Using a more ecologically valid test to measure hearing aid outcomes

Skylar McSorley, Yu-Hsiang Wu MD/PhD, Elizabeth Stangl AuD
University of Iowa

INTRODUCTION

—Hearing aid outcomes are most often measured in one of two ways: Self-report questionnaires and tests in laboratory settings.

—Studies have shown that these two methods often do not match.

—Our lab is looking for a better way to measure hearing aid outcomes that would create more consistency between these two methods.

—One possibility is that lab tests are too contrived, so using material that better reflects real world conditions in the lab would better reflect real world outcomes.

—The Lectures, Interviews, and Spoken Narratives (LISN) Test was developed to assess extended speech comprehension, similar to the way we listen to conversation in everyday life.

—Our goal was to understand if this new test is more ecologically valid than traditional tests like the American Four Alternative Auditory Feature (AFAAF) Test, a well-established and commonly used test of word recognition.

—The AFAAF has shown increased word recognition accuracy with hearing aids, so it is considered sensitive enough to detect the effects of amplification.

—If the AFAAF and LISN show comparable results in comparable conditions, then the LISN is an ecologically valid outcome measure that should also be able to show performance differences when individuals use amplification.

METHODS

Participants

—25 adults

—Ages 64-83 years (mean = 72.5 years; SD = 4.47 years)

—12 males and 13 females

—All had bilateral, mild-to-moderate sensorineural hearing loss and were experienced hearing aid users

Procedure

—Participants were fit with Starkey 1200 behind-the-ear hearing aids

—Each person completed 2 tests: the LISN (for listening comprehension) and the AFAAF (for word recognition)

—Tests were both conducted in background noise (+8 signal-to-noise ratio, or SNR) so that hearing aids would show the same behavior in both tests

Figure 1: Composite audiogram for participants

—Participants completed one version of each test in three conditions: unaided, with DNR processing off, and DNR processing on
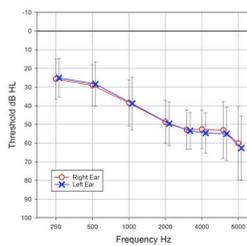
What is DNR processing?

Digital noise reduction is a programming strategy in hearing aids designed to detect noise in the background and reduce its intensity level. Previous research has shown mixed results as to whether it can reduce participants' levels of speech understanding and listening effort.

METHODS

LISN Administration

—Three versions of the LISN were administered, one for each condition: unaided, DNR off, and DNR on

—Each version included:
- 2 lectures, 2 interviews, and 2 narrative passages 2-4 min in length
- 6 multiple choice questions per passage, including 2 information, 2 inference, and 2 integration questions

—Participants answered 2 subjective questions about their comprehension and effort following each passage (shown in Figure 3)

Narrative Passage Sample: *"...In 1912 he was campaigning for a third term as president as an Independent candidate, the Bull Moose candidate, and he was in Milwaukee towards the end of the campaign and he was climbing into his car in front of his hotel on his way to give a speech. And as he got into the car, he turned and suddenly a man in the crowd pulled out a gun and shot him in the chest. And as Roosevelt later said, he felt like he'd been kicked by a horse and fell back into the car. Then he remembered what he'd been told once when he was in the army, to cough up blood and see if he'd been wounded internally. So he did and he did not cough up blood, so he knew that it was not a very serious wound..."*

Narrative Passage Topic: The assassination attempt on Theodore Roosevelt
Information: What was Teddy Roosevelt campaigning for when he was shot?
Integration: After being shot, which of the following did Teddy Roosevelt do first?
Inference: What was significant about John Schrank's death coinciding with Franklin Roosevelt's third presidential win?

AFAAF Administration

—Three versions of the AFAAF were also administered, one for each condition: unaided, DNR off, and DNR on

—Administration of each test version included:
- 5 practice words and 80 test words in 20 groups of 4
- A recording of test words inside the carrier phrase, "Can you hear ___ correctly?"
- Participants clicked on each word they heard from 4 similar choices on a computer screen

—Participants also answered the question from the LISN about their level of effort at random intervals during each version of the test

"Can you hear shoe correctly?"

shoe   Sue

stew   groom

Figure 2: Example of a trial from the AFAAF test

How much speech did you understand while listening to the passage?
0 10 20 30 40 50 60 70 80 90 100
Nothing / Everything

How hard did you have to work to understand the speech in the passage?
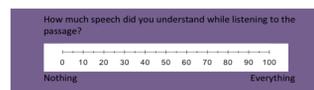0 10 20 30 40 50 60 70 80 90 100
Very, very hard / Not at all

Figure 3: Subjective listening questions for the LISN. The second question is very similar to the effort question used in the AFAAF as well.
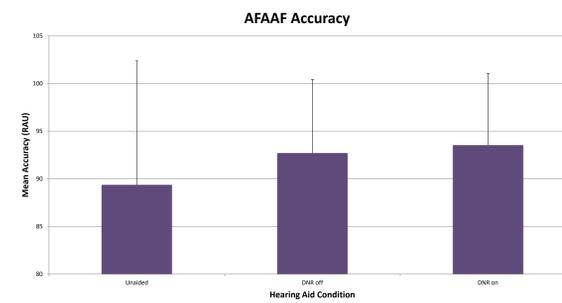
RESULTS

AFAAF Accuracy

Figure 4: In this graph, lower scores reflect higher accuracy because the AFAAF test is scored based on how many items each participant answers incorrectly. These data suggest a significant effect of amplification on word recognition accuracy. Data bars represent the mean level of accuracy in each condition on the word recognition tasks.

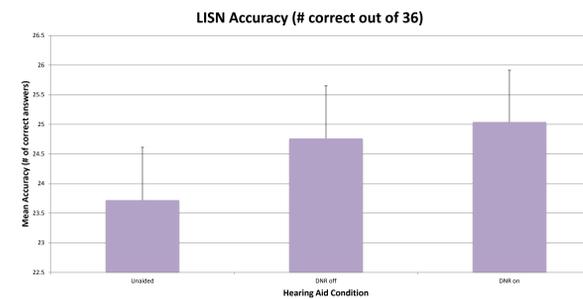**LISN Accuracy (# correct out of 36)**

Figure 5: On average, participants answered 1.32 more questions correctly in the DNR condition than the unaided condition. However, these data suggest that there is no significant relationship between hearing aid condition and accuracy on the listening comprehension questions of the LISN test. The data bars represent the number of questions each participant answered correctly.

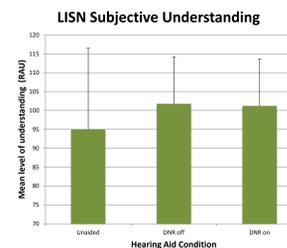**LISN Subjective Understanding**

Figure 6: These data suggest that there was a significant relationship between how subjects rated their level of understanding and hearing aid condition when we compare the DNR on condition to unaided. A lower value in this case represents that a lower portion of the passages were understood. The data bars represent the average level of understanding reported by the participants for each condition, in RAU.
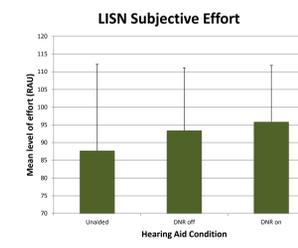
LISN Subjective Effort

Figure 7: These data suggest that there was a significant relationship between how subjects rated their level of effort and hearing aid condition when we compare the DNR on condition to the unaided condition. A lower value in this case represents a higher level of effort, meaning that with hearing aids on, subjects reported they did not have to work as hard to understand the passages The data bars represent the average level of effort reported by the participants for each condition, in RAU.

DISCUSSION

—We examined the sensitivity of the LISN test in detecting the effects of amplification and DNR processing on listening effort and comprehension.

—Our results indicated a significant increase in word recognition accuracy while wearing hearing aids on the AFAAF test. This is consistent with previous research.

—Results also indicated that the LISN results did not reflect a significant increase in accuracy on answering listening comprehension questions while participants were wearing hearing aids.

—At this point, the objective portion of the LISN is not sensitive enough, but the subjective portion of the test is sensitive to these differences.

—This suggests that a new way to use the LISN, such as administering only the subjective questions, might be more valid.

—It is possible that the benefit hearing aids provide for word recognition, as shown in the AFAAF results, does not translate to extended speech comprehension passages.

—Because listening comprehension requires different cognitive processes than word recognition, other factors could reduce or eliminate the benefit:
    Attention? Memory? Interest? Prior Knowledge? Context?

—The benefit that hearing aids provide at the phoneme and word levels may not translate to overall comprehension, which would explain why we do not observe a decrease in listening effort.

—Future studies should continue to evaluate the sensitivity of similar measures and explore the effects of higher level processing on listening effort.

CONCLUSIONS

—The objective portion of the LISN cannot detect the effects of amplification on listening comprehension.

—The subjective portion of the LISN could potentially be used to determine levels of understanding and effort more accurately.

—More research on the effects of DNR processing is necessary, especially with larger groups of participants.

—We are still looking for ecologically valid ways to measure listening effort that match up with results obtained in the lab.

CONTACT FOR MORE INFORMATION

— Yu-Hsiang Wu: yu-hsiang-wu@uiowa.edu

REFERENCES

— Best, V., Keidser, G., Buchholz, J. M., & Freeston, K. (2015). Development and preliminary evaluation of a new test of ongoing speech comprehension,. *International Journal of Audiology*, 1-8.
— Hornsby, B. (2012). 20Q: Hearing loss, hearing aids, and listening effort. *AudiologyOnline*. Retrieved from http://audiologyonline.com/articles/20q-hearing-loss-aids-and-788.
— Johnson, J., Xu, J., Cox, R., & Pendergraft, P. (2015). A Comparison of Two Methods for Measuring Listening Effort As Part of an Audiologic Test Battery. *American Journal of Audiology 24*, 419-431.
— Seeman, S. & Simms, R. (2015). Comparison of psychophysical and dual-task measures of listening effort. *Journal of Speech, Language, and Hearing Research*. Retrieved October 2, 2015 from http://jslhr.pubs.asha.org/.
— Sommers, M. S., Hale, S., Myerson, J., Rose, N., Tye-Murray, N., & Spehar, B. (2011). Listening Comprehension Across the Adult Lifespan. *Ear & Hearing, 32*(6), 775-781.
— Wu, Y., Stangl, E., Zhang, X., & Bentler, R. A. (2015). Construct Validity of the Ecological Momentary Assessment in Audiology Research. *J Am Acad Audiol., 26*, 872-884.
— Xu, J., & Cox, R. M. Recording and Evaluation of an American Dialect Version of the Four Alternative Auditory Feature Test. *University of Memphis*.

THE UNIVERSITY OF IOWA