Best application of head-related transfer functions for competing voices speech recognition in hearing-impaired listeners

LARS BRAMSLØW*, MARIANNA VATTI, RENSKJE K. HIETKAMP, AND NIELS HENRIK PONTOPPIDAN

Eriksholm Research Centre, Snekkersten, Denmark

When presenting separated speech sources over hearing aids, should the normal physical spatial cues be restored? The answer was sought by presenting speech sources to a listener via headphones, either directly or after application of generic head-related-transfer functions (HRTF) in different modes to simulate free-field listening. For the presentation of two competing voices, we have measured the relative monaural and binaural contributions to speech intelligibility using a previously developed competing voices test. Two consecutive tests, using 13 and 10 hearing-impaired listeners with moderate, sloping hearing losses were conducted, combining different HRTF modes and horizontal plane angles. We found that neither the monaural HRTF gain nor the binaural cues imposed through crosstalk do affect the speech recognition. The only factor improving the competing voices scores is a large spatial separation, with as little mixing of the two voices as possible.

INTRODUCTION

Many situations require listeners to attend to two equally important voices, e.g., a dinner situation, or answering questions while watching TV. In some cases, the two voices are available separately, e.g. streaming from two phone lines simultaneously. In a normal, physical acoustic situation, the two voices will always be mixed, but one might imagine a perfect separation algorithm. In this latter case, the question again comes up: Is the application of generic head-related transfer functions (HRTF) beneficial? Moreover, which HRTF contributions are important: the monaural open-ear gain (see Fig. 1) and/or the binaural cross-talk that provides natural interaural level cues (interaural level and time differences)?

According to Brungart and Simpson (2005), with normal-hearing listeners there is a advantage of roughly 5% by going from separate (dichotic) to binaural HRTF (termed ‘3D’ by the authors), in a word-based test known as Coordinate Response Measure (CRM; Bolia *et al.*, 2000).

The severity of the problem is typically larger for hearing-impaired (HI) listeners, and this study only concerns this group for hearing aid applications.

*Corresponding author: labw@eriksholm.com

Proceedings of ISAAR 2015: Individual Hearing Loss – Characterization, Modelling, Compensation Strategies. 5th symposium on Auditory and Audiological Research. August 2015, Nyborg, Denmark. Edited by S. Santurette, T. Dau, J. C. Dalsgaard, L. Tranebjærg, and T. Andersen. ISBN: 978-87-990013-5-4. The Danavox Jubilee Foundation, 2015.

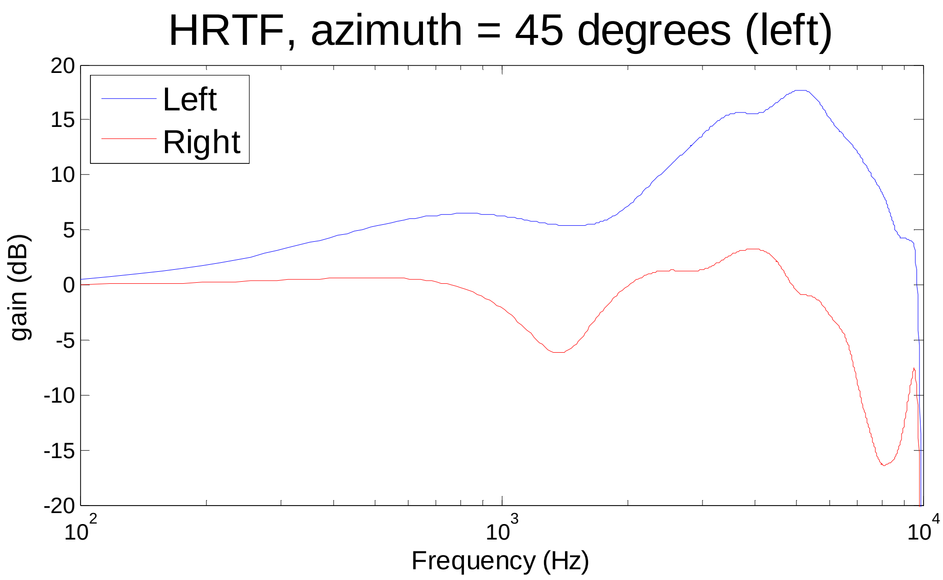

Fig. 1: Example of head-related transfer function (HRTF) from free-field frontal incidence (45 degrees azimuth) to blocked ear canal.

Aim of the study

To investigate the effect on competing voice situations from these components:

- The monaural component – due to the large gain applied by the open ear gain (OEG) contained in the HRTF. See Fig. 1 and, e.g., Fig. 3 (left pane).
- The binaural component – due to the interaural cues provided by the crosstalk from a sound source to the contralateral ear. See, e.g., Fig. 4 (right pane).
- The spatial separation effect from a co-located to a left-right configuration. See, e.g., Fig. 4 (right and left panes).

METHODS AND MATERIAL

Test method

The competing voices scenario was evaluated by using the Competing Voices Test (CVT), where two Danish Hearing In Noise Test (HINT; Nilsson *et al.*, 1994) sentences are played simultaneously in a spatial configuration. The Competing Voices Test was developed for these types of scenarios (Bramsløw *et al.* 2014; 2015). The task of the listener is to repeat sentences spoken either by the male or the female as prompted randomly (p = 0.5) by a sign on a PC monitor. The outcome measure was percent correct score, which was rau-transformed to provide better 'normal' distribution of the data (Studebaker, 1985).

Tests conducted

Two separate tests were conducted, which covered different aspects of the overall aim:

1. Test 1 evaluated the effect of applying generic binaural HRTF vs no HRTF in a spatially separated configuration. The HRTFs were measured on a Brüel & Kjær 4128 HATS manikin at the entrance of the blocked ear canal. The data shown here is the relevant subset from a larger experiment.
2. After test 1, it was speculated that normal binaural HRTF application (with crosstalk) was not the optimal. Therefore, test 2 evaluated the same contrast as test 1 plus the separate contributions from binaural HRTF (with crosstalk) and monaural HRTF (without crosstalk).

Spatial configurations via headphones

The possible spatial configurations are shown in Figs. 2, 3, 4 below.

Fig. 2: HRTF ‘Off’ spatial modes.

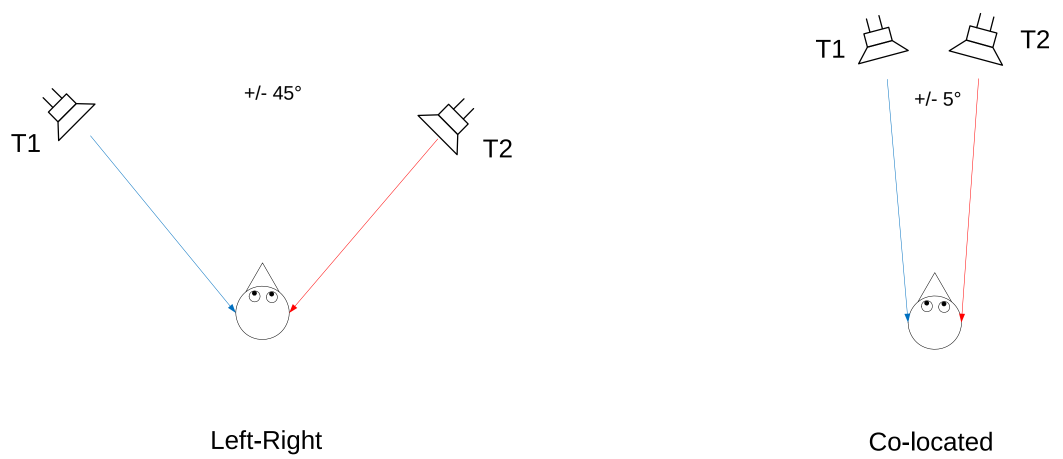

Fig. 3: HRTF ‘No crosstalk’ spatial modes.

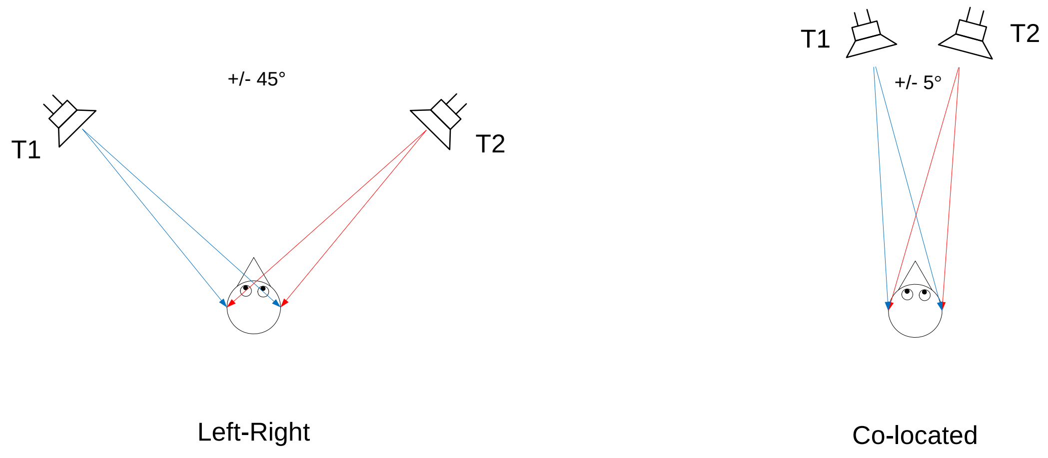

Fig. 4: HRTF 'Crosstalk' spatial modes. This corresponds to a normal physical loudspeaker setup.

The two tests covered these six possible configurations as listed in Table 1.

Conditions	Test 1		Test 2		
HRTF mode	**Off**	**Crosstalk**	**Off**	**No crosstalk**	**Crosstalk**
Left-Right	Separate	+/- 45 º	Separate	+/- 45 º	+/- 45 º
Co-located		+/- 5 º	Sum	+/- 5 º	+/- 5 º

Table 1: Overview of spatial modes and HRTF modes employed in the two tests.

Listeners

Both tests used elderly hearing-impaired listeners with moderate, sloping sensorineural hearing losses. Test 1 used 13 listeners and test 2 used 10 listeners, with an average age of 70 years. The hearing losses were compensated linearly according to the CAMEQ linear gain rule (Moore and Glasberg, 1998) and the listening level was set to most comfortable level during the initial training phase of the test.

RESULTS

The mean value scores are summarised in Table 2.

% CVT scores	Test 1		Test 2		
HRTF mode	**Off**	**Crosstalk**	**Off**	**No crosstalk**	**Crosstalk**
Left-Right	66.7 %	63.2 %	72.3 %	76.6 %	70.6 %
Co-located		50.9 %	54.1 %	76.7 %	54.6 %

Table 2: Overview of CVT scores from the two tests. The significantly different values in the two tests are underlined.

Both test 1 and test 2 were analysed using repeated-measures analysis of variance. In both cases, the following significant effects were found:

- Test person (test 1 and test 2). The values spread from app. 20 to 80% across test persons (not shown).
- Spatial mode (test 1 and test 2).
- HRTF mode (test 2).
- Spatial mode * HRTF mode (test 2). Note that test 1 was not a complete design of spatial mode and HRTF mode.

For test 2, the interaction of spatial mode is shown in Fig. 5. A post-hoc Tukey HSD test showed that:

Left-right: The scores in this spatial mode are not significantly different across the HRTF modes. Neither HRTF gain, nor crosstalk on/off affect the scores.

Co-located: ‘No crosstalk’ is significantly better than the two other conditions, because this does not have the large contralateral contribution as the other two HRTF modes do. The differences between ‘Off’ and ‘Crosstalk’ is the large gain from HRTF to both ears; however this does not affect the scores.

CONCLUSION

The best scores were obtained in the left-right (spatially separated) mode, as expected. In this spatial mode, there is no effect of HRTF, neither with or without HRTF. Thus, the HRTF (‘3D’ spatialised) advantage found in Brungart and Simpson (2005) was not replicated here for a hearing-impaired group.

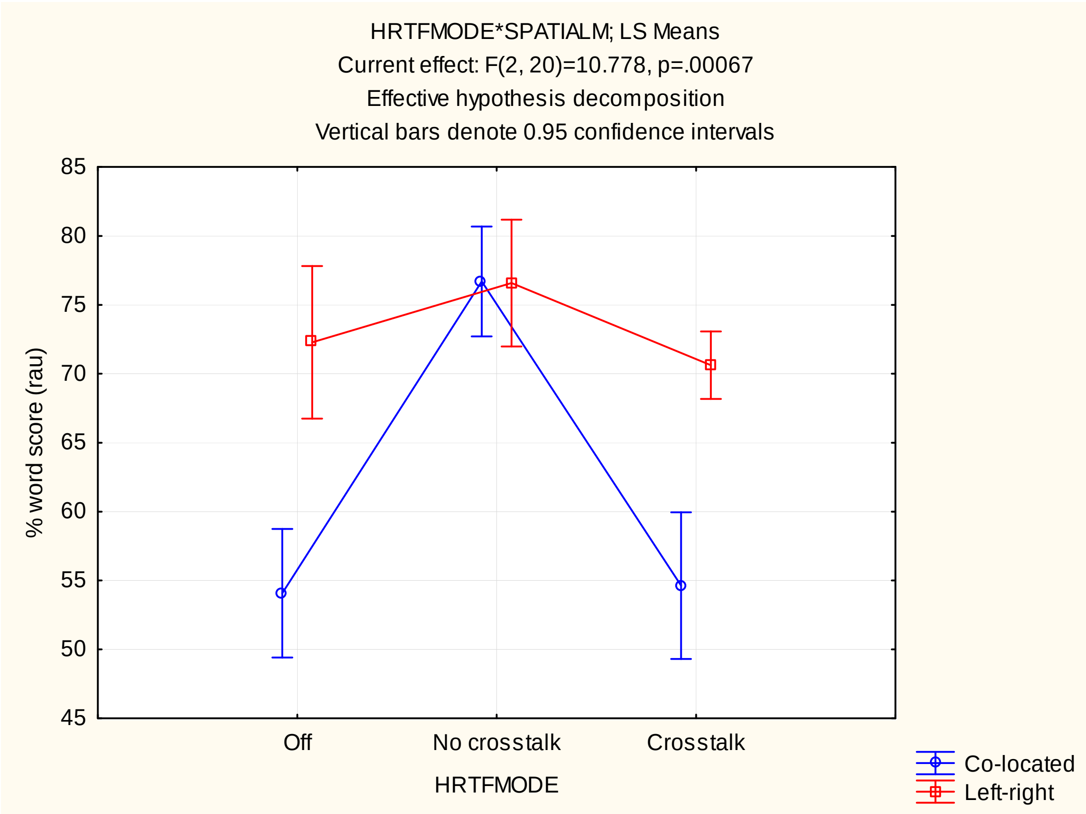

Fig. 5: Test 2 – Plot of the HRTF mode and spatial mode interaction.

Likewise, in the co-located mode, the effect of HRTF was due to removal of crosstalk, rather than due to the substantial gain difference (see Figure 1) added by the HRTF. The mode with no crosstalk is essentially separated regardless of angles.

The best presentation mode is thus spatially separated, without crosstalk. The normal transformation from free field to eardrum applied in hearing aids will satisfy this requirement.

REFERENCES

Bolia, R.S., Nelson, W.T., Ericson, M.A., and Simpson, B.D. (**2000**). “A speech corpus for multitalker communications research,” J. Acoust. Soc. Am., **107**, 1065-1066.

Bramsløw, L., Vatti, M., Hietkamp, R.K., and Pontoppidan, N.H. (**2014**). “Design of a competing voices test;” Poster presented at International Hearing Aid Conference (IHCON) 2014, Lake Tahoe, CA, USA. Available at: http://www.eriksholm.com/about-us/news/IHCON_2014.

Bramsløw, L., Vatti, M., Hietkamp, R.K., and Pontoppidan, N.H. (**2015**). "Binaural speech recognition for normal-hearing and hearing-impaired listeners in a competing voice test. Poster presented at Speech in Noise Workshop 2015, Copenhagen. Available at: http://www.eriksholm.com/about-us/news/2015/SPIN_2015.

Brungart, D.S. and Simpson, B.D. (**2005**). "Improving multitalker speech communication with advanced audio displays," In *New Directions for Improving Audio Effectiveness.* Neuilly-Sur-Seine, France, pp. 30/1-30/18.

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.

Nilsson, M., Soli, S.D., and Sullivan, J.A. (1994). Development of the Hearing In Noise Test for the measurement of speech reception thresholds in quiet and in noise," J. Acoust. Soc. Am., **95**, 1085-1099.

Studebaker, G.A. (**1985**). "A "rationalized" arcsine transform," J. Speech Hear. Res., **28**, 455-462.