Edinburgh, Scotland
EURONOISE 2009
October 26-28

Most comfortable level of speech under different illuminations

Daniel Menzel[a]
Arbeitsgruppe Technische Akustik, MMK, Technische Universität München, Germany

ABSTRACT

To determine if measurements of the level of speech necessary for most comfortable loudness are dependent on illumination, two experiments were performed under normal illumination and in total darkness: First, subjects were instructed to adjust the level of different speech sounds until a comfortable loudness was achieved. Then, following a two alternative forced choice paradigm, subjects had to rate speech sounds, which were presented at different levels, either as "too soft" or "too loud". In both cases, results indicate that the level necessary for most comfortable loudness is lower in total darkness by about 1 dB.

1. INTRODUCTION

A sound which elicits most comfortable loudness is judged to be neither too soft nor too loud. However, as known from previous studies (Fastl [4], Menzel [7], Viollon [8]), judgments of loudness and ratings of pleasantness can be influenced by images presented during the experimental task. It is therefore conceivable that the process of determining most comfortable loudness might also be affected by varying visual stimulation such as different illuminations of the surrounding room. For example, one might adjust the level of a radio in a living room for pleasurable listening, only to find the music too loud once the lights are turned down (as suggested by Haverkamp [6], p. 212).

In the present study, two experiments where performed to test for the occurrence of such influences of illumination: The level of speech necessary to produce most comfortable loudness was determined in total darkness and under normal lighting conditions first by loudness adjustment and then via a two alternative forced choice (2 AFC) procedure.

2. ADJUSTMENT OF MOST COMFORTABLE LOUDNESS

A. Participants and instructions

Sixteen normal hearing subjects (1f, 15m, 23 to 29 years, median 24 years) took part in this experiment. They had the task to adjust speech sounds to a comfortable loudness.

[a] Email address: menzel@mmk.ei.tum.de

The instructions stated that the speech should neither sound too soft nor too loud. Also, the subjects were asked not to close their eyes for longer periods of time. Each participant performed the adjustments in total darkness and under normal illumination.

B. Apparatus and procedure

The experiment took place inside a sound-proofed booth. The two visual conditions were realized by totally darkening the booth on the one hand and by using a conventional desk lamp with incandescent lighting on the other hand. The lamp was positioned near the rear wall of the booth to reduce glare for the subjects. In this configuration, an illumination of about 100 lux was measured in the frontal horizontal direction using a GretagMacbeth i1 spectral photometer.

Unprocessed recordings of German male (DEm) and female (DEf) as well as English male (ENm) and female (ENf) speakers from the EBU sound quality assessment material CD [2] were used as acoustic stimuli. The German speakers were recorded under anechoic conditions, the English speakers under near anechoic conditions (vocal booth). Short sequences with durations between 4 and 5.5 seconds were extracted from these recordings. The level-time-functions of the signals can be seen in Figure 1.

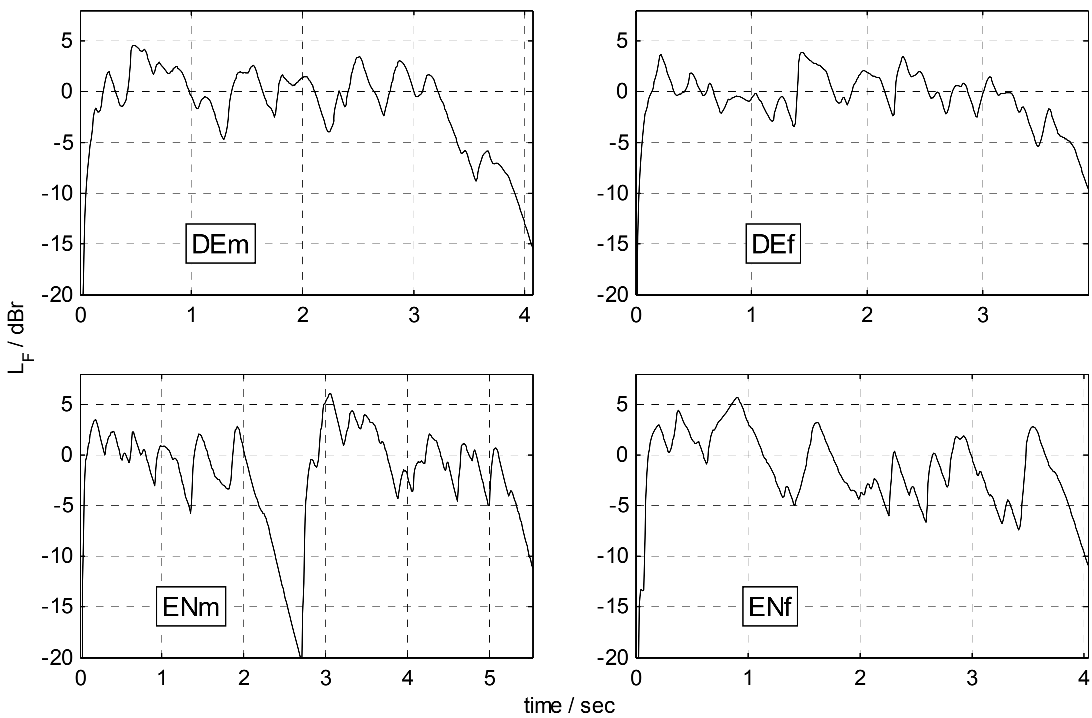

Figure 1: Level-time-functions of the four speech signals used in the experiments. Depicted is the unweighted level measured with time constant “fast” and normalised to the RMS level of the entire signal.

They were presented diotically via calibrated electrodynamic headphones (Beyer DT48) with free-field equalisation as described by Fastl and Zwicker ([5], p. 7). The initial levels of the speech sounds were chosen pseudo-randomly to be either low or high (for this paper, level is understood to mean RMS sound pressure level measured over the whole duration of the signal). A sound with low initial level started between 40 and 45 dB, a sound with high initial level started between 75 and 80 dB. These ranges of initial level were selected with the intention of producing speech which sounds either too soft or too loud (as indicated by Fastl [3]), so that a subsequent adjustment of loudness would always be necessary.

The level of the speech sounds could be varied by the subjects by turning the scroll-wheel of a computer mouse. Each rotational step of the wheel corresponded to a change in sound level of ±0.2 dB, depending on the direction of rotation. Clicking one of the mouse buttons enabled the subjects to listen to the stimulus at the currently adjusted level. This sequence of adjusting the level and listening to the stimulus was repeated by the participants until the goal of comfortable loudness was reached, which had to be indicated by pressing a button on a keyboard. Subjects had the opportunity to familiarise with this method during a short training sequence at the beginning of the experiment.

Each of the visual conditions constituted one session of the experiment. The visual condition for the first session was chosen randomly, so that half of the participants started in the "dark" environment, the other half in the "light" environment. In each session, all four speech sounds (DEm, DEf, ENm, ENf) were presented three times with each of the two initial levels (low or high), resulting in a total of 48 adjustments. All subjects were able to adjust repeated stimuli with an average error of less than ±3 dB.

C. Results

In Figure 2 (left) the adjusted RMS levels are shown as inter-individual medians of the intra-individual medians together with the corresponding interquartile ranges. The median values mostly are in the range of 55 to 65 dB, which corresponds to normal everyday speech levels (cf. hearing area in Fastl and Zwicker [5] p. 17), suggesting that for speech most comfortable loudness is related to realistic loudness. However, a large spread of the data can be seen with interquartile ranges covering more than 15 dB and an overall variability of more than 20 dB.

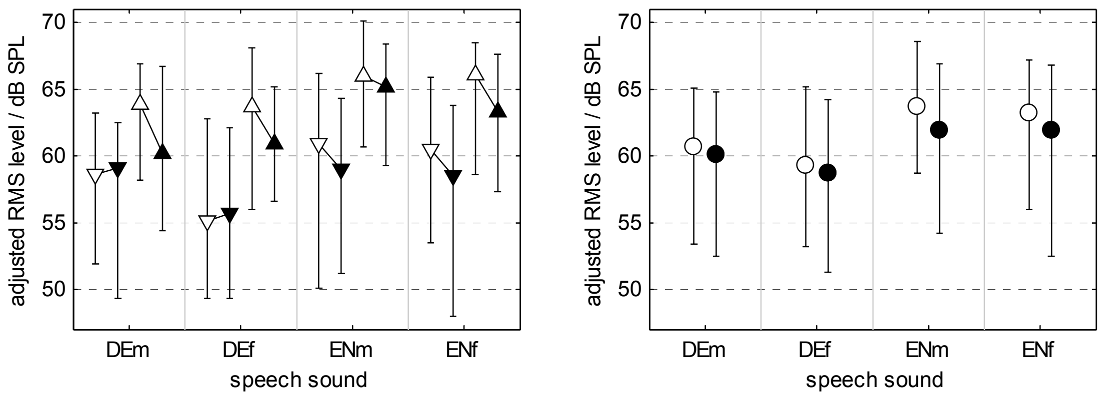

Figure 2: Medians and interquartile ranges of speech SPL adjusted for most comfortable loudness under varying illumination (white symbols: light condition, black symbols: dark condition). Left: Low and high initial levels are indicated by downward and upward pointing triangles respectively. Right: Data from low and high initial levels intra-individually pooled.

In the case of low initial levels (downward pointing triangles), the adjusted levels lie between 55 and 61 dB, while for high initial levels (upward pointing triangles) adjusted levels between 60 and 66 dB are observed, indicating a strong influence of initial level on the outcome of the adjustment itself. If a sound starts at a low level and subjects have to adjust its level upwards to reach the desired loudness, they tend to stop at a lower level than if the sound had initially started at a high level.

Levels adjusted in total darkness (black symbols) were for the most part lower than levels adjusted under normal illumination (white symbols) with median differences between -0.6 and 3.7 dB. For high initial levels, these differences were always positive, while for low initial levels, only the English speech signals yielded higher levels in the "light" condition.

Repeated measures analysis of variance shows highly significant main effects for visual condition [$F(1,15) = 10.36$; $p = 0.0057$], initial level [$F(1,15) = 33.92$; $p < 0.0001$], and speech signal [$F(3,45) = 27.09$; $p < 0.0001$] as well as a significant initial level by speech signal interaction [$F(3,45) = 3.34$; $p = 0.0274$].

Post-hoc comparisons according to Scheffé ($\alpha = 0.01$) suggest that significant differences exist between the speech signals DEm and ENm, DEf and ENm, and DEf and ENf. These differences could occur due to the differing original level of the speech signals, as differing speech levels generally also cause the timbre of speech to change. Therefore, subjects could be able to infer the original speech level by timbre and perform the adjustment accordingly.

The right half of Figure 2 shows medians and interquartile ranges calculated from intra-individually pooled data to remove the influence of initial level. Differences in median adjusted level between light and dark conditions of 0.6 to 1.8 dB can be seen, again with large variability but the same tendency for all speech signals.

To analyse the inter-individual distribution of the effect of the two visual conditions regardless of absolute SPL, overall intra-individual differences between median levels from pooled stimuli in the "light" condition (L_{light}) and in the "dark" condition (L_{dark}) were calculated. Figure 3 shows the histogram of the distribution of these differences.

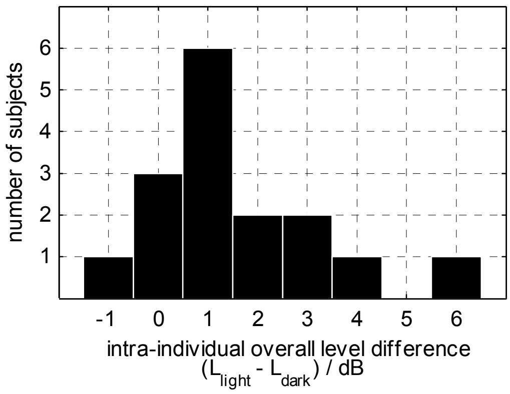

Figure 3: Distribution of intra-individual differences between median adjusted levels in light and dark conditions calculated from pooled data.

It can be seen that overall level differences between the two visual conditions span a range from -1 to 6 dB with a peak at 1 dB. One subject shows a negative difference ($L_{light} - L_{dark} < -0.5$ dB), three subjects show no difference ($|L_{light} - L_{dark}| < 0.5$ dB), and twelve subjects show a positive difference ($L_{light} - L_{dark} > 0.5$ dB).

For each speech sound at each adjusted level additionally the percentile loudness N_5 was calculated using the dynamic loudness model after Chalupper and Fastl [1]. As current loudness models do not take into account illumination changes, data from both visual conditions were pooled for this analysis. This is shown in Figure 4, again with low and high initial levels depicted as downward and upward pointing triangles. It can be seen that, as expected, effects of initial level are still visible. But differences in calculated loudness between the speech sounds (e.g. between English and German male speakers) are smaller than those expressed by RMS level (Figure 2, left), suggesting that N_5 is a better predictor for loudness of speech than RMS level (see also [3], [5]).

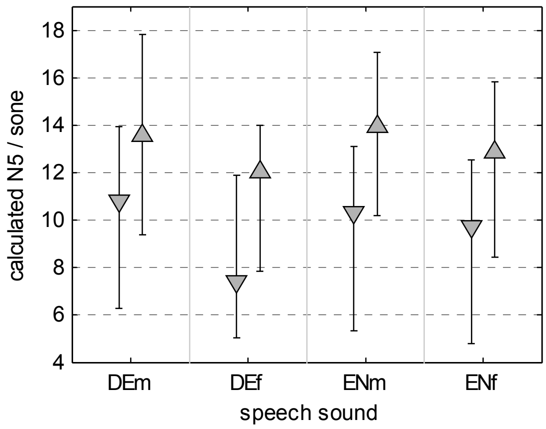

Figure 4: Medians and interquartile ranges of percentile loudness N_5 calculated from speech sounds adjusted to most comfortable loudness. Low and high initial levels are indicated by downward and upward pointing triangles respectively.

3. TWO ALTERNATIVE FORCED CHOICE

A. Participants and instructions

Fifteen normal hearing subjects (1f, 14m, 23 to 29 years, median 25 years) took part in this experiment. In a two alternative forced choice task they had to decide if speech sounds presented to them where either “too soft” or “too loud”. The subjects were again asked not to close their eyes for longer periods of time. Each participant performed the task in total darkness and under normal illumination.

B. Apparatus and procedure

The experimental setup was the same as in the previous experiment. Subjects could enter their decision if a sound is “too soft” or “too loud” by pressing one of two keys on a computer keyboard. The four speech sounds where presented with effective SPLs ranging from 50 to 70 dB in 1 dB steps in random order.

Each of the visual conditions constituted one session of the experiment. The visual condition for the first session was chosen randomly. In each session, all four speech sounds (DEm, DEf, ENm, ENf) were presented three times with all 21 levels for a total of 252 answers.

C. Results

Plotting the percentage of the answer “too loud” for all speech signals at all levels results in the graphs shown in Figure 5.

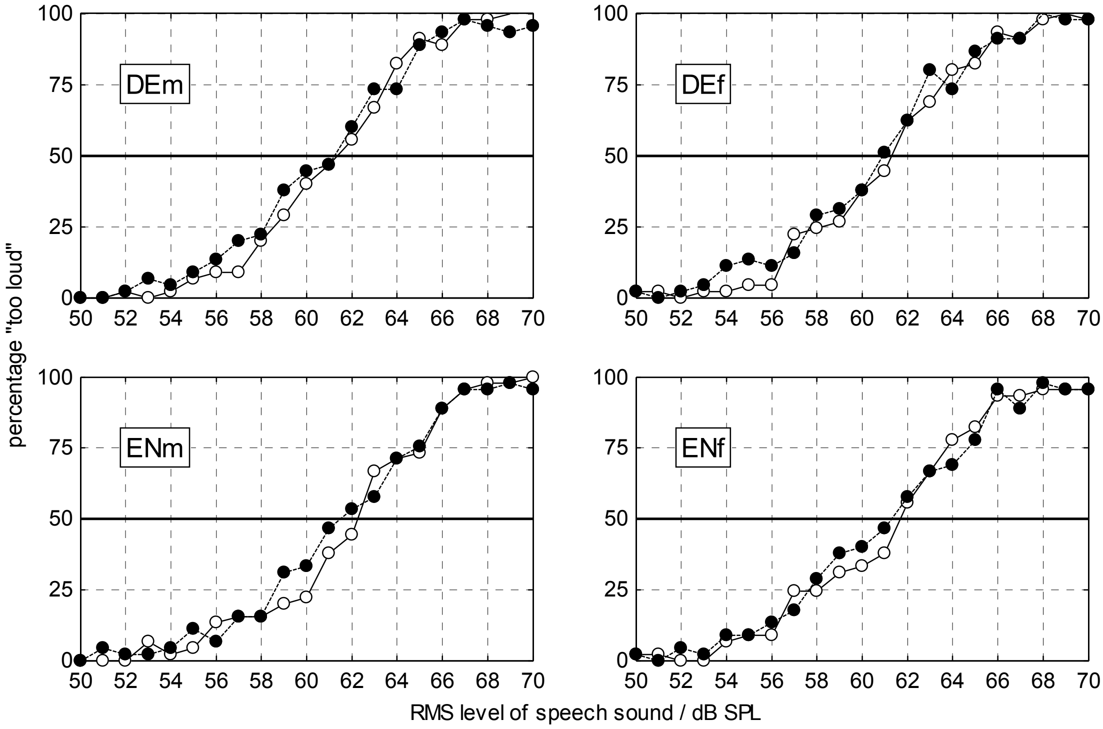

Figure 5: Percentage of answer "too loud" in 2 AFC task with four speech sounds presented at different SPLs. Measurements performed in total darkness are given by black symbols, measurements in normal illumination are shown with white symbols.

At levels up to 53 dB almost all subjects judge the speech signals as being too soft, while for levels over 68 dB the sounds are mostly judged "too loud". In all cases the 50% points lie between 61 and 62 dB. The differences between speech sounds as seen in Figure 2 (left) are not evident in these measurements. For levels below the 50% points, a small tendency can be seen to judge sounds in the "dark" condition as "too loud" more often than in the "light" condition.

All answers of the subjects where then pooled at each level for each visual condition. This is shown in Figure 6. It can be seen that the 50% points lie at $L_{50\%,dark}$ = 61.2 dB and $L_{50\%,light}$ = 61.7 dB, which is in good accordance with the data presented in Figure 2. Thus, in the dark visual condition, speech sounds seem to be rated "too loud" more often at the same SPL compared to the bright visual condition, which is in line with the previous experiment.

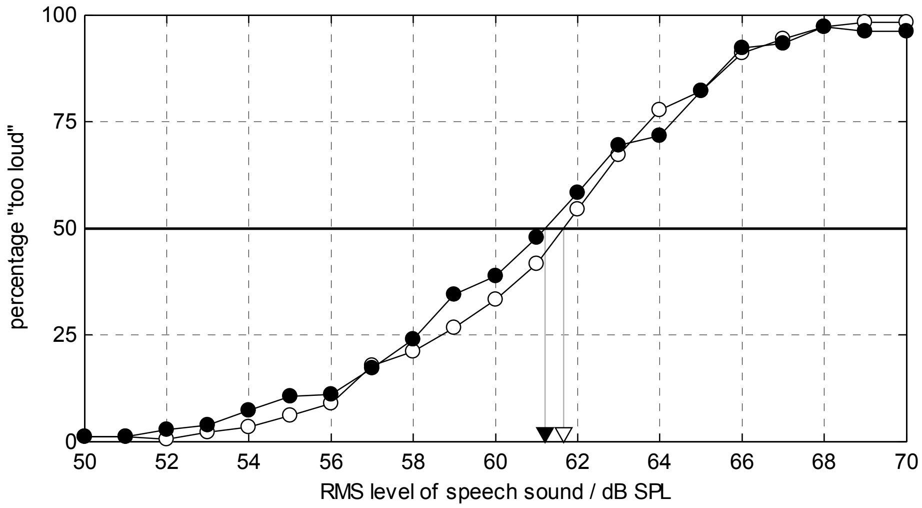

Figure 6: Percentage of answer “too loud” in 2 AFC task pooled for all four speech sounds. Measurements performed in total darkness are given by black symbols, measurements in normal illumination are shown with white symbols. The arrows indicate the respective 50% points at 61.2 dB (dark) and 61.7 dB (light).

4. DISCUSSION

Small effects of the illumination of the visual surrounding on the level of speech necessary to produce most comfortable loudness were found. In total darkness, adjusted levels were lower than in a light environment. As these effects are on average in the order of 1 dB, they are above the threshold of just noticeable level differences ([5], p. 180) and thus could explain, at least partially, situations as described by Haverkamp [6], in particular regarding the fact that some subjects showed much larger influences of up to 6dB.

However, it is not yet known if these changes in adjusted level are caused by a changing loudness perception dependent on illumination (while the subjective reference point of most comfortable loudness stays constant), by a shift in the subjective reference point of most comfortable loudness (while the loudness perception stays constant), or by a combination of both mechanisms.

In the first case, lower adjusted levels would signify that the speech sounds produced a higher loudness in total darkness. Possible explanations of this phenomenon could include the degree to which subjects were concentrated on the acoustic stimuli. In case of total darkness there was no external visual stimulation so that the importance of auditory sensations might increase. In turn, subjects may be inclined to assign a higher loudness to the speech signals when presented in darkness, resulting in a lower adjusted level to reach the same loudness as under normal illumination.

In the second case, lower adjusted levels would signify a lower subjective reference point of most comfortable loudness. This could be caused by the expectations a subject has in a certain environment. Darkness might be associated with night-time and in turn with expectations of quietness and less interfering noise, while during day-time (= light environment) more noise could be expected. Accordingly, the point of most comfortable loudness would have to be adapted to stay relatively constant with regard to expected interfering external noise.

5. CONCLUSIONS

Experiments were performed to evaluate the influence of illumination on measurements of the level of speech producing most comfortable loudness. It was found that speech signals were adjusted about 1 dB lower to reach most comfortable loudness in total darkness than in a light surrounding, with some subjects showing differences of up to 6 dB. Also, speech sounds were judged as being “too loud” more often in a dark surrounding than in normal illumination at the same SPL of the sound.

ACKNOWLEDGMENTS

The author wishes to thank Prof. Fastl for valuable suggestions on methodology and experimental procedure as well as helpful comments on earlier versions of the manuscript. This work is supported by Deutsche Forschungsgemeinschaft by grant FA140/5.

REFERENCES

[1] Chalupper, J., Fastl, H.: Dynamic loudness model (DLM) for normal and hearing-impaired listeners. Acta Acustica united with Acustica **88**, 378-386, 2002

[2] EBU: Tech 3253 - Sound Quality Assessment Material recordings for subjective tests. European Broadcasting Union, Geneva, 2008

[3] Fastl H.: Schallpegel und Lautstärke von Sprache. Acustica **35**, 341-345, 1976

[4] Fastl H.: Audio-visual interactions in loudness evaluation. In: Proceedings of the 18th International Congress on Acoustics, Kyoto, Japan, 1161-1166, 2004

[5] Fastl H., Zwicker E. *Psychoacoustics. Facts and Models*, 3rd ed. (Springer, Berlin, 2007)

[6] Haverkamp M.: *Synästhetisches Design - Kreative Produktentwicklung für alle Sinne*. (Hanser, München, Wien, 2009)

[7] Menzel, D., Fastl, H., Graf, R., Hellbrück, J.: Influence of vehicle color on loudness judgments. J. Acoust. Soc. Am. **123**, 2477–2479, 2008

[8] Viollon, S., Lavandier, C.: A Sound Judgment Depending On The Urban Visual Setting?. J. Acoust. Soc. Am. **105**, 981, 1999