ISSN 0280-6819

KONGL CAROLINSKA MEDICO CHIRURGISKA INSTITUTET

TECHNICAL AUDIOLOGY

Report TA No. 115
January 1987

LOUDSPEAKER FREQUENCY RESPONSE AND PERCEIVED SOUND QUALITY: COMPARISON BETWEEN MEASUREMENTS IN LISTENING ROOM, ANECHOIC ROOM AND REVERBERATION ROOM

Alf Gabrielsson, Björn Lindström and Ove Till

Karolinska Institutet
Dept. of Technical Audiology
KTH
S-100 44 Stockholm
Sweden
Tel: +46-8-11 66 60

Report TA115
January 1987

LOUDSPEAKER FREQUENCY RESPONSE AND PERCEIVED SOUND QUALITY: COMPARISON BETWEEN MEASUREMENTS IN LISTENING ROOM, ANECHOIC ROOM AND REVERBERATION ROOM

Alf Gabrielsson, Björn Lindström and Ove Till

ISSN 0280-6819
Report TA115
January 1987

Karolinska Institutet
Dept. of Technical Audiology
KTH
S-100 44 Stockholm
Sweden
Tel: +46-8-11 66 60

LOUDSPEAKER FREQUENCY RESPONSE AND PERCEIVED SOUND QUALITY: COMPARISON BETWEEN MEASUREMENTS IN LISTENING ROOM, ANECHOIC ROOM AND REVERBERATION ROOM

Alf Gabrielsson, Björn Lindström and Ove Till

ABSTRACT

In connection with an earlier report (Gabrielsson, Lindström & Till, 1986) a comparison is made between three different methods for measurement of loudspeaker frequency response with regard to how well the relationships between the frequency response and the perceived sound quality agree with stated hypotheses. The three methods were measurements in the listening room, in an anechoic room, and in a reverberation room. The results obtained by using the measurements in the listening room were in better agreement with the hypotheses than the results from the other two methods.

This work was supported by the Bank of Sweden Tercentenary Foundation and the Swedish Ministry of Health and Social Affairs, Commission for Social Research.

1 INTRODUCTION

In a previous report (Gabrielsson, Lindström & Till, 1986) we investigated the relationship between the perceived sound quality of 18 high fidelity loudspeakers and the frequency response of these loudspeakers in the listening room that was used in the original listening test (Gabrielsson & Lindström, 1985). Hypotheses concerning the effects of different frequency responses on six perceptual variables (clarity, fullness, spaciousness, brightness vs. dullness, softness vs. sharpness, and absence of extraneous sounds) and on overall evaluation (fidelity) were formulated on the basis of results from earlier listening tests. Practically all hypotheses were supported by the results in this investigation.

From the standpoint of perception psychology it is advantageous to be able to describe the stimulus as accurately as possible. It thus seems natural to study the frequency response in the actual listening room. On the other hand this is presently not a common procedure, and it also requires much more work (despite the fact that certain simplifying assumptions were made) than other, more or less standardized methods, such as measuring the frequency response in an anechoic room or in a reverberation room. It is well known that the predictive power of the last-mentioned alternatives is limited, that is, they do not allow very accurate predictions of how the sound quality of the loudspeaker will be perceived in a "normal" listening room. It is therefore of great interest to compare the results of measurements by these different methods (in listening room, anechoic room, and reverberation room) and, especially, to see how well they can be related to judgments of the perceived sound quality. This is then the purpose of the investigation reported here.

2 METHODS

The methods for measurement of the frequency response in the listening room were described in detail in the previous report (Gabrielsson, Lindström & Till, 1986). Briefly white noise was fed through each of the 18 loudspeakers and recorded on one channel of a tape recorder by an omnidirectional microphone in the listening position. The noise was also directly recorded on the other channel to provide a reference. After A/D conversion the frequency response of the loudspeakers was calculated from the crosscorrelation function between the recorded loudspeaker noise and the reference noise. The frequency response was estimated from an impulse response truncated to 40 msec. After Hanning windowing the frequency response was smoothed by using a moving average over the entire frequency range with a bandwidth of one octave. The smoothed frequency responses for all 18 loudspeakers are shown in Figure 1 (which is the same as Figure 8 in our previous report).

In the anechoic room the measurements of the frequency response were made on axis at a distance of one meter from the loudspeaker. A computer-generated periodic, bandlimited signal was applied to the speaker. The measured acoustic response was FFT transformed and compared with the computer-generated spectrum to calculate the transfer function. The frequency response is then calculated as the magnitude of the transfer function.

The measurements in the reverberation room were made by the Swedish National Testing Institute. In the reverberation room (7.40 x 4.40 x 6.13 m) there are five standardized loudspeaker positions, one of which can be chosen by the customer. The test signal is white noise fed through a 30 Hz wide bandpass filter. The acoustic power is registrated by a graphic level recorder.

3 RESULTS

3.1 Hypotheses

Hypotheses concerning the relationships between the frequency response and the perceptual dimensions were stated in the previous report. They were based upon results from many earlier listening tests summarized in Gabrielsson & Sjögren (1979) and from a recent investigation (Gabrielsson, Schenkman & Hagerman, 1985). The hypotheses are repeated here as follows:

Clarity is favored by a broad frequency range and by a certain (not too large) rise in the response at midfrequencies (about 250-1000 Hz) and the lower high frequencies (roughly 1-4 kHz). The more the emphasis (center of gravity) of the response is shifted toward lower frequencies, the worse clarity.

Fullness is likewise favored by a broad frequency range but with relatively more emphasis on lower frequencies.

Spaciousness seems to be favored by the same conditions as clarity. Narrow frequency range and/or marked resonances are negative factors for spaciousness.

Brightness vs. dullness: Brightness increases when the emphasis of the response moves toward higher frequencies. Conversely, dullness increases when the emphasis of the response moves toward lower frequencies.

Softness (gentleness) vs.sharpness: Softness is favored by a certain emphasis on lower frequencies, while sharpness is associated with more or less steeply rising responses toward higher frequencies or marked resonance peaks at higher frequencies.

Absence of extraneous sounds (e.g., hiss and the like) is related to a (relatively) decreased response at very high frequencies (roughly above 5 kHz).

The hypotheses deal solely with the effects of the frequency response, all other things being equal.

The perceptual data used for evaluating the hypotheses were given in Gabrielsson & Lindström (1985).

With regard to the fidelity scale the hypothesis is that fidelity is favored by a frequency response similar to that given above for clarity and spaciousness but somehow modified to reflect the influence of the other perceptual dimensions. For the rationale behind this and for other comments concerning the hypotheses, see the previous report.

3.2 General comparison between the three types of frequency responses

The frequency responses of the 18 loudspeakers measured in the listening room, the anechoic and the reverberation rooms appear in Figures 1-3.

Inspection of Figures 1-3 shows that the frequency response in the anechoic room (henceforth AR) is much more rugged than those from the listening room (LR) and the reverberation room (RR). Of course, this is a consequence of the fact that the AR response is unsmoothed, while smoothing is made both for the LR (octave smoothing) and RR response (bandpass filter 30 Hz). Examples of unsmoothed LR response were given in the previous report.

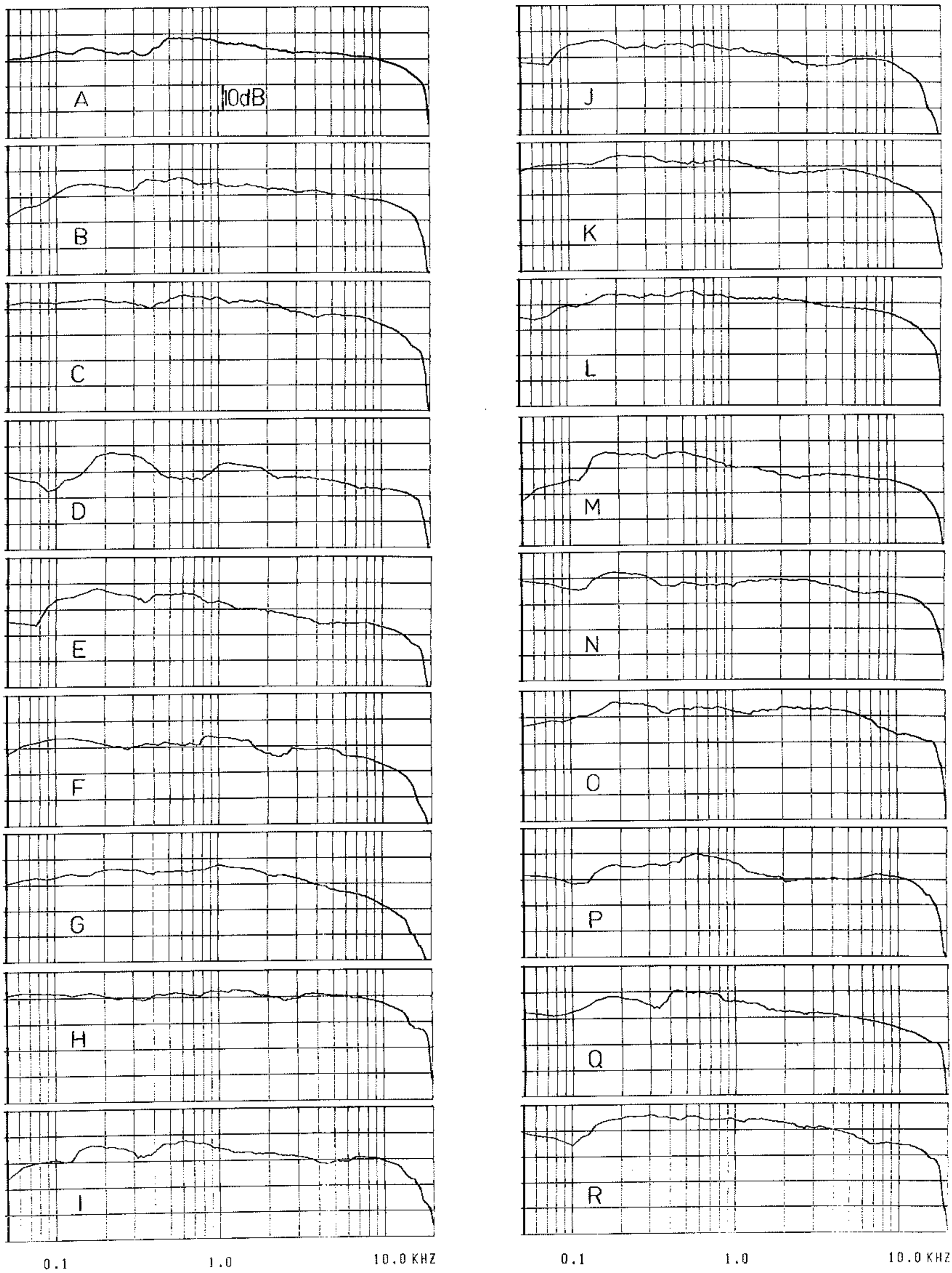

Figure 1. Frequency responses measured in listening room.

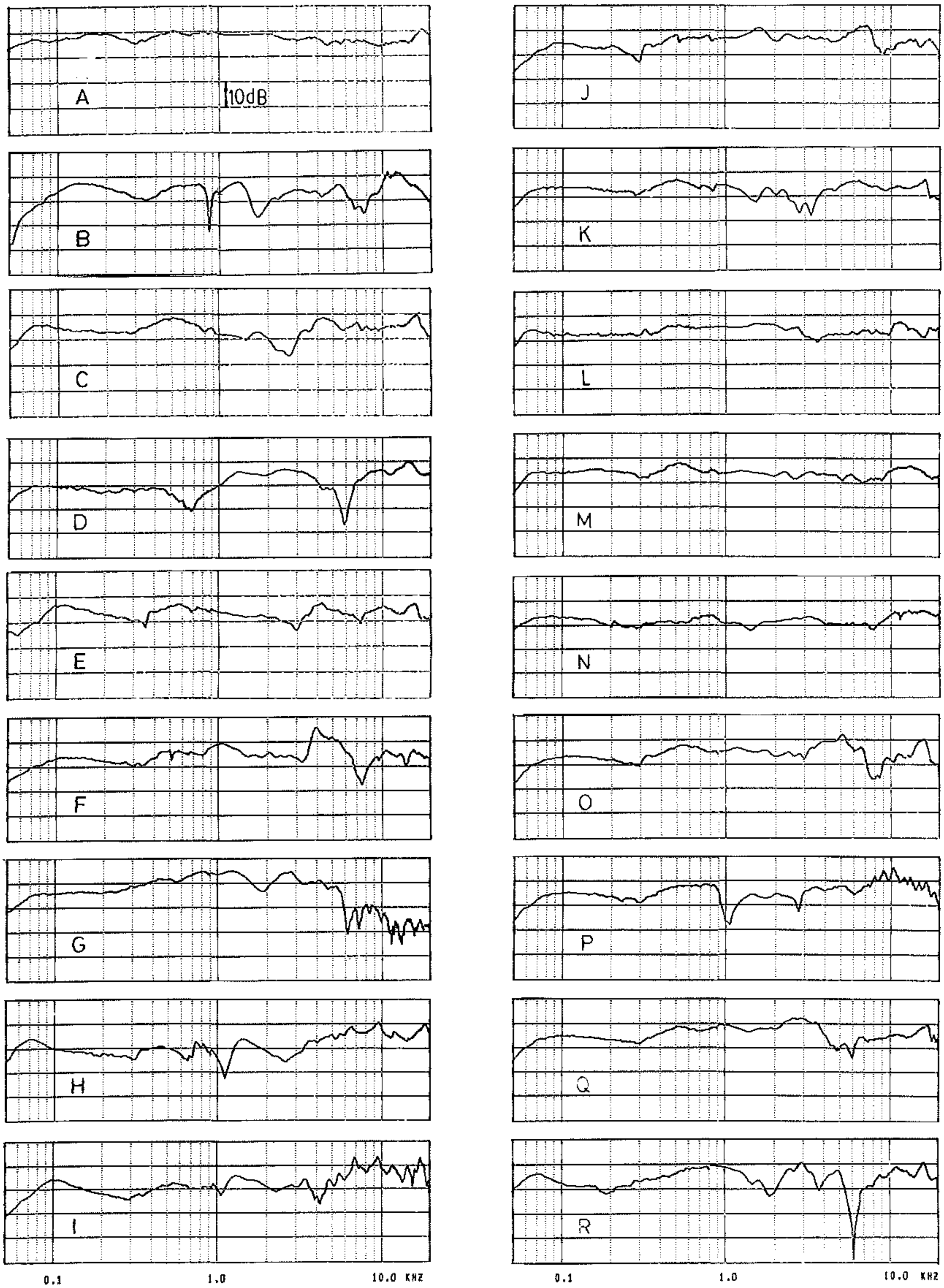

Figure 2. Frequency responses measured in anechoic room.

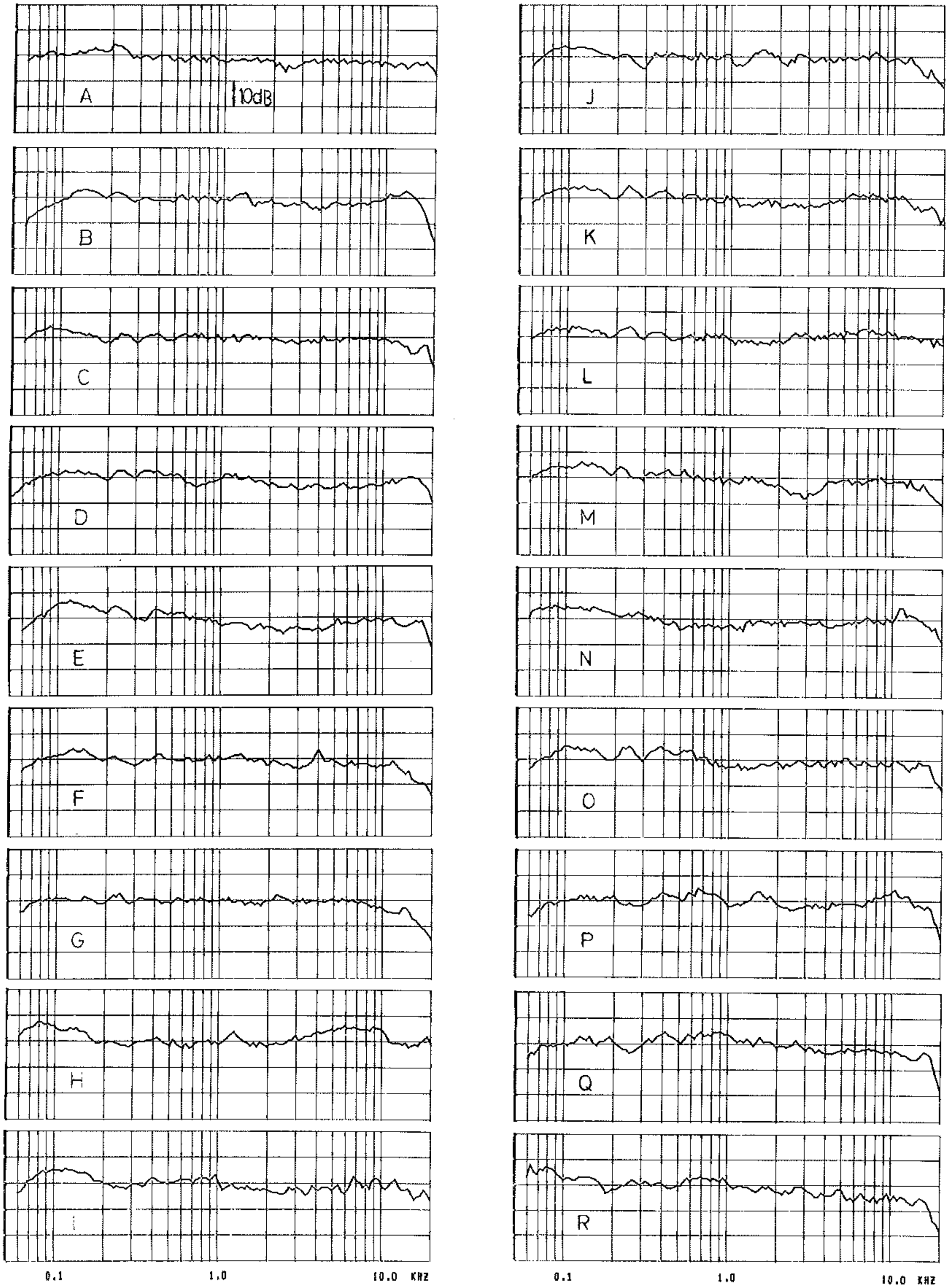

Figure 3. Frequency responses measured in reverberation room.

With regard to the overall contour of the response, most LR responses are more or less "symmetrical" in the sense that they show an elevated response (a "plateau") somewhere in the region 100-2000 Hz, and the response falls off, more or less continuously, on both sides of this maximum, especially toward the higher frequencies. In most of the RR responses there is also a downward slope toward higher frequencies but not as consistently as in the LR responses; furthermore, the RR responses are not as symmetrical as the LR responses. In the AR responses there are more examples of responses rising toward higher frequencies, even though they are sometimes interrupted by dips.

Among the AR responses those for loudspeakers A, L, M, and N are considerably smoother than the others. It is interesting to note that three of these loudspeakers, A, L, and N, were considered as the best in the listening test (Gabrielsson & Lindström, 1985). On the other hand loudspeaker M was judged as one of the worst loudspeakers in the test, so there is no simple relationship to be found here. Looking at the LR and RR responses for these four loudspeakers it is hardly possible to find any characteristics, which set them apart from the other loudspeakers as clearly as with the AR responses.

For the majority of the loudspeakers the LR and RR responses are more similar to one another than what the LR and AR, or AR and RR, responses are to each other (and this is not only a function of the smoothing in the LR and RR responses). Still there is hardly any loudspeaker for which the overall contours of the LR and the RR response are very similar to each other, but the similarity extends only to certain limited frequency regions.

The LR and AR responses are usually very dissimilar. Among the four above-mentioned loudspeakers with relatively smooth AR response it may be noted that loudspeaker A has a fairly close resemblance between the LR and AR responses, except at very high frequencies. For the other three loudspeakers (L, M, and N), however, there are obvious dissimilarities between the LR and AR responses.

On the whole the three types of responses are thus fairly or very dissimilar among themselves for the same loudspeaker. This also means that they will differ with regard to how well they are in agreement with the hypotheses concerning the relationships to the perceptual variables.

3.3 Relationship to subjective judgments

In order to study how well the relationships between each type of frequency response and the various perceptual dimensions agree with the hypotheses stated in 3.1, we use the same procedure as in the previous report, that is, to pick out the loudspeakers which were rated as the most different in each perceptual scale and look at their frequency responses. Table 1 shows the loudspeakers which were rated highest and lowest, on average over music programs Nos. 1-4, in each scale and listening session. For instance, in clarity the best loudspeaker in the first session was A, the worst was E. If there was no statistically significant difference between the respective loudspeakers, this is indicated by a dash. (Table 1 is the same as Table 1 in the previous report.)

	Clarity	Fullness	Spaciousness	Brightness
Session 1	A - E	D - B	A - E	A - E
2	A - F	-	A - F	A - F
3	A - M	L - A	N - M	A - M
4	A - Q	R - P	A - O	A - Q

	Softness	Abs. extr. sounds	Fidelity
Session 1	E - C	-	A - B
2	I - H	G - H	A - F
3	-	-	L - M
4	R - A	Q - A	A - O

Table 1. The loudspeaker rated highest (left) and lowest (right), on average over programs Nos. 1-4, in each scale and listening session.

This type of inspection was already made regarding the LR response in the preceding report. However, the results of this will be briefly repeated here for each scale to allow direct comparisons with the results regarding the AR and RR responses.

3.3.1 Clarity

In clarity loudspeaker A was considered best in all sessions in contrast to loudspeakers E, F, M, and Q, respectively

(Table 1). In the LR response (Figure 1) loudspeaker A has a relatively flat response with a maximum at 500-1000 Hz, a midhigh frequency region. In comparison with that loudspeakers E, F, M, and Q have maxima (or center of gravity) more toward lower frequencies, and their response toward higher frequencies decreases more rapidly than for loudspeaker A. These results are in line with the hypothesis concerning clarity. In the AR response loudspeaker A has again a relatively smooth response, which is fairly similar to its LR response except at very high frequencies (Figure 2). Loudspeakers E, F, and Q have more rugged responses. In comparison with A loudspeaker E has its emphasis toward lower frequencies, loudspeaker M at about the same frequency region, and loudspeakers F and Q rather toward higher frequencies. The results regarding the AR response are only partly in agreement with the hypothesis concerning clarity. In the RR response (Figure 3) loudspeaker A has a maximum around 200 Hz, which does not agree with the hypothesis. However, loudspeakers E, F, and M have their maxima at still somewhat lower frequencies (100-150 Hz) and fall off more rapidly toward higher frequencies than A (except F). On the other hand loudspeaker Q, whose RR response is fairly similar to its LR response as noted earlier, would perhaps be expected to be best in clarity with its maximum at 400-1000 Hz - but is thus not. The results concerning the RR response are then only in partial agreement with the hypothesis.

3.3.2 Fullness

With regard to fullness the LR response of the highest rated loudspeakers (D, L, and R) have more emphasis, to a varying extent, on lower frequencies than their respective opposites (loudspeakers B, A, and P, see Table 1), which is in accordance with the hypothesis concerning fullness. It is doubtful whether the same can be said regarding the AR responses. Although loudspeakers D and L extend to somewhat lower frequencies (below 100 Hz) than loudspeakers B and A, respectively, they at the same time have maxima at higher frequencies than in B and A. Loudspeaker R might be said to have more emphasis on lower frequencies than loudspeaker P. The results from the RR responses seem to agree with the hypothesis: loudspeakers D, L and R have relatively more emphasis on lower frequencies than their opposites (B, A, and P, respectively).

3.3.3 Spaciousness

Concerning spaciousness the LR response of loudspeaker A has relatively more emphasis on midhigh frequencies than loudspeakers E, F, and O, and loudspeaker N has a fairly flat response 150-3000 Hz in contrast to loudspeaker M with its marked boost at 150-700 Hz. The results are in accordance with the hypothesis concerning spaciousness. With regard to the AR response the comparison between loudspeakers A and E may perhaps be in line with the hypothesis but not the comparisons between A and F and between A and O, since loudspeakers F and O have more emphasis on higher frequencies than A (however, the sharp resonance peak in F may be considered as an argument in favor of the hypothesis). The comparison between loudspeakers N and M can hardly be said to support the hypothesis. With respect to the RR responses the comparisons between loudspeakers A and E and between A and O may be said to support the hypothesis in the sense that E and O have more emphasis on lower frequencies than A, but on the other hand loudspeaker A's RR response itself with its maximum around 200 Hz goes against the hypothesis. Nor can the comparisons between loudspeakers A and F and between N and M be said to support the hypothesis.

3.3.4 Brightness

In brightness loudspeaker A was throughout considered as the brightest, while loudspeakers E, F, M, and Q were rated as the dullest. This is in fact the same result as for the clarity scale (see Table 1). As said already under Clarity, the LR responses of loudspeakers E, F, M, and Q have their center of gravity toward lower frequencies than loudspeaker A, which is in accordance with the hypothesis concerning brightness vs. dullness. Inspection of the AR responses reveals that the comparison between loudspeakers A and E may support the hypothesis, while this is clearly not the case regarding the comparisons between A and F or between A and Q, and hardly for the comparison between A and M either. With regard to the RR responses already the maximum around 200 Hz for loudspeaker A is against the hypothesis. Disregarding this, the comparisons between A and E, A and F, and A and M mainly support the hypothesis (especially the A-E comparison), while loudspeaker Q may rather be expected to sound brighter than A.

3.3.5 Softness

In softness (gentleness) the highest rated loudspeakers were E, I, and R, in contrast to C, H, and A, respectively. The LR responses for E, I, and R have relatively more emphasis on lower frequencies (with some reservation for very low frequencies, below 100 Hz) and also decrease more toward higher frequencies than the responses for loudspeakers C, H, and A, respectively. These results are in agreement with the hypothesis concerning softness. With regard to the AR responses the comparison between loudspeakers E and C may perhaps be said to support the hypothesis, but hardly the comparison between I and H (in fact both of them would sound rather sharp considering the peaks at high frequencies, not the least in I), nor the comparison between R and A (this comparison is also difficult because of the very rugged response in R). With the RR responses, however, it seems that the hypothesis is supported in all three comparisons (perhaps with some hesitation regarding the comparison between loudspeakers I and H).

3.3.6 Absence of extraneous sounds

With respect to absence of extraneous sounds the LR responses of loudspeakers G and Q show a relatively lower response at very high frequencies, above 5 kHz, than loudspeakers H and A, as predicted by the hypothesis. The hypothesis is also supported by the results both from the AR responses (possibly with some reservation for the comparison Q-A) and the RR responses.

3.3.7 Fidelity

Regarding fidelity loudspeaker A was considered best in three sessions and loudspeaker L in one in contrast to loudspeakers B, F, M, and O. The hypothesis concerning fidelity is somewhat loose, since fidelity is considered to be a weighted function of all perceptual variables, but it should especially reflect the properties stated in the hypotheses concerning clarity and spaciousness. The LR responses for the worse loudspeakers B, F, and O have their center of gravity toward lower frequencies than loudspeaker A; F also decreases more rapidly at higher frequencies. Likewise loudspeaker M has much more emphasis on lower frequencies than loudspeaker L. These results are then in general agreement with the hypothesis. The rugged character of the AR responses for loudspeakers B, F, and O, in comparison with loudspeaker A, may be said to support the hypothesis. On the other hand they cannot be said to have their center of gravity toward lower frequencies than in loudspeaker A, rather the reverse. The responses for loudspeakers L and M are both relatively smooth, but M may have somewhat more emphasis on lower frequencies than L. With regard to the RR responses already the peak around 200 Hz in loudspeaker A is against the hypothesis, as noted earlier. Loudspeaker O has more emphasis on lower frequencies than A, and this also applies for loudspeaker M in comparison with L. Although these results are in line with the hypothesis, the interpretation of the comparisons between loudspeakers A and B and between A and F is more doubtful.

4 DISCUSSION

It is obvious that the relatioships between the LR responses and the perceptual variables are in much better agreement with the hypotheses in 3.1 than the corresponding relationships using the AR or RR responses. Among the last-mentioned alternatives the results for the RR responses are somewhat more in accordance with the hypotheses than the results for the AR responses.

That the LR responses are superior to the others should be no direct surprise. Although the measurement of the LR response was simplified in different ways (cf. our previous report), it still comes much closer to the actual stimulus situation for the listener than when the frequency response is measured in an anechoic room or in a reverberation room. However, the latter procedures are more convenient for many reasons. It would thus be preferable to be able to state the relationships between the AR response and the perceptual variables, as well as between the RR response and the perceptual variables - or between the AR and RR responses to the LR response, since the relationships between the LR response and the perceptual variables seem relatively well established. However, considering the results obtained in this investigation, this is certainly no simple task and requires much more research, before it can be handled with any reasonable success.

It must also be emphasized that the hypotheses in 3.1 were explicitly stated as dealing only with the frequency response, all other things being equal. However, in this case we actually examine the hypotheses under circumstanses, in which the last condition - all other things being equal - is not fulfilled. The eighteen loudspeakers do not differ only with regard to the frequency response but in other factors as well, and it is thus possible that the listeners' ratings are influenced not only by the differences in frequency response but also by other factors. Definite conclusions concerning these questions have to await the results from further investigations.

5 ACKNOWLEDGEMENTS

The authors want to express their gratitude to Gerhard Elger for much assistance in the physical measurements, and to Björn Hagerman and Åke Olofsson for valuable discussions.

6 REFERENCES

Gabrielsson, A. & Lindström, B (1985). Perceived sound quality of high-fidelity loudspeakers. Journal of the Audio Engineering Society, 33, 33-53.

Gabrielsson, A., Lindström, B. & Till, O. (1986). Loudspeaker frequency response and perceived sound quality: Measurements in listening room. Technical Audiology Reports No. 114. Stockholm: Karolinska Institute.

Gabrielsson, A., Schenkman, B. & Hagerman, B. (1985). The effects of different frequency responses on sound quality judgments and speech intelligibility. Technical Audiology Reports No. 112. Stockholm Karolinska Institute. (Revised version in press in Journal of Speech and Hearing Research)

Gabrielsson, A. & Sjögren, H. (1979). Perceived sound quality of sound-reproducing systems. Journal of the Acoustical Society of America, 65, 1019-1033.