Variations in “adequate” own-voice level used by speakers and preferred by listeners when communicating across a distance

Søren Laugesen, Niels Søgaard Jensen, Patrick Maas, and Claus Nielsen

Eriksholm, Oticon Research Centre, Kongevejen 243, DK-3070 Snekkersten, Denmark

This paper describes a follow-up study to two pilot experiments on hearing-aid users’ ability to control the level of their own voice. In particular, the role of the so-called intervener in the experiments was examined. The intervener’s task is to supervise the speakers into speaking at the adequate level for the occasion. The study reveals that there are indeed differences among interveners, but that in broad terms the variation among interveners is similar in magnitude to the individual intervener’s test-retest consistency. Also the test-retest consistency of the speakers was examined, and was found to be similar to the aforementioned intervener-variations. The between-speaker variation is about twice as large, though. This suggests that it is highly necessary to include several speakers in experiments on own-voice level control, whereas relying on a single intervener is only mildly problematic.

INTRODUCTION

One aspect of successful communication is using the adequate voice level for the occasion (Lane and Tranel, 1971). While own-voice level control may be trivial for the normally hearing, it has been found to be difficult for hearing-aid users (Nielsen and Laugesen, 2004; Jensen *et al.*, 2006). Two previous level-control experiments with hearing-aid users (one unpublished and one presented by Laugesen *et al.*, 2006) have raised the question about the variability in “adequate level for the occasion”, which is the subject of the present study.

METHOD

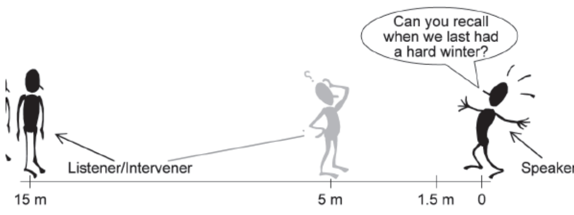

Fig. 1: Simple sketch of the experimental set-up.

The variations in adequate voice level were investigated by having each of four speakers asking a predefined question to each of four listeners/interveners across a range

Auditory signal processing in hearing-impaired listeners. 1st International Symposium on Auditory and Audiological Research (ISAAR 2007). T. Dau, J. M. Buchholz, J. M. Harte, T. U. Christiansen (Eds.). ISBN: 87-990013-1-4. Print: Centertryk A/S.

of distances. The speakers either spoke at the level they found adequate themselves (unsupervised condition), or at the level found adequate by the intervener (supervised condition). In both conditions the speakers repeated the question until a satisfactory version was produced (according to the speaker and intervener, respectively). A simple sketch of the set-up is shown in Fig. 1.

Thus, the main independent variables were SPEAKER (1, … 4), INTERVENER (1, … 4), and DISTANCE (1.5, 5, and 15 m). In the unsupervised condition, each combination was tested twice, described by the TEST/RETEST variable. In the supervised condition, each combination was also tested twice with the speaker either wearing a set of 3M 1310 hearing protectors or with open ears, described by the OPEN/ATT variable. The hearing protectors were of the ear-plug type, and were expected to create a considerable occlusion effect, which – together with the attenuation of the air-conducted component – would disturb the speaker's auditory feedback and thus provoke *in*adequate vocal level from the speaker (Lane and Tranel, 1971). This was introduced to increase the likelihood of the interveners actually having to make corrections to the voice level of the speakers.

Protocol

The order of trials used in the experiment was determined as a compromise between being practical for both experimenters and test subjects, and obtaining a reasonable degree of balancing and randomisation. In the resulting order of trials, the interveners were only active in three trials in a row (at three different distances) while the speakers were active in 15 trials in a row during a session. This was partly out of practical considerations (the speakers had to be equipped with a behind-the-ear (BTE) microphone on their right ear), and partly because randomisation was most important for the interveners. During a session, the intervener who was not active (but still present) was equipped with circumaural hearing protectors so as to make her/him unable to make note of how the other intervener performed the experimental task. With regard to the speakers, they always spoke across different distances in consecutive trials, and the order of the distances differed between test and retest. Finally, in the supervised trials, the speaker alternated as often as possible between the use of hearing protectors and open ears. However, it was necessary to accept two consecutive open-ear trials for each speaker in each session.

The entire experiment was carried out in one day, which was started by an introductory meeting for all eight test subjects, where the purpose of the experiment was briefly explained. The written instructions were given to speakers and interveners, respectively, and the ear-plug hearing protectors were distributed and tried out by the speakers. Finally, the experimental procedure was illustrated by means of a small role play, and discussed.

MATERIAL

As already mentioned, the present study was a follow-up on two more elaborate investigations with hearing-aid users. In order to keep time consumption to within reasonable limits, it was decided to reuse as much as possible from the most recent of the two studies (Laugesen *et al.*, 2006), and to recruit the test subjects among colleagues.

Speakers

The speakers were selected to have normal hearing, Danish as native language, normal vision (in order to be able to eyeball the distance to speak across), even distribution of gender, and no markedly dysphonic voices. The key part of the instructions for the speakers in the unsupervised condition was: *"Your voice should have a level that you think is adequate for the given situation. This means that the person in front of you should be able to understand what you are saying, while you neither speak unnaturally soft nor loud – according to your own perception – taking the surroundings and the distance into consideration"*. In the unsupervised condition the speakers addressed a dummy listener (the first author that served as lead experimenter) who did not provide any feedback to the speakers. For the supervised condition the speaker simply had to repeat the question, and adjust the voice level according to the intervener's instructions.

Interveners

The interveners were selected to have normal hearing, Danish as native language, and even distribution of gender. Furthermore, two interveners had an above-average knowledge of the human voice and logopaedics (expert listeners) while two were (in this respect) more naïve. Age, gender, expert/naïve status, and audiograms (actually obtained after the experiment) of the four interveners are shown in Fig. 2. Here normal thresholds are observed, with the exception of the right ear of intervener 2, which shows an unexpected, considerable high-frequency hearing loss.

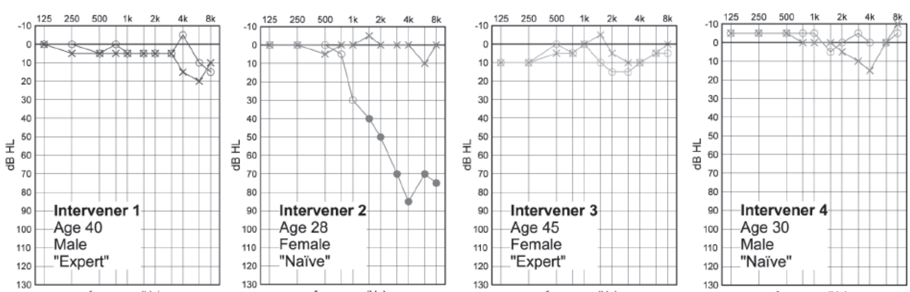

Fig. 2: Audiograms of the four interveners, as indicated. Crosses designate left ear and circles right ear. Filled circles designate masked thresholds. Age, gender, and expert/naïve status is also included.

The key part of the instructions for the interveners was: "*Your task is to decide whether or not the other person speaks at a level, which is adequate in the given situation, when he or she addresses you with the question. This means that you should be able to understand the question, while you at the same time think that the person is speaking at a level, which is neither unnaturally soft nor loud, taking the surroundings and the distance into account"*.

Location

The experiment was carried out in a quiet corridor that allowed speaker-listener distances up to 22 m. Measurements performed with a Tannoy System 1200 loudspeaker, a Brüel&Kjær 4144 condenser microphone connected to a Brüel&Kjær 5935 power supply, and the Brüel&Kjær PULSE system showed that between distances of 1 m up to 15 m the sound pressure level (SPL) due to the loudspeaker gradually drops by about 17 dB. At distances longer than 15 m, the SPL was basically constant. In a free field, the SPL drops by 23.4 dB from 1 to 15 m. Thus the SPL drops by less in the corridor than in a free field, but by much more than traditional "shoebox" room acoustics would predict. Similar observations were made by Wang *et al.* (2005).

Voice recordings

The speaker's voice was recorded through a Knowles EK3027CX electret microphone mounted in a BTE hearing-aid shell that also contained the microphone's power supply. The microphone signal was routed through a preamplifier to the digi001 front-end of a ProTools hard-disk recording system. Each experimental trial was recorded in its entity, and each recording subsequently edited to leave only the final valid utterance that had been accepted by the speaker (unsupervised condition) or the intervener (supervised condition). Finally, the power spectral density was determined across each whole utterance, the calibration response of the Knowles microphone was applied, and a single broad-band long-term average sound pressure level (re. 20 μPa) was formed.

BASIC RESULTS

First, the broad-band level data described above are analysed for effects of the experimental variables. This was done with mixed-model ANOVAs with SPEAKER and INTERVENER as random factors and DISTANCE, TEST/RETEST, and OPEN/ATT as fixed factors.

Supervised condition

In a model that only considers main effects there are highly significant effects (adopting a $p < 0.01$ limit) of the SPEAKER, INTERVENER, and DISTANCE variables, as shown in Fig. 3. In contrast, there is no effect of the OPEN/ATT variable ($p = 0.4$), which indicates that the interveners requested the same vocal level from the speakers, irrespective of whether the speakers wore ear-plugs or not – as one would hope for. The results in Fig. 3 show the expected increase in mean level with distance, as well as some degree of variation in the mean levels accepted by the interveners and produced by the speakers.

If the statistical model is expanded to include second-order interactions, SPEAKER and INTERVENER loose their significance, as is seen in table 1. Again, there is no effect of OPEN/ATT, but two significant interactions appear that are illustrated in Fig. 4. Higher-order models were also tested but no further significant effects emerged. The results in table 1 suggest that the significant main effects of SPEAKER and INTERVENER from the main-effects model actually belong to the two significant interactions in Fig. 4.

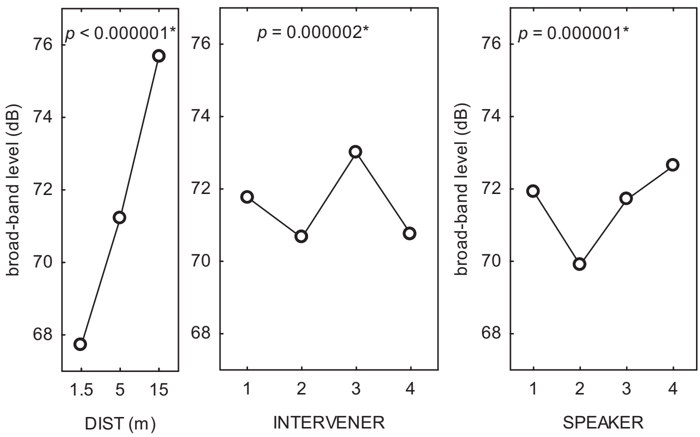

Fig. 3: Plots of the three significant effects in the main-effects ANOVA concerning the supervised condition, in terms of mean values and p-levels.

Effect	*p*-level
SPEAKER (rand)	0.06
DISTANCE (fix)	0.00007*
INTERVENER (rand)	0.19
OPEN/ATT (fix)	0.4
SPEAK*DIST (rand)	0.3
SPEAK*INT (rand)	0.000003*
SPEAK*O/A (rand)	0.4
DIST*INT (rand)	0.00008*
DIST*O/A (fix)	0.9
INT*O/A (rand)	0.2

Table 1: Results of a test of significance in a second-order mixed-model ANOVA. Significant effects ($p < 0.01$) are marked with an *.

The left-hand plot in Fig. 4 shows an interesting difference among the interveners. At 1.5 m distance, all four interveners seem to agree on the adequate level. However, as the distance goes up, interveners 1 and 3 request increasingly higher levels from the speakers than interveners 2 and 4. This observation immediately draws the attention to the audiometric profiles of the interveners (Fig. 2). However, this does not explain the results in Fig. 4 (left), since interveners 1 and 3 show thresholds well within the normal range, whereas intervener 2 – who has a one-sided hearing loss – requested the lowest levels of all. Finally, it is interesting that interveners 1 and 3 are the “expert interveners” as mentioned above, but this issue is not pursued any further. The results to the right in Fig. 4 once again show the difference in overall level requested by the four interveners, but mainly suggest a difference in pattern across the four speakers. Thus, interveners 1, 2, and 4 basically produced parallel patterns, whereas the pattern produced by intervener 3 deviates. In particular, the level requested from speaker 2 by intervener 3 is strikingly high.

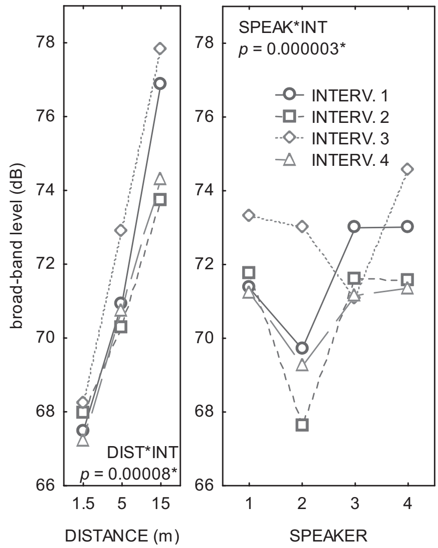

Fig. 4: Plots of the significant interactions from table 1.

Unsupervised condition

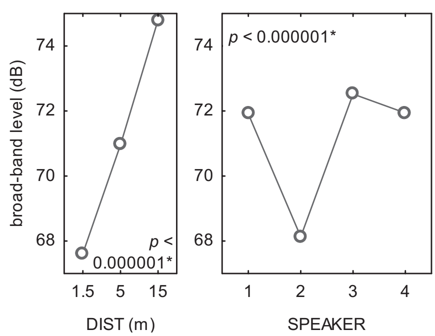

Fig. 5: Plots of the two significant main effects from the unsupervised condition.

In this condition only the two main effects illustrated in Fig. 5 are significant. Thus, there was no systematic effect of the TEST/RETEST variable and no significant interactions. A comparison of the results in Fig. 5 with the equivalent ones from the supervised condition in Fig. 3 shows a smaller increase in mean level with distance and a larger variation among the speakers in this unsupervised condition.

DETERMINATION OF KEY VARIABILITIES

In this section, all four variabilities that belong to the speakers and interveners are determined; partly from the present data and partly from the results of the two previous studies with hearing-aid users. These are briefly described below.

The two previous studies

In the *background-noise* test six hearing-aid users spoke in various levels of background noise, with different amplification and vent conditions. In the *distance test* (Laugesen *et al.*, 2006) seven hearing-aid users spoke across various distances (as above), with different amplification conditions. The same single intervener (no. 1 in this study) was used in both cases.

Between-intervener variation

This is estimated directly from the mean values shown in the INTERVENER plot in Fig. 3 (supervised), in terms of the standard deviation, which is 1.1 dB. Note that this simple approach disregards the two significant SPEAKER*INTERVENER and DISTANCE*INTERVENER interactions.

Within-intervener variation

For this study, a first estimate is found from the residual error in the statistical model that includes only the significant main effects and interactions. In this supervised condition, the residual error may be attributed to the intervener (since the speakers are controlled by the interveners, and distance was determined with great accuracy). The result is 1.2 dB.

A second estimate may be found from the differences between the OPEN and ATT conditions, which represent an approximate test-retest. Such results are shown in Fig. 6, individually for each intervener. Simple F-tests suggest that neither pair of variances are different, which means that all four interveners basically were equally reliable. Moreover, a common test-retest standard deviation can be determined, as indicated in Fig. 6.

Note that each test-retest difference, D, is the subtraction of two levels, L. Thus, $\mathrm{Var}\{D\} = \mathrm{Var}\{L_1–L_2\} = 2 \cdot \mathrm{Var}\{L\}$, which means that the within-intervener standard deviation is $\sigma_L = \sigma_D/\sqrt{2}$, where σ_D is the test-retest standard deviation.

Fig. 6: Raw intervener "test-retest" differences, with individual and common standard deviations as indicated.

A similar test-retest approach was applied to the data from the two previous studies (only intervener 1). The two "test-retest" difference standard deviations are 1.9 dB (background-noise) and 1.7 dB (distance), which are close to the 1.6 dB found for intervener 1 in this study. Final scaling by √2 and forming a common weighted within-intervener standard deviation across all three studies yields the result 1.3 dB.

Between-speaker variation

For this study, the between-speaker variation is again estimated directly from the mean values shown in the SPEAKER plots (supervised and unsupervised). Similar data are also available from the two previous studies and are included in the two plots in Fig. 7. As above, F-tests suggest that in each condition neither pair of variances are significantly different. Thus, a common weighted standard deviation can be determined for each condition. It is interesting that supervision reduces the between-speaker variation in all three studies. Examination of the results in Fig. 7 indicates that this is achieved mainly by bringing up the level from the softest speakers.

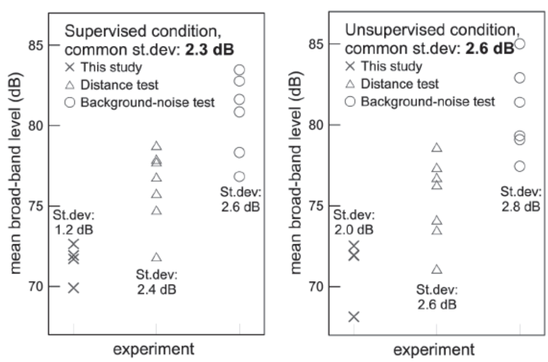

Fig. 7: Speaker-specific mean levels from all three studies, with study-individual and common standard deviations as indicated.

Within-speaker variation

This is only meaningfully estimated from the unsupervised condition, since the intervener in principle is in control of the vocal level in the supervised condition.

For this study, the estimate was found directly from the residual error in the statistical model that includes only the two significant main effects. In this unsupervised condition the residual error may be attributed solely to the speaker (since distance again was determined with great accuracy). The result is 0.9 dB, which was confirmed by a more direct evaluation of test-retest differences, similar to the one above.

Unfortunately, the two previous studies do not allow for estimates of the within-speaker variation.

Summary of variations

A summary of the determined variations is given in table 2. The key observation is that the between-speaker variation is about twice as large as the other variations.

Variation	Supervised	Unsupervised
Between-intervener	1.1 dB	
Within-intervener	1.3 dB	
Between-speaker	2.3 dB	2.6 dB
Within-speaker		0.9 dB

Table 2: Summary of key variabilities.

CONCLUSIONS

The results presented above show that there is considerable between-speaker variation in the overall level used for a given distance, both in the unsupervised and supervised conditions. The between-intervener variation is much smaller, and is similar in magnitude to the test-retest measures (within-speaker and within-intervener). There are, however, notable differences among the interveners, particularly at the long distances.

These results have implications for the design of experiments with own-voice level control. This is important because it has been found that while own-voice level control is trivial for the normally hearing, it is difficult for hearing-aid users. Furthermore, the results illustrate the unpredictability under which hearing aids operate. E.g., the level of “normal speech at 1 m” is often referred to when gain rules for hearing aids are considered. Assuming that “normal” also means “adequate”, the above results point out that in real life this level depends strongly on the speaker and (to a lesser extent) on who defines what is adequate.

REFERENCES

Jensen, N. S., Maas, P., Nielsen, C., Laugesen, S. and Flynn, M. (**2006**). “A first attempt at a comprehensive Own Voice Qualities (OVQ) questionnaire,” Poster presented at IHCON, Lake Tahoe, USA.

Lane, H. L., and Tranel, B. (**1971**). “The Lombard sign and the role of hearing in speech,” J. Speech Hear. Res., **14**(4), 677-709.

Laugesen, S., Nielsen, C., Maas, P., and Jensen, N.S. (**2006**). “Effects of amplification variations on vocal level and own-voice sound quality of hearing-aid users,” Podium presentation at IHCON, Lake Tahoe, USA.

Nielsen, C., and Laugesen, S. (**2004**). “Why own voice is different,” Instructional Course presented at the American Academy of Audiology Convention, Salt Lake City, USA.

Wang, L. M., Rathsam, J., Christensen, C. L., and Rindel, J. H. (**2005**). “Sound-pressure level distribution in a long, narrow hallway: Measurements versus results from a computer model with scattering from surface roughness and diffraction," J. Acoust. Soc. Am., **118**, 1998.