Towards a clinically viable spectro-temporal modulation test

Zaar, Johannes; Simonsen, Lisbeth Birkelund; Behrens, Thomas; Laugesen, Søren

Publication date:
2018

Document Version
Publisher's PDF, also known as Version of record

Citation (APA):
Zaar, J., Simonsen, L. B., Behrens, T., & Laugesen, S. (2018). *Towards a clinically viable spectro-temporal modulation test*. Poster session presented at International Hearing Aid Conference 2018, Tahoe, California, United States.

Towards a clinically viable spectro-temporal modulation test

Johannes Zaar[1], Lisbeth Birkelund Simonsen[2], Thomas Behrens[3], Torsten Dau[1], and Søren Laugesen[2]

1: Hearing Systems Group, Department of Electrical Engineering, Technical University of Denmark; 2: Interacoustics Research Unit; 3: Oticon A/S

Interacoustics Research Unit
Contact
Søren Laugesen
slau@iru.interacoustics.com

Introduction

The Spectro-Temporal Modulation (STM) test has shown **good predictive power for speech-in-noise outcomes** beyond the audiogram in several studies [1-3]. Thus, the STM test has potential diagnostic value for hearing-aid fitting.

In the STM test, the depth of spectro-temporal modulations applied to a wide-band carrier is varied, and a threshold is determined at which the test subject can just detect the difference between the STM stimulus and a reference without modulations. A schematic example of an STM stimulus is shown in figure 1.

However, in a recent study in a large clinical population [3], a substantial sub-group of the participants reached the test's upper bound, in the sense that even with the modulations set to maximum, they could not reliably discriminate the target stimulus from the reference.

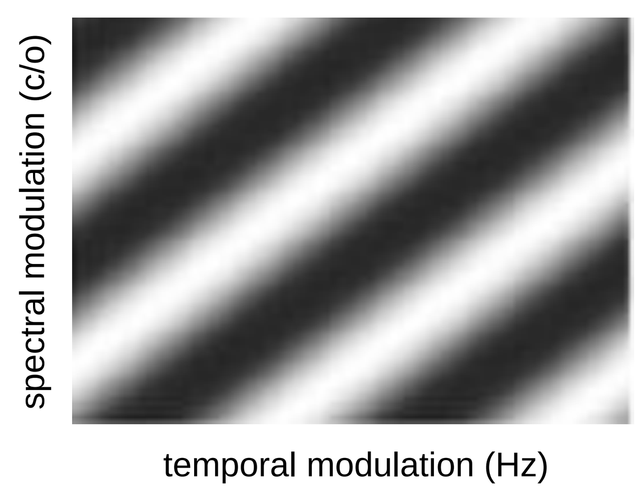

Figure 1. Schematic spectrogram of an STM test stimulus. Grey-scale indicates magnitude. c/o: cycles per octave.

Aims of this study

To **modify the STM test** in terms of stimulus parameters and procedure **to make it sensitive within the target population** of elderly people with hearing impairment. In particular, to carry out the STM test with full compensation for audibility [4].

To **extend the earlier experiments towards more realistic speech-in-noise scenarios**. Previous speech tests typically used headphone presentation of target and maskers that were either co-located steady-state noise [1,2,3], modulated noise [2,3], or babble noise [3] presented at high levels but without frequency-specific compensation for audibility [1,2].

Method and material

Participants

N = 13, age 61-82 years, hearing-loss configurations with modest asymmetry. Audiograms in terms of left/right-ear means of Hearing Threshold Levels (HTL) are shown in figure 2.

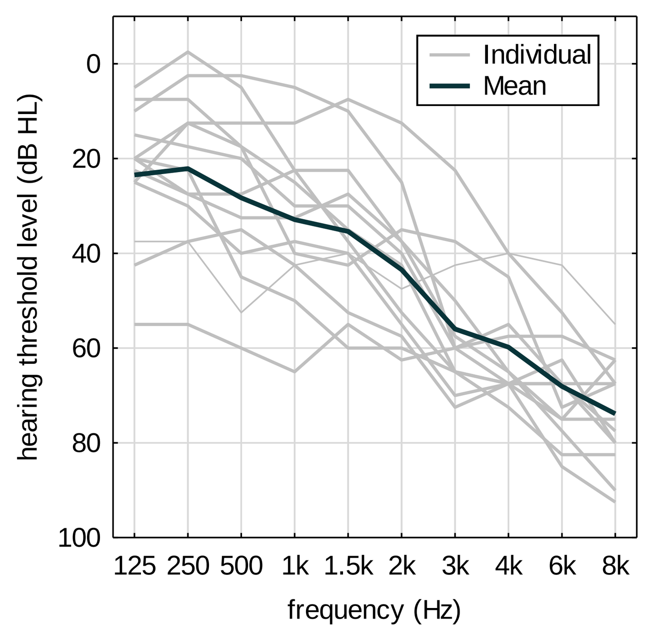

Figure 2. Left/right mean audiograms of all participants.

STM conditions

STM stimulus parameters were chosen to make the test easier, so as to avoid the upper-bound issue, and to emphasize speech-like properties:

- Wide carrier bandwidth, 354 – 5656 Hz (except one condition with 354 – 2000 Hz bandwidth).
- Low spectral ripple densities (1, 2, and 4 c/o), and 4 Hz upward-moving temporal ripple (main speech modulation frequency).
- Frequency-specific hearing-loss compensation scheme, starting at 65 dB broadband SPL and then ensuring at least 15 dB SL (sensation level) in all relevant 1/3-octave bands, according to the individual left/right-mean audiogram [4].
- 3-AFC test paradigm with 1 s stimulus duration, no level roving, and diotic presentation.
- In addition to using a noise carrier [1-3], two conditions with a tone-complex carrier (100-Hz spacing) were included.

The 6 conditions selected are reflected in figure 4 and table 1. All reported STM thresholds were averaged across 3 test runs.

Speech-in-noise test

The speech-in-noise test was set up in an IEC listening room ($T_{60} \approx 0.4$ s). Targets were 20-item lists of Danish HINT sentences [5] from a male talker, presented at a nominal level of 70 dB SPL(C). The four maskers presented from ±100° and ±155° (see figure 3) were different running male speech signals with speech-shaped steady-state noise (SSN) mixed in 6 dB below each speech masker. All speech signals and the SSNs were shaped to have the same long-term spectrum.

Compensation for hearing loss was achieved by raising (where necessary) all 1/3-octave band levels of the target to 15 dB SL [4], with a roll-off starting at 3 kHz down to 4 dB SL at 8 kHz to protect the loudspeakers' tweeters. As for the STM test, the SLs were based on the individual left/right-mean audiogram. This compensation was applied to both the target and all maskers. Thus, the participants were listening aided, with open ears (no hearing aids).

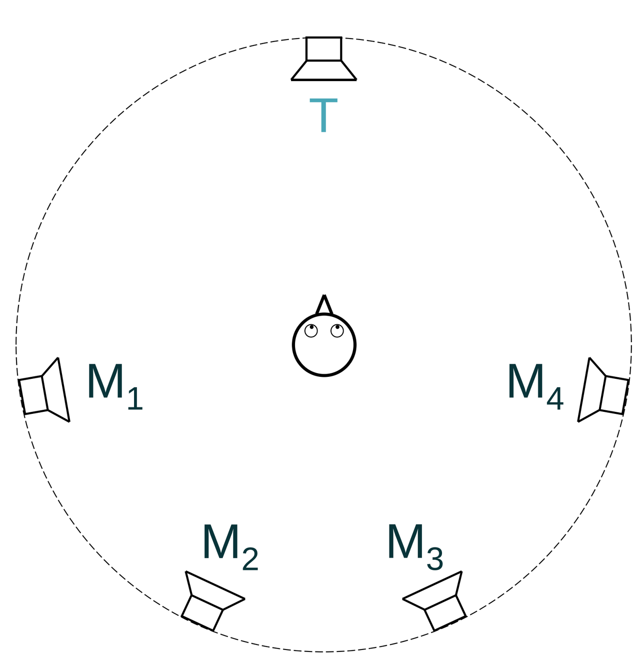

Figure 3. Sketch of the speech-in-noise test loudspeaker layout, showing target (T) in front and four maskers (M_1-M_4).

In addition to this *Spatial* condition, an *SSN* condition was tested with a single SSN masker presented co-located with the target. Masker levels were changed adaptively, tracking 50% correct sentences. The final Speech Reception Thresholds (SRTs) were determined from the word-scoring psychometric functions [6]. The SRTs are presented as the Signal-to-Noise Ratio at the centre of the set-up (test subject absent). Three lists were used for training, then two lists per test condition.

Protocol

Visit 1: Otoscopy, audiogram, and speech-in-noise testing. Order of conditions (*Spatial* first, then *SSN*) was the same for all test subjects, while the use of test lists was balanced across conditions.

Visit 2: STM, training and 3 test conditions: *noisy1co*, *noisy2co*, and *tonal1co*, in balanced order.

Visit 3: STM, refresher and 3 test conditions: *noisy4co*, *noisy2co2kHz* (narrower carrier bandwidth), and *tonal2co*.

All experiments were done under ethical approval from the Scientific Ethical Committee of the Capital Region of Denmark.

Results

The STM thresholds for each of the 6 different conditions are shown in figure 4. Note that in all runs, including the training runs, all participants managed to produce a valid STM threshold. The results for the two speech-in-noise conditions are shown in figure 5.

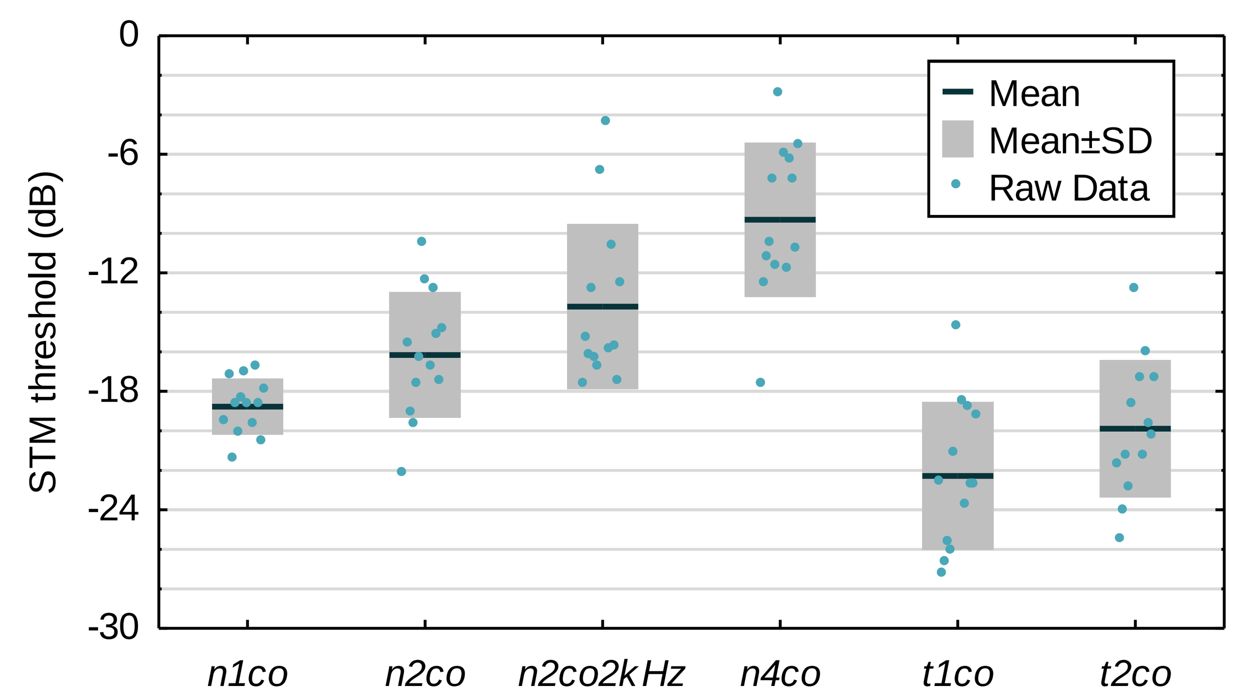

Figure 4. STM thresholds for all six test conditions. On the ordinate, 0 dB corresponds to full modulation (upper bound). SD: standard deviation; *n… : noisy…; t… : tonal….*

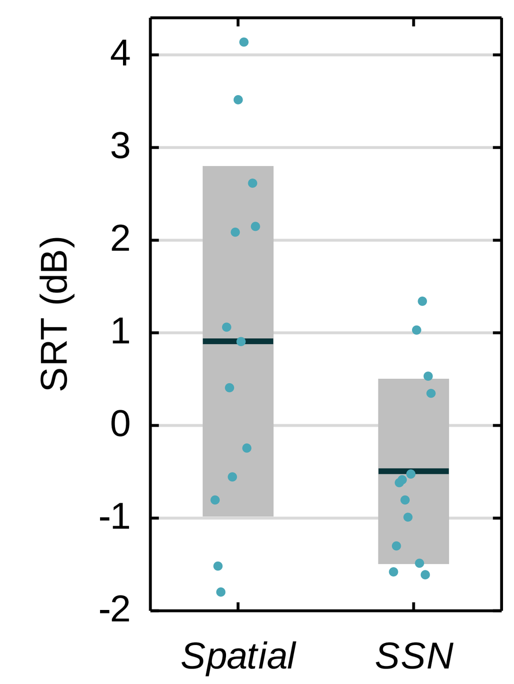

Figure 5. SRTs for the two speech-in-noise conditions.

Analysis

First, the potential predictor variables were individually correlated with the SRTs to select the preferred ones. The considered predictors were the 6 STM variants, *Age*, and three hearing-loss descriptors:

- *4PTA*: mean of left and right HTL values across 500, 1000, 2000, and 4000 Hz.
- *LFA*: mean across 125, 250, 500, and 1000 Hz, as in [3].
- *HFA*: mean across 2000, 3000, 4000, and 6000 Hz, as in [3].

All correlations are summarised in table 1. For the *Spatial* SRTs, the preferred STM predictor was the *tonal2co* condition, whereas the *noisy2co2kHz* was preferred for the *SSN* SRTs. In both cases, the *LFA* hearing-loss descriptor was preferred. The correlations with *Age* were small and counter-intuitive (negative); hence this variable was disregarded from further analysis.

Table 1. Individual predictor-outcome correlations, *N* = 13. Preferred predictors are highlighted (all $p < 0.05$).

	Age	*4PTA*	*LFA*	*HFA*	*noisy1co*	*n2co*	*n2co2kHz*	*n4co*	*tonal1co*	*t2co*
Spatial	–.28	.67	.75	0.06	.37	.59	.66	.62	.69	.73
SSN	–.30	.55	.67	0.06	.22	.47	.59	.53	.49	.51

Secondly, a multi-variate linear regression was conducted for each SRT outcome. This was done in a manual stepwise fashion, considering first the preferred hearing-loss descriptor, then adding the STM predictor. The results are illustrated in figure 6, left-most columns.

Discussion

The hearing-loss descriptors

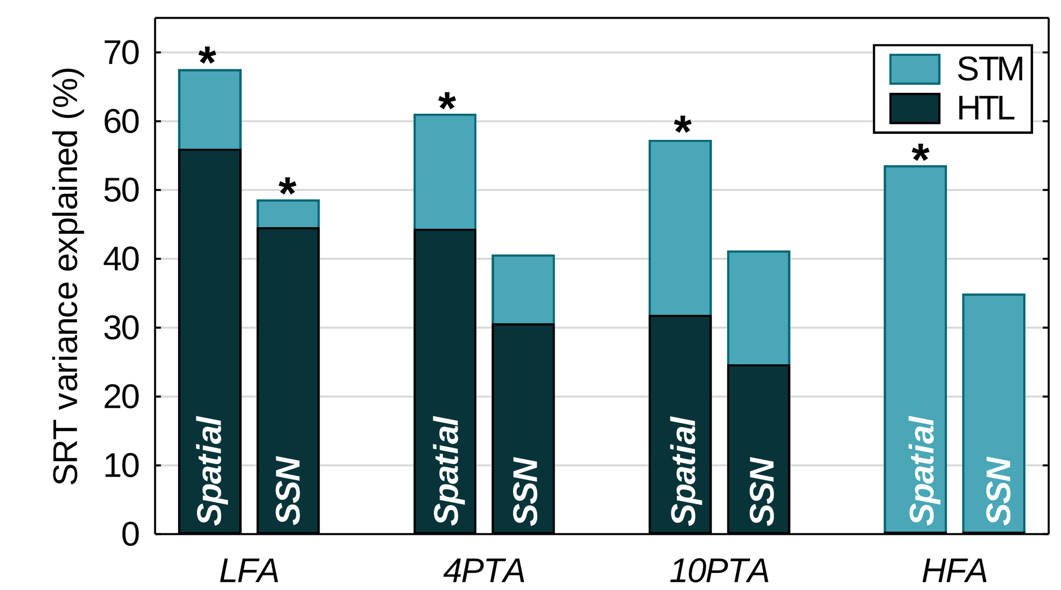

Figure 6. This study: amount of SRT variance successively explained by the hearing-loss (HTL) descriptors and the preferred STM predictor. Whole-model significance ($p < 0.05$) is indicated by *.

The higher correlations observed for *LFA* compared with *HFA* in table 1 are in good agreement with the results from [7,8] obtained with similar speech-in-noise set-ups. In contrast, [3] found *HFA* to be the best hearing-loss descriptor in a much larger dataset (*N* = 154), based on figure 7. Note that table 1 and [7] gave preference to the *LFA* predictor for both spatially separated and co-located SRT outcomes, indicating that the different result found in [3] is not due to a difference between separated and co-located maskers and neither to masker type. Given the weighty evidence from [3], multiple-regression results are also included for the *HFA* predictor in figure 6 (right-most columns). As expected, this leads to much more variance being explained by the STM predictor in our data, compared with the modest contribution from the STM predictor when used together with *LFA*. As two intermediate examples, results for the *4PTA* hearing-loss predictor as well as a *10PTA* predictor spanning all 10 audiometric frequencies measured in this study are also included in figure 6 (middle columns).

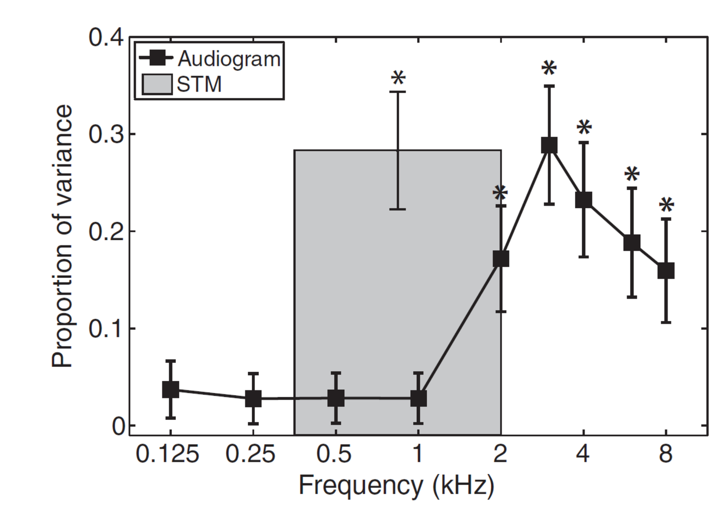

Figure 7. Bernstein et al. (2016) [3]: amount of speech-in-noise outcome variance explained by individual HTL values, *N* = 154.

Spatial versus SSN SRTs

Figure 5 shows that the *Spatial* SRTs on average were higher than the *SSN* SRTs and that the spread was wider. The higher mean SRT makes the *Spatial* condition more ecologically relevant [9], and the higher spread indicates that the *Spatial* condition brought out more individual difference among the participants. This agrees with the higher predictor-outcome correlations (table 1) and amounts of variance explained (figure 6), observed for the *Spatial* compared with the *SSN* SRTs.

Note on future clinical use

Diotic presentation of the STM stimuli was chosen here to match the way hearing loss was compensated for the speech-in-noise testing, i.e. not ear-specific. In potential future clinical applications, where the two ears ultimately are aided independently, ear-specific audibility compensation in the STM test should be considered.

Conclusions

Referring back to the study's aims it can be concluded that:

- The proposed modifications to the STM test appear to have **solved the upper-bound issue** reported in [3].
- The preferred **STM thresholds** on their own **have considerable predictive power for the SRT outcomes**, and can explain additional outcome variance beyond the hearing-loss descriptors. The actual amount depends on the degree to which the hearing-loss descriptor is tailored to the dataset.
- **The *Spatial* speech-in-noise outcome showed greater individual variation and allowed more variance to be explained** by the hearing-loss and STM predictors, compared with the co-located steady-state-noise (*SSN*) condition.

As an additional observation it was found that:

- Even with the non-compromising compensation for audibility applied in this study, the **hearing-loss descriptors have considerable predictive power for the supra-threshold SRT outcomes**.

Acknowledgements

The authors wish to acknowledge the support from the Oticon Foundation, as well as the contributions to this study from Laurel Carney, Golbarg Mehraei, Thomas Lunner, Elaine Ng, Alejandro Lopez Valdes, Gary Jones, Nicolas Le Goff, Raul Sanchez Lopez, and James Harte.

References

[1] Bernstein JGW, Mehraei G, Shamma S, Gallun FJ, Theodoroff SM, Leek MR. (2013). *Spectrotemporal modulation sensitivity as a predictor of speech intelligibility for hearing-impaired listeners*. Journal of the American Academy of Audiology, 24(4), 293–306. [2] Mehraei G, Gallun FJ, Leek MR, Bernstein JGW (2014). *Spectrotemporal modulation sensitivity for hearing-impaired listeners: Dependence on carrier center frequency and the relationship to speech intelligibility*. The Journal of the Acoustical Society of America, 136(1), 301–316. [3] Bernstein JGW, Danielsson H, Hällgren M, Stenfelt S, Rönnberg J, Lunner T. (2016). *Spectrotemporal modulation sensitivity as a predictor of speech-reception performance in noise with hearing aids*. Trends in Hearing, 20, 1-17. [4] Humes LE. (2007). *The contributions of audibility and cognitive factors to the benefit provided by amplified speech to older adults*. Journal of the American Academy of Audiology, 18(7), 590–603. [5] Nielsen JB, Dau T. (2011). *The Danish hearing in noise test*. International Journal of Audiology, 50(3), 202–208. [6] Rønne FM, Laugesen S, Jensen NS. (2017). *Selection of test-setup parameters to target specific signal-to-noise regions in speech-on-speech intelligibility testing*. International Journal of Audiology, 56(8), 559–567. [7] Neher T, Behrens T, Carlile S, Jin C, Kragelund L, Petersen AS, van Schaik A. (2009). *Benefit from spatial separation of multiple talkers in bilateral hearing-aid users: Effects of hearing loss, age, and cognition*. International Journal of Audiology, 48(11), 758–774. [8] Neher T, Laugesen S, Jensen NS, Kragelund L. (2011). *Can basic auditory and cognitive measures predict hearing-impaired listeners' localization and spatial speech recognition abilities?* The Journal of the Acoustical Society of America, 130(3), 1542–1558. [9] Smeds K, Wolters F, Rung M. (2015). *Estimation of signal-to-noise ratios in realistic sound scenarios*. Journal of the American Academy of Audiology, 26(2), 183–196.

Presented as poster B3-P-15 at the International Hearing Aid Research Conference, IHCON, 2018, August 15-19, Granlibakken Conference Center, Tahoe City, California, USA