Validation of the Spatial Fixed-SNR (SFS) test in anechoic and reverberant conditions

Niels Søgaard Jensen[1], Søren Laugesen[1], Filip Marchman Rønne[1], Christian Stender Simonsen[2], Jens-Christian Britze Kijne[2]

[1]Eriksholm Research Centre, Rørtangvej 20, DK-3070 Snekkersten, www.eriksholm.com

[2]Oticon A/S, Kongebakken 9, DK-2765 Smørum, www.oticon.com

Contact: Niels Søgaard Jensen, nsj@eriksholm.com

Scan to access poster pdf

Presented at the 7th Speech in Noise (SpiN) Workshop, Copenhagen, Denmark, January 8-9, 2015

The SFS test addresses the issues of 'SNR confounds' and testing at unrealistically low SNRs, which may occur when adaptive speech intelligibility tests are used in comparative studies of hearing aids

Introduction

Facts: Adaptive speech intelligibility tests such as the HINT [1] or the Hagerman (Matrix) test [2] are frequently used in assessment of hearing aids (HAs). These tests are easy to implement and perform, they are sensitive, and they provide reliable results.

Problem: Since different test subjects typically have different speech reception thresholds (SRTs), they may perform the test at different signal-to-noise ratios (SNRs). This may introduce a 'SNR confound' [3] if the HA functionality under test depends on the SNR. Furthermore, the test SNRs may be substantially lower than those encountered in everyday life [4]. This may be problematic if the HA functionality is optimised for higher and more realistic SNRs. If the test outcome is confounded by the SNR, the test validity is reduced despite the high reliability.

Solution: We propose the Spatial Fixed-SNR (SFS) test [5]. For each test subject, an individual test set-up is prescribed, which brings speech intelligibility at a given fixed SNR within the range of approximately 20-80% words or sentences correct (preferably the upper end of the range, above 50% correct). This means that all test subjects can be tested at the same SNR without floor or ceiling effects.

The SFS test

In a previous study [6], three appropriate 'SRT manipulators' for an adaptive speech-in-speech test were identified:

- Changing the spatial target-masker separation
- Changing the gender of the masker talkers
- Changing between scoring words and sentences correct

Each manipulator's effect on the SRT was assessed in a set-up implemented in an anechoic chamber (see Figure 1), and this led to the nomination of four test conditions to be included in a validation study of the SFS test (see Figure 2). The ideal SFS test flow in a comparison of different HAs is shown in Figure 3.

Figure 1. Set-up used for SFS test.

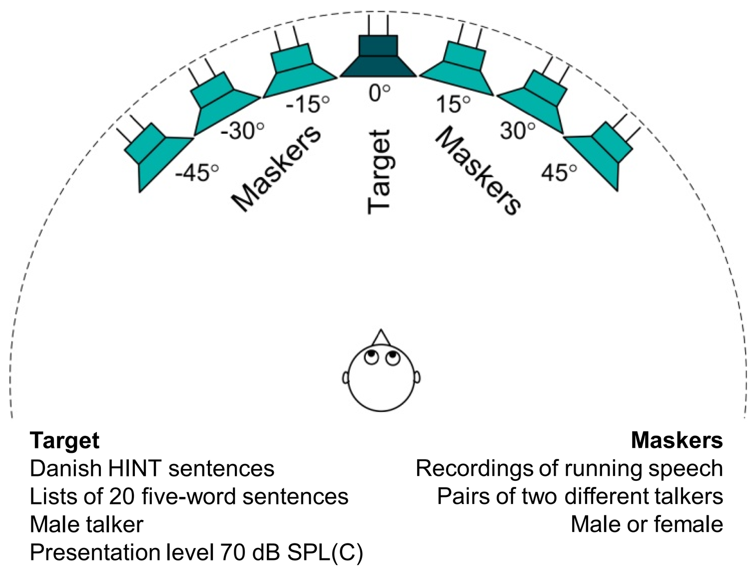


Figure 2. The four SFS test conditions and their nominal effects (re. baseline SRT).

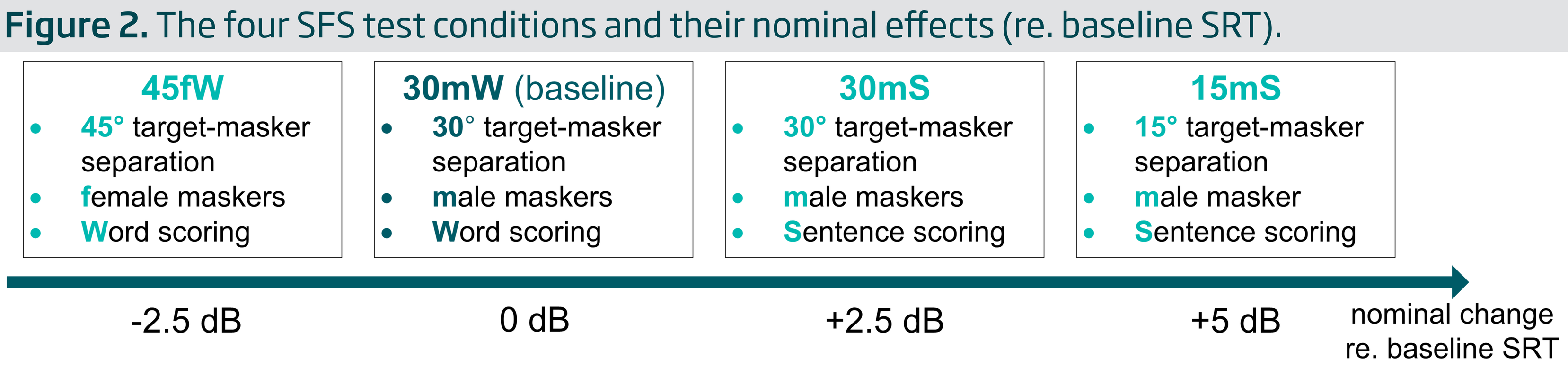
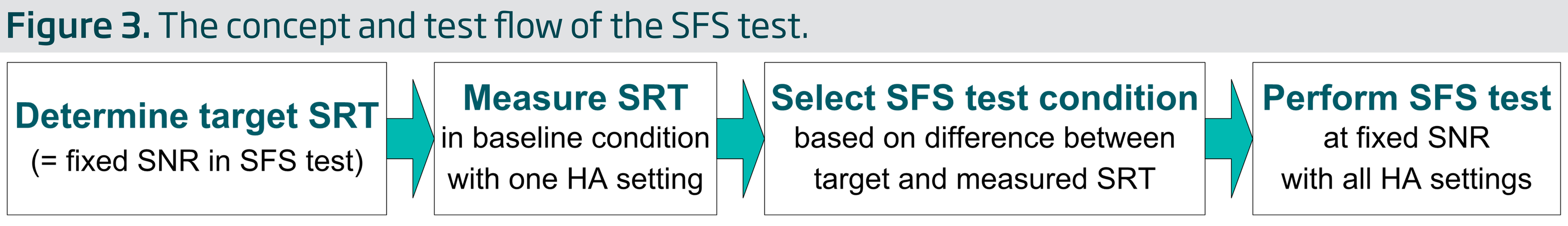


Figure 3. The concept and test flow of the SFS test.



Aim of validation studies

To validate the SFS test, with special focus on the questions:

- Do the four SFS conditions change test difficulty as expected?
- Can the SFS test measure an expected experimental contrast?
- What is the reliability of the SFS test?
- Does the SFS test run in both anehoic and reverberent conditions?

Methods

Two separate studies were carried out, validating the SFS test in two different environments, an anechoic chamber (ANECHO), and a soundtreated listening room (REVERB) with a reverberation time of around 0.3 seconds. The characteristics of test subjects and the experimental contrasts used in the two studies are listed in Table 1.

Table 1. Description of test subjects and experimental contrasts used in the two studies.

	ANECHO		REVERB	
Number of test subjects	26 (11 male, 15 female)		19 (10 male, 9 female)	
Age (years)	range: 42-79; mean: 65; SD: 11		range: 48-82; mean: 71; SD: 9	
PTA hearing loss (dB HL)	range: 29-66; mean: 46; SD: 12		range: 34-60; mean: 46; SD: 6	
Experimental contrast	Linear gain (LIN) vs. Aggressive Compression Limiting (CLM) (in Oticon Agil Pro miniRITE)		Reference 10 kHz bandwidth (REF) vs. Low-Pass at 2.5 kHz (LP) (in Oticon Alta Pro RITE)	
Target SNR	-5 dB (N=12)	+5 dB (N=14)	0 dB (N=5)	+3 dB (N=14)
Hypothesis	CLM better than LIN	LIN better than CLM	REF better than LP	

All hearing aids were fitted bilaterally. Whereas the experimental contrast used in the ANECHO study was expected to introduce a SNR confound, the experimental contrast in the REVERB study was expected to be significant, but independent of the SNR. The fact that two target SNRs were used in the REVERB study was only due to a non-optimal initial choice of target SNR for the first five test subjects.

The two studies used different test protocols. The differences reflected that the ANECHO study focused on assessment of the experimental contrast, whereas the REVERB study focused on assessment of the SRT manipulators.

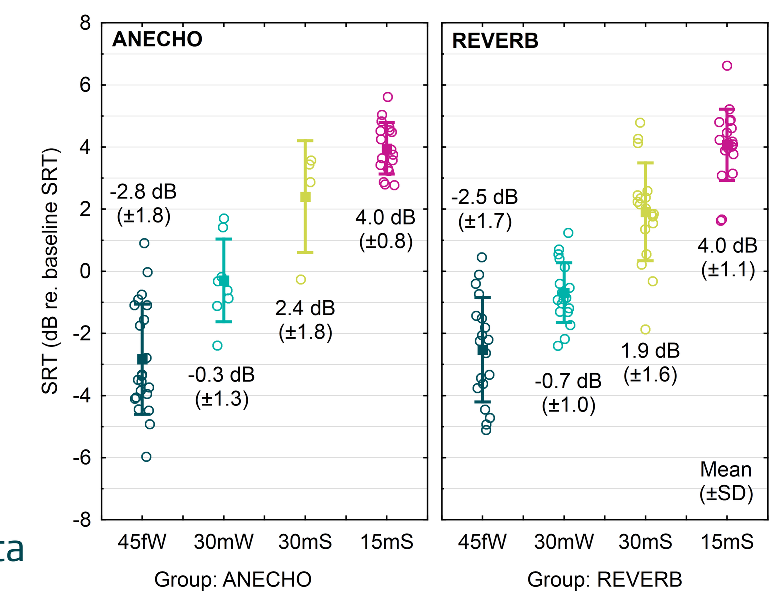
In both studies, test subjects were trained in the task before actual testing was initiated. The order of test conditions and HA settings was counterbalanced to avoid order effects.

Results

Effects of SRT manipulators

- Quite similar mean effects were observed in the two environments (Figure 4).
- Some deviations from nominal values, mainly in the 15ms condition (Figure 4).
- Slightly negative 30mW mean effects were most likely due to training (Figure 4).
- Use of SRT manipulators moved SRTs closer to the target SNRs (Figure 5).

Figure 4. The SRT-manipulator effects of the four SFS test conditions measured in the two test environments.



Effects of experimental contrasts

- Differences between HAs were detected with high stat. significance (Figures 5-6).
- The SNR confound in the ANECHO study was not as clear as expected, and only significant in SFS test data (Figures 5-6).

Figure 5. SRT measured adaptively in the baseline (left) and SFS test (right) conditions in the ANECHO study.

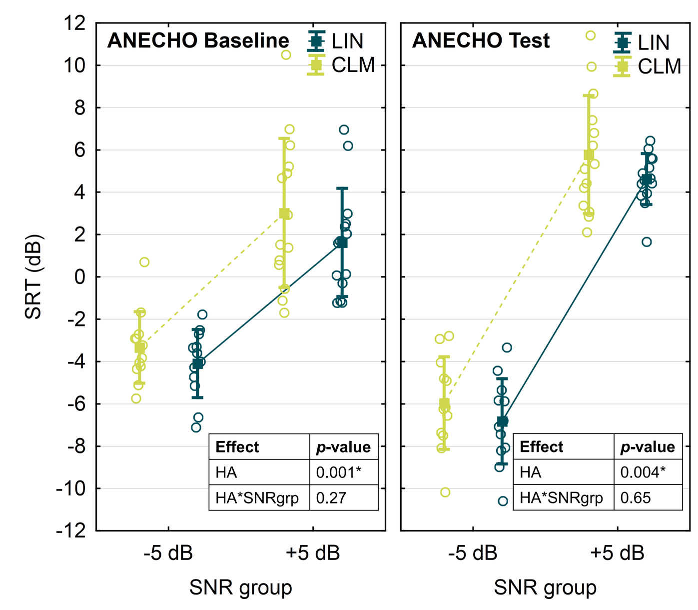
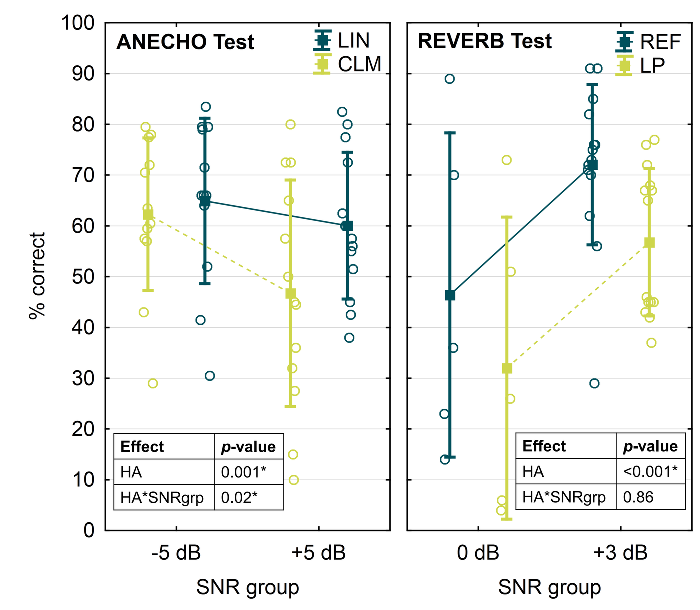


Figure 6. SFS test results obtained in the ANECHO (left) and REVERB (right) studies. 'SNR group' indicates the fixed test SNR.



Effects of hearing loss

- SRT increased with PTA, as expected (Figure 7).
- Almost parallel regression lines for the four SFS conditions were observed in the REVERB study (Figure 7), indicating no effect of hearing loss on the SRT manipulator effects.
- REVERB and ANECHO regression lines have quite similar slopes but are offset by approximately 3 dB (Figure 7). With REF and LIN gains being quite similar with the given stimuli, this indicates that the test is more difficult in reverberant conditions.

Figure 7. SRT plotted as function of PTA for the LIN (ANECHO) and REF (REVERB) settings, measured in the baseline condition, with regression lines added.

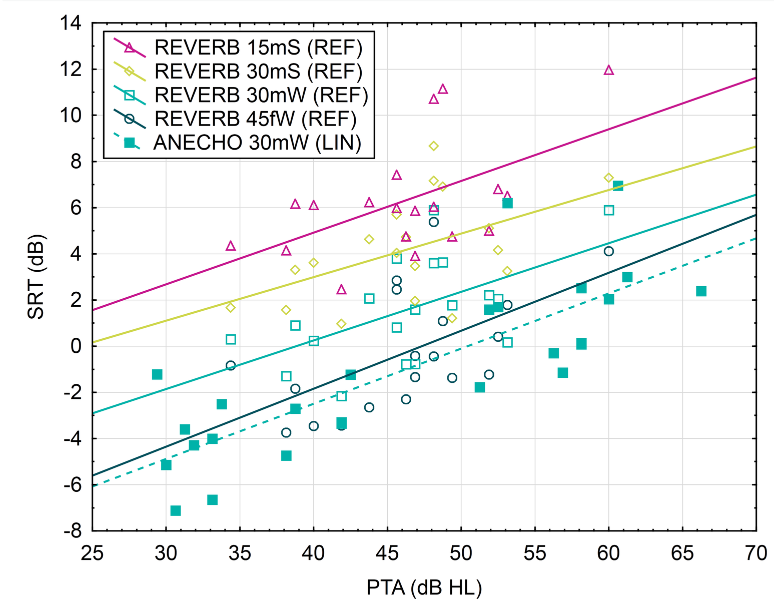
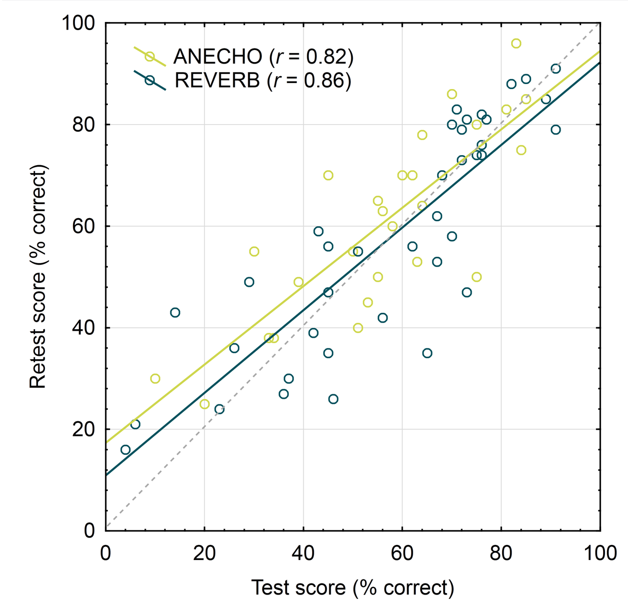


Figure 8. Retest SFS data plotted as function of test data in the two environments, with regression lines added.



Test-retest reliability

- Strong correlations between test and retest SFS data, r > 0.8 in both studies (Figure 8).
- Test-retest standard deviations were 8.6% (ANECHO) and 8.5% (REVERB).

Discussion

- Use of the SRT manipulators allowed participants in both studies to perform the SFS test at the target SNRs with the vast majority of data points within the desired range of 20-80% (and none at floor or ceiling). The observed overall 3-dB difference between SRT results in the two environments is likely due to detrimental effects of reverberation.
- The observed test-retest SD of 8.5% corresponds to a SRT test-retest SD of 0.6 dB (based on an estimation of the slope of the underlying psychometric function). This is equal to or better than the test-retest SD reported for the most common adaptive SRT tests (e.g., 0.92 dB for the Danish HINT with hearing-impaired listeners [7]).
- The SFS test is only relevant for intra-individual hearing-aid comparisons, not for absolute performance assessments. Furthermore, HA functionality must not be confounded with the SRT manipulators. For example, changing the target-masker separation is not a valid SRT manipulator when testing directional hearing aids.
- The observed training effects call for proper training of participants and appropriate counterbalancing of test conditions when using the SFS test in comparative studies.

Conclusions

With reference to the aim, the validation studies led to the following conclusions:

- The SRT manipulators generally changed the SRT as expected.
- The experimental contrasts were measured with appropriate statistical significance.
- The SFS test-retest SD was around 8.5%, which is very satisfying.
- The SFS test runs in both anechoic and reverberant conditions - but reverberation makes the test more difficult.

References

[1] Nilsson et al. (1994), J. Acoust. Soc. Am. 95(2)
[2] Hagerman et al. (1995), Scand. Audiol. 24(1)
[3] Bernstein (2011), Proceedings of 3rd ISAAR
[4] Smeds et al. (2012), Poster presented at IHCON
[5] Laugesen et al. (2013), Proceedings of 4th ISAAR
[6] Laugesen et al. (2013), Poster presented at SpiN Workshop
[7] Nielsen & Dau (2011), Int. J. Audiol. 50(3)

Acknowledgements

We would like to thank our colleagues in Oticon, Maria Grube Sorgenfrei, Anne Specht Petersen, Lise Bruun Hansen, and Jacob Anders Roug for their contributions to the planning and execution of the validation studies.

ErikshholmResearchCentre
PART OF OTICON