Technical University of Denmark

Predicting effects of non-linear frequency compression and impulse-noise suppression on consonant perception

Zaar, Johannes; Schmitt, Nicola; Derleth, Ralph-Peter; Dau, Torsten

Publication date:
2016

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

Predicting effects of non-linear frequency compression and impulse-noise suppression on consonant perception

Johannes Zaar[1], Nicola Schmitt[2], Ralph-Peter Derleth[2], and Torsten Dau[1]

[1]*Hearing Systems Group, Department of Electrical Engineering, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark*

[2]*Sonova AG, CH-8712 Stäfa, Switzerland*

Contact: jzaar@elektro.dtu.dk

BACKGROUND AND OBJECTIVE

- The audibility and integrity of high-frequency speech information is difficult to measure using classical speech tests.
- As many consonants are defined by high-frequency bursts/noise (Li et *al.*, 2010; 2012), consonant perception tests yield detailed information on the audibility and the integrity of high-frequency speech cues:
  - *Inaudible* consonant cues => reduced recognition scores and random confusions (high entropy, cf. Scheidiger and Allen, 2013)
  - *Distorted* consonant cues => reduced recognition scores and systematic confusions (low entropy)
- A recent study (Schmitt et *al.*, 2016) showed that a consonant perception test was sensitive to effects of high-frequency amplification and non-linear frequency compression (NLFC) in hearing-impaired (HI) listeners.
- Recently, a consonant perception model was presented (Zaar and Dau, 2016) that accounts well for consonant recognition and confusion scores obtained with normal-hearing (NH) listeners in conditions of additive stationary masking noise.
- The present study measured the perceptual effects of consonant-cue distortions induced by different hearing-aid (HA) signal processing strategies in NH listeners and evaluated to what extent the model of Zaar and Dau (2016) could account for the considered conditions.

SPEECH MATERIAL

- 12 consonant-vowels (CVs) were used: /ba, ga, da, pa, ka, ta, sa6, sa9, ʃa3, ʃa5, fa, tsa/, spoken by a female native German professional speaker.
- The CVs were obtained from the vowel-consonant-vowel speech material collected by Schmitt et *al.* (2016) by cropping the initial vowel.
- The CVs /sa6, sa9/ and /ʃa3, ʃa5/ represent different versions of /sa/ and /ʃa/, spectrally shaped to exhibit spectral peaks at 6 kHz/9 kHz and 3 kHz/5kHz, respectively.

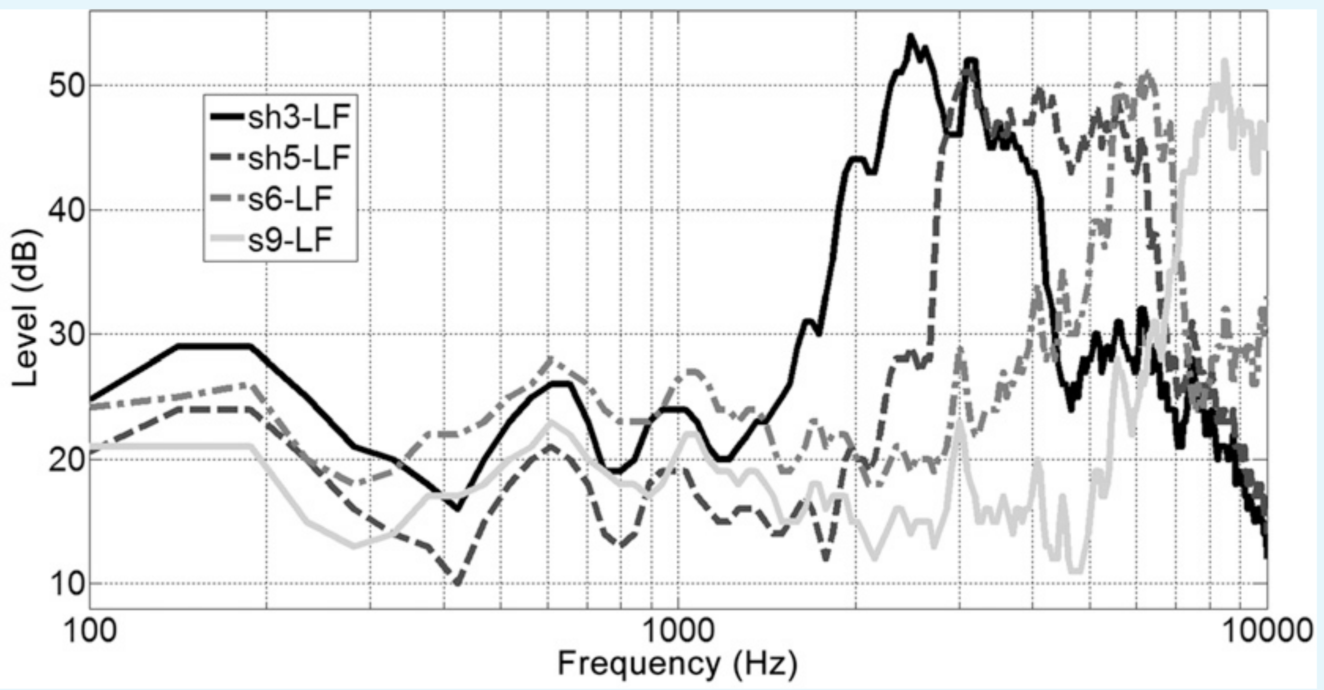

Fig. 1: Spectra of spectrally shaped /sa/ and /ʃa/ stimuli. Reprint from Schmitt et *al.* (2016).

EXPERIMENTAL CONDITIONS AND STIMULI

- Five conditions were considered:
  - ***Unaided***: natural listening condition without HAs
  - ***Default***: default HA setting defined by the fitting software
  - ***NLFC***: strongest possible setting of non-linear frequency compression (Phonak SoundRecover)
  - ***INS***: strongest possible setting of impulse-noise suppression (Phonak SoundRelax)
  - ***NLFC&INS***: combination of *NLFC* and *INS* using the settings described above
- Phonak Naída V90-RIC HAs were employed assuming a moderate to severe hearing loss.
- The CVs were mixed with stationary speech-shaped noise (long-term spectrum of female speech) at an effective SNR of 8 dB.
- The noisy CVs were played from a frontally positioned loudspeaker and recorded using a KEMAR dummy head with/without HAs in a sound-attenuating room (speech level at KEMAR: 70 dBA).
- The recordings were equalized to compensate for the amplification applied in the conditions with HAs.
- 10 adult NH native German listeners were binaurally presented with the diotic stimuli via Sennheiser HD 650 headphones at 60 dB SPL.
- Each of the 60 stimuli (12 CVs X 5 conditions) was presented 8 times to each listener.

GRAND AVERAGE RECOGNITION SCORES

Condition	*Unaided*	*Default*	*NLFC*	*INS*	*NLFC&INS*
% correct	95.9	93.7	55.3	92.3	56.2
Std in %	8.1	7.3	36.2	10.8	34.3

Table I: Recognition scores averaged across CVs and corresponding standard deviations.

CONFUSION MATRICES

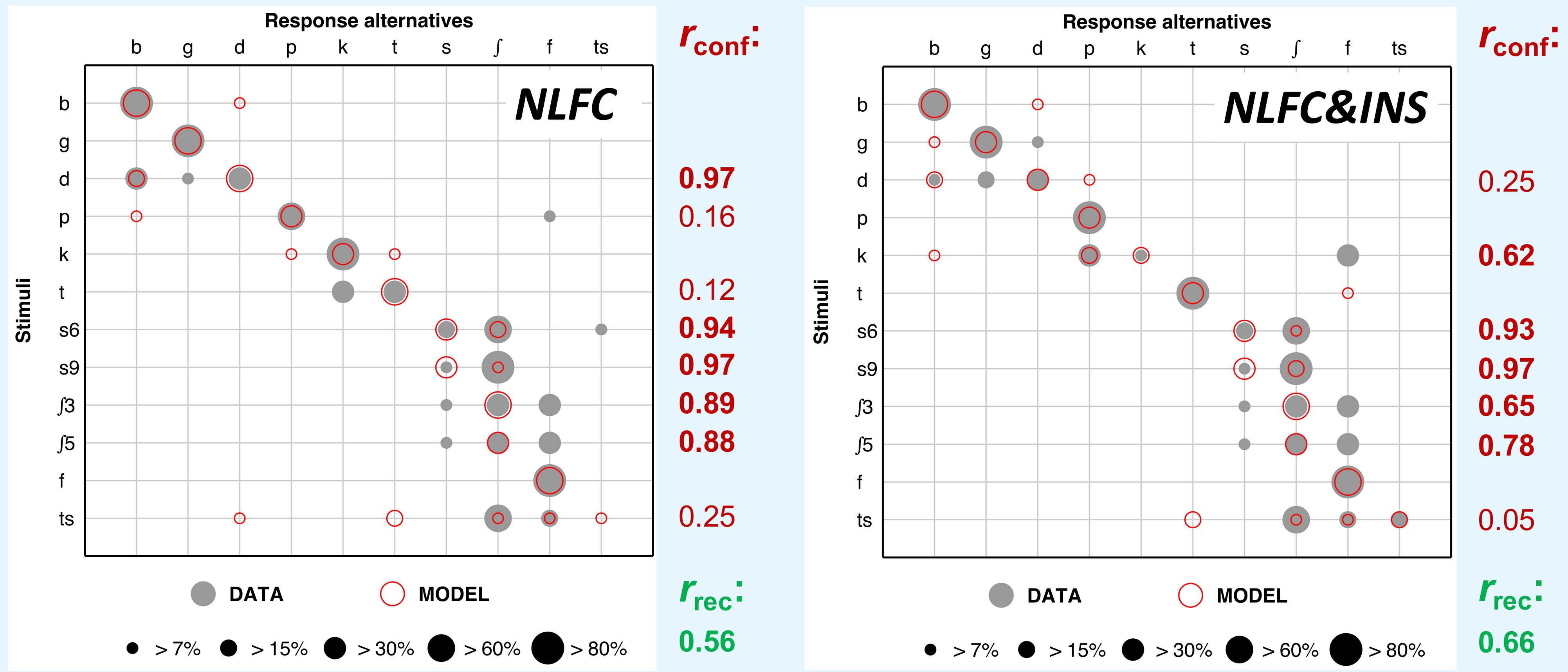

Fig. 3: Confusion matrices measured (gray circles) and predicted (red circles) in the *NLFC* (left) and *NLFC&INS* (right) conditions. Additionally, Pearson's correlation between measured and predicted recognition scores (r_{rec}) and confusion scores (r_{conf}, only obtained if P_e>20%) are shown (bold: p<0.05).

MODELING

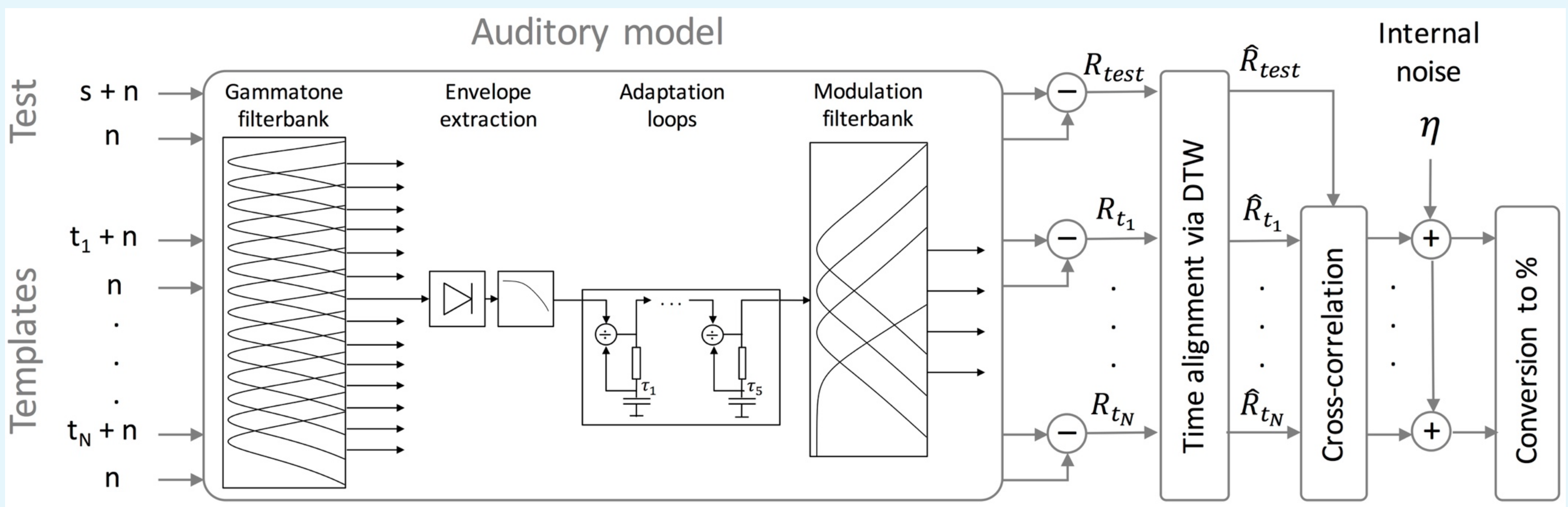

Fig. 2: Scheme of the consonant perception model. Reprint from Zaar and Dau (2016).

- The experimental data obtained in the *NLFC* and the *NLFC&INS* conditions were predicted using the consonant perception model of Zaar and Dau (2016).
- The experimental stimuli and the corresponding "noise alone" were considered as test signals.
- The stimuli from the *unaided* condition and the corresponding "noise alone" signals were considered as templates in the model.

CONCLUSIONS

- The consonant perception model of Zaar and Dau (2016) accounts well for HA-processing induced consonant morphs.
- The model could therefore be useful for evaluating HA processing strategies, particularly when combined with simulations of individual hearing impairment.

ACKNOWLEDGEMENTS

This research has been funded with support from the European Commission under Contract No. FP7-PEOPLE-2011-290000.

REFERENCES

Li, F., Menon, A., and Allen, J.B. (2010): "A psychoacoustic method to find the perceptual cues of stop consonants in natural speech," J. Acoust. Soc. Am. 127 (4): 2599–2610.

Li, F., Trevino, A., Menon. A., and Allen, J. B. (2012). "A psychoacoustic method for studying the necessary and sufficient perceptual cues of American English fricative consonants in noise," J. Acoust. Soc. Am. 132 (4): 2663–2675.

Scheidiger, C., and Allen, J. B. (2013). "Effects of NALR on consonant-vowel perception," 4th International Symposium on Auditory and Audiological Research (ISAAR-2013), Nyborg Denmark.

Schmitt, N., Winkler, A., Boretzki, M., and Holube, I. (2016): "A phoneme perception test method for high-frequency hearing aid fitting," J. Am. Acad. Audiol. 27 (5): 367–379.

Zaar, J., and Dau, T. (2016): "Predicting consonant recognition and confusions in normal-hearing listeners", J. Acoust. Soc. Am., *under review*.