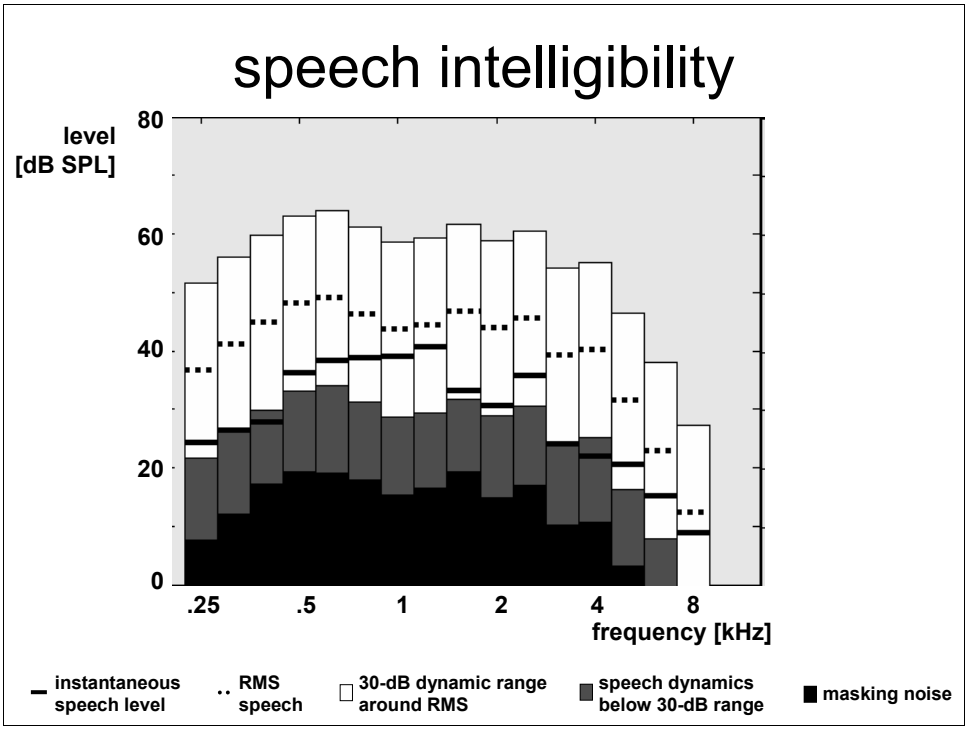
# Intelligibility trade-off between multi-band peak-limiting and masking by noise

Gaston L.M. Hilkhuysen
*GHilkhuysen@hotmail.com*

To make a long story short, this title is about the best I can do. Before going into the details of the current study, I want to start with some classic material.



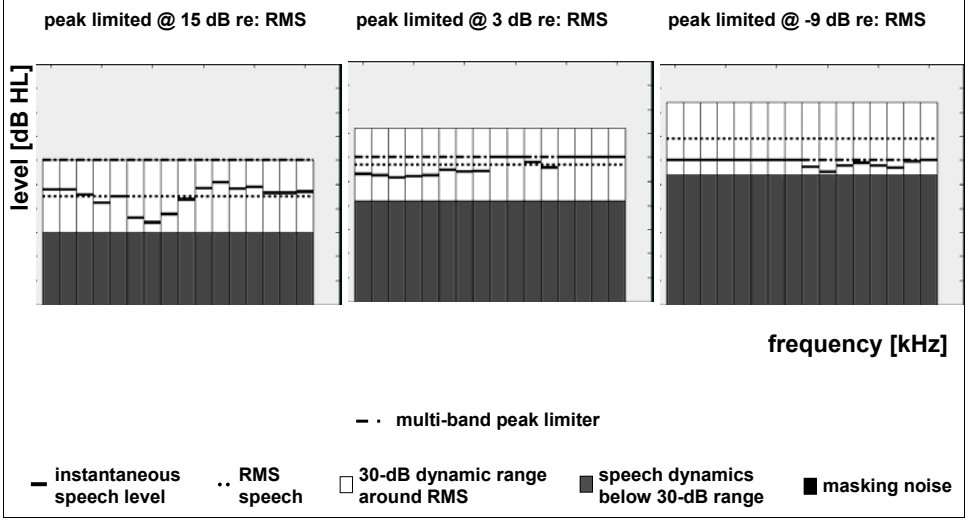

I hope you recognized an animation of the Articulation Index by French and Steinberg published in 1947, currently implemented in an ANSII standard known as the Speech Intelligibility Index. It fascinates me that the contribution of speech dynamics to intelligibility is uniformly distributed over a 30 dB dynamic width, centered around the RMS. Dynamics in the top of the 30 dB dynamic range contribute as much as dynamics at the bottom, and speech dynamics just outside this range do not contribute at all. Connected to this issue is the 30-dB range contributing to intelligibility. Studebaker defied this assumption and some findings in a previous study that I did also indicated a larger dynamic width contributing to intelligibility.

## research questions

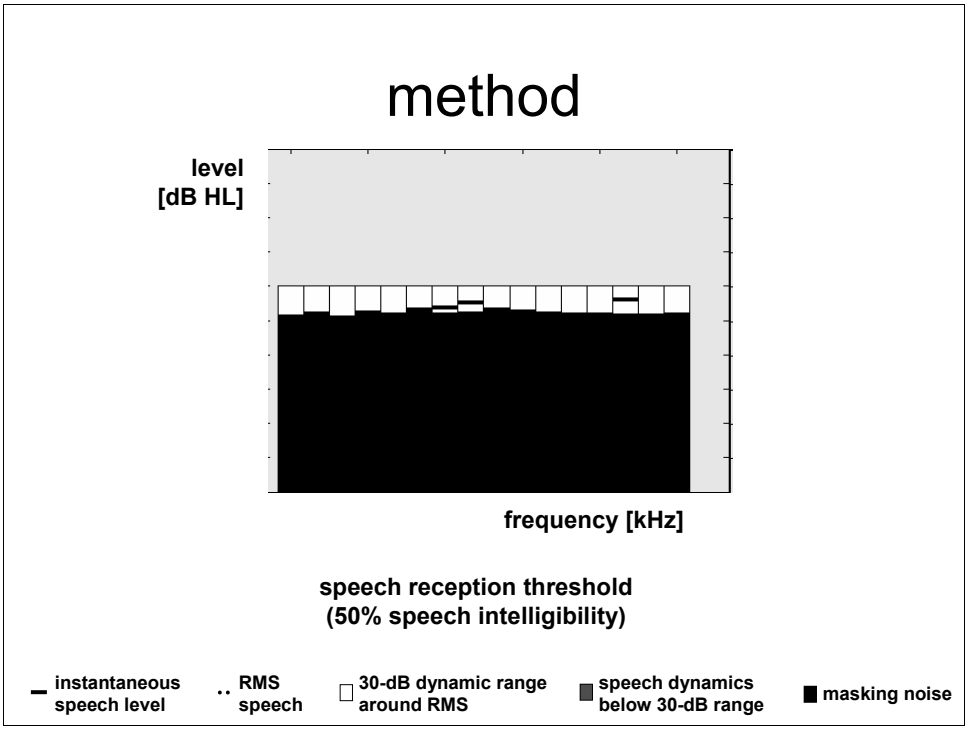
1. Do speech dynamics contribute uniformly to intelligibility?
2. Can speech dynamics below the 30-dB width defined in the ANSI-SII contribute to intelligibility?

With this in mind, I tried to answer the following two questions in the current study.

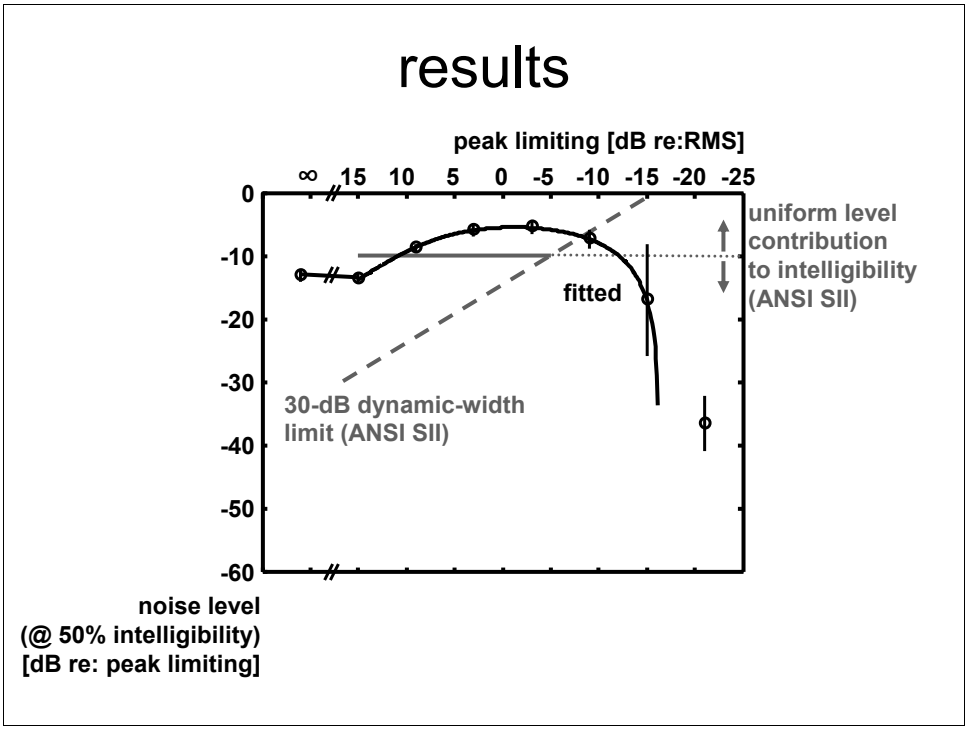
# method




I spectrally shaped the speech in order to have as much speech dynamics above hearing thresholds as possible. Speech dynamics from the top were removed with a multi-band peak limiter. To acquire different amounts of peak limiting, while keeping enough speech dynamics above the hearing thresholds, the absolute peak-limiting level was fixed, and the speech was amplified. This flattens the speech dynamics at different levels relative to the RMS of speech. Speech processed in such a way, sounds like this.

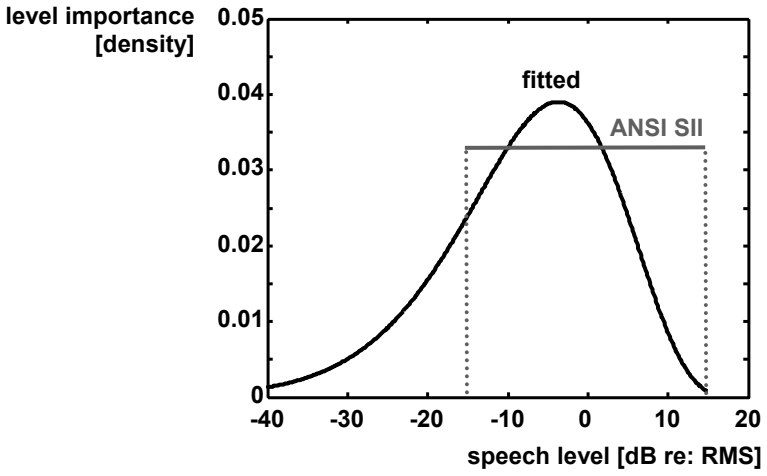



Using this spectrally shaped speech, I measured speech reception thresholds: speech dynamics at the bottom of the dynamic range were masked by noise up to a level that resulted in 50% speech intelligibility. This is a classic method developed by Plomp. For each peak-limiting condition, the noise level was adjusted to a level that results in 50% intelligibility. Hence we listen thru a dynamical slit to the speech; we move this slit over the speech dynamics and adjust its width in order to attain 50% speech intelligibility.




This Figure is going to show the results. Peak-limiting levels relative to the RMS of non-processed speech are represented on the x-axis. From left to right, the amount of peak limiting increases. The ordinate indicates the height of the dynamical slit, resulting in 50% speech intelligibility that was acquired by measuring speech reception thresholds. If peak-limiting and masking by noise are equally detrimental to speech intelligibility independent of their positions within the 30-dB range, the slit's height should be constant over conditions and points should be located on a horizontal line. Points below the titled dashed line indicate measurements that include speech dynamics below the ANSI's 30-dB dynamic width contributing to intelligibility. Results show that down to 3 dB peak limiting, noise masking is more detrimental than peak limiting. The contrary holds for peak limiting below -3 dB. Intelligibilities can be measured below -15 dB re: RMS. These data can be modeled with a gamma distribution, leading to the following curve. The curve misses a point. Due to time restrictions, I will not focus on this issue, but it could make a nice question for the discussion afterwards.

# results




Here you see the density function of the estimated intensity importance. It is clearly asymmetric. The top of the function is located 3 dB below the RMS of speech. With this function it is hard to continue talking about THE dynamic range of speech. Evidently, speech dynamics below the 30-dB currently used in the ANSI standard can contribute to intelligibility.

# discussion

conclusions

- the contribution of speech dynamics to intelligibility is not uniformly distributed
- speech dynamics below 30-dB dynamic width can contribute to intelligibility

remarks

- non-linear processing in SII
- fine-structure matters
- speech spectrally masking speech

Two conclusions of this study, and some final remarks. I recognize that I used a nonlinear process, while the SII only holds for linear processes. Besides that, fine structure improves intelligibility. That is part of this study, but not of this presentation. Finally, speech might be masking speech. Peak limiting could reduce this masking.

**I'll make it easier for you… Go ahead and shoot, you'd be doing me a favor**

Some other classic materials were presented. I hope you noticed the excerpts from the second best film ever made: "Casablanca". To encourage your questions, I finish with small extract from this film.

# Intelligibility trade-off between multi-band peak-limiting and masking by noise

Gaston L.M. Hilkhuysen
*GHilkhuysen@hotmail.com*

Thank you for your interest. Please don't hesitate to pose questions, or to comment. E-mails are also welcomed!