Spatial unmasking in aided hearing-impaired listeners and the need for training

TOBIAS NEHER, THOMAS BEHRENS, LOUISE KRAGELUND, AND ANNE SPECHT PETERSEN

Oticon A/S, Research Centre Eriksholm, Kongevejen 243, 3070 Snekkersten, Denmark

Even though spatial unmasking (SU) in hearing-impaired subjects has been the subject of a number of studies, very little research seems to have been carried out under aided conditions, especially not for more complex speech-on-speech masking situations.

As part of an earlier pilot study into aided SU, a group of test subjects were found to exhibit substantial training effects across different visits, despite some initial training. A new training program was therefore designed based on some findings from the perceptual learning and training literature. Nine elderly hearing-aid users with mild-to-moderate, sloping hearing losses were systematically trained in a speech-on-speech SU task. All subjects were bilaterally fitted and only tested with their own hearing aids.

Using a new speech corpus suitable for speech-on-speech SU assessment, performance was then determined at two subsequent visits. Whilst there were substantial differences between test subjects, half of them showed SU as large as 10 dB. Moreover, performance across the two visits was found to be much more stable. These results hint at the need for thorough training when elderly hearing-aid users are to be tested under complex listening conditions.

INTRODUCTION

It is well known that spatial hearing can help listeners achieve better speech intelligibility in complex listening situations such as cocktail parties where there are other talkers interfering with a target signal (e.g. Bronkhorst, 2000). In the laboratory, the contribution of spatial cues to speech intelligibility is usually quantified using an SU test paradigm. In such a paradigm, a target signal is presented together with one or more masker signals, whereby the target and masker(s) exhibit either the same or different spatial cues. For each condition, the Target-to-Masker Ratio (TMR) corresponding to a given level of speech recognition is then determined and the difference in TMR taken as a measure of the release from masking afforded by the availability of spatial differences.

Whilst considerable research has dealt with the effects of hearing loss on SU [see (Durlach *et al.*, 1981) for an overview], most of that work has used non-representative test signals. Moreover, few studies have been conducted under aided conditions. Consequently, this study was designed to quantify spatial release from speech-on-speech masking in elderly hearing-aid users. In addition, it aimed to find out if systematic training can lead to more stable subject performance over time. This additional aim was motivated by a previous study, as part of which a group of hearing-aid users (with

hearing losses very similar to the ones of the subjects tested in this study) had been briefly task-trained immediately before the first SU measurement. Since that type of training was found inadequate to prevent substantial performance variability across different visits, a new training program was devised and tested.

EXPERIMENTAL CONDITIONS

Physical test set-up

All training and testing was carried out under anechoic conditions. Four loudspeakers were positioned in the horizontal plane at 0°, ±50° and 180° (cf. Fig. 1). The distance to the listening position was about 1.6 m. Below the frontal loudspeaker an LCD screen was hung that was used for displaying instructions. The subject was seated in a custom-made chair equipped with a head rest that was small enough, so that sound reaching a subject's ears from behind was not obstructed. The chair was adjusted as necessary to ensure that the subject's head was located precisely in the middle of the test set-up and that the subject was seated comfortably. The subject was instructed to move as little as possible whenever measurements were made, which was also checked by the experimenter with the help of a video monitoring system.
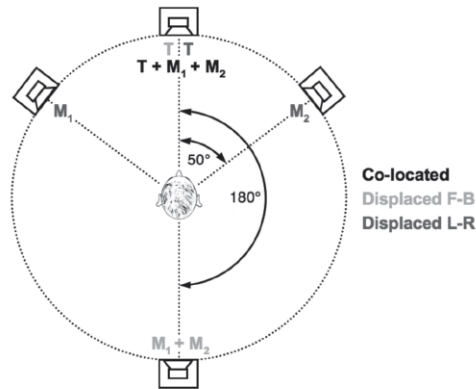

Fig. 1: Physical test set-up and spatial test conditions.

Spatial test conditions

To quantify SU, speech intelligibility was measured for three spatial test conditions: (1) co-located, (2) displaced F-B, and (3) displaced L-R (cf. Fig. 1). In the co-located condition, three speech signals were presented simultaneously from the frontal loudspeaker. One of the signals served as the target (T) and the other two as maskers (M_1 and M_2). In the displaced F-B condition, the target stayed in front and the maskers were presented from behind. In this condition, only high-frequency monaural spatial cues (e.g. Middlebrooks and Green, 1991) were therefore potentially available to the subjects to spatially separate the target from the maskers. In the displaced L-R condition, the target came still from in front, whilst the maskers were presented from the left and right loudspeaker, respectively. In this condition, both interaural and monaural spatial cues were therefore

potentially available to the subjects. To allow the subjects to identify the target signal in the co-located condition, the first word of the target sentence (the "call sign") was displayed on the LCD screen. The subjects' task was to repeat all five words of the target sentence (the call sign scores were not included when estimating speech recognition).

Speech material

The speech material used was a modified version of the DANTALE II speech corpus (Wagener *et al.*, 2003). A detailed description of this new corpus can be found in the companion paper by Behrens *et al.* (2007). Briefly, it consists of a large set of Danish sentences, spoken by each of five trained female talkers, that all follow the form "name verb numeral adjective object". A typical utterance would be "Michael had eight yellow houses".

Test subjects and hearing aids

Nine experienced, bilateral hearing-aid users with symmetrical, sloping hearing losses (cf. Fig. 2) participated in this study. In terms of hearing loss averaged across the audiometric frequencies of 0.5, 1, 2, and 4 kHz ("4FA-HL"), the subjects ranged from 41 to 58 dB HL ($\mu = 48$ dB HL). In terms of age, they ranged from 51 to 80 yrs ($\mu = 69$ yrs). All subjects used Oticon Syncro hearing aids that were either of the ITE, ITC or CIC type. These hearing aids have an output bandwidth of ca. 7 kHz. BTE devices were deliberately excluded to ensure that the hearing-aid microphones were located as close to the ear-canal entrances as possible. Subjects were only tested with their own hearing aids. Prior to all testing the noise reduction and directionality systems were disabled, as these could have modified the masker signals in the displaced conditions. In terms of insertion gain, there was very good correspondence between prescribed and fitted values up to 4 kHz; at 6 kHz four subjects received amplification that was about 5 dB lower than prescribed.
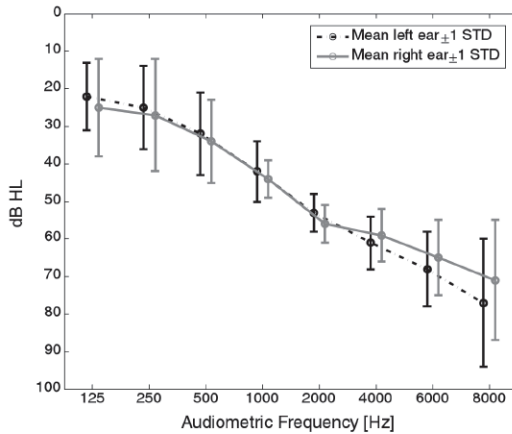
Fig. 2: Audiometric data of test subjects.

Estimation of spatial unmasking

In order to estimate SU, 50%-correct speech intelligibility thresholds were measured for the three spatial conditions. SU was then calculated by taking the difference between the TMR corresponding to the 50%-correct threshold estimate obtained for the co-located condition and the TMR corresponding to the 50%-correct threshold estimate obtained for either the displaced F-B (SU_{F-B}) or the displaced L-R (SU_{L-R}) condition:

$$SU_{F-B} = TMR_{Co-loc} - TMR_{Displ\_F-B}, [dB] \qquad \text{(Eq. 1)}$$

$$SU_{L-R} = TMR_{Co-loc} - TMR_{Displ\_L-R}, [dB] \qquad \text{(Eq. 2)}$$

The individual 50%-correct threshold estimates were extracted from psychometric functions that had been derived using the method of constant stimuli. Before the start of the actual SU measurement, all subjects completed a brief "task brush-up" that included a few "easy" TMRs, so that they could get used to the different conditions again. Next, 30 "pre-trials" were run per condition to get an indication of where each subject's threshold was likely to lie. The resultant data were then used to derive psychometric functions with the help of a maximum-likelihood estimation procedure. From these functions, a few suitably placed TMRs were extracted and another block of 30 trials was run per condition. Only the data from the pre-trials and subsequent trials were used to estimate the final thresholds.

Training program

The design of the training program was based on a gradual build-up of the task complexity, concluding with the actual SU task. It consisted of seven steps, which are summarised in Table 1. At the start of each step, the subject was provided with verbal and written instructions regarding the details of the subsequent stimuli. After each stimulus presentation, the experimenter provided verbal feedback to let the subject know if a given response was correct or not. If the subject had made a mistake, the experimenter gave instructions in terms of what aspect of the stimulus to pay attention to and played the same stimulus up to two times more, so that the subject could correct his response. Such feedback provision is known to play an important role in the successful outcome of training programs, including audiological ones (e.g. Sweetow and Palmer, 2005).

Step	Signal(s)	What's new?	No. of trials
1	T	L/S introduced using one sentence; Task: Locate T	8
2	T	Task: Repeat T, locate T	12
3	T	Sentence changes across trials	12
4	$T + M_1$	M1 introduced; Task: Repeat T, locate M_1	12
5	$T + M_1 + M_2$	M2 introduced; Task: Repeat T, locate $M_1 + M_2$	12
6	$T + M_1 + M_2$	Target call sign changes; Task: Repeat T	15
7	$T + M_1 + M_2$	TMRs change	45

Table 1: Training program details. For each step, the signals presented, the new elements introduced, and the number of trials run are indicated (L/S = loudspeakers).

Experimental time course

The training study comprised three separate visits per subject (cf. Fig. 3). At the first visit, all subjects were trained according to the program outlined in Table 1. At the second and third visit, SU performance was measured as described above. The decision to carry out the training program and first SU measurement at different visits was motivated by some research findings related to perceptual learning. Ortiz and Wright (2005) investigated the effect of a rest period on the performance of subjects who were trained in terms of an interaural time difference discrimination task. They were able to show that performance improved significantly if subjects were allowed to rest for about 9 hrs before the actual measurements were carried out compared to when the measurements were made immediately after the training program. Since for this study it was of interest to determine the longer-lasting effects of the training program on performance stability, a 1-week gap was chosen as the time interval between each pair of visits.
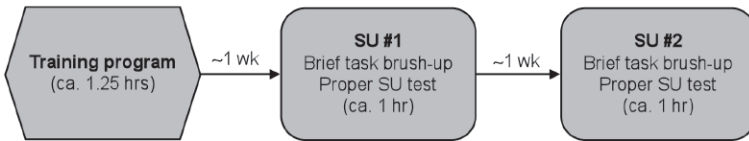

Fig. 3: Experimental time course.

RESULTS

Even though nine test subjects had participated in this study, SU was immeasurable for one of them, as speech intelligibility could not be estimated in the co-located condition. The results reported below are therefore based on the data from the remaining eight subjects.

Spatial unmasking performance

Fig. 4 displays the mean and raw SU data of the hearing-impaired (HI) subjects as a function of displaced condition (F-B or L-R) and measurement number (#1 or #2). The SU data of nine (trained) normal-hearing (NH) subjects (25-46 yrs; $\mu = 37$ yrs) are also shown for comparison [see (Behrens *et al.*, 2007) for more details]. On average, the HI group obtained ca. 8 dB less SU than the NH group, which is broadly in line with data reported by Marrone *et al.* (2007). In absolute terms, however, both subject groups exhibited several dB more SU than the ones of Marrone *et al.*, which could have been due to more training or smaller hearing losses, for example. Furthermore, SU_{L-R} was generally larger than SU_{F-B}, most probably because of the availability of interaural spatial information. Nevertheless, for the HI group there was also much more variation in SU_{L-R} than in SU_{F-B}. Whilst one subject obtained basically no SU_{L-R}, four other subjects exhibited around 10 dB of SU_{L-R} and came therefore fairly close to NH performance. This finding is somewhat in contrast to other such data that were obtained using headphone-presented stimuli and that showed less SU in similarly impaired listeners (Kalluri and Edwards, 2007). The large variation in SU_{L-R} furthermore indicates

that the HI subjects differed substantially in terms of their ability to exploit left-right spatial information to achieve release from masking, but that they were more homogeneous in terms of the extent to which they could access high-frequency monaural cues. The NH group, on the other hand, exhibited little variation in $SU_{L\text{-}R}$ and much variation in $SU_{F\text{-}B}$. This suggests that these subjects were all able to exploit left-right spatial information, but that they differed substantially in their ability to exploit front-back spatial cues, possibly because these cues are less robust or potent compared to left-right spatial information.
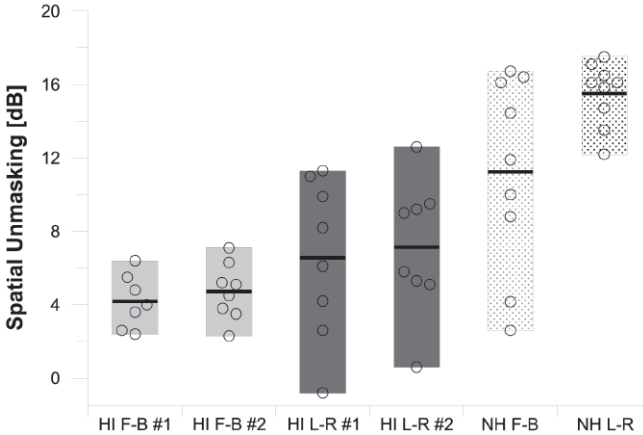

Fig. 4: Mean and raw SU data from eight HI and nine NH subjects plotted as a function of displaced condition (F-B or L-R) and measurement number (#1 or #2).

Effects of training

The efficacy of the training program was assessed by computing training effect and test-retest standard deviation (STD) estimates. For both SUF-B and SUL-R there was an average improvement of just 0.5 dB across the two visits (cf. Table 2). The test-retest STDs corresponding to the two SU estimates were 2.1 and 1.5 dB, respectively. These results compare favourably with those from the earlier SU pilot test, which had produced training effects of 0.9 and 1.6 dB and test-retest STDs of 2.9 and 2.1 dB, respectively. Hence, it can be concluded that subject performance in both SUF-B and SUL-R became more stable as a result of the new training.

	$SU_{F\text{-}B}$	$SU_{L\text{-}R}$
Mean #1 [dB]	4.2	6.6
Mean #2 [dB]	4.7	7.1
Training Effect [dB]	0.5	0.5
Test-Retest STD [dB]	2.1	1.5

**Table 2**: Training effect and test-retest STD estimates.

Effects of hearing loss

The effects of hearing loss on SU performance were examined by computing product-moment correlations between the different SU measures as well as 4FA-HL and indices of low-frequency and high-frequency hearing ability. To that end, the average hearing loss across the audiometric frequencies of 0.125, 0.25, 0.5, and 0.75 kHz ("4FA-HL low") as well as 3, 4, 6, and 8 kHz ("4FA-HL high") was calculated for each subject. The results are shown in Table 3, where correlations significant at the p < 0.05 level have been marked with an asterisk. There are strong (negative) correlations for SUL-R but not for SUF-B. Moreover, SUL-R seems to be pre-dominantly correlated with hearing ability in the lower frequencies. One could speculate that this was because the subjects with worse low-frequency hearing thresholds were less able to exploit interaural time difference information, which is known to dominate spatial perception in the low frequencies (Wightman & Kistler, 1992).

	4FA-HL	4FA-HL low	4FA-HL high
SU_{F-B} #1	-0.01	0.05	-0.42
SU_{F-B} #2	-0.02	-0.15	0.36
SU_{L-R} #1	-0.92*	-0.82*	-0.61
SU_{L-R} #2	-0.94*	-0.81*	-0.60

Table 3: Product-moment correlations between the various SU estimates and 4FA-HL, '4FA-HL low', and '4FA-HL high' (* = p < 0.05).

To further illustrate the negative correlation between degree of hearing loss and SUL-R, Fig. 5 displays scatter plots of 'SUL-R #2' and 4FA-HL or '4FA-HL low', together with lines of best fit.

(a)   (b)
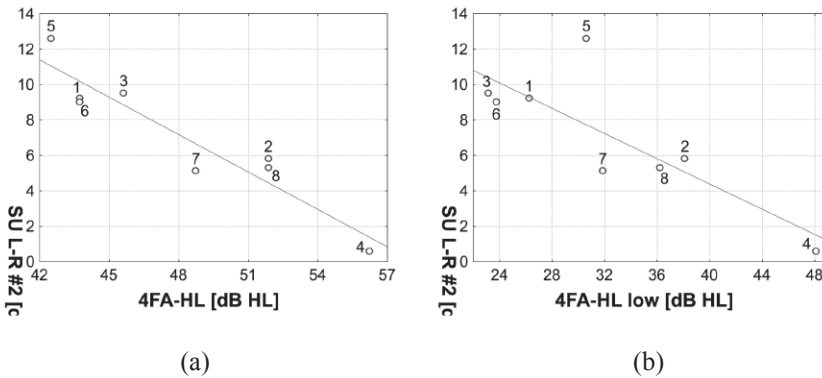
Fig. 5: Scatter plots of 'SU_{L-R} #2' and (a) 4FA-HL or (b) '4FA-HL low'.

(No systematic relationships between the various SU measurements and age, gender, hearing-aid type or high-frequency insertion gain were found.)

CONCLUSIONS

The results reported above show that even though, on average, elderly hearing-aid users obtain significantly less spatial release from speech-on-speech masking than NH subjects, some can come fairly close to NH performance. Furthermore, the results point towards a strong relationship between SUL-R and (low-frequency) hearing loss. SUF-B, in turn, is generally smaller than SUL-R and also more variable between NH and within HI subjects, which can probably be traced back to the (high-frequency) spatial cues being less robust or potent. In terms of changes in subject performance over time, this study indicates that thorough training can help stabilise the performance of elderly hearing-aid users that are tested under complex listening conditions. In view of the relatively small number of test subjects used, however, further research is needed to substantiate these findings as well as to separate the effects of hearing loss and aiding on SU performance.

REFERENCES

Behrens, T., Neher, T., and Burmand Johannesson, R. (**2007**). "Evaluation of speech corpus for assessment of spatial release from masking," Proceedings of the International Symposium on Audiological and Auditory Research, Helsingør, Denmark, Aug. 29-31.

Bronkhorst, A. W. (**2000**). "The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions," Acta Acust. Acust. **86**, 117-128.

Durlach, N. I., Thompson, C. L., and Colburn, H. S. (**1981**). "Binaural interaction in impaired listeners: A review of past research," Audiol. **20**, 181-211.

Kalluri, S., and Edwards, B. (**2007**), "Impact of hearing impairment and hearing aids on benefits due to binaural hearing," 19th International Congress on Acoustics, Madrid, Spain, Sept. 2-7.

Marrone, N. L., Mason, C. R., and Kidd, G. Jr. (**2007**). "Spatial release from speech-on-speech masking in adult hearing aid users," 153rd Meeting of the Acoustical Society of America, Salt Lake City, Utah, June 4-8.

Middlebrooks, J. C., and Green, D. M. (**1991**). "Sound localization by human listeners," Ann. Rev. Psychol. 42, 135-159.

Ortiz, J. A., and Wright, B. A. (**2005**). "Effects of different amounts of brief training and rest on the generalization of learning from interaural-level-difference to interaural-time-difference discrimination," J. Acoust. Soc. Am. **117**, 2561.

Sweetow, R. and Palmer, C. V. (**2005**). "Efficacy of individual auditory training in adults: A systematic review of the evidence," J. Am. Acad. Audiol. **16**, 494-504.

Wagener, K., Josvassen, J. L., and Ardenkjær, R. (**2003**), "Design, evaluation and optimization of a Danish sentence test in noise," Int. J. Audiol. **42**, 10-17.

Wightman, F. L., and Kistler, D. J. (**1992**). "The dominant role of low-frequency interaural time differences in sound localization," J. Acoust. Soc. Am. **91**, 1648-1661.