ISSN 0280-6819

TECHNICAL AUDIOLOGY

Report TA No. 116
March 1987

INFLUENCE OF NONLINEAR DISTORTION ON SPEECH INTELLIGIBILITY: HEARING IMPAIRED LISTENERS

Ann-Cathrine Lindblad

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

Report TA116
March 1987

INFLUENCE OF NONLINEAR DISTORTION ON SPEECH INTELLIGIBILITY: HEARING-IMPAIRED LISTENERS

Ann-Cathrine Lindblad

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

ISSN 0280-6819
Report TA116
March 1987

INFLUENCE OF NONLINEAR DISTORTION ON SPEECH INTELLIGIBILITY: HEARING-IMPAIRED LISTENERS

Ann-Cathrine Lindblad

ABSTRACT

Well-founded specifications on maximum tolerated nonlinear distortion in hearing aids ought to be set up to improve speech intelligibility and perceived sound quality.

The effect of nonlinear distortion on speech intelligibility was investigated for 16 subjects with moderate to severe sensorineural hearing losses. Quadratic and cubic distortions were introduced on speech materials by computer programs at two distortion levels each. Each type of distortion was tested separately with nonsense rhyme word lists, testing vowels or initial voiceless consonants.

Strong nonlinear distortion on vowels decreased the intelligibility scores considerably in comparison with an undistorted condition, e.g. by 18% at the higher level of quadratic distortion, probably by introducing "new formants" at multiple or difference frequencies of the original formants, shown plausible by logically explainable confusions. Vowels with low first formants were most sensitive because of the low frequencies of their harmonic distortion. Short vowels suffered considerably more than long vowels suggesting that the formant transition pattern gets badly damaged by the distortion, forcing the ear to rely more on the stationary part of the vowel.

For voiceless consonants there was a general reduction in intelligibility for subjects with flat losses showing a significant difference from the results of subjects with sloping losses. Among the subjects with sloping losses there were those who could make use of the transposition effect from distortion products at low frequencies (+60% for some consonants). For some subjects, whose frequency selectivity was tested, the benefit of transposition was correlated to having better frequency selectivity at lower frequencies than at 4 kHz.

This work was supported by the National Swedish Board for Technical Development and the Ministry of Health and Social Affairs.

INTRODUCTION

Nonlinear distortion influences perceived sound quality and speech intelligibility in sound-reproducing systems. Today there is still disturbing nonlinear distortion in many hearing aids and there is a need for a well-founded recommendation of maximum allowed nonlinear distortion. From such a recommendation technical and economical optimizations of hearing aids could be made. Knowledge of the effects of the various components of the nonlinear distortion is necessary for working out the recommendation. One problem hereby is that the fixed relationship between intermodulation and harmonic distortion for an ideal nonlinearity is destroyed in the hearing aid, since it contains several filtering parts and may contain more than one nonlinearity too. Another complication is the intricate time, frequency and intensity pattern of speech. We need to know for which type of speech sounds and in which frequency areas the nonlinear distortion is most critical. The methods of measuring harmonic distortion and intermodulation up to now use sinewaves (IEC 118-0, 1983, The Nordic Committee on Disability, 1986), which might be preferable for simplicity, but they do not always give fair judgements of the effects on sound quality. For setting up perceptually relevant methods the present knowledge of the perceived effects of the distortion is definitely insufficient.

Ideally the nonlinear distortion in a hearing aid should seldom be detected and it should not influence speech intelligibility negatively. Tolerances for maximum allowed distortion should be set as near to this ideal as regarded possible. Furthermore there should be specified a technically and perceptually relevant method of measuring the distortion. We attempt to collect the knowledge needed to obtain this.

For normal-hearing people two different mechanisms seem to determine the detection thresholds for nonlinear distortion (Günthersen, 1982):
a) direct perception of the distortion products - this mechanism works for simple stationary signals
b) fusing of signal and distortion products to form a new percept - which works for nonstationary signals.
For complex, stationary signals either mechanism can determine the threshold depending on properties of the signal, such as frequency and presentation level. No doubt complex nonstationary signals, and especially speech, are the most common to be presented through a hearing aid. This means that fusing of signal and distortion products is the dominating mechanism when detecting nonlinear distortion in a hearing aid.

In previous studies at our department we have investigated detection thresholds for various sound stimuli through some broadband and some narrowband systems. The report TA83 (Gabrielsson et al, 1976) includes both normal-hearing and hearing-impaired subjects listening to music, chat, voice and sinewave stimuli through a broadband system. Two reports regard speech stimuli through narrowband systems: TA87 (Gabrielsson et al, 1978) with normal-hearing and TA105 (Lindblad, 1982) with hearing-impaired subjects. The results indicate that no more distortion can be allowed in systems for hearing-impaired people than in systems for normal-hearing, since many hearing-impaired persons have as good detection thresholds for complex stimuli as the more sensitive persons among the normal-hearing.

In the investigations reported in TA87 and TA105 we used a method for measuring the distortion based on the correlation between the input and the output speech signals, TA90 (Olofsson, 1978). Used with a standardized speech spectrum as input signal and with a weighting according to perceptual data this might be the method looked for.

Jirsa and Norris (1982) studied effects of nonlinear distortion on speech intelligibility in hearing aids with the same low harmonic distortion, less than 2%, but with various, higher, degrees of intermodulation, different in different frequency areas. The selected hearing aids were chosen to have the same gain, frequency and noise characteristics and the same second and third harmonic distortion levels. The hearing aid with the highest intermodulation at frequencies below 1 kHz gave the significantly lowest speech intelligibility scores in quiet for sensorineural hearing losses, which the authors explained by upward spread of masking. When a competing noise at a S/N of 5 dB was added, the significant difference disappeared. However, the rank order between the hearing aids, according to intelligibility scores, was the same in all the subtests.

Singer (1981) generated synthetic intermodulation, f2-f1, with the difference frequencies 250, 500, and 750 Hz and two degrees of intermodulation, 12% an 30%. He explained the decrease in speech intelligibility for normal-hearing subjects with masking effects from the intermodulation noise on the second formant of each word.

In our department studies of effects on speech intelligibility from nonlinear distortion for normal-hearing were reported in the reports TA64 and TA64A (Ingelstam et al, 1970 and 1971). Both quadratic and cubic distortion decreased the speech intelligibility at the individual most comfortable levels, but strong quadratic distortion gave an increase in intelligibility at low sound pressure levels. In 1972 we also reported to the Swedish Board of Technical Development, but only in Swedish, about tests with mildly hearing-impaired persons. Up till then we used logatoms, but considered it impossible to go on with moderately and severely hearing-impaired persons without having a speech test better suited for them. The present study deals with speech intelligibility for hearing-impaired subjects with moderate to severe hearing losses. The speech test material and test procedure were developed for this purpose.

METHODS

The effect of the nonlinear distortion was evaluated by comparing the results of speech intelligibility tests with distorted and undistorted rhyme words. Quadratic and cubic distortion were tested separately.

Subjects

Sixteen subjects, 9 males and 7 females, with moderate to severe sensorineural hearing losses were tested. Their mean loss (0.5, 1., 2., 4. kHz) ranged from 49 to 95 dB, mean 68 dB. The slopes of the audiograms, defined as the difference between the mean loss from 2 kHz and upwards and the mean loss below 2 kHz, ranged from -6 to 44 dB, mean 20 dB. Figure 1 shows mean audiograms for the subjects divided into two groups with audiogram slopes flatter or steeper than 15 dB respectively. The subjects were 25 to 68 years old, median 40. They were paid for their participation.

Speech material

The speech material and the test procedure were designed especially to meet the needs at testing moderately to severely hearing impaired subjects. The test was constructed several years ago by the author but has not been formally reported. The speech material was recorded with a female speaker (the author, fundamental 190-240 Hz) and consisted of CVC combinations without linguistic meaning, testing either vowels or voiceless consonants.

The test consonants were the Swedish consonants /P, T, K, S, SJ, TJ, F, H / in initial position, see Table 1. These consonants also formed the set of phonemic initials in the response alternatives. Consequently, there were eight response alternatives for each test stimulus, producing a low probability of guessing correctly. The vowel and the final consonant were the same in all the response alternatives of a test stimulus to avoid extra cues. Every Swedish vowel, both long and short, was represented twice after each initial consonant in the total consonant material forming six lists of 45 or 46 words. All CVC combinations in the test followed the rules of the Swedish language. They were chosen not to give real Swedish words. However, to meet the rules mentioned above for the composition of the test stimulus and response alternatives some real Swedish words (7%) had to be allowed. The words thus allowed were very neutral and not very common words. According to an earlier test of the speech material, they did not attract any extra responses by being words.

The vowel lists tested all the Swedish long and short vowels. Examples of the test vowels are shown in Table 1. There were two lists of 45 words testing the nine long vowels and two lists of 40 words testing the eight short vowels. (Short E and Ä become the same sound Ä in most Swedish dialects.) When long vowels were tested the response alternatives consisted of all the long vowels, and correspondingly they consisted of the short vowels at tests of short vowels. Consonants and combinations of consonants, possible in initial position in Swedish words, were divided into groups, e.g. voiced and voiceless clusiles, and fricatives. Each group was represented twice in front of each vowel. With that exception, grouping the initial consonants to keep down the number of lists, the same rules were applied as for the consonant lists. No real words were needed.

The speech material had been designed to be used in a computer controlled test procedure. To make the response alternatives familiar to the subject, they were shown on a monitor in front of the subject before the test stimulus was heard. The response alternatives were presented in random order to avoid systematic errors. They were numbered. Thus the subject responded by pressing the key with the number corresponding to what he perceived. The test stimulus was presented three times in immediate succession instead of using a carrier phrase. The program included a feedback option to be used for training: after a wrong response the text "The correct response was ..." was displayed on the monitor and the test stimulus was repeated, after a correct response the text "You were right" was displayed.

Nonlinear distortion

The sound stimuli were stored in the computer, where also the nonlinear distortion was generated and added to the undistorted signal. The systems thus simulated were $x + ax^2$ and $x + ax^3$, where x was the original signal and $a>0$. The coefficient, a, was chosen individually for each stimulus to obtain the same distortion level for all the stimuli. Otherwise for a fixed coefficient a, the distortion level would vary according to the individual levels of the original sounds. For the vowel test the whole test word was distorted. For the consonant test the consonant was cut out by means of a sound editing program. The distortion was generated for the cut out consonant part and added to the test word. Distorting the whole word would have caused a much higher distortion level on the following, stronger vowel, which might have given rise to masking effects but also to technical problems generating enough distortion on weak consonants. The cut was made at the level minimum before the onset of the vowel or slightly after that. It was checked that the consonant could not be identified by transitions in the undistorted, final part of the stimulus word. For voiceless fricatives (and H) the spectral transition is much less important than the relatively long unvoiced interval at the onset, see Furui (1986). Revoile et al (1987) show that for hearing-impaired listeners the perception of the voiceless consonants P, T, K relies mainly on the burst, not on the transition part. This should mean that the exact time of the cutting (on the transition part), i.e. where the distortion ends, is not critical.

At the maximum of the original signal (maximum RMS value, integration time 35 ms) the distortion levels were set to 25% and 50% for quadratic distortion and to 10% and 20% for cubic distortion. The relation between the quadratic and cubic distortion levels was chosen to make the distortions about equally detectable according to the results reported in TA105 and TA87. The increase in total sound pressure level, +2 dB for the long vowel lists at 20% cubic distortion, was compensated for the list as a whole. For all the other conditions, i.e. combinations of distortion level and type of distortion, and types of sound the increase in total sound pressure level was less than 1 dB and not compensated for. The distortion levels as such are not important for this investigation. The point is that they were the same for all the test sounds and that they were high enough to show effects.

Test equipment

The test stimuli were fed from a Prime 400 computer via a 12 bit D/A-converter and an anti-aliasing filter to a psychoacoustical test station at a sound insulated cell. The sampling frequency was 30 kHz and the bandwidth 10 kHz, but the signal level was very low above 7 kHz at the undistorted condition. The test station included amplifiers, attenuators and instruments to set the listening level. The subject listened monaurally with his normally aided ear through TDH-39 headphones with MX-41/AR cushions. He was seated at a monitor and a keyboard in the cell. The operator controlled the test from a terminal at the test station. The maximum allowed speech level in the headphones was 120 dB SPL (max RMS, 35 ms).

Test procedure

The test was performed individually in three or four sessions, each of them one and a half to two hours long. All subjects started with a few particular training lists to familiarize with the test routine and the simple quasiphonetic spelling. The subject's most comfortable listening level was also established during the training lists and, equipment permitting, verified to be about 10 dB below an uncomfortable loudness level. The tests were performed at the subject's most comfortable level.

The test design was intended to be counterbalanced as to list number, order of distortion level and type of distortion.

Before the test schedule was fixed two subjects were tested for learning effects. Six lists each with feedback and with the same type of distortion were tested at the higher distortion level. No learning effects were observed after the first two lists.

An example of one subject's test schedule is shown in Table 2. With a few exceptions for the most severely hearing-impaired the consonant tests were run first. Either quadratic or cubic distortion was tested in the first session, the other type of distortion in the second session. The distortion block was either preceeded or succeeded (or both) by an undistorted list. (Totally three undistorted lists were presented.) A distortion block started with a training list at one of the distortion levels and with feedback. A test list at the same distortion level and without feedback followed immediately. Next a list at the other distortion level was presented. All the training lists had feedback, the test lists had not. In the middle of each session there was a coffee-break.

In the vowel test each subject listened to one undistorted list with short vowels and one with long vowels. They tested one level of distortion on short vowels and the other level of the same type of distortion on long vowels. This arrangement was made to avoid repeating the limited number of lists too many times getting learning effects.

Six of the subjects later on participated in another experiment at the department. That experiment included frequency selectivity measurements, Psychoacoustical Tuning Curves, i.e. measurements of the noise level required to mask a pure tone at 10 dB SL with the noise centered at a lower or a higher frequency or at the same frequency as the tone. The measurements were performed at three frequencies: with a 1.0 kHz tone masked by 50 Hz noise bands at 0.625, 1.0 or 1.25 kHz; with a 2.0 kHz tone masked by 1.25, 2.0 or 2.5 kHz; and with a 4.0 kHz tone masked by 2.5, 4.0 or 5 kHz. For each of the tone frequencies the difference between the required masking level at the lower (and the higher) noise band frequency and the required masking level at the tone frequency was calculated and used in the analyses. Thus one measure of the frequency selectivity at upward spread of masking, PTCL, and one at downward spread of masking, PTCH, were established at each frequency.

Seven of the subjects had participated in the investigation of detection thresholds for nonlinear distortion, report TA105.

EFFECTS OF NONLINEAR DISTORTION ON SPEECH SOUND SPECTRA

Sending two pure tones with the frequencies f1 and f2 through a quadratic nonlinearity creates harmonic distortion at the frequencies 2f1 and 2f2, and intermodulation distortion at f2-f1 and f1+f2 (see Appendix). For a cubic nonlinearity the harmonic distortion falls at 3f1 and 3f2 and the intermodulation distortion at 2f1-f2, 2f2-f1, 2f1+f2 and 2f2+f1.

A narrowband noise is a reasonably simple idealization of a voiceless consonant. Quadratic distortion on such a noise gives rise to a noise band higher in frequency, between the double frequencies of the original bandlimits, and a low frequency band with the original bandwith. Also a DC component is created which is presumed to be filtered out. Cubic distortion gives one noise band higher in frequency, between three times the original bandlimiting frequencies, and one noise band between the difference frequencies mentioned above and overlapping the original band. For both kinds of distortion the noise band at higher frequencies might fall outside the upper bandlimit of the system, e.g. the hearing aid. Naturally this is more likely for cubic distortion and for the distortion of original sounds with fairly high frequencies. The distortion falling at the original frequencies at cubic distortion can increase the sound energy at these frequencies at a positive distortion coefficient.

For a vowel the formant pattern is essential for the discrimination. Fant, 1983, concludes that the most important parameters for the perception of vowels are the first formant, F1, and a second formant equivalent, $F2' = F2 + 0.5*(F3-F2)*(F2-F1)/(F3-F1)$. He also adopts the theory of Chistovich & Lublinskaja, 1979, that two formants placed closer than 3.5 Bark may be substituted for a single formant at some weighted mean frequency. If, on the contrary, the two formants are more than 3.5 Bark apart subjects normally match a single formant to one of the formants. This is the case for a large range of amplitudes. Their theory of the frequency processing mechanism should definitely be taken into account when considering the vowel intelligibility at nonlinear distortion: the first stage is a peripheral mechanism capable of finding formants for a large range of relative intensities. At a higher level there is some kind of integration and comparison process which gives the 3.5 Bark critical measure and a center of gravity measure, F2′ or, correspondingly, F1′.

Strong quadratic distortion creates a low-frequency bump caused by difference frequencies but also extra formants at the double and sum frequencies of the original formants. The extra formant at the double frequency of the original first formant can be expected to be very destructive. It is relatively strong because the first formant is considerably stronger than the higher formants for most vowels, and for many vowels it falls within 3.5 Bark from the original first or second formant. For hearing-impaired people with broader than normal auditory filters this might be even more disastrous. In the Appendix an I-sound with and without distortion is shown. The spectrum of the distorted I resembles very much the spectrum of an Ä-sound, which is also shown for comparison.

Some of the vowels, originals and distorted ones, were fed into a wave digital filter realization of a one-dimensional electrical analog model of the peripheral hearing. The results, where for example the distorted I replicates the undistorted Ä-sound, justified the discussions in terms of spectra. Further work will be done feeding distorted sounds into the model. To keep the amount of work on a reasonable level the discussions are based on checked-up general spectra for the phonemes and from these simply derived distorted spectra. Using distance measures based on frequency spectrum differences would have required analyses of all the test sounds and the results would still not indicate what frequency areas or distortion components are most important. Besides, Nord (1980) testing various distance measures did not succeed in achieving a single physical distance measure that is independent of formant number and vowel spectrum. Thus the extra work would not be worth-while at this stage.

It might be superfluous to mention that the nonlinear distortion works on the spectrum only. The temporal relations of the speech signal are not affected by the distortion.

RESULTS AND DISCUSSION

The differences between speech intelligibility scores at distorted and undistorted conditions (distorted minus undistorted) were analysed and means and standard deviations were calculated, Table 3. Table 4 shows the results of the undistorted lists for reference.

Consonants

For consonants there were no significant differences between the results at any of the distorted conditions and the undistorted condition as regards the whole group. However, there was a correlation, but significant only at $p<0.1$, between audiogram slope and positive difference scores at the highest distortion levels. All the subjects with relatively flat audiogram slopes had negative or zero difference scores for these distortion levels. Among the subjects with more sloping audiograms, however, there were subjects with quite large negative difference scores but also subjects with quite large positive difference scores. This discrepancy in performance made it reasonable to divide the subjects into two groups: one flat group with audiograms sloping less than 15 dB getting purely negative results with distortion and one sloping group to be analyzed further.

For the flat group there were significant decreases in intelligibility scores for voiceless consonants at the higher distortion levels (Table 3). There were also significant differences between these results of the flat group and the corresponding results of the sloping group.

For the sloping group the results with the high level of distortion did not differ significantly in any direction from the results at the undistorted conditions, which is not astonishing regarding the large spread of positive and negative results for individuals, +12% - -11%. Neither was there any correlation between slope and difference scores within the sloping group. Other factors than audiometric ones must have determined the positive or negative effects of the distortion, if not random. Since the difference between one subject's two test results obtained with 50-word-lists has to exceed 14 - 16% to be significant even at 10% significance level (Hagerman, 1976), it is almost impossible to judge from the individual percentages which subjects had true benefit from the distortion. In this case where the difference is taken between the result of one test list with distortion and the mean of three test lists without distortion the significant discrimination difference is multiplied by a factor $\sqrt{4/3} \times 1/\sqrt{2} \times \sqrt{50/45} = 0.86$. ($s^2 diff = s^2/3 + s^2$, instead of $s^2 diff = s^2 + s^2$, and 45 words instead of 50.) This gives a significant discrimination difference of 12 - 14% for 10% significance level. (The subject's standard deviation for the three consonant lists, 8% as shown in Table 4, is in reasonable agreement with Hagerman's results.)

To judge if there might be any transposition effects causing the positive difference scores for some of the subjects with sloping audiograms confusion matrices were computed. The matrices were computed for individuals and for groups. The sloping group was then divided into one group with positive difference scores (sloping plusgroup) and one with negative difference scores (sloping minusgroup). These group divisions were made differently for the two types of distortion. Naturally a subject placed in a minusgroup may have positive difference scores for individual phonemes and even obvious transposition effects although they are masked in his total difference score by larger negative effects on other phonemes. The group division was made to systemize the data treatment.

The group results of the confusion matrices for the various consonants for the highest level of quadratic distortion are put together in Table 5. The corresponding results for cubic distortion are shown in Table 6. To facilitate the discussion of the figures and to make it possible to relate them to spectra, a schematical representation of the consonant spectra is shown in Figure 2.

At quadratic distortion the consonants S, SJ, TJ, with broad original spectra at high frequencies and no lower components, gained considerably in intelligibility (17 - 34%) for the sloping plusgroup (the subjects with positive difference scores for the highest distortion level). Obviously the original spectra of these consonants were not clearly audible for these subjects, but the low frequency difference spectra were. Also the sloping minusgroup seems to have used transposition effects for the TJ-sound. They have gained 32% which is about the same as for the sloping plusgroup. The reason might be that TJ is the sound with the highest lowfrequency limit. The F-sound does hardly suffer from the quadratic distortion and for the sloping plusgroup there is some gain. This might be explained by the distinct and widely separated original noise bands causing distinct difference frequency bands. The distinct narrow band sound P seems relatively unaffected. The remaining consonants T, K and H are broad noise bands whose difference frequency components are likely to compete with the lowest components of the original sound. This is clearly demonstrated by the H-sound, having a broad original band 500 - 1200 Hz, losing about 20% for all the three groups. Note that the difference scores as a whole for the flat group and the sloping minusgroup are much more alike than those of the two sloping subgroups. The only exception is the TJ-sound as mentioned above.

Reading Table 5 a rule of thumb can be used for which absolute difference scores are significant: around 17% should be required for at worst 10% significance level with the quite large standard deviations in question. Actually this corresponds to the 17.6% change that would be needed for a mean change of one response to the five or six presentations per phoneme in each test list. If applying Hagerman's reasoning about significant difference scores for the individual on each phoneme, correspondingly 34 - 39% should be needed. According to this criterion all the subjects in the sloping plusgroup have significant difference scores for one or more phonemes. For the group the differences are significant for SJ-, TJ- and F-sounds and also for the H-sound (negative effect).

Regression analyses for the six subjects, whose frequency selectivity was tested, show positive correlations between absolute scores at all conditions and PTCLs and PTCHs at all the tested frequencies. Some of these correlations were significant, $p<0.1$ (correlation>0.73). Similar regressions on difference scores show positive correlations to the 1000 Hz and 2000 Hz PTCLs and PTCHs, but the correlations to the 4000 Hz PTC-values have consistently changed, for both upward and downward spread of masking and for all conditions of distortion on consonants. They have all become negative. This means that good frequency selectivity gives the best absolute intelligibility scores, but those with bad frequency selectivity for high frequencies can get more positive effects from the distortion. Among those tested for PTCs four belonged to the sloping group. They all form illustrative examples. One of them had generally good PTCs. For him quadratic distortion only made it worse. Another one had a moderately bad PTC at 4000 Hz and some better PTC-values at lower frequencies. He had positive diffence scores and a confusion matrix at high quadratic distortion that indicates use of transposition effects. The third one had a really bad PTC at 4000 Hz and moderately bad PTCs at lower frequencies. She obviously used transposition effects at the high level of quadratic distortion. Finally the fourth one had a very bad PTC at 4000Hz, but good PTCs at all lower frequencies. He could use transposition effects also at the lower level of quadratic distortion and gained 19% with that distortion level. His confusion matrices are shown in Table 7 as an example. The slope of his audiogram is 29 dB and his four-frequency-mean 57 dB, both parameters relatively average in the group. The results of these four subjects demonstrate that a hearing aid transposing high frequency sounds to lower frequencies where the hearing aid user has better frequency selectivity can be advantageous to a hearing aid with more high frequency gain. - Just hearing a speech sound is not enough. It has to be identified as well.

The effects of cubic distortion according to Table 6 are more diffuse. There are some significant increases, K and TJ, for the sloping plusgroup, and one significant decrease, T. (It should be noted that the sloping plusgroup here consists of those with positive difference scores for both levels of cubic distortion. It is a subgroup of the plusgroup at quadratic distortion, since all the subjects with positive difference scores at cubic distortion also had positive difference scores at quadratic distortion.) There are also some significant difference scores for individual subjects and phonemes according to the limits derived from Hagerman (1976), but not as many as for quadratic distortion.

For these composite phoneme spectra the distortion patterns from the cubic distortion are even more complicated than for the quadratic distortion. It becomes more difficult for the ear to discriminate but it also makes it more difficult for us to sort out the logical connections to know what effects to expect. For the TJ-sound, however, it is obvious that the information moving to lower frequencies is valuable for all subjects with sloping audiograms in the same way as at quadratic distortion. The changes are significant for all the three groups: increases for the sloping groups, +19% and +29%, and a decrease for the flat group, -18%.

A test of correlations between the TJ-sound difference scores and the frequency selectivity at 4000 Hz (n=6) reveals absolute values of the correlations between 0.70 (negative) for the mean PTC at 4000 Hz and 0.79 for the difference between the 4000Hz PTC mean and the PTC mean for the lower tested frequencies. This correlation is significant, $p<0.1$, for $n=6$. Again we have found evidence that moving the spectrum from a frequency area with unfavourable frequency selectivity to a better one might be a good idea.

Summary: The general effect of nonlinear distortion on consonants is negative. Positive effects for individual consonants may occur when the distortion products fall in frequency areas where the listener has got better frequency selectivity than for the original sound, provided the perceived sound does not become too complex. The total positive or negative effect then depends on how many phonemes that gain or lose and to what extent. The explanation for the consistently negative effect for subjects with flat audiograms might be that they perceive more of the distortion products, which makes the sound more complex, but also their generally less frequency dependent frequency selectivity. However, it should not be forgotten, that the subjects with flat audiograms often have better absolute scores for the original consonants than the subjects with sloping audiograms. The increased absolute scores for phonemes with good transposition effects for subjects with sloping audiograms seldom exceed those of the original phonemes for the subjects with flat audiograms. There is a loss of resolution that can not be compensated for.

Vowels

In contrast to the consonants the vowels suffer significantly at the higher distortion levels also when all the subjects are regarded as one group. The difference score for 50% quadratic distortion is -18% and for 20% cubic distortion -6%, see Table 3. For quadratic distortion, however, the table indicates a large, significant difference between the flat and the sloping groups. As was the case for the consonants the difference scores are worse for the flat group, -30% versus -11% for the sloping group. For the higher level of cubic distortion the negative difference score for the flat group is not significant at 10% significance level but the slightly worse difference score for the sloping group is. The division into flat and sloping groups was made after the test. Since each subject was only tested with either long or short vowels at a specific distortion condition, the group results unfortunately got various proportions of long and short vowels. If long and short vowels had suffered equally from the distortion, this would not have mattered. But the short vowels generally suffer more, as will be shown later. At the highest degree of quadratic distortion there is a majority of lists with short vowels for the flat group making the result more negative and also enhancing the difference to the sloping group. For the highest degree of cubic distortion the opposite has happened. There is a majority of lists with short vowels for the sloping group, which might have made the difference score for the sloping group relatively too negative. As a whole each condition shows the most negative result for the group with the largest proportion of lists with short vowels. Therefore there is reason to believe that there is no essential difference between the groups except probably at the highest degree of quadratic distortion. In the following the data are analyzed for all the subjects as one group regardless of audiogram shape.

In Table 8 and Table 9 the scores for short and long vowels respectively are shown for the whole group. The score at each distorted condition is compared to the score at the undistorted condition for the same subgroup. Showing the undistorted score also for the whole group might give an idea of how representative a particular subgroup might be. The difference scores exceeding 10% significance level for the subgroup are marked. Applying Hagerman's calculation methods, the limit for that significance level ranges from 14% to 21% for word lists of 50 to 30 words.

Generally the short vowels suffer more from the distortion than the long ones as can be seen by comparing the difference scores for the long and short vowels at the same condition. The difference scores for the short vowels are mostly worse. The difference is significant, $p<0.01$. There are two possible reasons for that general result. One is that the formant transitions at the beginning of the vowels might get so seriously damaged by the distortion products that more of the identification relies on the stationary part of the vowel than normally. The other reason might be that the distortion products influence the perception of the formants in a more favourable way for long vowels. The long vowels have mostly lower first formants than the corresponding short vowels which for example makes the harmonic distortion generated by the first formants fall at lower frequencies. A combination of both reasons seems likely. Since long and short vowels were tested separately with the response alternatives consisting of either long or short vowels, no conclusions can be drawn regarding confusions between long and short vowels.

Quadratic distortion applied on short vowels gives the conditions where the effects of the distortion are most obvious. The difference scores of I, Y, U, and O are -58% to -38% at the highest distortion level, see Table 8. As can be seen in Figure 3 these are the vowels with the lowest first formants. The double frequencies of these formants all fall within 3.5 Bark from the original frequencies. Thus when the distortion products at a high distortion level create an extra formant, this extra formant and the original formant together are perceived as a new formant with a higher frequency than the original one according to the center of gravity theory. This makes I and Y resemble Ä, which is also confirmed by the confusion matrices. Likewise U and O become Å. However, for these vowels the confusions might be caused by a combined effect of lowered first and second formants. (Actually for O already the original first and second formants fall within 3.5 Bark. Now they are joined by extra formants in between.)

At quadratic distortion and long vowels the effects are less marked than for short vowels but are still the same. U: and E:, having low first formants suffer significantly, -17% and -23% respectively.

For cubic distortion and short vowels the negative effects of the distortion are still strongest for the same short vowels as at quadratic distortion, although the difference scores are less negative. U, I and Y get significantly negative difference scores, -16%, -16% and -24%, again indicating that the low frequency of the first formant creates the problem. However, three times the first formant frequency never falls within 3.5 Bark from the first formant. Therefore there is never any fusion with the first formant. Instead it changes the perception of the second and higher formants. For some vowels the changes in the confusion patterns indicate that the difference frequency intermodulation products from the original first and second formants are responsible: U becomes O to a much larger extent than before because of an impression of a lower first formant, A sometimes becomes Å, Ä or Ö probably because the intermodulation falls around the first formant frequency of these vowels.

At cubic distortion almost all the long vowels get positive difference scores at the lower distortion level although only two vowels, Å: and E:, get significant increases, 16% and 14% respectively. At the higher distortion level the significant increase in difference score remains only for Å: . There are two more vowels, Ä: and A:, that keep their positive scores. Coincidence or not these are the only three vowels where the intermodulation from difference frequencies from the first and second formant fall at lower frequencies than the original first formant. Probably it was possible for the subjects to learn these "new long vowels" and to some extent substitute the old references for new ones. What gives some strength to that assumption is that the vowel Å: which has the highest, and significant, difference score gets the formant pattern that is most different from that of other vowels, both original and distorted ones. Its first formant falls at a lower frequency than any other formants. Generally in this test it is difficult to know to what extent the subjects rely on their old references and to what extent they have been able to establish and use new ones . Long training periods might have given the answer (although not possible because of the limited speech material).

Summary: Nonlinear distortion has a disastrous effect on short vowels especially on those with first formants at low frequencies. This is true at quadratic distortion on long vowels too. At cubic distortion on long vowels some vowels can actually gain somewhat in intelligibility depending on difference frequency intermodulation falling at perceptually favorable frequencies. No conclusions on effects of audiogram shape can be drawn from this investigation.

Agreement with other investigations

This investigation only partly overlaps the investigations mentioned in the introduction, but when it does the results are not incompatible. Günthersen's theory of fusion of the signal and the distortion products to form a new percept for complex nonstationary signals seems as probable here as at the detection task. Note the positive effects of transposition of consonants at high frequencies and of the new low frequency formants of some long vowels. - The correlation was very low between results in my detection threshold test, report TA105, and speech intelligibility scores in this test for the seven subjects who participated in both investigations. This is not astonishing as intelligibility relies on discrimination and identification abilities, but not necessarily on detection of the distortion.

Gordon-Salant (1985) made a test of recognition of CV syllables in multitalker babble. In contrast to what happens in white noise, the mid-place consonants suffered more than the front- and back-place consonants in multitalker babble. It is most likely that the slow amplitude modulations of the babble and masking effects from the vowels make the difference. In my test, however, the two mid-place consonants also performed worse than expected: T was a general loser, S might have had higher scores considering possibilities of positive transposing effects. It might be far-fetched but not totally improbable that the similar results are due to the same mechanism - consonant or distortion components, thus with the time patterns of speech, filling in the gaps in the spectrum in a critical way. - The earlier results on speech intelligibility at our department are confirmed as to transposition effects and are built out further. - The results of the studies by Jirsa & Norris and by Singer, testing effects of intermodulation or intermodulationlike distortion only, were interpreted as masking effects from below 1 kHz and on the second formant. My investigation, using both harmonic and intermodulation distortion with their natural level relations for a flat reproduction system, confirms the importance of the low frequency area especially at low first and second formants, but also introduces the idea of fusions with new formants created by harmonic distortion or intermodulation. - The effect of noise on the signal at the input of the system, before the nonlinearity, remains to be investigated.

CONCLUSIONS

Nonlinear distortion on speech signals introduces harmonic distortion and intermodulation, sometimes with so complicated patterns, and so spread out in frequency, that its influence on speech intelligibility is worse than difficult to predict or explain. For some phonemes and distortions, however, the distortion products form more distinctive patterns. Depending on their frequency localization both absolutely and in relation to the original phoneme, and also depending on the properties of the listener's hearing loss, the intelligibility becomes either reduced or improved.

In my investigation, testing vowels or voiceless initial consonants, with a flat frequency curve and a high signal-to-noise ratio,

- the lower distortion level used did not give any significant changes in total consonant or vowel intelligibility,

- the higher distortion levels at both quadratic and cubic distortion caused significant decreases in consonant intelligibility for subjects with flat audiograms, but not for subjects with sloping audiograms of which some could use the low frequency distortion positively,

- vowels suffered more than consonants and short vowels more than long ones,

- the higher distortion levels at both quadratic and cubic distortion caused significant decreases in vowel intelligibility,

- confusion matrices combined with formant and distortion patterns for the vowels makes it evident that the frequency area below approximately 1 kHz is critical especially for vowels with first and/or second formants at low frequencies,

- not only certain consonants but also some vowels can gain from the distortion at favourable conditions,

- the results indicate how to proceed the work towards a specification of maximum tolerated nonlinear distortion in hearing aids.

ACKNOWLEDGEMENTS

I would like to express my gratitude for valuable discussions to Björn Hagerman, Bertil Johansson and Åke Olofsson. My thanks also to the helpful subjects.

This work was supported by the National Swedish Board for Technical Development and the Ministry of Health and Social Affairs.

REFERENCES

Chistovich, L.A. & Lublinskaya, V.V. 1979. The "center of gravity" effect in vowel spectra and distance between the formants: Psychoacoustical study of the perception of vowel-like stimuli". Hearing Research 1, 185-195.

Fant, G. 1983. Feature analysis of Swedish vowels - a revisit. Dep. of Speech Comm., Royal Inst. of Technology, Stockholm, STL-QPSR 2-3, 1-19.

Furui, S. 1986. On the role of spectral transition for speech perception. J Acoust Soc Am 80(4), 1016-1025.

Gabrielsson, A., Nyberg, P-O., Sjögren, H. & Svensson L. 1976. Detection of amplitude distortion by normal hearing and hearing impaired subjects. Dep. of Technical Audiology, Karolinska institutet, report TA83.

Gabrielsson, A., Lindblad, A-C., Lindström, B. & Olofsson, Å. 1978. Detection of nonlinear distortion in telephone systems. Dep. of Technical Audiology, Karolinska institutet, report TA87.

Gordon-Salant, S. 1985. Recognition of digitized CV syllables in multitalker babble. Audiology 24, 241-253.

Günthersen, C. 1982. Perception of nonlinear distortion. The Acoustics Laboratory, Technical University of Denmark, report no. 34.

Hagerman, B. 1976. Reliability in the determination of speech discrimination. Scand Audiol 5, 219-228.

Ingelstam, R., Johansson, B., Pettersson, A. & Sjögren, H. 1970. The effect of non-linear amplitude distortion, an investigation by variation of the quadratic and the cubic components. Dep. of Technical Audiology, Karolinska institutet, Stockholm, report TA64.

Ingelstam, R., Johansson, B., Pettersson A. & Sjögren, H. 1971. Non-linear distortion in hearing aids. Dep. of Technical audiology, Karolinska Institutet, report TA64A.

Jirsa, R.E. & Norris, T.W. 1982. Effects of intermodulation distortion on speech intelligibility. Ear and Hearing 3(5), 251-256.

Lindblad, A-C. 1982. Detection of nonlinear distortion on speech signals. Dep. of Technical Audiology, Karolinska institutet, report TA105.

Nord, L. 1980. Analysis and prediction of difference limen data for formant frequencies. The testing of two distance measures. Working papers, Phonetics Lab, Dep. of General Linguistics, Lund University, no. 19. Papers given at the Swedish-Danish symposium in Lund April 25-26 1980, 49-53.

Olofsson, Å. 1978. A generalization of the concept of total harmonic distortion. Dep. of Technical Audiology, Karolinska institutet, report TA90.

Revoile, S., Picket, J.M., Holden-Pitt, L.D., Talkin, D. & Brandt, F.D. 1987. Burst and transition cues to voicing perception for spoken initial stops by impaired- and normal-hearing listeners. J Speech and Hearing Res 30, 3-12.

Singer, J. 1981. Some effects of intermodulation distortion on speech intelligibility. J Audit Res 21, 201-206.

IEC 118-0, 1983. Hearing aids. Part 0: Measurement of electroacoustical characteristics. 35-39, 49-50.

The Nordic Committe on Disability 1986. Requirement specifications of hearing aids.

The Swedish Board of Technical Development 1972, project STU 71-434/U335, slutrapport, bilaga 3.

Table 1. Test sounds.

Basic spelling in Swedish	Phoneme	Example in Swedish	Example of similar sound if possible in English
P	/p /	paj	pie
T	/t /	te	tea
K	/k /	katt	cat
S	/s /	sång	song
SJ	/ʃ /	sjal, skjuts skär, stjärna	shawl
TJ	/ç /	tjärn, kyrka	church - without the initial t-sound, German: nicht
F	/f /	fot	foot
H	/h /	huva	hood
A	/ɑ:/ /ɑ /	stav hatt	staff stuff
O	/u:/ /u /	bur burr	? ?
U	/ʉ:/ /ʉ /	bo bom	food foot
Å	/o:/ /o /	båt lott	nought not
E	/e:/ (short E-sound equals /ɛ / in most Swedish dialects)	sed	German: See
I	/i:/ /i /	liv ritt	leaf lid
Y	/y:/ /y /	nys nyss	German: über German: Hütte
Ä	/ɛ:/ /ɛ /	där, bänk bädd, bedd, ärr	there - the first part of the diphthong ɛə bed, man
Ö	/ø:/ /ø /	bör, bön dörr, kött	bird, French: jeu the

Table 2. Example of test schedule.

	list type	distortion	test/training	feedback
Session 1:				
	consonants	-	pretraining	yes
	consonants	-	pretraining	yes
	consonants	10% cubic	training	yes
	consonants	10% cubic	test	no
	consonants	20% cubic	test	no
	consonants	-	test	no
Session 2:				
	consonants	-	pretraining	yes
	consonants	-	test	no
	consonants	50% quadratic	training	yes
	consonants	50% quadratic	test	no
	consonants	25% quadratic	test	no
	consonants	-	test	no
Session 3:				
	long vowels	10% cubic	training	yes
	long vowels	10% cubic	test	no
	short vowels	20% cubic	test	no
	short vowels	-	test	no
(Session 4:)				
	(short vowels	-	training)	yes
	long vowels	-	test	no
	long vowels	50% quadratic	training	yes
	long vowels	50% quadratic	test	no
	short vowels	25% quadratic	test	no

Table 3. Means and (±) standard deviations of differences between speech intelligibility scores with and without nonlinear distortion. 16 subjects with sensorineural hearing losses tested with one list per condition at most comfortable level.
* indicates significant difference between distorted and undistorted condition ($p<0.05$).
>-< indicates significant difference between groups ($p<0.02$).

Difference in score, %, mean and sd		All the subjects, N=16	Flat group, slope <15 dB N= 6		Sloping group, slope >15 dB N=10
<u>Initial consonant</u>:					
Quadratic dist.,	25%	1 ± 7	1 ± 6		1 ± 8
	50%	2 ±11	-10* ± 9	>-<	3 ± 9
Cubic dist.,	10%	-3 ± 7	-5 ± 9		-1 ± 6
	20%	-3 ±11	-13* ± 6	>-<	3 ± 8
<u>Vowel</u>:					
Quadratic dist.,	25%	-3 ±10	-1 ± 8		-5 ±11
	50%	-18* ±15	-30* ±14	>-<	-11* ±10
Cubic dist.,	10%	1 ± 9	-1 ±10		2 ± 8
	20%	-6* ± 8	-5 ±10		-7* ± 8

Table 4. Absolute speech intelligibility scores at undistorted condition. 16 subjects with sensorineural hearing losses listening at most comfortable level.

Absolute score, %:		All the subjects, N=16	Flat group, slope <15 dB N= 6	Sloping group, slope >15 dB N=10
Initial consonant (3 lists/subject)	mean, sd	60 ±19	74 ±17	52 ±16
	range	17-93	37-93	17-84
	intraindividual sd	8	10	6
Long vowel (1 list/subject)	mean, sd	62 ±18	74 ±11	55 ±18
	range	20-95	35-95	20-72
Short vowel (1 list/subject)	mean, sd	57 ±22	73 ±11	48 ±16
	range	26-86	60-86	26-82
All vowels		60 ±20	73 ±16	51 ±17

Table 5. Initial consonants. Intelligibility scores at 50% quadratic distortion and at undistorted condition, and the difference scores (%).

	P	T	K	S	SJ	TJ	F	H
Flat group, N=6,								
distorted	56	52	71	53	63	65	85	68
undistorted	65	74	75	71	67	75	79	87
difference	-9	-22	-4	-18	-4	-10	+6	-19
Sloping minusgroup, N=4,								
distorted	33	46	75	45	46	82	50	68
undistorted	37	71	75	57	63	50	57	88
difference	-4	-25	0	-12	-17	+32	-7	-20
Sloping plusgroup, N=5 (1 dropout),								
distorted	39	46	54	63	86	50	59	57
undistorted	34	47	56	46	60	16	39	78
difference	+5	-1	-2	+17	+26	+34	+20	-21

Table 6. Initial consonants. Intelligibility scores at 20% cubic distortion and at undistorted condition, and the difference scores (%).

	P	T	K	S	SJ	TJ	F	H
Flat group, N=6,								
distorted	41	58	68	69	69	57	70	59
undistorted	65	74	75	71	67	75	79	87
difference	-24	-16	-7	-2	+2	-18	-9	-28
Sloping minusgroup, N=6,								
distorted	33	50	79	74	74	56	31	74
undistorted	37	60	73	65	66	37	56	88
difference	-4	-10	+6	+9	+8	+19	-25	-14
Sloping plusgroup, N=4,								
distorted	48	27	65	39	57	61	41	73
undistorted,	32	57	46	26	41	32	29	65
difference	+16	-30	+19	+13	+16	+29	+12	+8

Table 7. Confusion matrices of subject using transposition effects at quadratic distortion. Initial consonants.
^ indicates a positive difference score probably limited by a 100% score.

Undistorted condition. 3 lists.

RESPONSE \ STIMULUS	P	T	K	S	SJ	TJ	F	H	total
P	10	.	.	.	.	.	4	2	16
T	6	14	5	3	.	.	3	.	31
K	1	3	12	.	.	1	.	.	17
S	.	.	.	4	.	4	.	.	8
SJ	.	.	.	1	13	8	.	.	22
TJ	.	.	.	.	4	4	.	.	8
F	.	.	.	8	.	.	8	2	18
H	.	.	.	1	.	.	2	13	16
total	17	17	17	17	17	17	17	17	
Correct,%	59	82	71	24	76	24	47	76	57

25% quadratic distortion. 1 list.

RESPONSE \ STIMULUS	P	T	K	S	SJ	TJ	F	H	total
P	3	.	2	.	.	.	.	.	5
T	2	5	1	.	.	.	.	.	8
K	1	.	3	.	.	.	.	.	4
S	.	.	.	5	.	.	.	.	5
SJ	.	.	.	.	4	1	.	.	5
TJ	.	.	.	.	2	5	.	.	7
F	.	.	.	1	.	.	5	1	7
H	.	.	.	.	.	.	.	5	5
total	6	5	6	6	6	6	5	6	
correct,%	50	100	50	83	67	83	100	83	76
dist.-undist.	-9	+18^	-21	+59	-9	+59	+53^	+7	+19^

The responses for S, TJ and F, which are quite scattered at the undistorted condition, concentrate and most often become correct at both levels of quadratic distortion. This is true also for P at the higher distortion level. The difference scores are quite typical for the sloping plusgroup except for SJ where this subject gets no gain presumably due to his high undistorted score.

50% quadratic distortion. 1 list.

RESPONSE \ STIMULUS	P	T	K	S	SJ	TJ	F	H	total
P	5	1	1	.	.	.	1	.	8
T	.	4	1	.	.	.	.	.	5
K	.	1	3	.	.	.	.	.	4
S	.	.	.	5	.	3	.	.	8
SJ	.	.	.	.	4	.	.	.	4
TJ	.	.	.	.	1	3	.	.	4
F	.	.	.	1	1	.	5	3	10
H	.	.	.	.	.	.	.	2	2
total	5	6	5	6	6	6	6	5	
correct.%	100	67	60	83	67	50	83	40	69
dist.-odist.	+41^	-15	-11	+59	-9	+26	+36	-36	+12^

10% cubic distortion. 1 list.

RESPONSE \ STIMULUS	P	T	K	S	SJ	TJ	F	H	total
P	3	.	.	.	.	.	1	.	4
T	2	3	.	.	.	.	2	.	7
K	1	3	6	.	.	.	.	.	10
S	.	.	.	3	.	1	.	.	4
SJ	.	.	.	.	4	1	.	.	5
TJ	.	.	.	.	1	3	.	.	4
F	.	.	.	1	.	.	2	1	4
H	.	.	.	1	.	.	1	5	7
total	6	6	6	5	5	5	6	6	
correct,%	50	50	100	60	80	60	33	83	64
dist.-odist.	-9	-32	+29^	+36	+4	+36	-14	+7	+7^

Note that K gives 100% correct responses at both distortion levels. S and TJ also get considerably better at the lower level of cubic distortion.

20% cubic distortion. 1 list.

RESPONSE \ STIMULUS	P	T	K	S	SJ	TJ	F	H	total
P	4	.	.	.	.	.	1	.	5
T	1	3	.	1	.	.	.	.	5
K	1	2	6	.	.	.	2	.	11
S	.	.	.	2	.	2	1	.	5
SJ	.	.	.	.	4	1	.	.	5
TJ	.	.	.	.	2	3	.	.	5
F	.	.	.	2	.	.	1	.	3
H	.	.	.	1	.	.	.	6	7
total	6	5	6	6	6	6	5	6	
correct,%	67	60	100	33	67	50	20	100^	63
dist.-odist.	+8	-22	+29^	+9	-9	+26	-27	+24^	+6^

Table 8. Short vowels. Intelligibility scores at distorted and at undistorted condition, and the difference scores (%).
* indicates significant difference, p<0.1.
< indicates score at guessing level.

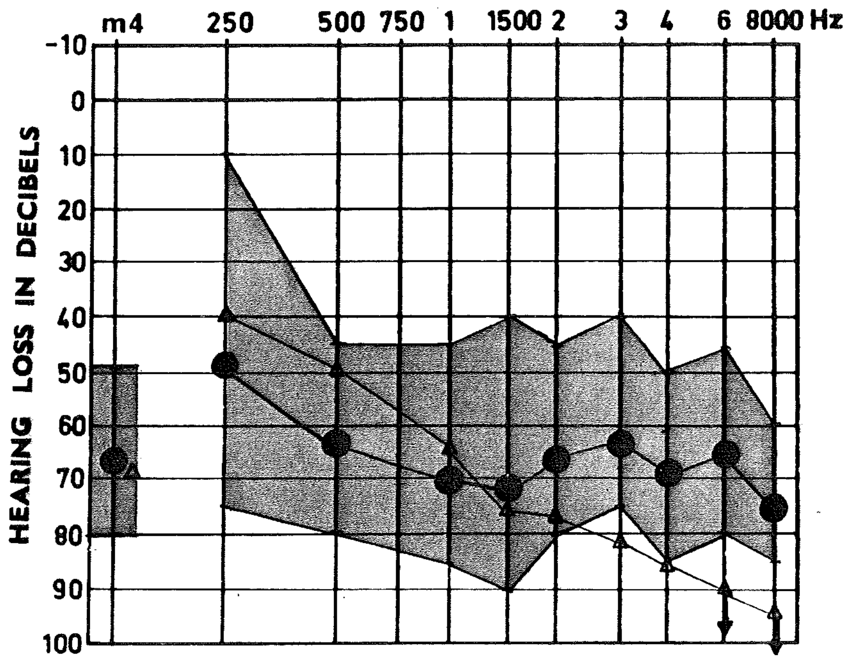
	A	O	U	Å	I	Y	Ä	Ö
Undistorted, all subjects, N=16	84	49	50	45	70	54	63	44
Quadratic dist., 25%, N=8	83	54	30	45	43	33	52	54
undist. same group	83	38	53	50	60	48	45	40
difference	0	+16	-23*	-5	-17*	-15	+7	+14
Quadratic dist., 50%, N=9	80	24	13<	38	22	22	69	40
undist. same group	87	63	51	44	80	64	80	54
difference	-7	-39*	-38*	-6	-58*	-42*	-11	-14
Cubic dist., 10%, N=6	73	54	30	57	60	30	60	47
undist. same group	73	54	47	53	63	57	57	47
difference	0	0	-17	+4	-3	-27*	+3	0
Cubic dist., 20%, N=10	84	38	36	48	58	28	58	46
undist. same group	90	45	52	40	74	52	66	43
difference	-6	-7	-16*	+8	-16*	-24*	-8	-3

Table 9. Long vowels. Intelligibility scores at distorted and at undistorted condition, and the difference scores (%).
* indicates significant difference, p<0.1.
< indicates score at guessing level.

	A:	O:	U:	Å:	E:	I:	Y:	Ä:	Ö:
Undistorted, all subjects, N=16	83	80	53	66	70	66	29	63	53
Quadratic dist., 25%, N=8	93	75	53	65	75	53	33	83	45
undist. same group	85	88	63	78	68	63	30	78	50
difference	+8	-13	-10	-13	+7	-10	+3	+5	-5
Quadratic dist., 50%, N=7	80	57	17	63	46	54	9<	31	51
undist. same group	77	69	34	54	69	66	20	46	51
difference	+3	-12	-17*	+9	-23*	-12	-11	-15	0
Cubic dist., 10%, N=10	92	82	42	84	86	60	44	66	52
undist. same group	88	80	52	68	72	70	32	62	50
difference	+4	+2	-10	+16*	+14*	-10	+12	+4	+2
Cubic dist., 20%, N=6	77	73	37	83	60	57	17	77	33
undist. same group	73	80	53	63	67	60	23	63	57
difference	+4	-7	-16	+20*	-7	-3	-6	+14	-24*

Figure 1. Mean audiograms for the flat group and the sloping group. Shaded areas show the ranges of audiograms for the two groups. The corresponding four-frequency-means are indicated to the left in each figure. Minisymbols show the means of the other group for comparison.

FLAT GROUP, N=6



SLOPING GROUP, N=10

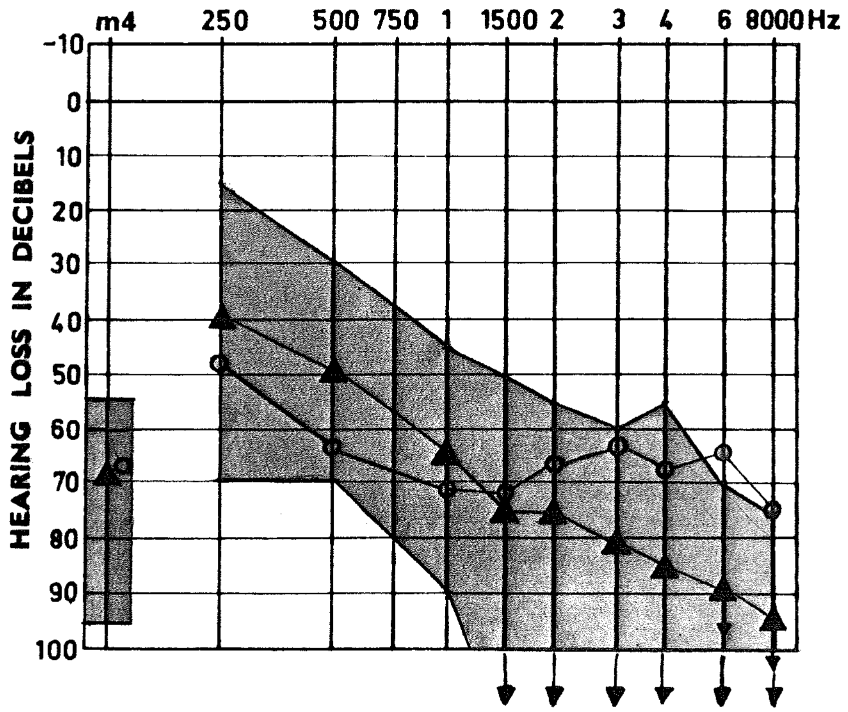
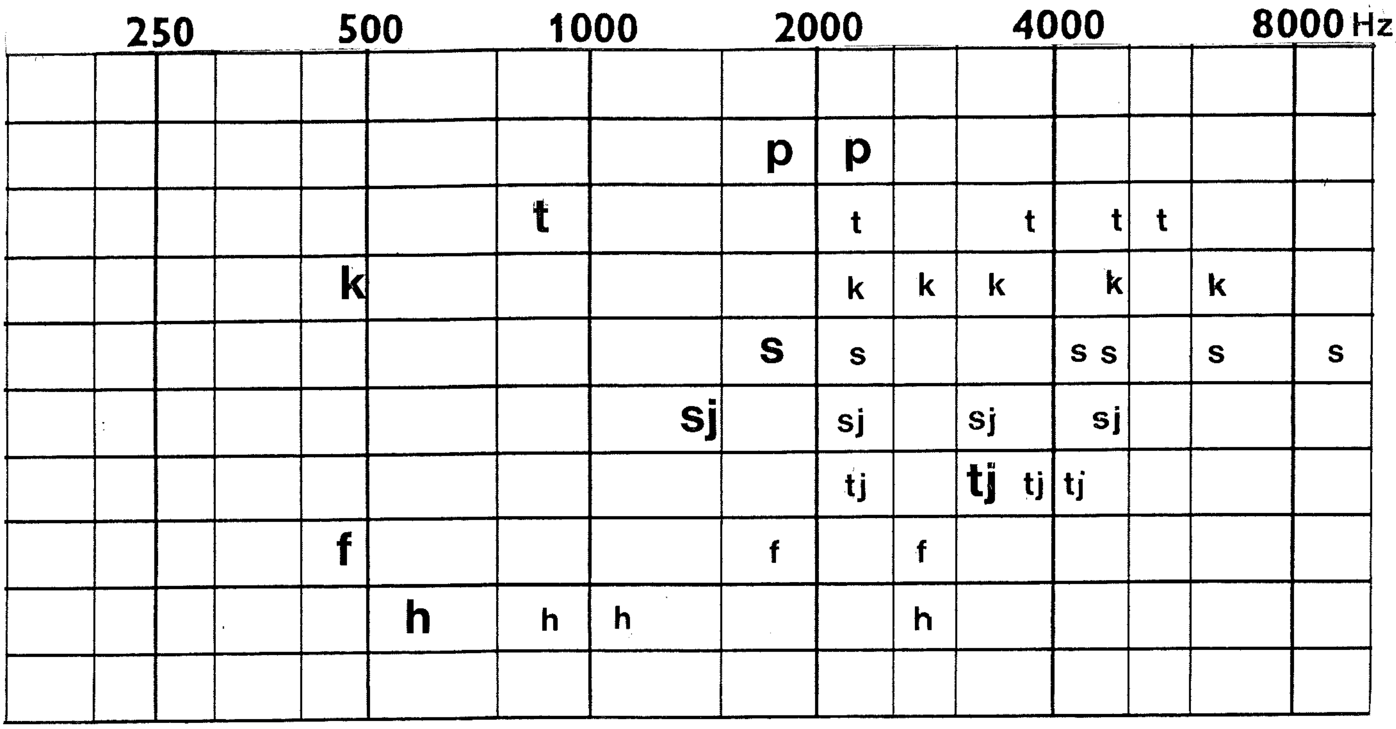


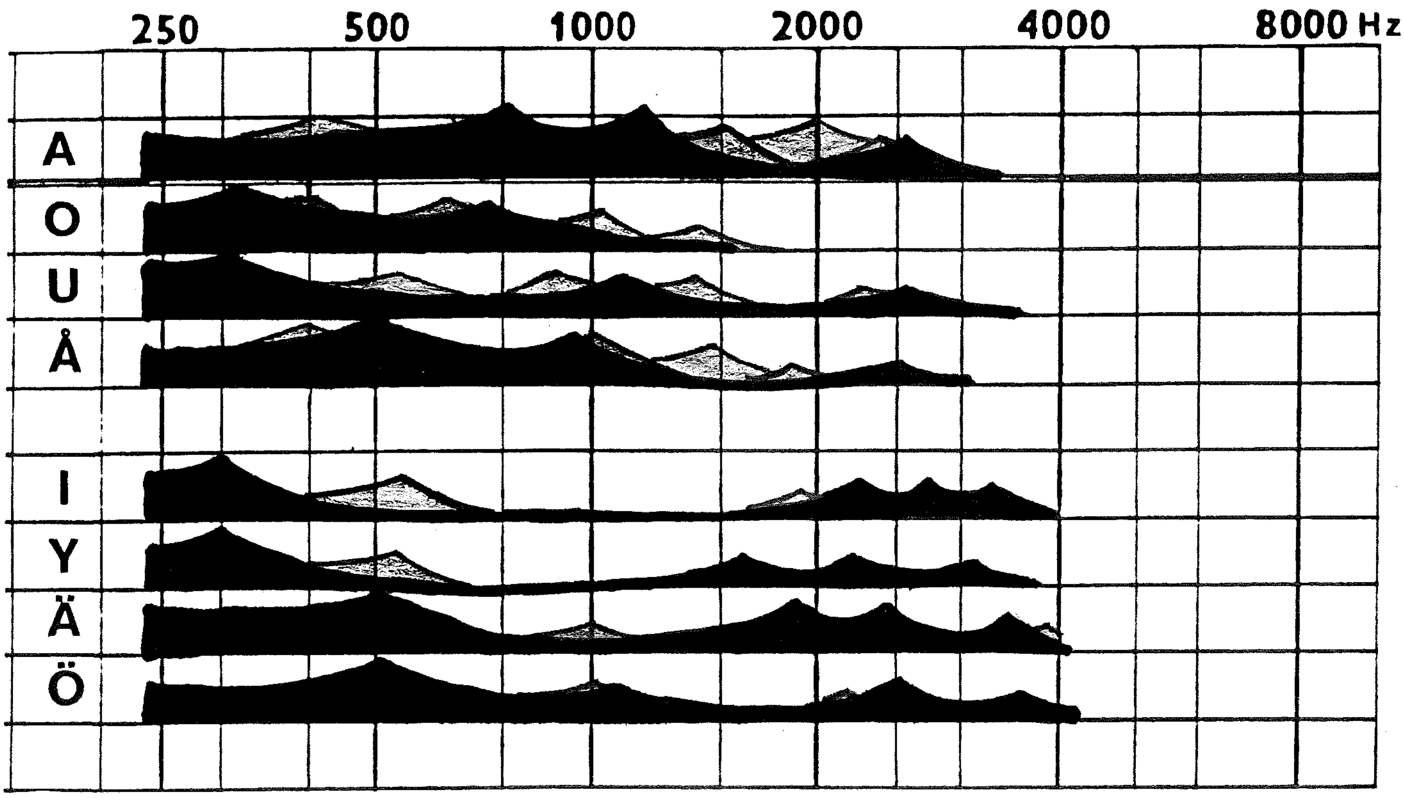
Figure 2. Schematical frequency representation of the Swedish consonants in the test.
The largest symbol represents the largest peak in a consonant spectrum.



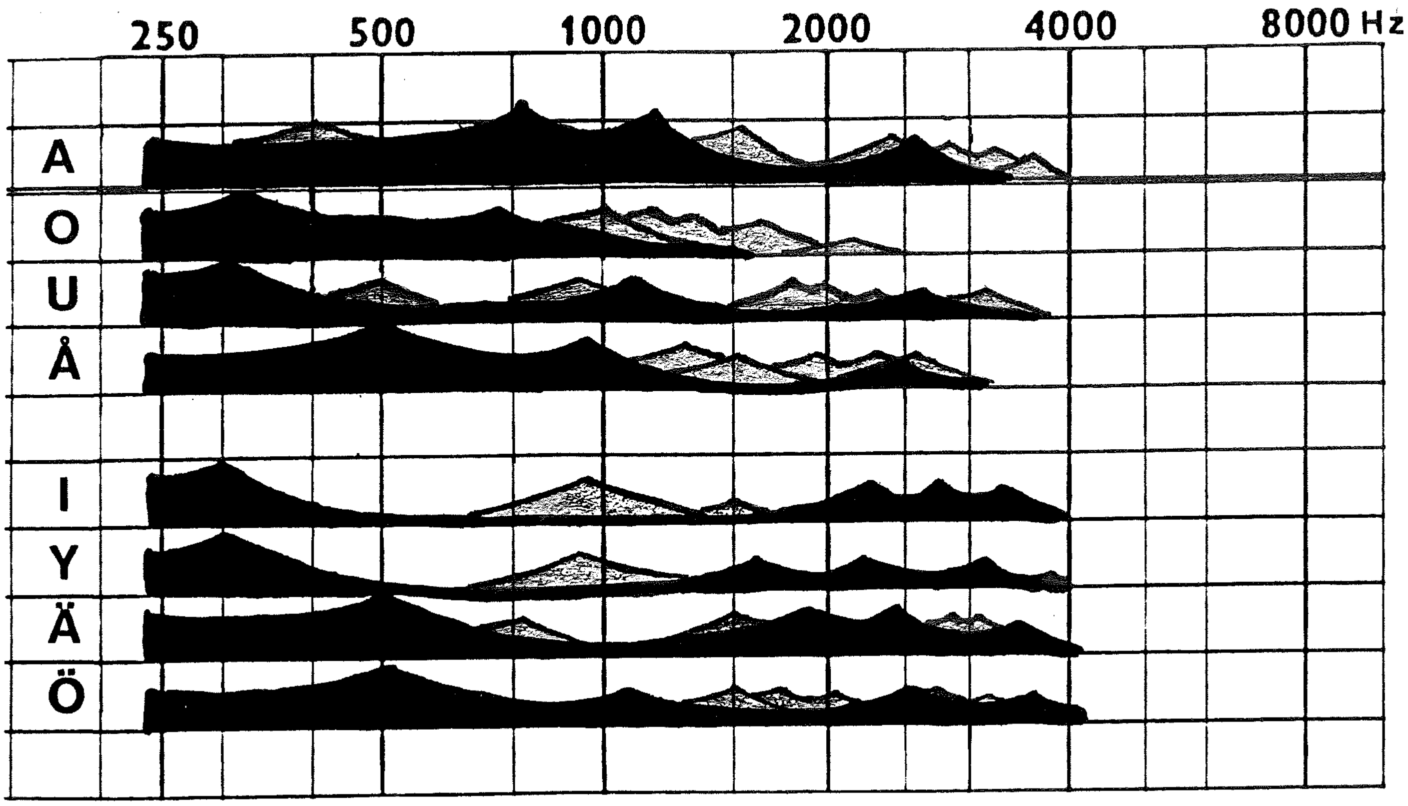
General data according to G. Fant.

Figure 3. Schematic formant patterns of Swedish short vowels. Distortion products from quadratic and cubic distortion added as shaded areas.

QUADRATIC DISTORTION



CUBIC DISTORTION



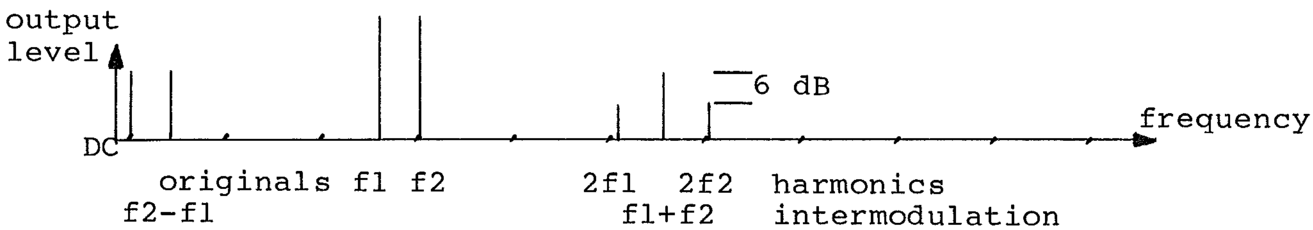
Basic data from G. Fant.

APPENDIX

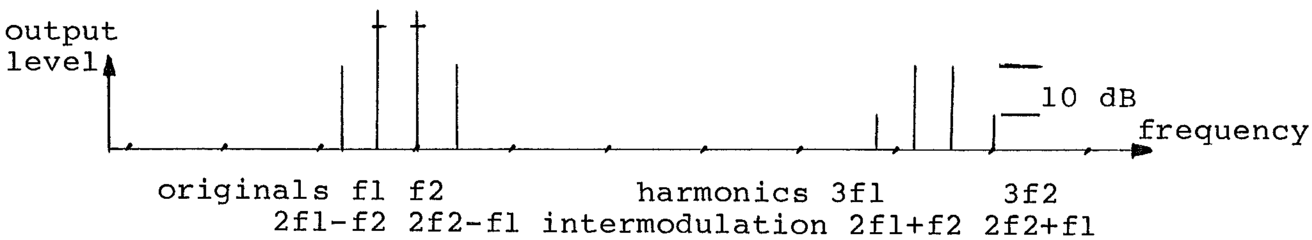
Effects of nonlinear distortion on sound spectra. Simple examples to show the basic principles.

Distortion on two pure tones at the frequencies f1 and f2:

Quadratic distortion

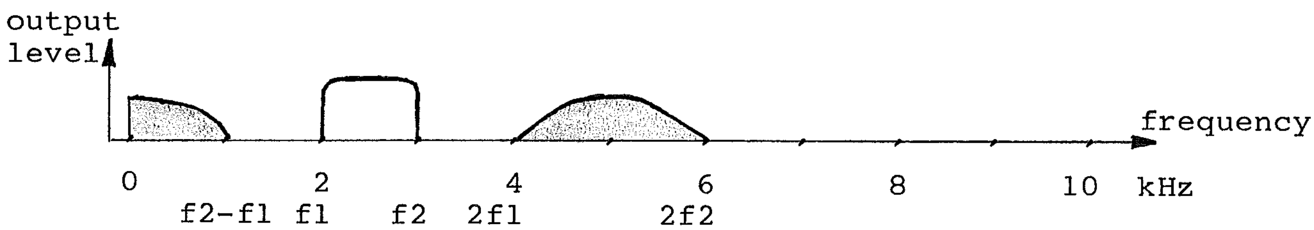


Cubic distortion

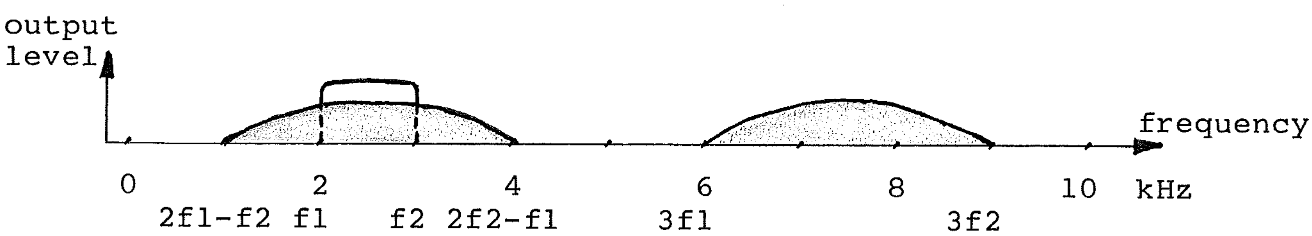


Distortion on a narrowband noise, "voiceless consonant", between the frequencies f1 and f2. The non-shaded area between f1 and f2 represents the original signal:

Quadratic distortion



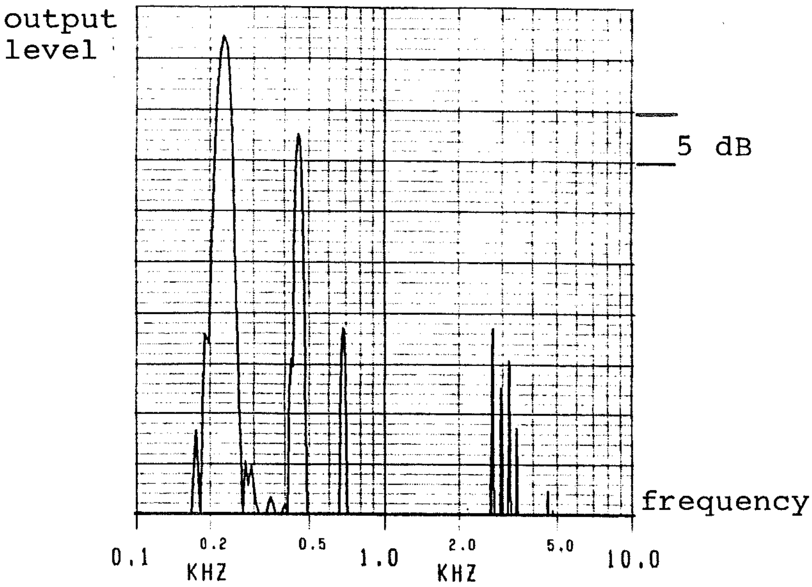
Cubic distortion



Examples of recorded vowel spectra:

Undistorted I-sound

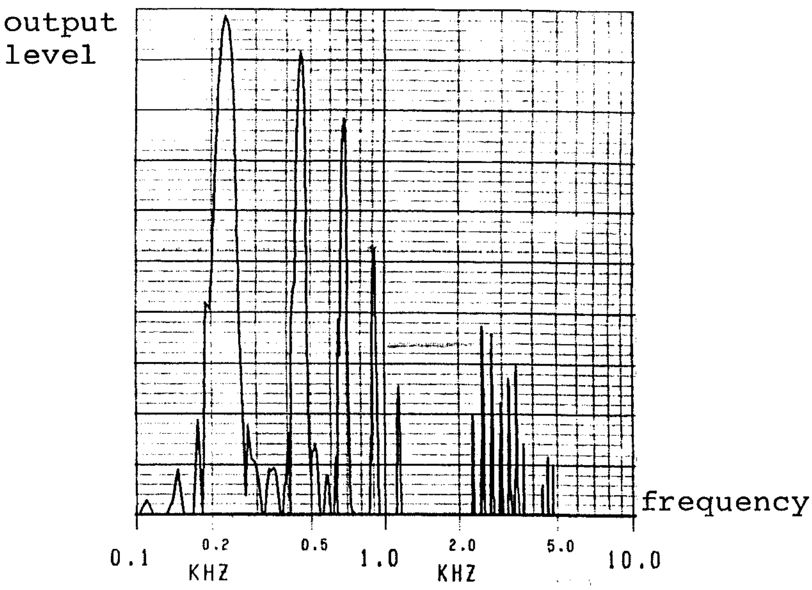
(Sw. DIST)



Distorted I-sound
strong quadratic distortion

Note the influence of the distortion products on the multiples of the fundamental, F0.

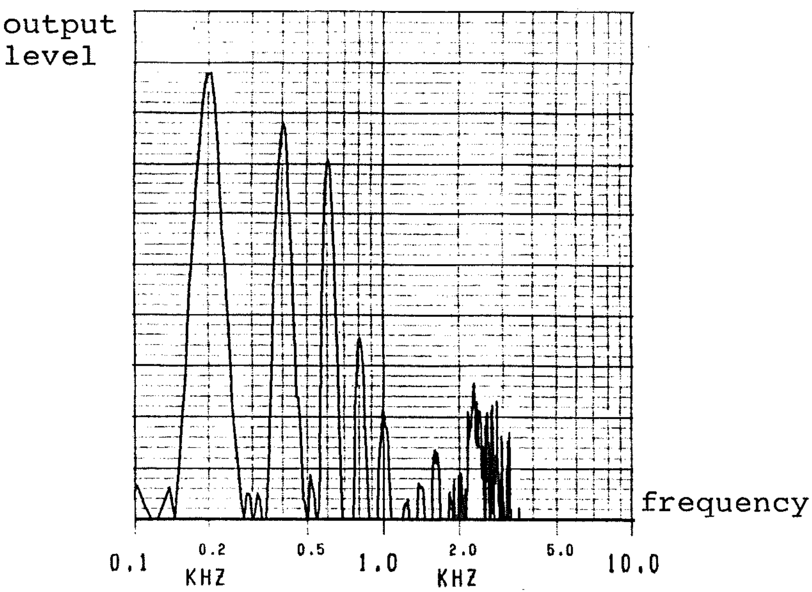
(Sw. DIST)



Undistorted Ä-sound

The level relations between the multiples of F0 up to 5*F0 are almost identical for the distorted I and the Ä.

(Sw. RÄLD)



TA93 1979 Assessment of perceived sound quality in high fidelity sound-reproducing systems.
A Gabrielsson, S-Å Frykholm, B Lindström.

*TA92 1978 Statistical treatment of data from listening tests on sound-reproducing systems.
A Gabrielsson.

TA91 1978 Taluppfattbarhet för högtalarsystem vid Arlanda utrikesterminal.
B Hagerman, A-C Lindblad.

TA90 1978 A generalization of the concept of total harmonic distortion.
Å Olofsson.

*TA89 1978 Ett programpaket för mätning, syntes och analys av linjära tidskontinuerliga och tidsdiskreta system.
Å Olofsson.

TA88 1978 Reliability in the determination of speech reception threshold (SRT).
B Hagerman.

TA87 1978 Detection of nonlinear distortion in telephone systems.
A Gabrielsson, A-C Lindblad, B Lindström, Å Olofsson.

*TA86 1977 Adjective ratings and dimension analyses of perceived sound quality of headphones.
A Gabrielsson, S-Å Frykholm, H Sjögren.

*TA85 1977 Adjective ratings and dimension analyses of perceived sound quality of hearing aids. III.
A Gabrielsson, H Sjögren.

*TA84 1977 Automatic gain control and hearing aids. The influence of different attack and release times on speech intelligibility for hearing impaired with recruitment.
T Ahren, S Arlinger, C Holmgren, L Jerlvall, B Johansson, A-C Lindblad, H Sjögren.

*TA83 1976 Detection of amplitude distortion by normal hearing and hearing impaired subjects.
A Gabrielsson, P-O Nyberg, H Sjögren, L Svensson.

*TA82 1976 Preferred listening levels and perceived sound quality at different sound levels in "high fidelity" sound reproduction.
A Gabrielsson, H Sjögren.

*TA81 1976 Preliminärrapport slingor.
R Ingelstam.

TA80 1975 Emulator för kommunikation mellan minidator typ Honeywell 316 och stordator typ IBM 360.
Bo Jacobsson.

TA79 1975 Mätning av amplitud- och faskurvor för hörapparater med hjälp av minidator.
Å Olofsson.

TA78 1975 Styr- och registreringsenhet för psykoakustisk mätplats.
Å Olofsson, L Persson, H Sjögren.

*TA77 1975 Adjective ratings and dimension analyses of perceived sound quality of hearing aids. II.
A Gabrielsson, H Sjögren.

*TA76 1975 Similarity ratings and dimension analyses of perceived sound quality of hearing aids.
A Gabrielsson, H Sjögren.

*TA75 1974 Adjective ratings and dimension analyses of perceived sound quality of hearing aids. I.
A Gabrielsson, H Sjögren.

TA74 1974 Assessment of comfort and discomfort levels for pure tone, a methodological study.
A Gabrielsson, B Johansson, Bodil Johansson, A-C Lindblad, L Persson.

*TA73 1973 Hörapparaternas driftkostnader.
B Johansson.

TA72 1973 Adjective ratings and dimension analyses of perceived sound quality of sound reproducing systems.
A Gabrielsson, U Rosenberg, H Sjögren.

TA71A 1974 Mätsond för mätning på magnetiska fält inom tonfrekvensområdet.
R Ingelstam.

* indicates that the TA-report is out of stock, but is available at a cost of SEK 35.

LIST OF REPORTS FROM THE DEPARTMENT OF TECHNICAL AUDIOLOGY, KAROLINSKA INSTITUTET

TA115 1987 Loudspeaker frequency response and perceived sound quality: Comparison between measurements in listening room, anechoic room and reverberation room.
A Gabrielsson, Björn Lindström and Ove Till

TA114 1986 Loudspeaker frequency response and perceived sound quality: Measurements in listening room.
A Gabrielsson, B Lindström and O Till

TA113 1986 Relations between speech intelligibility and psychoacoustical modulation transfer function (PMTF).
B Hagerman, Å Olofsson and A-C Lindblad

*TA112 1985 The effect of different frequency responses on sound quality judgments and speech intelligibility.
A Gabrielsson, B N Schenkman and B Hagerman

TA111 1985 Hearing aid measurements on occluded-ear simulator compared to simulated in-situ and in-situ measurements.
U Olsson

TA110 1984 Improvement of induction loop field characteristics using multi-loop systems with uncorrelated currents.
Å Olofsson

TA109 1984 Questionnaires on desirable properties of hearing aids.
B Hagerman and A Gabrielsson

TA108 1983 Audio versus audio-visual test - implications for hearing aid frequency responses.
A-C Lindblad, M P Haggard and J R Foster

TA107 1983 Clinical measurements of speech reception threshold in noise.
B Hagerman

TA106 1983 Assessment of perceived sound quality in eighteen High Fidelity loudspeakers.
A Gabrielsson, B Lindström and G Elger

TA105 1982 Detection of nonlinear distortion of speech signals by hearing impaired listeners.
A-C Lindblad.

*TA104 1982 Measurement of speech reception threshold - a comparison between two methods.
B Hagerman.

TA103 1981 Problems and methods in judgments of perceived sound quality.
A Gabrielsson.

TA102 1981 Scaling of perceptual dimensions in sound reproduction.
A Gabrielsson, B Lindström.

TA101 1981 Sentences for testing speech intelligibility in noise.
B Hagerman.

TA100 1981 Modulation transfer function and speech transmission index - a method to predict speech intelligibility in Swedish? - The Swedish contribution to the IEC 29B/WG15 multilanguage test.
B Hagerman, A-C Lindblad.

TA99 1980 En metod för subjektiv utvärdering av "kompaktanläggningar".
B Lindström.

TA98 1980 Clinical assessment of perceived sound quality in hearing aids.
A Gabrielsson, B Hagerman, C Berg, A Ovegård, L Änggård.

TA97 1980 Statistical analysis of fricative sounds.
Å Olofsson.

TA96 1980 A method for suppression of periodic noise signals.
S-Å Frykholm, Å Olofsson.

*TA95 1980 Hörapparaters frekvenskurvor: IEC-coupler, KEMAR, människa.
U Olsson.

TA94 1980 Some ideas in computerized acoustical impedance measurements.
S-Å Frykholm.

* indicates that the TA-report is out of stock, but is available at a cost of SEK 35.