Technical Audiology
Karolinska institutet
KTH
S-100 44 STOCKHOLM

Report TA No. 87
November 1978

DETECTION OF NONLINEAR DISTORTION IN TELEPHONE SYSTEMS

Alf Gabrielsson
Ann-Cathrine Lindblad
Björn Lindström
Åke Olofsson

ABSTRACT

This investigation deals with the detection of non-linear distortion in telephone speech. Three Swedish sentences, read by a male and a female, were recorded and transmitted by different (simulated) telephone systems introducing certain amounts of quadratic or cubic amplitude distortion. One of the systems represented a local call between two modern telephone sets, while another system represented a long distance call with a restricted frequency range. The subjects (in all 24 subjects) listened to pair presentations of a certain sentence, in which one member of the pair was undistorted and the other member distorted, and judged which of the two presentations had the worst sound quality. Threshold values were computed corresponding to 75% and 60% correct detection. The threshold values were generally higher for quadratic distortion than for cubic distortion. As seen before the line and the receiver the threshold values were higher for the narrow-band system than for the broad-band system. At the outputs of the systems the situation was more varying. The threshold values differed between the three sentences, and there were also some complex interactions and marked inter-individual differences in detection ability. Some suggestions concerning maximum allowed distortion are given.

INTRODUCTION

For the design of high quality telephone systems it is necessary to consider the effects of nonlinear distortion on telephone speech. Nonlinear distortion can be expected to impair speech intelligibility as well as the perceived sound quality in general. To protect telephone listeners from these negative consequences, limits for the maximum allowed nonlinear distortion should be stated. Estimations of such critical limits could be made from data on threshold values for the detection of nonlinear distortion in telephone speech.

However, there are very few reports on this topic, for instance by Ahlborg (1967), referred to in CCITT Green Book. Some data for detection of amplitude distortion in speech for headphone listening are given in Gabrielsson et al. (1976).

The investigation reported in this paper was designed to deal directly with the detection of nonlinear distortion in telephone speech. Three Swedish sentences were recorded and transmitted by different (simulated) telephone systems introducing certain amounts of quadratic or cubic amplitude distortion. Each sentence was presented undistorted as well as distorted for immediate comparison by the subjects. Threshold values for detection of the distortion were computed for all stimulus conditions.

METHODS

Speech material

Three Swedish sentences were read by one male and one female and recorded in an anechoic chamber.Considered al-

together the sentences were phonetically balanced.
The sentences were:

1) Den vackra sommaren lider mot sitt slut
(= The beautiful summer is coming to an end)

2) Den tomma flaskan stod överst på en bred hylla
(= The empty bottle stood at the top of a broad shelf)

3) Räkningen skulle betalas var fjärde månad
(= The bill was to be paid every fourth month)

The density functions of the levels for the sentences, read by the male speaker are shown in Figure 1. Note in particular the differences at the highest levels.

Simulated systems

Two different systems were investigated, called system A and system B.

Each system consisted of two linear filters separated by a nonlinearity. The nonlinearities were of the forms $x + ax^2$ (quadratic distortion, X^2 distortion) and $x + ax^3$ (cubic distortion, X^3 distortion), $a>0$. The first filter simulated the transmitting function of a telephone and the second filter simulated a telephone exchange and the receiving parts of a telephone including the earphone. (In a real telephone system this would mean that the distortion is created in the microphone and the microphone amplifier.) The transfer function of the second filter also included the inverse of the transfer function of the receiver earphone which was actually used in the listening tests. The frequency curve of this earphone is shown in Figure 2. The frequency characteristics of system A and B are shown in Figure 3 and 4, respectively. The system A transmitting filter corresponded to a microphone curve as specified by the Swedish Telecommunications Administration for telephone sets with "linear" microphones.

In system B the transmitting filter represented a frequency response which was set by the U.K. Post Office in 1947 as the design objective for best performance under adverse conditions. System B also contained a filter representing a connection with local cables of maximum length (4 km Ø 0.4 mm) at both ends and two long-distance circuits in series (as measured from Stockholm to Gothenburg and back).

The receiving part in both systems contained a filter corresponding to the standard telephone receiver used in Sweden.

Calculation of nonlinear distortion

Total harmonic distortion is measured, according to IEC publication 268, as the ratio of the RMS output voltage of the total harmonics to the total RMS output voltage for a sinusoidal input.

An equivalent formulation is that the total harmonic distortion is the square root of that part of the output signal variance which is uncorrelated to the sinusoidal input signal. This is demonstrated in Appendix. The advantage of this formulation is that the definition can be generalized to cases where the input signal is arbitrary.

For calculations of the distortions involved in this investigation we have used the latter formulation but instead of sinusoidal input signals we used the actual speech signals. In mathematical forms this could be expressed as

$$d^2 = 1 - \rho^2$$

where $d \cdot 100$ is the distortion in percent and ρ is the correlation between the input and output signals of the nonlinearity. The coefficients giving the desired distortions of 5%, 10%, 20%, 40% and 60% in the quadratic case and 2%, 5%, 10%, 20% and 40% in the cubic case were calculated separately for the various speakers, systems and sentences. The distortions were thus calculated before the second filter of each system.

Generation of stimulus material

Three sentences in Swedish were recorded from one male and one female speaker. Each recording was lowpass filtered at 5 kHz, sampled and converted to digital form at 30 kHz and stored in a computer system. The choice of sampling frequency was determined by the wish that the highest frequency produced by the simulated nonlinearity should not exceed half the sampling frequency to avoid folding effects.

For each of the four linear filters the corresponding time discrete pulse response was calculated.

The signals at the input of the nonlinearity were obtained by convolution of the recorded speech signals with the pulse responses of the transmitting filters of each of the two systems.

At the nonlinearity the density functions of the signals were estimated and used as input to a computer program which calculated the amount of nonlinearity which should be added to the undistorted signal to obtain the desired distortions.

From each signal the squared and cubed signals were calculated. Each of them and the undistorted signal were separately convolved with the pulse responses of the receiving filters. The final stimuli were obtained by adding the linear filtered signal to the filtered squared and cubed signals with the chosen different weights and then scaling the results to get the same RMS value for all the stimuli. Figure 5 gives an example of level density functions of a sentence with and without distortion.

Listening conditions

The test subject was seated in a sound-insulated booth, holding a telephone handset against his "telephone ear", pressing keys to indicate his judgements. The handset was the same as that normally used in the "Dialog" table set but with the telephone receiver replaced by a Standard Electric 4026 A earphone, that is an earphone with distortion below 1% (quadratic) and 0.5% (cubic), and a hard earcap.

An optimum listening level was determined in pilot tests. For both systems, A and B, 82 dB SPL was chosen, also corresponding to the results of Gleiss (1974). In the narrowband system the stimuli were also presented at a low level, 67 dB SPL, denoted BL in the following.

Subjects

24 subjects, 12 males and 12 females, age 16-29 years took part in the experiment. They were randomly divided into three groups (see Procedure). All subjects were checked for normal hearing (less than 20 dB hearing loss 250-8000 Hz/ISO R 389) and were paid for their participation.

Procedure

Basically the procedure meant that each subject listened to a number of pair presentations of a certain sentence. Within each pair distortion was added to one member, either the first or the second, and the subject should judge which of them sounded "worst" (see further the instruction below).

On the basis of preliminary experiments on the authors and some other subjects five distortion levels were chosen for each type of distortion. For X^2 distortion these five levels corresponded to 5, 10, 20, 40 and 60% distortion, and for X^3 distortion to 2, 5, 10, 20 and 40% distortion.

The number of stimulus conditions was 3 sentences x 2 voices x 2 distortion types x 5 distortion levels x 3 telephone systems. In addition there were two possible positions of the distorted stimulus, either first or second within each stimulus pair as described above. Furthermore, each stimulus pair had to be judged many times by each subject to ensure satisfactory reliability of the judgements.

To reduce the work for the subjects it was decided that each subject should only listen to one of the three sentences. The 24 subjects were therefore randomly divided into three groups of eight members each (four males and four females), and it was also randomly decided which group should listen to which sentence.

For each single subject there were 1 sentence x 2 voices x 2 distortion types x 5 distortion levels x 3 telephone systems x 2 positions of the distorted stimulus x 7 replications (= repeated judgements of the same stimulus pair) = 840 cases to judge. The 120 stimulus pairs were recorded in blocks of ten stimuli. Each block contained only one specific voice, one telephone system and one type of distortion. All distortion levels were represented twice in random order in each block. Seven tapes with different random orders for stimuli and blocks were used. The time interval between the members in a stimulus pair was about 1.0 s and the time interval between stimulus pairs was about 4 s. After every block there was a longer interval of about 20 s.

In all the experiment required four experimental sessions of about one hour each for each subject.

The instruction was recorded on tape and was also given in written form to the subject as follows:

> "This experiment deals with perceived sound quality of telephones. We have recorded on tape how voices sound in different types of telephones, and you will listen to these recordings through the telephone receiver beside you.
>
> You hold the receiver close to your usual "telephone ear" (the same ear must be used all the time). In the receiver you will hear a certain sentence read by a man or a woman. The sentence is always presented twice in immediate succession like this (some demonstrations were given) ....
>
> From a technical point of view the two reproductions within each pair are of different quality, that is, one is better and one is worse. How big the difference is varies from case to case. We are interested to know if you also can hear which of the two reproductions is the worst one. For each presentation you shall therefore judge which of the reproductions you think is the worst one, the first one or the second one. You do this by pushing one of the buttons in front of you: button No. 1 if the first reproduction sounds worst, button No. 2 if the second reproduction sounds worst.
>
> Listen carefully to the two reproductions within each pair and try to decide which is the worst one. Give your answer after the presenta-

> tion when the red light is turned on (it goes out when you have pushed one of the buttons). In some cases it may be easy to decide which reproduction is worst, in other cases it may be difficult. But you always have to give your answer soon after the end of the presentation. The next presentation will appear in about 5 s and then you must be prepared to listen again.
>
> The worst reproduction appears equally often in the first position as in the second position so you cannot conclude which one is the worst one on basis of the position. Further, the order of the recorded cases is random so you cannot find a system in that either. You shall trust your spontaneous impression about which reproduction is the worst one! If you sometimes feel uncertain and think you are solely guessing which is the worst one, you may safely continue to do so. Many earlier experiments clearly show that you "guess" much better than you believe yourself!
>
> The experiment is rather long and may be monotonous to you - but there is need for many judgements to get reliable data. By your participation you help us to investigate if you can hear differences between telephones of different technical quality. After the experiment you can get more information about the purpose and the results".

Some further practical points were included in the instruction. After the instruction at least 30 preliminary trials were made as practice before the experiment started. In the middle of each session there was a break of about 5 minutes.

The judgement procedure used here corresponds to what is often called a two-alternative forced-choice (2AFC) procedure (Green & Swets, 1966).

Data treatment

For each single subject his/her number of correct judgements was computed for each distortion level at each of 12 listening conditions (2 voices x 2 distortion types x 3 telephone systems; the position variable was neglected). The threshold value was defined as the distortion level corresponding to 75% correct judgements. These individual threshold values appear in Table I as well as threshold values for the male group, the female group and the whole group of males and females. The group threshold values were obtained by cumulating the correct judgements of all members in the same group and finding the 75% limit in this cumulated distribution (the group threshold values are thus not equal to a simple average of the given individual threshold values).

A separate analysis was made to study the possible effects of the position of the distortion (in the first or in the second position within a stimulus pair, see Procedure). Table II presents the group threshold values for the first and the second position of the distorted stimulus as well as irrespective of the position (the last mentioned values are the same as the group threshold values in Table I).

Table III shows the percentage of correct judgements cumulated over the respective eight subjects, at each distortion level for all 12 listening conditions. This table supplements the information given in Tables I-II and may be used, if wanted, to compute threshold values differently defined than by the 75% criterion used here (compare also Table VI.)

The threshold values in Table I and II were subjected to different forms of analysis of variance according to procedures described in Kirk (1968) using the BMD08V computer program. However, since these analyses only confirmed results which are obvious from direct visual inspection of Tables I and II, the details of these analyses are not given here.

The distortion levels in this investigation were defined as $\sqrt{1 - \rho^2}$ for the actual speech signals (see Methods). The corresponding percentages of distortion using different definitions of distortion levels are given in Table IV.

RESULTS

Inter-individual variations

As seen in Table I there is in most cases a noticeable variation between the individual threshold values within the same listening condition. Since the 2AFC procedure avoids the problem of individual differences in "decision criterion" (Green & Swets, 1966), the observed variations here may be ascribed to "real" individual differences in detectability (including differences in attention, motivation etc). Which subject has the lowest or highest threshold value varies with listening conditions. (However, subject No. 30 has almost always higher threshold values than the other subjects within the same group.)

Range of threshold values

The lowest threshold values that appear in Table I correspond to 5-8% harmonic distortion and in general occur for X^3 distortion of the male voice in system A. The highest threshold values correspond to more than 60% harmonic distortion and in general occur for X^2 distortion in system B and BL.

Effects of distortion types

Threshold values for X^3 distortion are lower, often considerably lower, than threshold values for X^2 distortion. This effect was strongly statistically significant in all analyses and is easily observable in Table I.

Effects of telephone systems

Threshold values for the broadband system (system A) are generally lower than for the narrowband system at either sound level (system B and BL). This effect was strongly statistically significant in all analyses and is also seen in Table I. However, there is no unambigious difference in threshold values between the two sound levels used for the narrowband system (B and BL) - the difference between them varies with subjects and listening conditions as seen in Table I.

For one of the sentences, sentence No. 1, there was a significant interaction between type of distortion and telephone system meaning that the difference in threshold value between system A on one hand and system B and BL on the other hand was bigger for X^2 distortion than for X^3 distortion. This interaction was in its turn accompanied by a significant three factor interaction distortion type x telephone system x voice, meaning that the bigger difference between A and B/BL at X^2 distortion than at X^3 distortion was more marked for the female voice than for the male voice. As seen for sentence No. 1 in Table I the difference between system A and systems B/BL is especially big for X^2 distortion of the female voice (more than half of the threshold values for B/BL are higher than 60%).

Effects of voices and sex of listeners

For sentence No. 1 the threshold value for the male voice was significantly lower than for the female voice, mainly due to the high threshold values occuring at X^2 distortion of the female voice noted above (significant interaction voice x distortion type). For this sentence there was also a significant interaction between voice and sex of listener: distortion on the male voice seems to be more readily detected by the male listeners (subjects 10-13) than by the female listeners (subjects 14-17), see Table I.

For sentence No. 2 and 3 the effects were more varying and not statistically significant.

Effects of sentences

There was a significant difference between the threshold values for the three sentences: sentence No. 2 has the lowest average threshold value, sentence No. 3 the highest average threshold value, and No. 1 falls in between. In other words, the distortion is more easily detectable in sentence No. 2 than in the other two sentences (as read by the present speakers).

As noted from the preceding paragraphs the results for sentence No. 1 are more complex than for the other two sentences in the sense that more interactions between different variables occur for this sentence.

Effects of position of distorted stimulus

As seen in Table II there is a clear tendency to lower threshold values when the distortion is placed on the second member (position 2) in a stimulus pair. The effect is statistically significant and is most evident for sentences No. 2 and 3. There are exceptions, however, and a detailed analysis would be rather complex.

Threshold values expressed in distortion at the output of the system

All results in this report have thus far been based on measurements of the distortion after the nonlinearity but before the second filter. In the practical situation this corresponds to measurement of the distortion caused by the nonlinearities in the microphone and the microphone amplifier, before the telephone line.

Obviously the second filter (line + receiver) strongly affects the perceived distortion. To estimate the distortion at the output of the second filter, the time varying correlations between the signals in each presented pair of sentences were calculated. The averaging time in these calculations could be choosen to maximally 67 ms. The corresponding time varying distortion estimates were calculated according to Appendix formula (7). For each of the stochastic variables thus obtained the 95% level was calculated and used as a conservative measure of the maximum distortion. The mean value was calculated, too. These distortion values were plotted versus the scores of Table III to test if any simple relation existed. Examples are shown in Figures 6 and 7. Tendencies to concentrations of the data points are shown, but the ranges are too wide to justify descriptions with simple curves.

A more detailed picture of the detection thresholds expressed in distortion at the output of the system is given by means of Tables V and VI. Table V shows the mean values of the distortion at the output of the system corresponding to the 5, 10.....60% X^2 distortion and the 2, 5.....40% X^3 distortion as measured before the second filter. By combining these values with the cumulative percentage correct judgements in Table III, it is possible to compute threshold values expressed in distortion at the output of the system. The resulting threshold values are given in Table VI. To facilitate comparisons Table VI presents threshold values expressed both in distortion before the second filter and after the second filter (= at the output). Furthermore this is done for two different definitions of threshold values, namely the value corresponding to 75% correct judgements and also the value corresponding to 60% correct judgements. All data refer to the whole group of eight subjects within each sentence.

Since the distortion was lower after the second filter than before it, all threshold values are lower expressed in distortion after the second filter than before the second filter.

The second filter of system B/BL was more narrow than that of system A thus eliminating more distortion. The threshold values measured before the second filter are higher for system B/BL than for system A as expected. When the distortion is measured after the second filter ("at the ear"), the situation is more varied.

Threshold values for X^3 distortion are generally lower than for X^2 distortion.

The comments made here hold for the threshold values computed both according to the 75% criterion and to the 60% criterion.

DISCUSSION

The detectability of nonlinear distortion in telephone speech apparently depends on many factors, stimulus factors as well as listener factors. Obviously the choice of sentences and speakers is of great importance, and there are marked inter-individual variations in the detection of the distortion. Too much generalization from the present results should therefore be avoided, awaiting results from further research.

When comparing results from different reports and when suggesting recommendations for maximum allowed nonlinear distortion it is extremely important to notice the exact definition of the nonlinear distortion used in the actual reports. In the present report the distortion was defined as the square root of that part of the output signal variance which is uncorrelated to the input signal. Corresponding values using two other definitions were given in Table IV. Furthermore there is a considerable difference between threshold values expressed in distortion before the second filter and threshold values expressed in distortion at the output of the system.

The threshold values reported by Ahlborg (1967) are higher than in the present report, when using the definition for distortion of a sinusoidal voltage whose RMS value equals the RMS value of the speech signal. This one is the only definition usable for direct comparison. There are, however, several differences between the tests: The nonlinearities are defined in somewhat different ways. The nonlinearities Ahlborg used approximated the quadratic and cubic functions only within a certain amplitude range; outside this range pure saturation occured. The cubic term was negative. The instructions were different. The listeners in Ahlborgs test were experienced to recognize distortion etc.

The threshold values for the case of the female voice transmitted by system A (expressed in distortion at the output of the system) agree fairly well with the threshold values for detection of X^2 and X^3 distortion of a recorded female voice listened to by headphones as reported by Gabrielsson et al. (1976).

The consequences of the present results for recommendations about maximum allowed nonlinear distortion in telephone speech may be discussed as follows. Since the threshold values for quadratic distortion are generally higher than those for cubic distortion, the maximum allowed cubic distortion should be set lower than the maximum allowed quadratic distortion. As regards differences between broadband systems (system A here) and narrowband systems (B/BL) the limit must be set lower for broadband systems if the distortion is measured before the second filter. However, for distortion measured at the output of the system no clear difference between broadband and narrowband systems can be established from the present results. It is also noted that there is no unambigious difference in detectability of distortion at the two different listening levels of the narrowband system (B and BL), neither when the distortion is measured before the second filter, nor when it is measured at the output.

To get some kind of general suggestions from the present data it may be convenient to refer to the cumulative distribution of correct judgements in Table III. From these data threshold values corresponding to 75% correct detection and to 60% correct detection were computed as given in Table VI. Although the 75% criterion is a common one for threshold values, it seems preferable here to adopt the lower 60% criterion not to allow the more sensitive listeners to detect the distortion. As can be seen in Table III there are sometimes "regressions" in the percentage correct judgements (that is, the percentage correct judgements is <u>lower</u> for the following higher distortion level). Such "regressions" occur for percentages between 39% and 61% and may be taken as an indication of random error around the 50% value that should result from pure guessing. A 60% threshold criterion would therefore approximately reflect a transition from pure guessing to "real" detection of the distortion.

Adopting the 60% criterion and especially noting the cases with the lowest occurring threshold values in Table VI it seems that the maximum allowed quadratic distortion would be of the order 12 - 16% and the maximum allowed cubic distortion of the order 3 - 6%, when the distortion is measured before the second filter. The corresponding values for the distortion measured at the output of the system would be 7 - 9% for quadratic distortion and 2 - 4% for cubic distortion.

It may be argued that the 60% criterion is too severe. If another criterion is wanted the information given in Tables III - VI provide necessary information for computing the corresponding limits for maximum allowed distortion from the present data. It can also be noted that in practice a telephone subscriber would rarely have the opportunity to compare a distorted system with an undistorted one in rapid succession as made in this test (and here it also seems easier to detect the distortion when the undistorted system is presented first). Without such a direct comparison the attitude to the effect of nonlinear distortion may be more lenient, but on the other hand the listening time in a real conversation is considerably longer than the time used here, which may act in the opposite direction. A definite answer to the question of the most appropriate limit for maximum allowed nonlinear distortion in telephone circuits can thus not be given without further investigations.

ACKNOWLEDGEMENTS

The authors express their gratitude to Norman Gleiss for valuable discussions concerning methods and results, to Lennart Persson, responsible for the control device, and to Ann Pettersson, who served as experimenter. The major part of this project was supported by the Swedish Telecommunications Administration. Some general parts were also supported by the Swedish National Board for Technical Development.

REFERENCES

Ahlborg, H. (1967). Konstruktion av anordning som ger en definierad olinjär distorsion, samt bestämning av hur denna påverkar transmissionskvaliteten på en telefonförbindelse. Rep. No. 284/67, Swedish Telecommunications Administration.

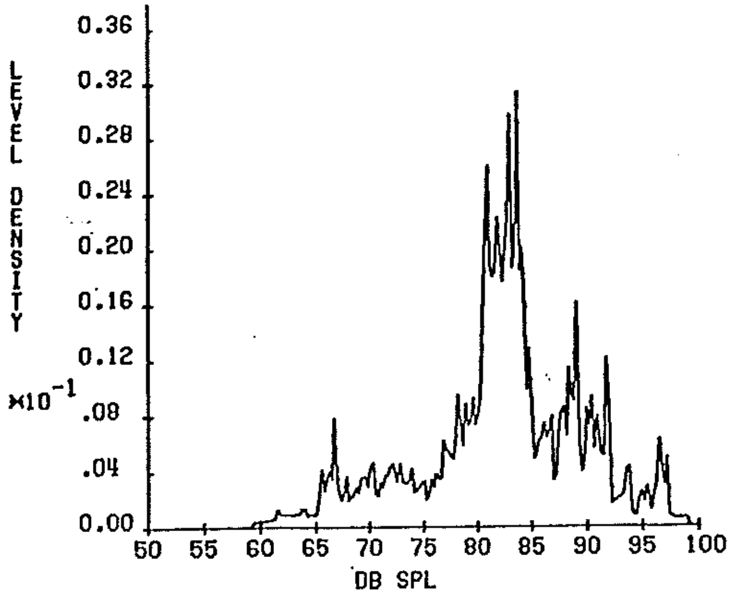
CCITT Green Book, Volume V - Question 13/XII, Annex.

Gabrielsson, A., Nyberg, P.O., Sjögren, H., Svensson, L. (1976). Detection of amplitude distortion by normal hearing and hearing impaired subjects. Rep. No. 83, Technical Audiology, Karolinska institutet, Stockholm.

Gleiss, N. (1974). Preferred listening levels in telephony. Tele (English edition) No. 2, 1974.

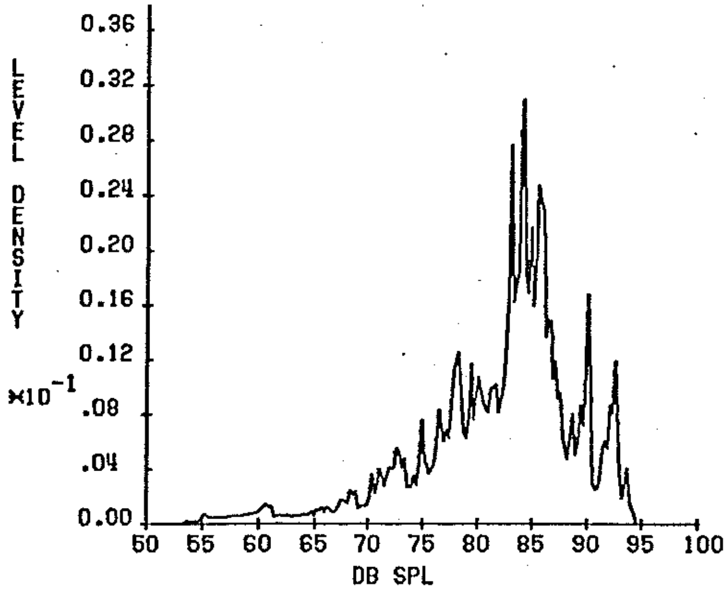
Green, D.M. & Swets, J.A. (1966). Signal detection and psycho-physics. New York: Wiley.

Kirk, R.E. (1968). Experimental design. Procedures for the behavioural siences.
Belmont, California: Brooks/Cole Publishing Company.



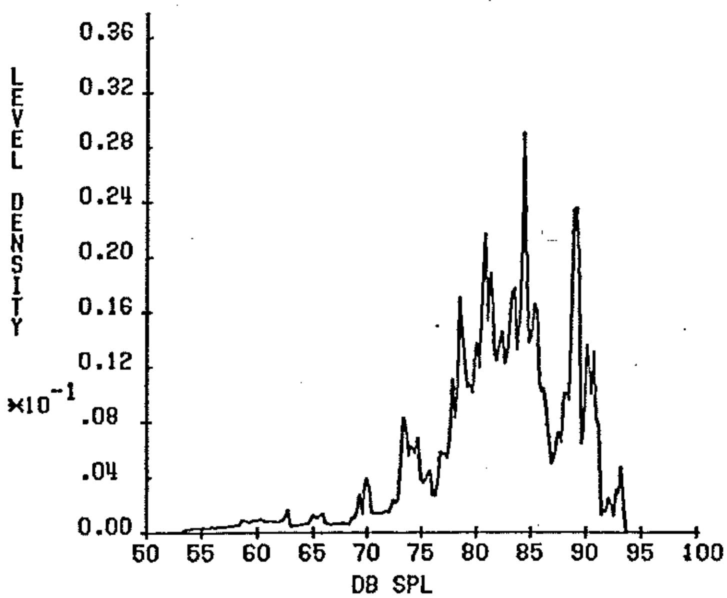
Sentence No 1

Mean	= 82.0	dB
St.dev.	= 7.44	dB
Max	= 99.0	dB
Min	= 59.0	dB



Sentence No 2

Mean	= 82.0	dB
St.dev.	= 7.06	dB
Max	= 94.3	dB
Min	= 53.2	dB

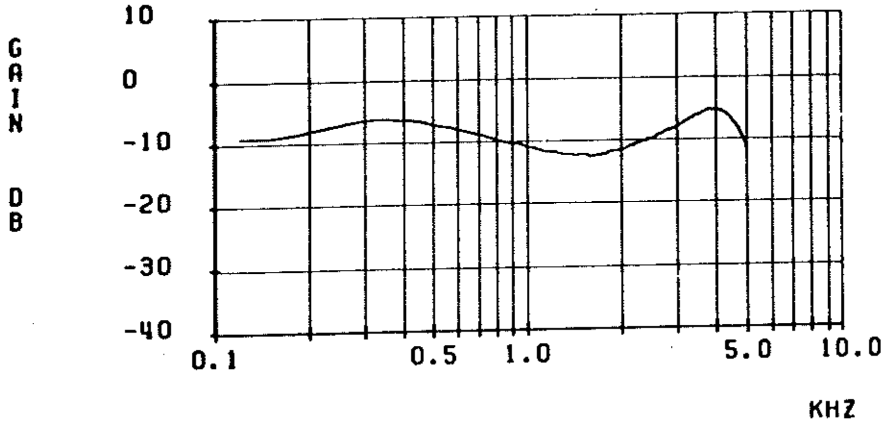


Sentence No 3

Mean	= 82.0	dB
St.dev.	= 6.76	dB
Max	= 93.4	dB
Min	= 53.0	dB

Figure 1 Level density of undistorted speech material. Averaging time 20 ms. Male speaker.

a)



b)

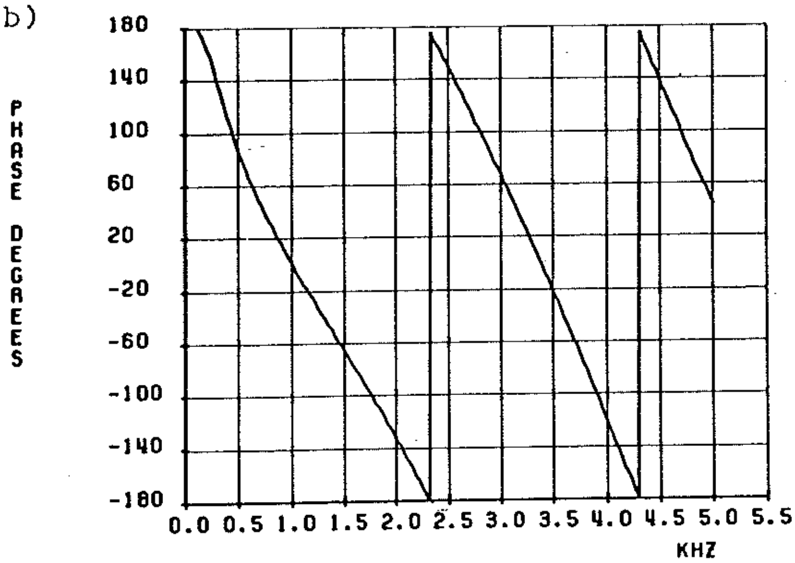


Figure 2 Reference earphone
a) frequency curve
b) phase curve

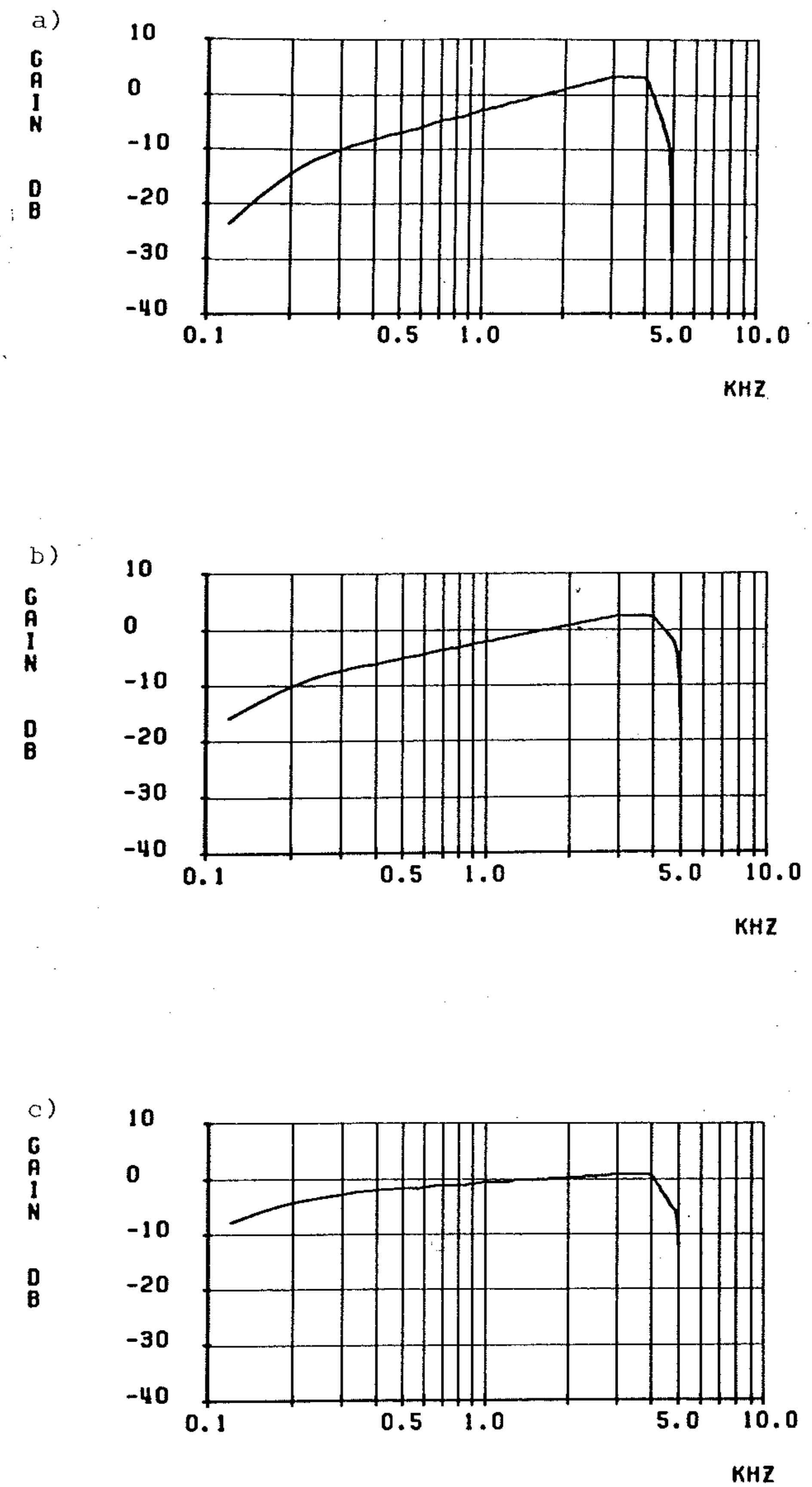


Figure 3 System A, frequency response
a) total
b) sending part
c) receiving part

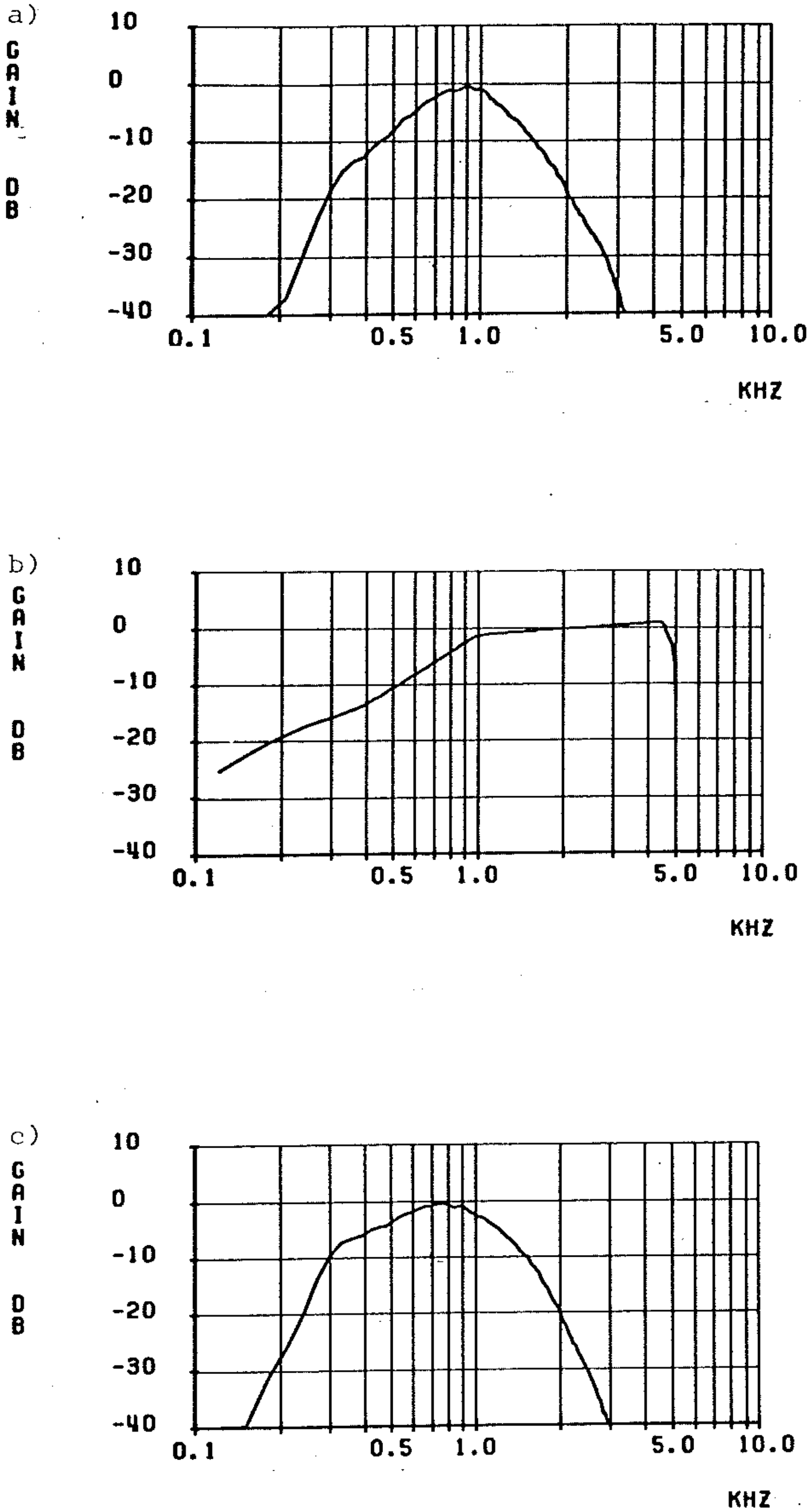
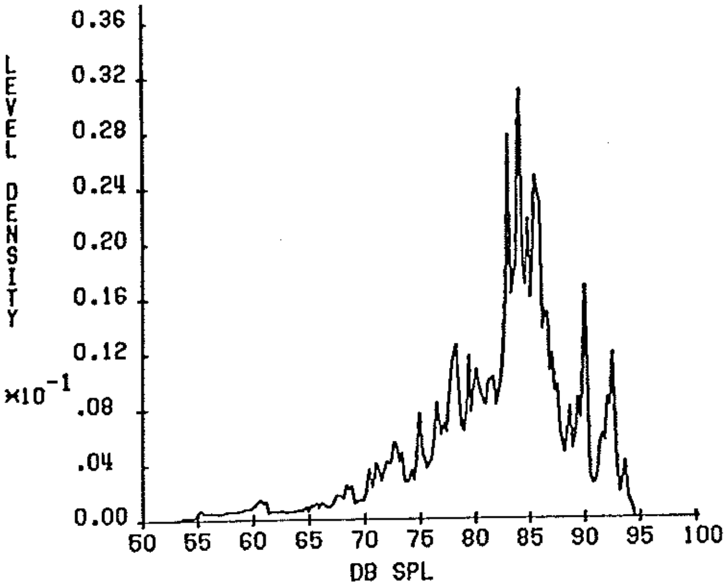


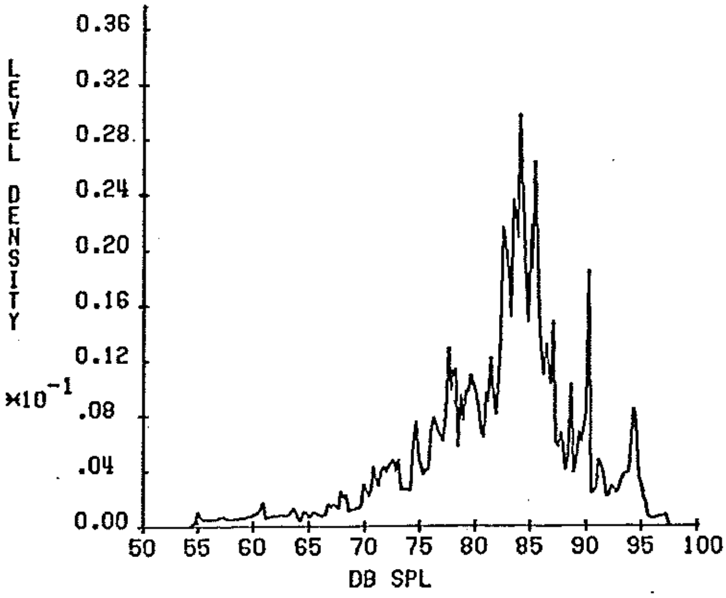
Figure 4 System B, frequency response
a) total
b) sending part
c) receiving part (line + receiver)

a)



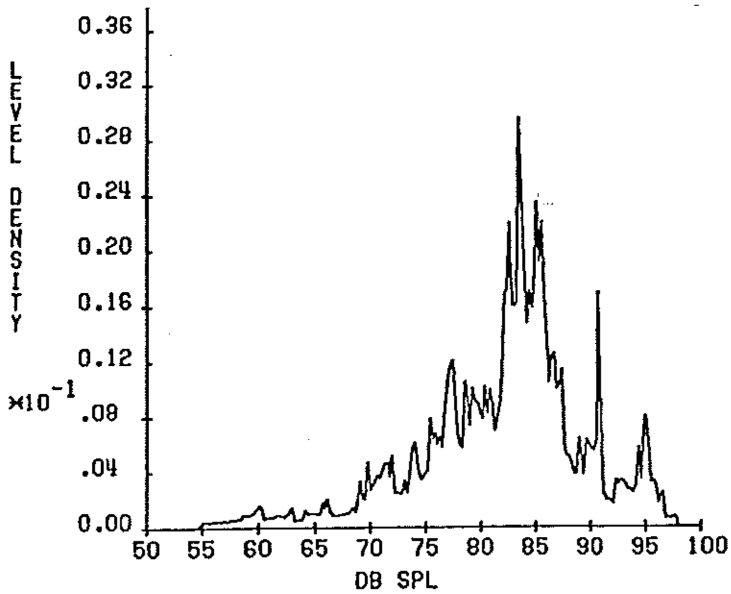
Mean = 82.0 dB
St.dev. = 7.06 dB
Max = 94.3 dB
Min = 53.2 dB

b)



Mean = 82.0 dB
St.dev. = 7.26 dB
Max = 97.5 dB
Min = 54.1 dB

c)



Mean = 82.0 dB
St.dev. = 7.52 dB
Max = 97.8 dB
Min = 54.6 dB

Figure 5 Level density of speech material. Sentence No 2 with male speaker.
a) undistorted and with
b) 40% quadratic distortion and
c) 20% cubic distortion

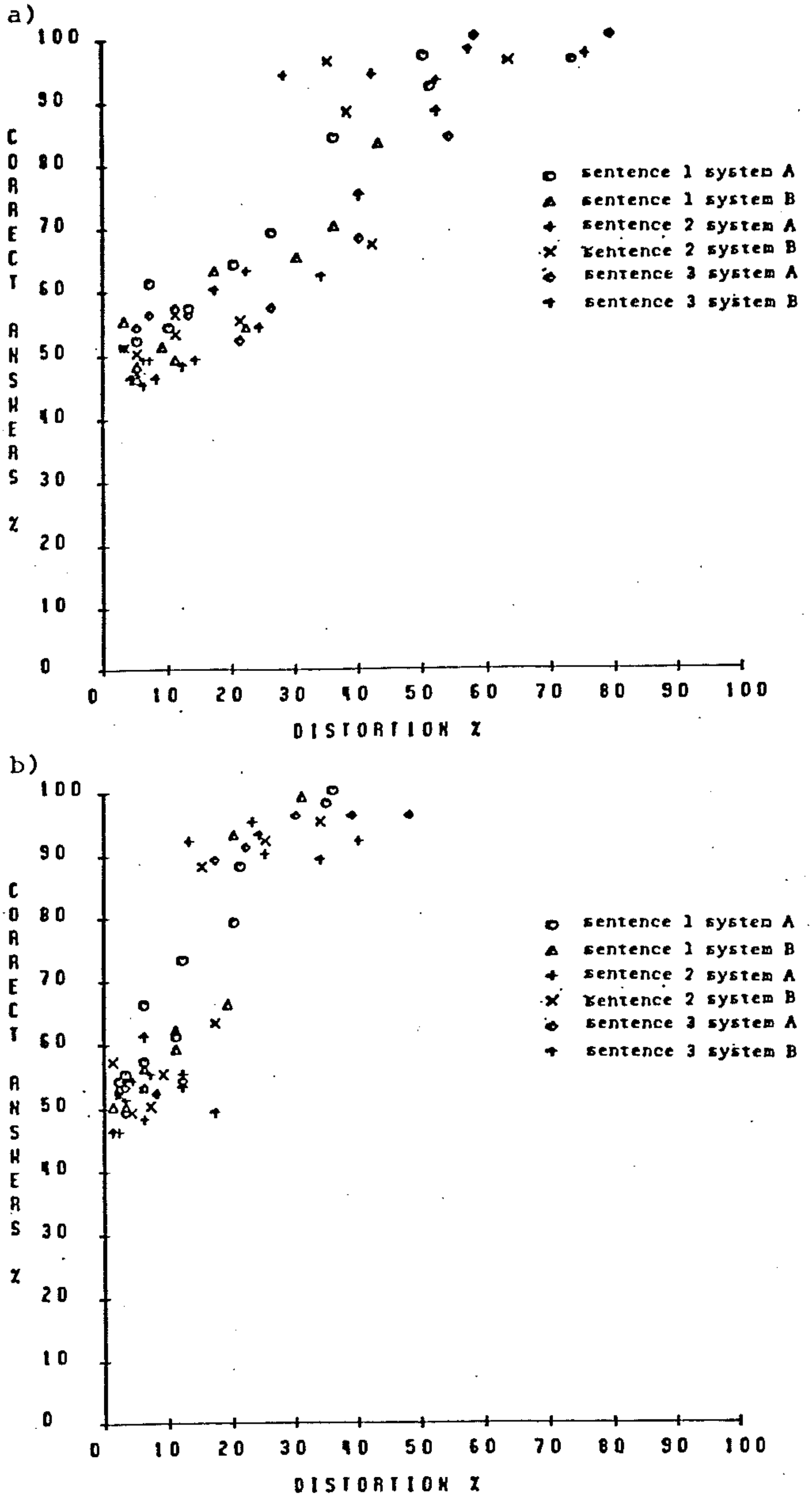


Figure 6 Number of correct answers relative peak values of the distortion at the output of the system. Averaging time 20 ms.
a) quadratic distortion
b) cubic distortion

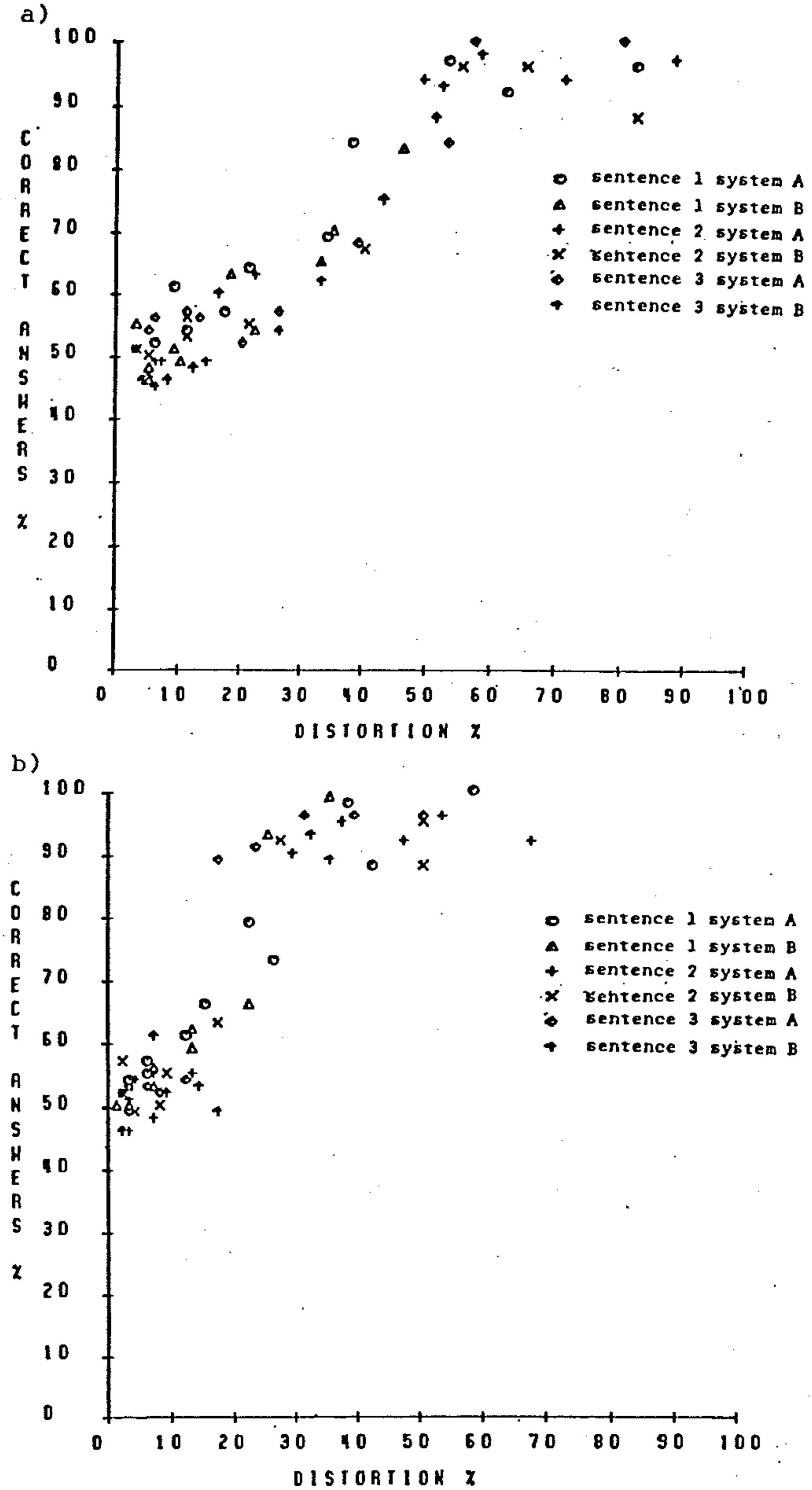


Figure 7 Number of correct answers relative peak values of the distortion at the output of the system. Averaging time 67 ms.
a) quadratic distortion
b) cubic distortion

Subject		Male voice: Quadratic distortion % A	B	BL	Male voice: Cubic distortion % A	B	BL	Female voice: Quadratic distortion % A	B	BL	Female voice: Cubic distortion % A	B	BL
Sentence No. 1													
	10	18.3	35.0	25.0	6.3	15.0	17.0	26.0	>60.0	55.0	17.0	>40.0	32.5
	11	6.3	17.0	30.0	6.3	26.0	26.0	15.8	50.0	48.3	7.5	28.3	23.3
	12	23.3	18.8	30.0	8.8	22.5	16.3	30.0	50.0	>60.0	17.5	27.5	16.3
	13	18.8	35.0	42.5	5.6	14.2	12.5	22.5	50.0	>60.0	13.0	22.5	28.3
Male group		17.5	25.0	31.1	6.5	17.1	17.0	22.7	55.0	60.0	14.2	31.1	25.7
	14	16.3	51.7	23.3	13.8	22.5	22.5	23.3	>60.0	>60.0	12.5	22.5	26.0
	15	31.7	55.0	36.7	18.3	30.0	27.5	32.5	>60.0	>60.0	30.0	35.0	30.0
	16	35.0	>60.0	52.5	30.0	30.0	26.0	52.5	>60.0	>60.0	19.0	18.3	15.0
	17	47.5	>60.0	55.0	15.0	32.2	31.3	51.3	>60.0	>60.0	27.5	22.5	15.0
Female group		32.5	60.0	45.7	18.2	29.6	27.6	44.3	>60.0	>60.0	24.0	25.9	21.8
Whole group		25.4	51.0	40.0	11.2	25.4	23.1	30.9	>60.0	>60.0	18.0	28.6	24.0
Sentence No. 2													
	20	15.6	42.5	46.0	7.1	28.3	26.0	28.3	19.0	31.3	14.2	16.1	15.0
	21	50.0	55.0	46.0	8.1	35.0	19.0	43.3	50.0	>60.0	19.0	28.3	15.8
	22	17.5	46.0	17.5	12.5	26.0	28.3	17.0	57.5	56.7	11.3	>40.0	32.5
	23	9.4	42.5	30.0	7.5	18.8	18.3	17.2	32.2	31.3	11.3	14.2	9.2
Male group		15.0	47.4	41.3	8.1	26.3	24.3	23.3	38.3	34.3	14.2	18.3	16.1
	24	19.2	37.5	43.3	7.5	37.5	23.3	37.5	30.0	35.7	18.8	17.0	>40.0
	25	15.6	37.5	30.0	7.5	30.0	30.0	26.0	28.3	31.3	15.0	13.0	16.8
	26	16.1	33.8	28.3	7.1	18.8	28.3	28.3	31.3	30.0	15.0	15.0	14.2
	27	15.0	55.7	50.0	6.9	28.3	32.9	31.3	22.5	32.2	15.0	15.6	14.2
Female group		16.3	43.1	40.0	7.3	29.0	29.6	30.4	28.7	32.3	15.6	15.2	16.4
Whole group		15.8	45.6	40.7	7.7	27.8	27.6	28.0	32.0	32.9	15.0	16.5	16.3
Sentence No. 3													
	30	48.3	>60.0	47.5	>40.0	>40.0	>40.0	35.0	>60.0	57.5	>40.0	>40.0	>40.0
	31	26.0	46.0	43.3	8.1	28.3	32.2	51.3	>60.0	47.5	7.5	22.5	31.3
	32	36.7	52.5	>60.0	7.9	32.5	37.5	8.5	50.0	>60.0	22.5	28.3	34.0
	33	26.0	50.0	50.0	7.1	32.9	32.2	50.0	>60.0	>60.0	16.1	28.3	32.9
Male group		30.9	50.9	46.7	8.9	34.1	36.4	44.4	60.0	60.0	17.7	32.0	33.9
	34	34.0	57.5	>60.0	7.5	32.9	30.0	50.0	>60.0	58.3	9.5	26.0	35.7
	35	30.0	47.5	37.5	7.8	26.0	31.7	46.0	57.5	38.3	14.2	32.2	32.2
	36	32.9	46.0	34.0	6.9	32.2	23.3	46.0	48.3	50.0	16.5	31.3	33.8
	37	38.3	48.3	30.0	7.5	35.7	32.5	30.0	>60.0	47.5	13.0	31.3	22.5
Female group		34.7	49.5	37.5	7.5	32.1	30.0	44.4	60.0	51.1	14.2	30.7	32.1
Whole group		33.3	50.0	40.0	8.1	32.9	33.5	44.4	60.0	55.9	15.6	31.1	32.9

Table I Threshold values for individual subjects, male group, female group and whole group at all stimulus conditions within each sentence.

	Male voice						Female voice					
	Quadratic distortion, %			Cubic distortion, %			Quadratic distortion, %			Cubic distortion %		
	A	B	BL	A	B	BL	A	B	BL	A	B	BL
Sentence no. 1												
Males												
Position 1	16.4	18.4	35.6	7.9	14.4	20.7	27.2	>60.0	>60.0	13.0	31.1	23.9
Position 2	18.2	46.0	27.1	4.7	22.9	16.7	17.3	46.8	57.6	21.0	32.0	26.8
Total	17.5	25.0	31.1	6.5	17.1	17.0	22.7	55.0	60.0	14.2	31.1	25.7
Females												
Position 1	28.3	54.3	46.8	18.7	27.7	31.8	46.4	>60.0	>60.0	25.8	29.9	31.0
Position 2	36.7	>60.0	43.8	16.0	34.4	25.0	42.2	>60.0	>60.0	18.5	20.0	9.2
Total	32.5	60.0	45.7	18.2	29.6	27.6	44.3	>60.0	>60.0	24.0	25.9	21.8
Whole group												
Position 1	22.3	45.0	42.7	13.6	22.2	26.3	33.9	>60.0	>60.0	17.3	30.5	26.9
Position 2	28.3	60.0	35.7	6.3	28.4	18.6	26.3	>60.0	>60.0	19.2	26.3	9.3
Total	25.4	51.0	40.0	11.2	25.4	23.1	30.9	>60.0	>60.0	18.0	28.6	24.0
Sentence no. 2												
Males												
Position 1	18.1	43.3	37.5	9.7	29.1	27.5	31.1	34.3	36.0	15.8	20.0	19.0
Position 2	5.0	49.2	44.4	4.4	20.0	20.0	15.7	43.3	30.0	11.4	17.7	12.5
Total	15.0	47.4	41.3	8.1	26.3	24.3	23.3	38.3	34.3	14.2	18.3	16.1
Females												
Position 1	16.8	48.0	44.4	8.5	29.2	32.0	33.6	29.2	34.4	17.1	16.9	18.0
Position 2	15.0	33.3	32.0	3.2	28.6	25.0	9.0	28.0	23.3	12.5	12.2	14.0
Total	16.3	43.1	40.0	7.3	29.0	29.6	30.4	28.7	32.3	15.6	15.2	16.4
Whole group												
Position 1	17.4	46.3	41.3	9.1	29.2	30.4	32.9	31.0	34.9	16.6	17.6	18.4
Postion 2	10.8	45.0	40.0	3.8	25.0	22.9	12.9	33.3	26.0	12.0	15.5	13.3
Total	15.8	45.6	40.7	7.7	27.8	27.6	28.0	32.0	32.9	15.0	16.5	16.3
Sentence no. 3												
Males												
Position 1	32.5	56.0	45.0	10.0	35.4	38.6	44.4	>60.0	>60.0	15.8	33.3	36.0
Position 2	26.7	46.7	50.0	6.0	30.0	32.5	44.4	40.0	50.0	24.0	30.0	30.0
Total	30.9	50.9	46.7	8.9	34.1	36.4	44.4	60.0	60.0	17.7	32.0	33.9
Females												
Position 1	38.5	43.3	45.0	8.7	31.3	33.3	42.5	60.0	53.3	15.6	30.0	31.4
Position 2	26.7	52.3	25.0	3.0	33.3	19.2	46.0	60.0	50.0	11.3	31.3	32.9
Total	34.7	49.5	37.5	7.5	32.1	30.0	44.4	60.0	51.1	14.2	30.7	32.1
Whole group												
Position 1	36.2	49.1	45.0	9.2	33.1	35.9	43.5	60.0	60.0	15.7	31.3	33.8
Position 2	26.7	50.5	32.0	4.0	32.5	27.3	45.3	60.0	50.0	15.4	30.9	31.8
Total	33.3	50.0	40.0	8.1	32.9	33.5	44.4	60.0	55.9	15.6	31.1	32.9

Table II. Threshold values for male group, female group, and whole group at different positions of the distorted stimulus and irrespective of position (= total).

Sentence no. 1

	System	Quadratic distortion 5	10	20	40	60%	Cubic distortion 2	5	10	20	40%
Male voice	A	61%	57	69	92	96	55	66	73	88	100
	B	46	51	63	65	83	53	53	59	66	99
	BL	48	50	62	75	96	50	52	62	71	99
Female voice	A	52	54	64	84	97	54	57	61	79	98
	B	55	48	49	54	70	50	50	56	62	93
	BL	53	53	59	56	62	50	58	61	71	93

Sentence no. 2

Male voice	A	49%	49	94	93	97	51	55	92	90	92
	B	50	53	55	67	96	52	49	55	63	95
	BL	52	51	60	74	99	46	39	54	63	96
Female voice	A	49	48	63	94	98	46	48	55	95	96
	B	51	47	56	88	96	57	49	50	88	92
	BL	50	49	49	89	93	49	52	51	89	91

Sentence no. 3

Male voice	A	56%	56	57	84	100	49	52	89	96	96
	B	46	46	60	62	88	52	54	52	49	89
	BL	51	46	49	75	84	52	39	44	51	87
Female voice	A	54	57	52	68	100	53	53	54	91	96
	B	51	45	48	54	75	46	54	61	53	93
	BL	54	57	46	54	80	45	52	54	46	91

Table III. Percentage of correct judgements summed over all eight subjects within the same sentence.

Definition	Voice	Sentence	System	Quadratic distortion, %					Cubic distortion, %				
I				5%	10%	20%	40%	60%	2%	5%	10%	20%	40%
II	Male	1	A	10	20	38	66	81	5	11	16	22	28
			B	8	16	31	57	77	3	7	11	17	24
		2	A	7	14	28	53	73	3	6	10	17	24
			B	8	15	30	55	75	3	6	10	16	23
		3	A	7	15	29	55	75	3	7	12	18	25
			B	7	13	26	49	70	3	6	11	17	24
	Female	1	A	7	13	26	50	71	2	5	9	15	22
			B	5	10	21	41	61	2	4	8	14	22
		2	A	7	14	27	52	72	3	6	10	16	23
			B	6	12	23	46	66	2	5	9	15	23
		3	A	7	13	26	50	70	3	6	11	17	25
			B	6	12	24	47	67	3	6	10	17	25
III	Male	1	A	1	2	4	8	14	0.1	0.1	0.3	0.6	4
			B	1	2	3	6	11	0.0	0.1	0.1	0.3	0.8
		2	A	1	2	5	10	18	0.1	0.2	0.4	0.9	2
			B	1	2	4	9	15	0.0	0.1	0.2	0.5	1
		3	A	1	3	5	11	19	0.1	0.3	0.6	1	3
			B	1	2	5	10	16	0.1	0.2	0.4	1	3
	Female	1	A	1	2	4	9	15	0.0	0.1	0.3	0.6	2
			B	1	2	3	7	12	0.0	0.1	0.2	0.5	2
		2	A	1	3	5	11	19	0.1	0.2	0.5	1	3
			B	1	2	4	7	13	0.0	0.1	0.3	0.6	2
		3	A	2	3	6	14	23	0.2	0.4	0.9	2	5
			B	1	3	6	12	21	0.2	0.4	0.8	2	5

Table IV: Distortions in percent for different definitions of distortion.

Definition I $p = \sqrt{1 - \rho^2}$ for the actual speech signals

Definition II Distortion of a sinusoidal voltage whose peak value equals the peak value of the speech signal

Definition III Distortion of a sinusoidal voltage whose RMS value equals the RMS value of the speech signal

Voice	System	Quadratic distortion					Cubic distortion				
		5	10	20	40	60%	2	5	10	20	40%
						Sentence no. 1					
Male	A	4	7	15	29	45	2	3	6	10	19
	B/BL	3	5	10	19	30	1	2	3	6	13
Female	A	3	5	10	20	30	1	1	2	6	11
	B/BL	1	2	5	9	16	0.9	1	2	2	6
						Sentence no. 2					
Male	A	4	8	16	32	49	2	3	5	11	21
	B/BL	3	7	13	26	40	1	2	4	7	17
Female	A	3	6	12	23	35	0.8	2	3	8	16
	B/BL	2	3	6	15	22	0.3	0.9	2	5	9
						Sentence no. 3					
Male	A	4	9	18	37	57	2	4	8	14	27
	B/BL	3	5	11	22	34	1	2	5	10	21
Female	A	4	8	15	29	44	2	3	7	13	27
	B/BL	2	4	7	15	25	1	2	3	7	17

Table V Mean values of the distortion at the output of the system. Averaging time 20 ms.

		75% criterion				60% criterion			
		Sentence No. 1							
Voice	System	Quadratic distortion % Before second filter	Quadratic distortion % After second filter	Cubic distortion % Before second filter	Cubic distortion % After second filter	Quadratic distortion % Before second filter	Quadratic distortion % After second filter	Cubic distortion % Before second filter	Cubic distortion % After second filter
Male	A	25	19	11	7	13	9	3	3
	B	51	25	25	8	18	9	11	3
	BL	40	19	23	7	18	9	9	3
Female	A	31	16	18	5	16	8	9	2
	B	>60	>30	29	4	48	12	17	2
	BL	>60	>30	24	3	53	14	8	2
		Sentence No. 2							
Male	A	16	13	8	4	12	10	6	3
	B	46	30	28	11	28	18	16	6
	BL	41	27	28	11	20	13	17	6
Female	A	28	16	15	6	18	11	11	4
	B	32	11	17	4	23	7	13	3
	BL	33	12	16	4	26	9	12	3
		Sentence No. 3							
Male	A	33	31	8	7	22	20	6	5
	B	50	28	33	17	20	11	26	13
	BL	40	22	34	17	29	16	25	10
Female	A	44	32	16	10	30	22	12	8
	B	60	25	31	13	46	18	24	9
	BL	56	23	33	13	45	17	26	10

Table VI Threshold values for whole group before and after the second filter.