Determining Threshold Level for Speech


Index terms: speech audiometry

DOI: 10.1044/policy.GL1988-00008
About This Document

These guidelines were prepared by members of the ASHA Committee on Audiologic Evaluation: Sandra Gordon-Salant, chair; Martin S. Robinette, past chair; Carmen C. Brewer; Margaret F. Carlin; John D. Durrant; Thomas A. Frank; Gregg D. Givens; Michael P. Gorga; Carol Kamara, ex officio; Robert H. Margolis; Laura Ann Wilber; and Gilbert H. Herer, then vice president for clinical affairs and current president-elect. These guidelines replace previous ASHA “Guidelines for Determining the Threshold Level of Speech” (Asha, 1979). The guidelines were approved by the ASHA Legislative Council in November 1987 (LC 2-87).

****

Background

Historically, the first speech tests were spoken or whispered messages presented at measured distances between the talker and the listener. These tests provided a gross estimate of an individual's ability to hear speech. Clinical speech audiometry developed from a need to quantify this ability. The Western Electric 4A (later 4C) test, a phonographic recording of spoken digits, was the first widely used recorded auditory test for determining hearing losses for speech (Fletcher & Steinberg, 1929). Later, Harvard's Psycho-Acoustic Laboratory (PAL) developed word lists that serve as a basic model for today's clinical measurement of speech threshold (Hudgins, Hawkins, Karlin, & Stevens, 1947). Central Institute for the Deaf (CID) published Auditory Tests W-1 and W-2 (Spondaic Word Lists), which were modifications of the PAL lists (Hirsh, Davis, Silverman, Reynolds, Eldert, & Benson, 1952).

In 1978, the ASHA Legislative Council approved the “Guidelines for Determining the Threshold Level for Speech,” which were published in 1979 (Asha, 1979). These guidelines presented a recommended set of speech threshold procedures based on consensus. The guidelines specified spondaic words (spondees) as standard test material. Consequently, the term spondees threshold (ST) was recommended for reporting a speech threshold rather than the more general term speech reception threshold. The advantage of using spondees threshold was that the term specified the test material. When a speech threshold was obtained with material other than spondaic words, the term speech reception threshold was recommended for reporting and notation of the test material was required. These guidelines also recommended that the speech threshold be obtained using an ascending technique in 5-dB steps paralling the ASHA Guidelines for Manual Pure-Tone Threshold Audiometry (Asha, 1978). The ASHA Guidelines for Determining the Threshold Level for Speech included a list of spondaic words of which 28 were from the CID W- 1 and W-2 lists and eight were other alternatives from the original PAL lists. This modified list represented an effort to increase phonetic dissimilarity and improve homogeneity of audibility.

Several concerns have been raised about the adequacy of the 1979 ASHA Guidelines for Determining the Threshold Level for Speech. These concerns have included the fact that the recommended speech threshold procedures were not based on experimental evidence (Ventry, 1979), a clear-cut 50% speech threshold criteria could not be found using the procedures (Olsen & Matkin, 1979), the procedures were too time-consuming (Olsen & Matkin, 1979), and the guidelines did not specifically define procedures for speech awareness.
Therefore, a new set of Guidelines for Determining the Threshold Level for Speech has been developed. The guidelines define common terminology and recommend a speech threshold procedure which has been supported in terms of length, reliability, and validity by research (Seattle, Forrester, & Ruby, 1977; Huff & Nerbonne, 1982; Martin & Stauffer, 1975; Robinson & Koenige, 1979; Tillman & Olsen, 1973; Wall, Davis, & Myers, 1984; Wilson, Morgan, & Dirks, 1973). The spirit of these guidelines, as in prior guidelines, is not to mandate a single way of accomplishing the clinical process. The intention is to suggest a standard procedure that will improve interclinician and interclinic comparison of data that, in the final analysis, will benefit the people we serve.

Scope

Speech threshold audiometry is the procedure used in the assessment of an individual's threshold of hearing for speech. There are differing opinions regarding the clinical utility of this measure. For a discussion of the utility of the test, the reader is referred to Wilson and Margolis (1983).

The basic purpose of a speech threshold is to quantify an individual's hearing threshold level for speech. Clinically, the primary purpose of a speech threshold is to serve as a validity check for the pure tone audiogram. Experimental evidence obtained from 100 ears with varying degrees, configurations, and types of hearing loss (Wilson et al., 1973) indicates that minimal differences exist between thresholds for spondaic words and averages of pure tone thresholds (0.3–3.1 dB). Pure tone averages were calculated using three methods based on thresholds at 500, 1000, and/or 2000 Hz; X loss at the three frequencies (Fletcher, 1929), X loss at the two frequencies with best sensitivity (Fletcher, 1950), and X loss at 500 and 1000 Hz minus 2 dB (Carhart, 1971). Further, correlation coefficients among the speech threshold and pure tone averages were extremely high (0.95–0.98). The high correlation between thresholds for speech and pure tones, as well as the minimal differences between them, confirm the value of using the speech recognition threshold to validate the pure tone average, assuming that pure tone thresholds are obtained using the preferred clinical procedure (Asha, 1978).

Although for most clinical cases there will be good agreement between the three-frequency pure tone average and the speech recognition threshold, in cases with a sloping loss agreement may be better between the speech recognition threshold and the two-frequency pure tone average (Fletcher, 1950). Disagreement between the speech recognition threshold and pure tone average is an indication of inconsistency in test results. This inconsistency may provide an early indication of pseudohypacusis. It may also be due to test variables such as equipment malfunction or misunderstanding of the instructions by the patient. Other factors which can contribute to a discrepancy include developmental level, irregular auditory sensitivity (Berlin, Wexler, Jerger, Halperin, & Smith, 1978; Roeser, 1982), or the presence of a cognitive, language or central auditory disorder. The speech threshold can also be useful during evaluation of difficult-to-test individuals. It may be used in audiological rehabilitation, particularly in hearing aid evaluations.

If one chooses to measure a speech threshold, then these guidelines present a standardized measurement method that has been statistically validated and should yield a speech recognition threshold that agrees closely with the pure tone average. The purpose of these guidelines is to recommend a simple, rapid, statistically based descending procedure for determining the speech recognition threshold and to
define common terminology associated with its use. In some cases, however, such as suspected pseudo-hypacusis, an ascending procedure may be more appropriate and can be used. The procedures described in these guidelines are usable in a variety of clinical circumstances; however, certain individuals such as young children and mentally disabled, uncooperative, or neurologically disabled individuals may require modification of the procedure. Modifications should be noted in recording and reporting the results.

**Speech Recognition Threshold (SRT).** The speech recognition threshold is the minimum hearing level for speech (see ANSI S3.6-1969 standard or subsequent superseding standards) at which an individual can recognize 50% of the speech material. A recognition task is one in which the subject selects the test item from a closed set of choices. The individual should repeat or in some other way indicate recognition of the speech material 50% of the time. The term speech recognition threshold is synonymous with speech reception threshold. Speech recognition threshold is the preferred term because it more accurately describes the listener's task. Spondaic words are the usual and recommended test material for the speech recognition threshold, Spondaic words are two-syllable words with equal stress on both syllables. It should be noted that other test materials can be used. If so, then the test material should be noted in recording and reporting the results.

**Spondee Threshold (ST).** Spondee threshold also refers to a speech recognition threshold obtained with spondaic words. This is not a preferred term because it does not describe recognition or reception of the material.

**Speech Detection Threshold (SDT).** The speech detection threshold is the minimum hearing level for speech at which an individual can just discern the presence of a speech material 50% of the time. The listener does not have to identify the material as speech, but must indicate awareness of the presence of sound. The material used to obtain a speech detection threshold should be noted in recording and reporting the results.

**Speech Awareness Threshold (SAT).** The speech awareness threshold is a commonly used synonymous term for speech detection threshold. Speech detection threshold is the more accurate term because it specifies the listener's task.

**General Considerations for Clinical Determination of the Speech Threshold**

**Instrumentation and Calibration.** Speech threshold audiometry shall be accomplished with a speech audiometer or diagnostic audiometer capable of transducing speech as defined and calibrated according to the American National Standard Specifications for Audiometers (ANSI S3.6-1969, or subsequent superseding standards). The reader is referred to *Asha* (1987) for a review of recommended procedures for calibration of speech signals delivered via earphones.

**Test Environment.** The test environment shall meet the criteria for background noise in audiometric rooms as specified by the American National Standard Criteria for Permissible Ambient Noise During Audiometric Testing (ANSI 53.1-1977 or subsequent superseding standards).
**Test Material.** Spondaic words are the standard test materials for determination of the speech recognition threshold. An alphabetical list of 36 spondaic words, appropriate for adults, is presented in Appendix A. Appendix B presents 15 homogeneous words from Auditec recordings of CID W-1 (Young, Dudley, & Gunter, 1982). These words are homogeneous in terms of audibility in those specific recordings. Recommended procedures for determining homogeneity of other tape-recorded lists are described in Wilson and Margolis (1983) and Young et al. (1982). Numerous variables associated with monitored live voice testing renders exact specification of homogeneity impossible with this procedure.

Appendix C presents a list of spondaic words appropriate for most children aged 5 to 12 years. These words can be represented pictorially, and if so, would be appropriate for children under age 6, depending on their capabilities.

There may be circumstances or individuals that require a modification of the standard word list, in which case alternative test materials would be used. The selection of such materials should be made with consideration of the person under test. Some factors to consider are age, language facility, and the physical condition of the client. Different lists of spondaic words may be required for different clinical populations and to accommodate certain physical or developmental limitations. It should be recognized, however, that the use of speech stimuli with less homogeneity than spondaic words may compromise the reliability of this measure.

When circumstances or individual capabilities prevent determination of a speech recognition threshold, the speech detection threshold may be determined instead. The speech detection threshold (in dB) should be consistent with the best pure tone threshold (in dB) between 250 and 4000 Hz (Olsen & Matkin, 1979) and should also be obtained at levels 8–9 dB weaker than the speech recognition threshold (Chaiklin, 1959). The type of speech material is not as critical for this measure because it reflects detection and not recognition. Some common materials are speech babble, running speech, or familiar words. Nevertheless, specification of the type of material helps to ensure test-retest reliability and may be useful information for future hearing evaluations.

**Response Mode.** The usual response mode for obtaining the speech recognition threshold is repetition of the stimulus item. For many patients it is not possible to obtain verbal responses, necessitating the use of alternative response modes. Many alternatives are acceptable but must convey recognition of test items from a closed set of choices. For example, response modes can take the form of picture pointing, signing, or visual scanning. In addition, if a picture-pointing task is used for obtaining the speech recognition threshold in children, then the clinician should be cautious in choosing the number of response items. Too few items increases the probability of chance performance and too many items may be distracting and increase response time (e.g., between 8 and 12 words usually is appropriate).

For assessing the speech detection threshold, a number of response modes can be used to convey signal detection. Usually, these response modes are non-verbal. Whenever a response mode other than repetition of a spondaic word is used, it should be specified in recording and reporting the results.
Recorded vs. Live Voice Presentation of the Test Material. Either a recorded or a monitored live voice technique can be used to obtain the speech threshold. Recorded presentation of the test material is the preferred procedure. The use of recorded material standardizes the composition and presentation of the test list. It allows for better control of the intensity of the test items and ensures that the speech pattern of the recorded talker will be consistent to each client. In other words, recorded presentation is preferred because the stimuli are consistent to each patient tested with a given set of recorded test materials. It must be remembered that phonographic recordings and tape recordings become worn after extensive use, introducing distortion and noise into the test system. The audiologist must be alert to such problems and periodically replace phonograph records or tape recordings. Further, cartridge needles must be replaced periodically, and tape playback units must be cleaned and demagnetized regularly. Many of these problems may be obviated by use of digitized speech recordings.

The use of recorded test materials may limit flexibility of the test procedure in terms of selection of test words and rate of presentation. Certain clinical situations arise which favor use of monitored live voice presentation. The disadvantages of live voice presentation are the problems and difficulty in monitoring the test words to a consistent hearing level. In addition, it is impossible to present each spondaic word in the same manner to every client, even if each syllable of each spondaic word peaks at 0 dB on the VU-meter. When monitored live voice is used, it should be noted with the test results.

Recording of Results. The speech recognition threshold or speech detection threshold shall be recorded in dB HL. The results should be recorded for each ear on the same form that contains the client's results for pure tone audiometry. Additional space should be available to report other pertinent information that describes the test situation, such as alternative materials or response modes.

Masking of Nontest Ear. When the obtained speech recognition threshold (or speech detection threshold) in one ear exceeds the apparent speech recognition threshold (or speech detection threshold) or a pure tone bone conduction threshold at either 500, 1000, 2000 or 4000 Hz in the contralateral ear by 40 dB or more, masking should be applied to the nontest ear. Of course, the criteria for the use of masking should consider the signal's spectrum and transducer. The appropriate masker for a speech stimulus must have a wideband spectrum (e.g., white noise or speech-spectrum noise). The level of effective masking used should be sufficient to eliminate reception by the nontest ear without causing overmasking and should be recorded on the same form as that used to record audiometric results.

Determination of Speech Recognition Threshold
The basic procedure consists of instructions, familiarization, a single series of descending threshold determination, and calculation of threshold hearing level.

Instructions. The specific wording of the instructions must be phrased in language appropriate to the client and should
1. Orient the client to the nature of the task,
2. Specify the client's mode of response,
3. Indicate that the test material is speech and specify that the client should respond with only words from the test list, and
4. Stress the need for the client to respond at faint listening levels and encourage the client to guess.

Familiarization. The first step in the testing procedure is to familiarize the client with the exact spondaic words in the word list. This necessary step ensures that
1. The test vocabulary is familiar to the client,
2. The client can auditorily recognize each test word, and
3. The client's responses can be accurately interpreted by the clinician.

The audiologist may read the test list to the client in a face-to-face situation or present the test list through a speech audiometer. In either case, visual cues should be eliminated during familiarization with the test words. The client repeats or in some other way demonstrates recognition of each word on the list.

The audiologist should emphasize that the client is to respond only with words from the test list. If the client has any difficulty understanding or responding to any spondaic word, then that word should be eliminated from the test list. Any spondaic word that the audiologist has any difficulty understanding should be eliminated from the test list.

Familiarization with the test list is essential to control for the effects of prior knowledge of test vocabulary on the speech recognition threshold (Tillman & Jerger, 1959). It should not be eliminated from the procedure.

Descending Threshold Determination. The method described in this document closely follows the procedures originally used in PAL Auditory Test Number 9 (Hudgins et al., 1947) and CID W-2 recordings (Hirsh et al., 1952). The basic procedure, which was described by Wilson et al. (1973), involves a preliminary phase and a test phase. During the test phase, different spondaic word pairs are presented at 2 dB decrements until at least five of the last six test words are missed.

Threshold for the hearing level of speech is calculated to estimate the 50% point on the psychometric function. When variations in the standard technique are used, the nature of the variation should be recorded with the results. The current procedure is a 2 dB descending approach but can be modified to a 5 dB descending approach. (Footnotes detail modifications for using a 5 dB rather than a 2 dB descending approach.)

Descending Technique
1. Preliminary Phase to Obtain Starting Level
   a. Set the hearing level to 30–40 dB above the estimated speech recognition threshold and present one spondaic word to the client.\(^1\) If the response is correct, then descend in 10 dB decrements, presenting one spondaic word at each level until the client responds incorrectly. If the client does not respond correctly to the first spondaic word at the first level, then increase the level in 20 dB steps until a correct response is obtained (Martin & Stauffer, 1975). Then initiate the 10 dB decrements (Martin & Stauffer, 1975).

\(^1\) A lower starting level may be required in certain cases of sensorineural hearing loss where tolerance problems may exist.
1975). In cases where a speech recognition threshold cannot be estimated prior to testing, the procedure for determining a starting level reported by Martin and Stauffer (1975) may be used.²

b. When one word is missed, present a second spondaic word at the same level. Continue this process of descending in 10 dB steps until a level is reached at which two consecutive words are missed at the same hearing level.

c. Increase the level by 10 dB (above the level at which two spondaic words were missed). This defines the starting level.

2. Test Phase

a. Present two spondaic words at the starting level and at each successive 2 dB decrement.³

b. Continue this process if five out of the first six words are repeated correctly. If this criterion is not met, then increase the starting level by 4–10 dB.

c. The descending series is terminated when the client responds incorrectly to five of the last six words presented.⁴

Calculation of Threshold. Threshold is calculated by subtracting the total number of correct responses from the starting level and adding a correct factor of 1.⁵ This calculation is based on a statistical precedent (Spearman, 1908) for estimating threshold at the 50% point of the psychometric function. The calculation procedure is the Spearman-Karber method (Finney, 1952) and is given by the formula:

\[
30\% = \frac{1}{2} + \frac{d}{n} - \frac{r}{n}
\]

where \(T\) = threshold, \(i\) = initial test intensity, \(d\) = dB decrement between steps, \(r\) = correct responses, and \(n\) = number of words per dB step. While the equation may appear cumbersome, it may be reduced to the simple form:

\[
SRT = \text{starting level} - \# \text{correct} + \text{correction factor}
\]

Figure 1 presents a worksheet for tallying responses and calculating the speech recognition threshold for both 2 dB and 5 dB step sizes, using this formula.

Speech recognition thresholds determined with this procedure and calculation method provide standardization and therefore reduce variability in threshold estimates. An ascending procedure, which may be warranted in certain instances, is described in Footnote ⁶.

---

² In the Martin and Stauffer modification, the subject is presented the first spondaic word at 50 dB HL, rather than at 30–40 dB above an estimated speech recognition threshold.

³ An acceptable alternative is to present five words in 5 dB steps.

⁴ If five words are presented in 5 dB steps, then the descending series is terminated when all words at a single intensity are not correctly recognized.

⁵ For presentation in 5 dB steps, the correction factor is +2 dB (Wilson et al., 1973).
Determination of the speech detection threshold involves a detection task that is similar to the one used in pure tone threshold audiometry. Stimulus familiarization is unnecessary. Test technique should follow the one used for pure tone threshold estimation, including appropriate modification for the client's capabilities (see ASHA Guidelines for Manual Pure Tone Threshold Audiometry, 1978).

References


---

6 In an ascending procedure, start at a level where five out of five or two out of two items are missed and stop the procedure where five out of five or two out of two items are correct. Scoring is identical to the descending procedures.


Appendix A — Alphabetical List of Spondaic Words

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>airplane</td>
<td>drawbridge</td>
<td>hot dog</td>
<td>padlock</td>
</tr>
<tr>
<td>armchair</td>
<td>duck pond</td>
<td>ice cream</td>
<td>pancake</td>
</tr>
<tr>
<td>backbone</td>
<td>eardrum</td>
<td>inkwell</td>
<td>playground</td>
</tr>
<tr>
<td>baseball</td>
<td>earthquake</td>
<td>mousetrap</td>
<td>railroad</td>
</tr>
<tr>
<td>birthday</td>
<td>eyebrow</td>
<td>mushroom</td>
<td>stairway</td>
</tr>
<tr>
<td>blackboard</td>
<td>greyhound</td>
<td>northwest</td>
<td>sunset</td>
</tr>
<tr>
<td>cookbook</td>
<td>hardware</td>
<td>nutmeg</td>
<td>toothbrush</td>
</tr>
<tr>
<td>cowboy</td>
<td>headlight</td>
<td>oatmeal</td>
<td>whitewash</td>
</tr>
<tr>
<td>doormat</td>
<td>horseshoe</td>
<td>outside</td>
<td>woodwork</td>
</tr>
</tbody>
</table>

Half Lists of Spondaic Words

<table>
<thead>
<tr>
<th>List A</th>
<th>List B</th>
</tr>
</thead>
<tbody>
<tr>
<td>airplane</td>
<td>ice cream</td>
</tr>
<tr>
<td>baseball</td>
<td>mousetrap</td>
</tr>
<tr>
<td>blackboard</td>
<td>northwest</td>
</tr>
<tr>
<td>cowboy</td>
<td>oatmeal</td>
</tr>
<tr>
<td>drawbridge</td>
<td>pancake</td>
</tr>
<tr>
<td>duck pond</td>
<td>playground</td>
</tr>
<tr>
<td>eardrum</td>
<td>railroad</td>
</tr>
<tr>
<td>horseshoe</td>
<td>sunset</td>
</tr>
<tr>
<td>hotdog</td>
<td>whitewash</td>
</tr>
<tr>
<td></td>
<td>headlight</td>
</tr>
<tr>
<td></td>
<td>backbone</td>
</tr>
<tr>
<td></td>
<td>inkwell</td>
</tr>
<tr>
<td></td>
<td>birthday</td>
</tr>
<tr>
<td></td>
<td>mushroom</td>
</tr>
<tr>
<td></td>
<td>nutmeg</td>
</tr>
<tr>
<td></td>
<td>outside</td>
</tr>
<tr>
<td></td>
<td>padlock</td>
</tr>
<tr>
<td></td>
<td>stairway</td>
</tr>
<tr>
<td></td>
<td>toothbrush</td>
</tr>
<tr>
<td></td>
<td>woodwork</td>
</tr>
</tbody>
</table>

These spondaic word lists are revisions of the Central Institute for the Deaf W-1 Auditory Tests, emphasizing the criteria of dissimilarity and homogeneity of audibility (Asha, 1979).
Appendix B — Streamlined Spondaic Word List for Adults

inkwell  woodwork grandson mousetrap
playground baseball eardrum drawbridge
sidewalk workshop toothbrush padlock
railroad doormat northwest

Spondaic words suggested by Young et al. (1982).
Appendix C — Children's Picture Spondaic Word List

- cupcake
- toothbrush
- popcorn
- flashlight
- airplane
- bathtub
- fire truck
- bluebird
- baseball
- ice cream
- mailman
- toothpaste
- cowboy
- shoelace
- snowman
- reindeer
- hotdog
- football
- sailboat
- seesaw

These 20 spondaic words include those recommended by Frank (1980).