It’s been almost thirty years since I administered a word recognition test (back then it was called speech discrimination). In many ways, not much has changed. The majority of audiologists continue to administer 25-word WR test lists via monitored live voice (MLV) at a single sensation level to determine if the WR scores are consistent with the audiometric configurations of each ear, the type of hearing loss, and if the results are symmetrical. Our otolaryngology colleagues use the results of the WR test as one variable to decide if the patient should be referred for an MRI. In fact, the American Academy of Otolaryngology has published guidelines to determine when asymmetry in cases of SNHL warrant an MRI (more on that later).

The question is, by following these long employed methods are we doing so with best and evidence-based practices in mind? I would say that if a half-list of a 50-word NU-6 or CIDW-22 list is administered at a fixed sensation level above the PTA or SRT and referenced data on what constitutes an abnormal WR test is not consulted, then the answer is a resounding NO!

Let’s take a closer look at the following variables:

- What’s the best level at which to administer WR tests?
- What test material should be used?
- How to save time while assuring validity?
- Test interpretation

**Word Recognition Test Level**

Generally, WR testing is completed at a fixed sensation level above the SRT (sometimes PTA). That SL is either SRT +30 or SRT +40. Unfortunately, testing at the SRT plus a fixed sensation level may not achieve the best score for that ear and that patient. And that should be the goal of the WR test...to find the best possible WRS regardless of the presentation level, provided it does not exceed the patient’s tolerance level for speech.

It all comes down to audibility. Depending upon the configuration of the audiogram, the words may not be presented at a level that achieves adequate audibility. By increasing the presentation level and, hence, audibility of the words, could the WRS actually improve? Many studies have shown that to be the case.

We all know about the PI-PB function. In fact, it was once a common metric employed in patient’s suspected of having a retrocochlear disorder. Providing the patient maximum audibility to achieve their best possible WRS should be the goal of every WR test, however,
there just is not enough time to administer the WR test at more than one presentation level per ear. So what is the single “best” presentation level to maximize the WRS?

It’s not the STR + 30 or 40dB method! According to Dr. Ben Hornsby, the SRT +40dB presentation level was based on early research conducted with people with normal hearing or, at worst, a mild to moderate hearing loss. But it has been clearly demonstrated that this method does not guarantee that the WRS measured represents the best score.

In an article published in the *Journal of the American Academy of Audiology* in 2009, Leslie Guthrie and Carol Mackersie investigated presentation levels required to maximize word recognition scores. They evaluated several test levels that included the more traditional SRT SL method as well as a fixed intensity level method. Of the several methods they evaluated, the two that yielded the highest mean scores were UCL-5dB and the 2KHz Sensation Level methods.

Of the two, the 2KHz SL method seems more suitable in very busy clinics since it does not require the audiologist to make an additional measurement (in the UCL-5dB method the LDL for speech must be determined). The 2KHz SL method is based on the 2KHz pure tone AC threshold plus a variable sensation level as follows:

- 2KHz threshold <50dB HL:  25dB SL
- 2KHz threshold 50-55dB HL:  20dB SL
- 2KHz threshold 60-65dB HL:  15dB SL
- 2KHz threshold 70-75dB HL:  10dB SL

The article is available by clicking on the following link:

https://www.dropbox.com/s/9ios9rddln0fzzs/JAAA_20_06_05.pdf?dl=0

**Test Material and Efficient Use of Time**

Empirically, it seems that most audiologists use one of the NU-6 monosyllabic word lists developed at Northwestern University when administering a WR test. The NU-6 test is comprised of four lists of 50 words each. Some audiologists may utilize the CID W-22 lists when doing the WR test.

Extensive research time was invested into the development of both the NU-6 and CID W-22 tests with the intent that an entire list of 50 words be presented to each ear. Due to time constraints imposed upon audiologists, especially in busy otolaryngology practices, most WR tests are conducted by using a half-list, or only 25 words. In other words, the first half of a 50-word list is presented to the right ear and the second half of the 50-word list is presented to the left ear. Makes sense, correct? Well, it might if the degree of difficulty of the first and second half of a 50-word NU-6 list was the same, but it isn’t. In fact, it has been documented that the first half of these lists is populated by more difficult words compared to the second half of the lists. This is a problem.
Word Recognition Testing
"Repeat After Me..."

For example, let’s say you are testing a patient with the following WR test results: 76% for the right ear and 92% for the left ear. Is this poorer score in the right ear due to a true abnormality or a result of the more difficult words populating the first half of the list? There really is no way to answer this question without administering a full 50-word test. There are several other questions that need to be considered based upon these scores but more on those later.

In a study by Rintelman and colleagues in 1974, they compared performance between ears when a half-list was used (25 words). They found the average difference between ears was 16%. The point is if you are administering only a half NU-6 list (25 words), the test itself is not valid. Therefore, it is not possible to interpret the findings with any level of confidence that it is accurate and truly representative of the patient’s performance.

So is there any way that this issue can be overcome? One way is by using a shortened version of the NU-6 lists that are ordered by how difficult the words are which functions as a WR “screening test”. Fortunately, each of the four 50-word NU-6 lists has been re-ordered based upon the difficulty of the words, starting from the hardest to the easiest. This task was undertaken in 2003 by Hurley and Sells. The objective of their research “was to develop a test methodology that would identify those patients requiring a full 50 item word recognition test and allow abbreviated testing of patients who do not need a full 50 item word recognition test”. Their research resulted in the re-ordering of the NU-6 lists from the most difficult to the easiest word and the development of 10-word and 25-word screening tests.

Their findings can be summarized as follows:

1) The four NU-6 50 word lists were equivalent in item difficulty;
2) The four NU-6 10 word and 25 word lists were equivalent in item difficulty; and
3) The 10 word and 25 word screening tests have hit rate (HR) values of 93 to 100%, false alarm rate (FAR) values of 0 to 20% and A' values of 0.946 to 1.00 depending upon word configuration and pass criterion.

They concluded that “the four NU-6 10 word and 25 word screening tests differentiated listeners with impaired word recognition ability who required a full 50 word test from listeners with unimpaired word recognition ability who required only a 10 word or 25 word test”. With clinical research data to support your position, this sounds like an excellent way to reduce test time while maintaining validity and accuracy. What more can one ask for?!!

So here’s how it works...In the case of the 10-word screening list, the patient may only miss one word. Provided no more than one word is missed, this patient would be expected to score 96% or better on a 50-word list. However, if two or more words are missed, then it is necessary to expand the test to 25 words. The patient may miss a maximum of three words in order to pass the 25-word screening test. This result would also be indicative of a 96% or better score on a 50-word list. Finally, if more than three words are missed, the full 50-
word list is required. In the Hurley and Sells study, approximately 25% of their patients did not need the full 50-word list. They estimated this approach can save around one hour of test time for every ten patients tested. I would say that is significant!

To summarize, WR tests are invalid if administered using only half of the standardized 50-word lists. Doing so can have implications regarding diagnosis and decisions concerning amplification. Horsnby and Mueller (2013) recommend either administering a full 50-word list or using the aforementioned approach from the Hurley and Sells article. Either of the above will ensure a valid measure of word recognition.

**Efficient Use of Clinical Time: Test Material and Recorded Word Lists**

Time is precious and always a consideration in a busy clinical practice. Audiologists need to be efficient while at the same time ensuring test validity. The administration of the word recognition test is a portion of the basic audiologic assessment that can have a significant impact on the time necessary to complete the assessment.

One way to save time is to follow the approach mentioned in the previous section; that being the Hurley and Sells Auditec NU-6 Ordered by Difficulty list (Version II). The data from their 2003 study is strongly supportive of this method, saving time while at the same time yielding valid results.

But this requires the use of recorded word lists! UGH! Doesn’t that add time to the assessment compared to Monitored Live Voice (MLV)? Not necessarily!

Hornsby and Mueller (2013) report that a 50-word recorded word list presented via CD media, takes approximately three minutes per ear, or six minutes for both ears. Therefore, conducting half lists via CD, would take only three minutes, or a 50% time savings. Over ten patients, that equals a time savings of thirty minutes!

This leads to a very important issue in the administration of WR tests, that being whether to present word lists from commercially available recorded lists or via MLV. Most audiologists know it is better to use recorded lists, however, the majority continue to use MLV in order to save time. Well, that issue has been addressed by storing the Auditec and other commercially available word lists in wave file format either on the hard drive of a PC controlling an audiometer (e.g., the Interacoustics Equinox/Affinity, MedRx A2D/Stealth) or into the memory of newer stand-alone clinical and diagnostic audiometers (e.g., the Interacoustics AD629/AC40; Grason-Stadler GSI Audiostar Pro).

The use of wave files stored on a computer or newer advanced clinical audiometers provide the audiologist with a method to use standardized recorded lists at a rate as fast as using MLV as the presentation method. How so? By setting up your WR test protocol so that the next word is presented immediately after the previous word is electronically scored by the audiologist. This allows the words to be presented as fast as the patient can respond without having to pause and restart a CD recording. Fast, efficient and VALID! This should
Word Recognition Testing
"Repeat After Me..."

be a no-brainer, provided you have a PC based audiometer or one of the newer stand-alone audiometers.

Interpreting Word Recognition Tests

So now that we have answered questions about test level, test validity and test time, let’s address the issue of interpretation of WR tests.

As mentioned earlier in this article, WR tests are administered for the purpose of determining if the ears perform symmetrically or not and impact management decisions about additional testing or amplification. This is one of the triggers ENTs use to determine if a patient needs to be referred for additional testing (MRI) to rule out a retrocochlear disorder, such as an acoustic tumor. Did you know that the American Academy of Otolaryngology published a guideline for this purpose? The guideline states that a patient should be referred for additional tests if the difference between ears on the WR test is >15%. 15%! Does this degree of asymmetry really suggest a potential abnormality? Assuming your test methods and material ensure validity, the answer is NO! A 15% difference in most cases is probably not a significant asymmetry.

So what does constitute a significant asymmetry when conduction WR testing? Again, referring to the Hornsby and Mueller Audiology OnLine presentation, there are statistical methods to determine what is a significant asymmetry or not. They refer to an article published in 1978 by Thornton and Raffin. They developed a statistical model against which one can plot WR scores to determine what is statistically significant based on “critical difference tables”.

Essentially, it is quite simple. The audiologist finds one of the two WR scores in one column and then finds the “normal” WR score range for the other ear. For example, let’s say the right ear WR score is 76% and a left ear WR score of 58%. The question, is this a significant difference? Using the Thornton and Raffin table, as long as the left ear score fell between 58% and 92%, the result would not be considered significantly different. Using the AAO guidelines in this example would lead the physician to unnecessarily refer this patient for an MRI! And the above result interpreted without the benefit of the critical difference tables may lead to a recommendation for amplification that may not be valid. According to Hornsby and Mueller, this method can also be used to determine if WR is changing over time.

Finally, to better interpret WR test data, I recommend using a chart that was developed by Linda Thibodeau in 2007 known as the SPRINT which is short for SPeech Recognition INTerpretation. This is a chart that incorporates the Thornton and Raffin data from 1978 along with data obtained in a 1995 study by Dubno et al.

The Dubno data provides a statistical method to determine if a WR score is considered to be close to PB Max. This is done by plotting the PTA against the WR score for the same ear. If the intersected point falls within a specific region, the WR score is considered
Word Recognition Testing
"Repeat After Me..."

disproportionately low, therefore requiring a retest at a higher presentation level.

The two WR scores can also be plotted on the SPRINT chart to determine if the difference in WR scores between the right and left ear falls within the 95% critical difference range as per the Thornton and Raffin data.

This is an exceptional tool that can be used to assist in the interpretation of WR test results. There should be no question about what constitutes asymmetry or if you have obtained a PB max score.

**Conclusions**

After taking the Audiology OnLine course presented by Hornsby and Mueller in 2013, I realized how much I didn’t know about WR testing and the many variables that impact the clinical value of this test. I strongly recommend that you obtain the transcript of this course (unfortunately, the recording of this course is no longer available on AOL). It will open your eyes and cause you to reconsider how you administer and interpret the WR test.

I will close by providing the actual summary of the Hornsby/Mueller presentation. It is as follows:

**THINGS TO DO:**

- Always use recorded materials. We recommend the shortened-interval Auditec recording of the NU-6
- Choose a presentation level to maximize audibility without causing loudness discomfort. That’s either UCL-5 or the 2000Hz –SL method.
- Use the Thornton and Raffin data to determine when a difference really is a difference (included in Thibodeau’s SPRINT chart)
- Use the Judy Dubno data to determine when findings are “normal” (included in Thibodeau’s SPRINT chart)
- Use the Ordered-by-Difficulty version of the Auditec version of the NU-6 and use the 10-word and 25-word screenings.

**THINGS NOT TO DO:**

- Live-voice testing
- Use a presentation of SRT+30 or SRT+40
- Make random guesses regarding when two scores are different from each other
- Conduct one-half list per ear testing (unless using the Ordered-by-Difficulty screening)
Word Recognition Testing
"Repeat After Me..."

References:


About e3 Diagnostics:

Expertise, Excellence, Everyday: The e3 Difference

e3 Diagnostics is comprised of thirty-six (36) full service offices nationwide - your network of local experienced experts, working in the field of Audiology instrumentation for decades. Our experience allows us to set the standard for the highest level of expertise in the hearing and balance field. The e3 Team makes a commitment to apply our knowledge of the industry to assist our customers from assessment to purchase, installation and training, to calibration, maintenance and repair.

We strongly believe you should have a choice, and we know how the small details make a difference to you! Our experts have a deep understanding of you and your needs. With e3 Diagnostics, you get more than just a partner; we are a family of people, dedicated to helping you provide only the very best to your patients.

To find out more about e3, visit [www.e3diagnostics.com](http://www.e3diagnostics.com).

e3 Diagnostics
3333 N. Kennicott Ave
Arlington Heights, IL 60004-1429
Toll-free: (800) 323-4371
Fax: (847) 459-7646