

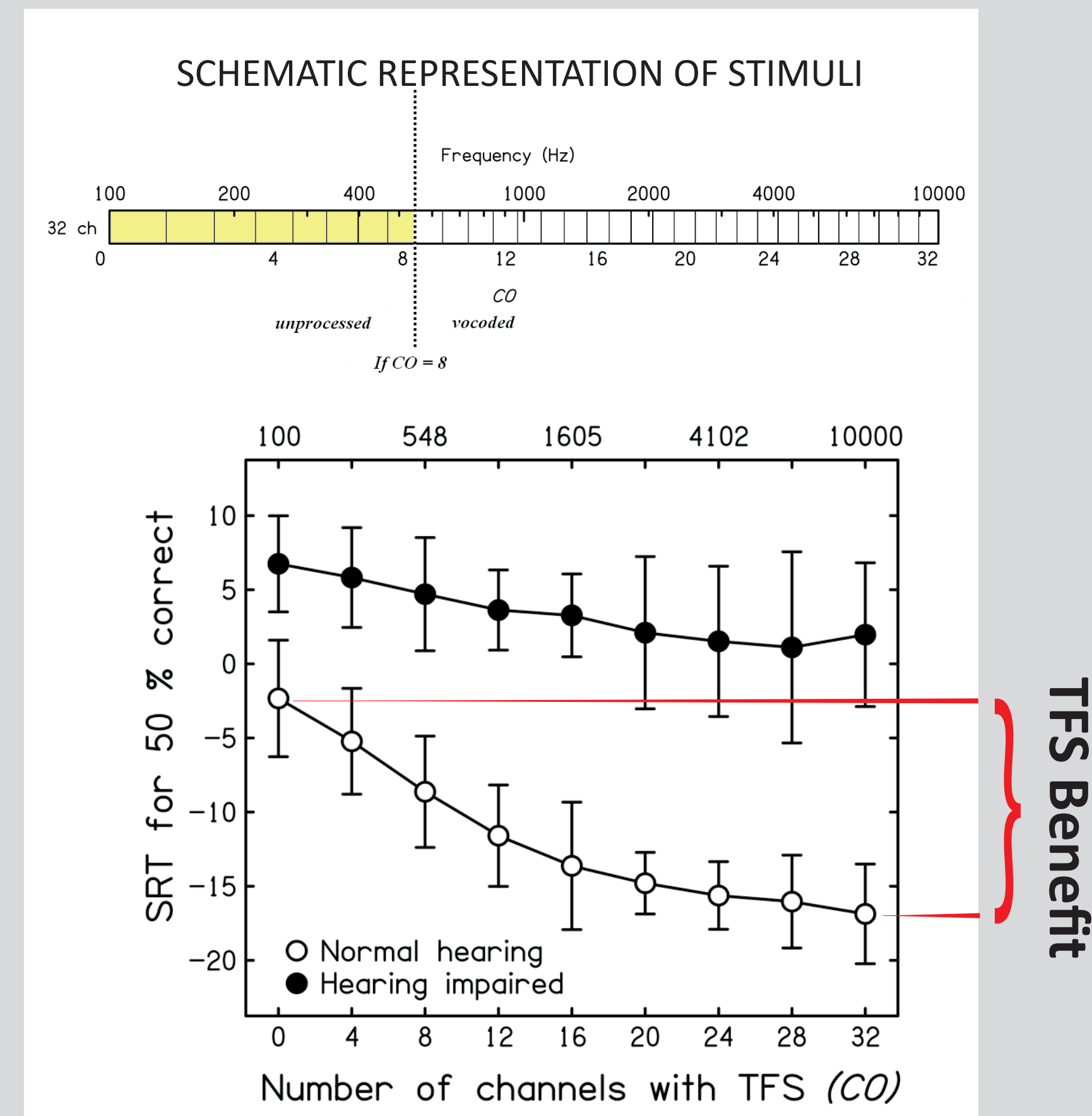
# Effect of speech material on the benefit of temporal fine structure information in speech for normal-hearing and hearing-impaired subjects

Thomas Lunner<sup>1</sup>, Renskje K. Hietkamp<sup>1</sup>, Martin R. Andersen<sup>1</sup>, Kathryn Hopkins<sup>2</sup>, & Brian C.J. Moore<sup>2</sup>

<sup>1</sup>Oticon Research Centre Eriksholm, Snekkersten, Denmark; <sup>2</sup>Department of Experimental Psychology, University of Cambridge, United Kingdom

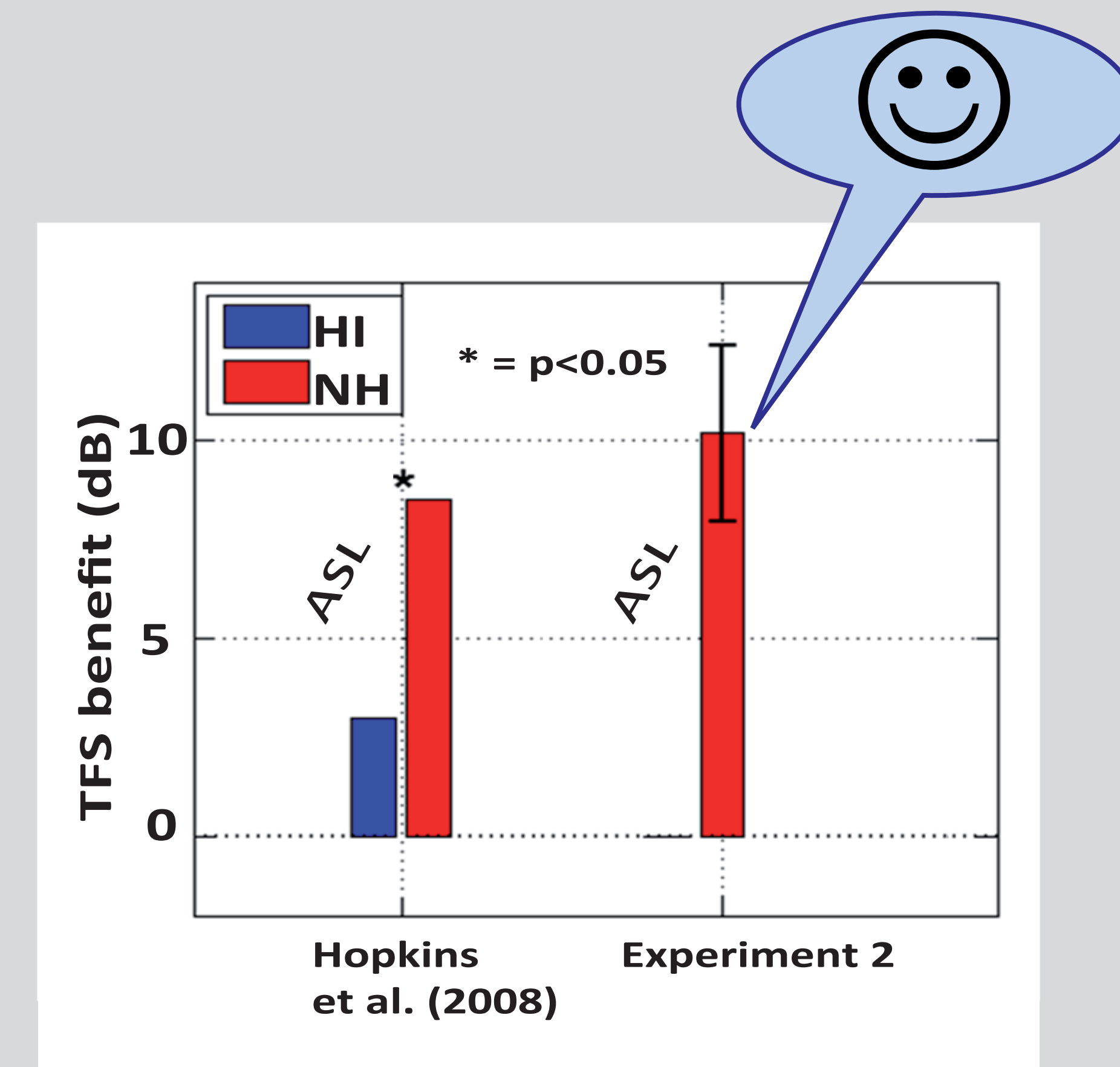
## Introduction

- In recent years, several researchers have presented evidence suggesting that the ability to process TFS information may be adversely affected by cochlear hearing loss.
- Hopkins et al. (2008) used vocoder processing to manipulate the amount of TFS information that was present in a speech signal with a competing-talker background. TFS information was progressively added from low frequencies.
- Hopkins et al. (2008) concluded that moderate cochlear hearing loss leads to a reduced ability to use TFS information to understand speech in a competing talker.
- The initial goal of the present study was to assess whether the findings could be replicated using a new language, new speech material, new background talker, and different hearing-impaired subjects.



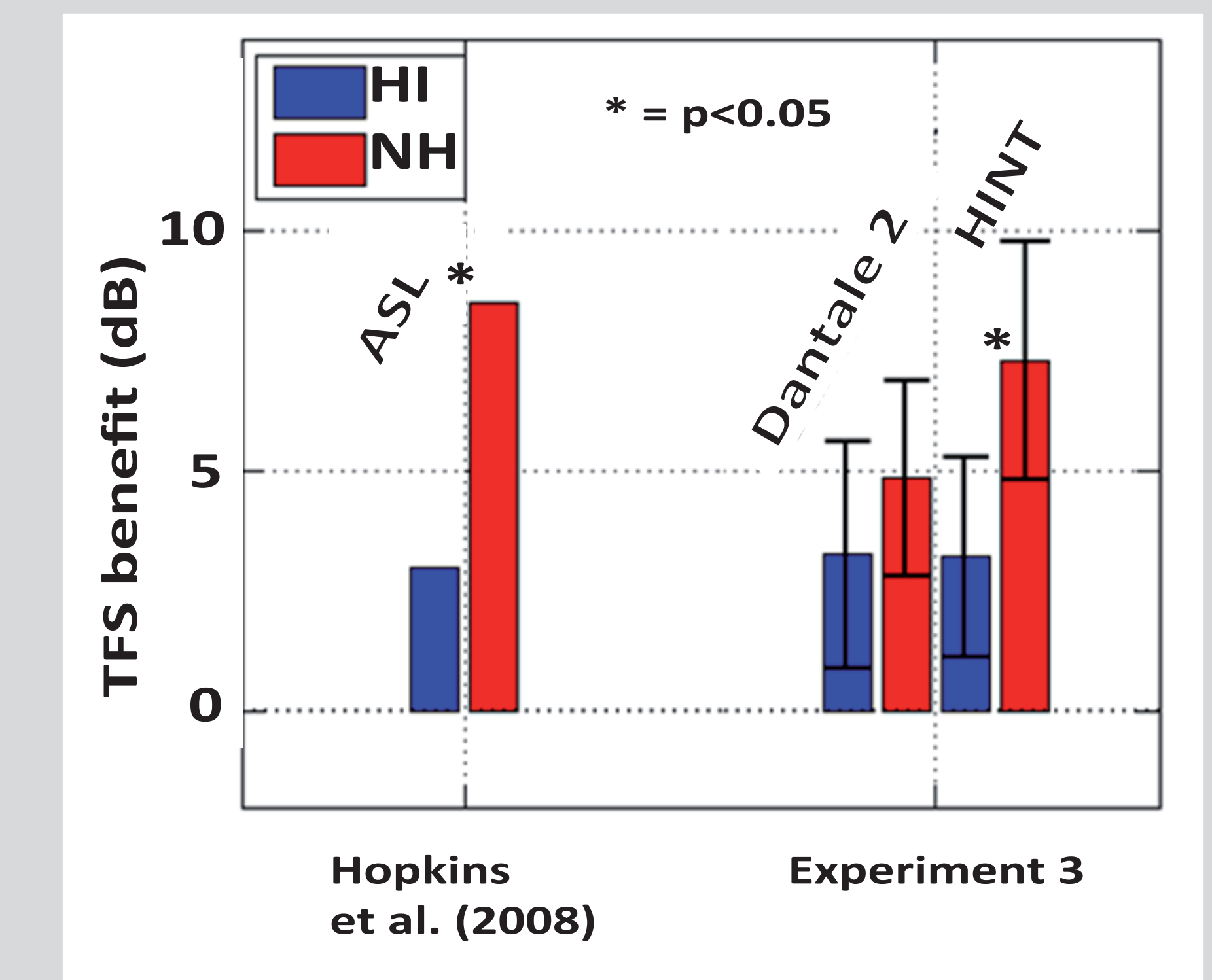
## Experiment 2

- Experiment 2 was similar to experiment 1, but the English-language ASL sentences (MacLeod and Summerfield, 1990) were used for the target speech instead of the Dantale 2 sentences, and only normal-hearing subjects were tested (N=5).
- The SRT<sub>50</sub> difference between unprocessed and fully vocoded processing was larger than for experiment 1, suggesting that the subjects benefited more from the additional TFS information that was available in the unprocessed signals for the more naturally spoken ASL sentence material.



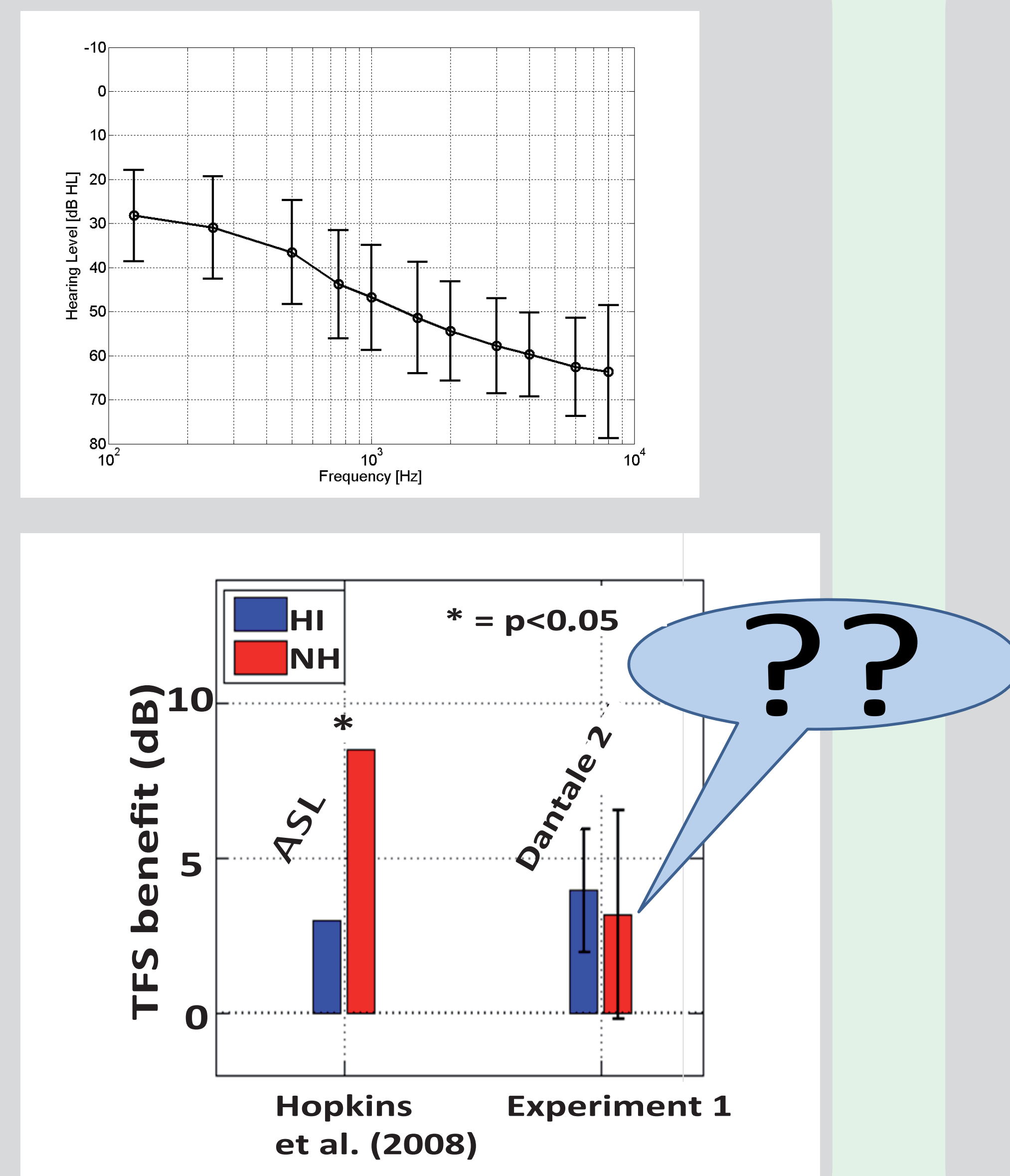
## Experiment 3

- In experiment 3, we repeated experiment 1 and included the more naturally spoken Danish HINT sentences (Nielsen and Dau, 2009).
- Six normal-hearing and 11 hearing-impaired subjects were tested.
- The results obtained using the Dantale 2 sentences again failed to show a significant difference between normal-hearing and hearing-impaired subjects in the benefit of adding TFS information, while the results obtained using the HINT sentences showed a greater benefit for the normal-hearing subjects, consistent with the results of Hopkins et al. (2008).



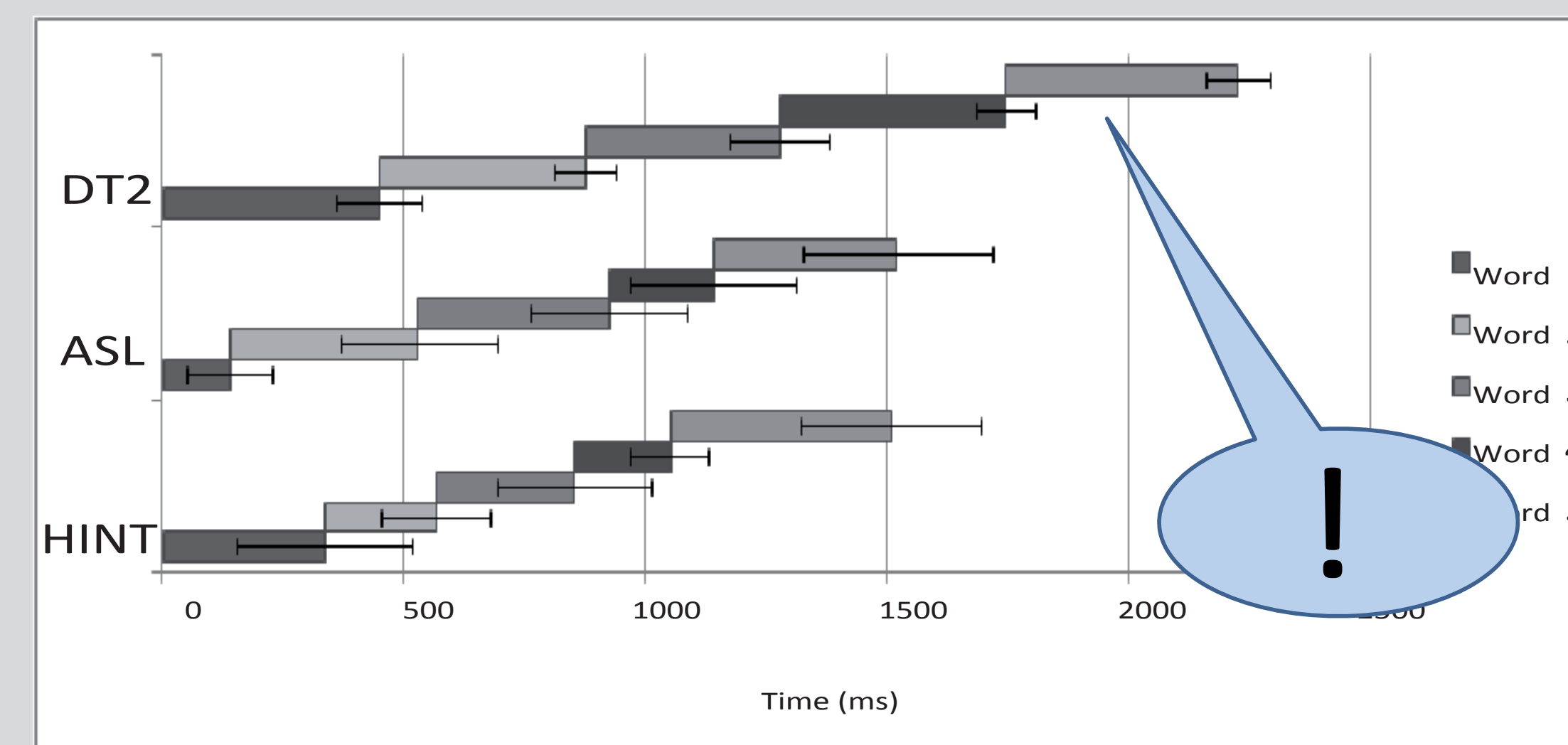
## Experiment 1

- Nine normal-hearing (NH) and 20 hearing-impaired (HI) subjects were tested. Target- and competing-talker signals were either processed with a 32-channel tone vocoder, to remove TFS information, or left unprocessed, and so contained all of the original TFS. The target speech was based on the Danish Dantale 2 sentences (Wagener et al., 2003). The competing background was a randomly chosen portion of continuous prose spoken by a Danish female. To ensure audibility, amplification according to the Cambridge formula (Moore and Glasberg, 1998) was applied.
- **To our surprise**, the results differed from those of Hopkins et al. (2008); the SRT<sub>50</sub> difference between unprocessed and fully vocoded processing was smaller than previously found, and was similar for the normal-hearing and hearing-impaired subjects.



## Speech material analysis

- Speech materials were analyzed to characterize the differences between the Dantale 2, ASL and HINT speech materials. Analysis was carried out using the speech analysis software Praat (Boersma and Weenink, 2009).
- The Dantale 2 sentences were longer, and had longer words and more syllables per word, than the ASL and HINT sentences. Dantale 2 words were less variable in duration than the ASL and HINT words.
- More syllables per word would provide more glimpsing opportunities and make guessing easier for the Dantale 2 sentences. In addition, guessing may be facilitated by Dantale 2 being a closed set corpus. The lower variability of word duration of the Dantale 2 corpus would make it easier for the subject to predict where one word ends and the next word begins.
- The lower variability of word duration of the Dantale 2 corpus would also make it easier to perceptually segregate the target from the background, by **reducing informational masking** (Arbogast et al., 2002).



## Discussion

- If a speech signal is represented by a spectrogram, it is apparent that the signal has a sparse representation; the energy is high in only a few spectro-temporal regions (Darwin, 2009). When the speech of two talkers is mixed, there is often **relatively little overlap between the spectro-temporal regions dominated by one talker and the regions dominated by the other talker**. In this situation, the ability to identify the speech of a target talker in the presence of a background talker is not limited by energetic masking, but rather by informational masking.
- It is possible that **TFS information is mainly useful for reducing informational masking**, by providing cues that aid the perceptual segregation of the target and the background. When such segregation is made easier by the predictable structure of the speech, as for the Dantale 2 sentences, TFS information may no longer be required.



## Conclusions

- When naturally spoken sentences were used, the results were similar to those reported by Hopkins et al. (2008); normal-hearing subjects showed more benefit from adding TFS information than subjects with cochlear hearing loss.
- When highly stylized and predictable sentences were used, informational masking may have been small and TFS information was of little benefit.

### Dantale 2 sentences:

Michael ejer tyve gamle ringe (Michael owns twenty old rings),  
Linda låner otte hvide kasser (Linda borrows eight white boxes)

### ASL sentences:

He hit his head,  
The yellow leaves are falling

### Danish HINT sentences:

Mødet sluttede efter tre timer (The meeting ended after three hours),  
Vinduet vendte ud mod gaden (The window faced the street)

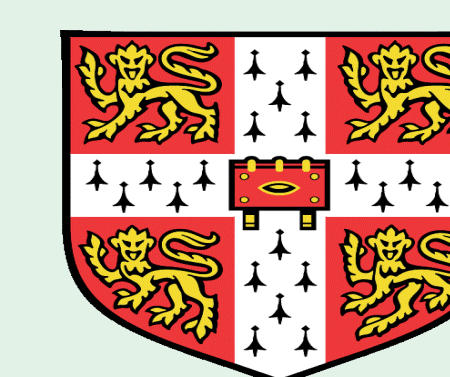
## Contact information

Thomas Lunner - [tlu@oticon.dk](mailto:tlu@oticon.dk)  
Eriksholm, Oticon Research Centre,  
Kongevej 243, 3070 Snekersten, Denmark

## References

Arbogast, T. L., Mason, C. R., and Kidd, G., Jr. (2002). "The effect of spatial separation on informational and energetic masking of speech," J. Acoust. Soc. Am. 112, 2086-2098.  
Boersma, P., and Weenink, D. (2009). "Praat: doing phonetics by computer (Version 5.1.3.1)," Retrieved 4 April, 2010, from <http://www.praat.org/>  
Darwin, C. J. (2009). "Listening to speech in the presence of other sounds," in The Perception of Speech: From Sound to Meaning, edited by B. C. J. Moore, L. K. Tyler, and W. D. Marslen-Wilson (Oxford University Press, Oxford).

Hopkins, K., Moore, B. C. J., and Stone, M. A. (2008). "Effects of moderate cochlear hearing loss on the ability to benefit from temporal fine structure information in speech," J. Acoust. Soc. Am. 123, 1140-1153.  
MacLeod, A., and Summerfield, Q. (1990). "A procedure for measuring auditory and audio-visual speech-reception thresholds for sentences in noise: rationale, evaluation, and recommendations for use," Br. J. Audiol. 24, 29-43.  
Moore, B. C. J., and Glasberg, B. R. (1998). "Use of a loudness model for hearing aid fitting. I. Linear hearing aids," Br. J. Audiol. 32, 317-335.  
Nielsen, J. B., and Dau, T. (2009). "Development of a Danish speech intelligibility test," Int. J. Audiol. 48, 729-741.  
Wagener, K., Josvassen, J. L., and Ardenjaer, R. (2003). "Design, optimization and evaluation of a Danish sentence test in noise," Int. J. Audiol. 42, 10-17.



UNIVERSITY OF  
CAMBRIDGE

oticon  
Research Centre Eriksholm