance peaks and then rapidly declines with higher intensity. Interpreting these patterns allows the audiologist to accurately classify the type and severity of hearing loss

and distinguish between cochlear and retrocochlear etiologies, guiding subsequent medical management.

Advantages and Limitations in Diagnosis

Speech audiometry offers numerous advantages over pure-tone testing, primarily because it evaluates functional communication ability, which is the patient's primary complaint. It provides face validity, meaning the test results directly relate to the patient's ability to understand speech in daily life, making the findings easily understandable and relevant to the patient. It also serves as an essential tool for determining candidacy for and predicted benefit from amplification. For instance, if a patient has a WRS below 50%, their predicted benefit from standard hearing aids may be limited, potentially leading the clinician to recommend alternative interventions, such as cochlear implants or assistive listening devices.

Despite its clinical utility, speech audiometry is not without limitations. A primary challenge is the dependence on the patient's subjective cooperation, cognitive status, and linguistic abilities. For young children, individuals with severe cognitive impairments, or those who do not speak the test language, traditional speech audiometry is often impossible or unreliable, necessitating the use of alternative, objective measures such as Auditory Brainstem Response (ABR) or Otoacoustic Emissions (OAEs). Furthermore, standard WRS testing is typically performed in quiet conditions, which fails to capture the complexity of real-world listening environments where background noise is pervasive.

To address the quiet-condition limitation, modern audiology heavily incorporates **speech-in-noise (SIN) testing**. SIN tests present speech stimuli (words or sentences) while simultaneously introducing a standardized background noise (e.g., babble or cafeteria noise). The result is typically expressed as a Signal-to-Noise Ratio (SNR) loss, indicating how much louder the speech needs to be than the noise for the patient to achieve 50% recognition. This measure provides a far more accurate prediction of communication difficulty in noisy environments and is essential for counseling patients about realistic expectations for hearing aid performance in complex acoustic settings. Thus, while traditional speech audiometry remains crucial for establishing foundational thresholds and recognition ability in quiet, its effectiveness is maximized when supplemented by these advanced, ecologically relevant SIN protocols.