ARCHIVES OF ACOUSTICS
32, 3, 561–578 (2007)

IDENTIFICATION OF SPECTRALLY ENHANCED POLISH VOWELS(1)

Edward OZIMEK

Adam Mickiewicz University
Institute of Acoustics
Umultowska 85, 61-614 Poznań, Poland
e-mail: ozimaku@amu.edu.pl

(*received April 10, 2007; accepted July 28, 2007*)

The role of spectral enhancement in identification of Polish vowels is investigated in this study. Identification scores of natural (unprocessed) vowels (Experiment 1) and spectrally enhanced vowels (Experiment 2) were measured for normal-hearing subjects (NHS) and for subjects with hearing impairment of cochlear origin (HIS). The NHS group demonstrated almost 100% vowel-identification scores for natural vowels. Instead, for the HIS group, an overall average identification scores were 85% for natural vowels, 96% for spectrally enhanced vowels. It has been shown that the spectral enhancement applied has improved the average identification score for subjects with sensorineural hearing losses. However, the degree of improvement clearly depended on the subjects, the type of vowel, and the talker.

Keywords: spectral enhancement, Polish vowels, identification score.

1. Introduction

Identification of speech sounds is affected in a particular way by the perception of their spectral envelope for which a spectral contrast, i.e., the difference in amplitude between peaks and troughs of the successive formants, plays an important role. In normal-hearing subjects, peak-to-trough differences are well transformed by the auditory filters and are encoded in the discharge patterns of the auditory nerve fibers. For steady-state vowels, for example, neuron responses are exhibited in terms of both discharge rate (SACHS and YOUNG, [28]) and phase locking to the harmonics of the vowel (YOUNG and SACHS, [37]; DELGUTTE and KIANG, [9]).

Abnormalities in the cochlear function usually cause a broadening of the auditory filter bandwidth, and, as a consequence, a worsening of frequency resolution. Poor frequency resolution results in a smear of internal auditory representation (perceptual spec-

(1) The paper is devoted to honor the 50th anniversary of the scientific activity of Professor Andrzej Rakowski. It is based on the experimental data obtained in the Institute of Acoustics UAM by the group comprising: E. Ozimek, A. Sęk, A. Wicher, E. Skrodzka and J. Konieczny.

trum) of the input sounds and worsenes their identification. It was found that frequency resolution decreased with increasing sensorineural hearing loss (PICK *et al.*, [27]), although it was also indicated that the auditory filters could change the perceptual spectrum in such a way that speech identification was not significantly affected until the threshold was more than 30 dB above normal (EVANS, [10]).

In some cases an impaired cochlea exhibits additionally a failure of suppression (WIGHTMAN *et al.*, [36]) which may further reduce the spectral contrast. MOORE and GLASBERG [18] suggested that suppression plays a major role in preserving and enhancing contrast in the internal representation of the vowel-like sounds. According to SIDWELL and SUMMERFIELD [29], in subjects with modest hearing impairment, the suppression mechanism is not entirely absent with a sensitivity loss of 20–40 dB and the internal representation of vowels may be enhanced by this mechanism. However, the impoverished internal representation of vowels does not always imply poor abilities to identify them, since OWENS *et al.*, [23] reported 94% accuracy for identification of vowels by hearing-impaired subjects, in phonetically balanced word lists. Only at severe hearing impairment is the vowel identification really poor.

A question arises to what extent the poor frequency resolution and the changes in the internal representations of vowels lead to a reduction in their identification by hearing-impaired subjects? It is assumed that these factors affect at least partly the vowel recognition ability of those subjects. Some authors have shown, however, that subjects with sensorineural hearing loss sometimes exhibit a reasonably good vowel identification ability (OWENS *et al.*, [22]; NÁBĚLEK and DAGENAIS, [21]). There have been several attempts to reconcile the apparent inconsistency between an abnormal representation of spectral peaks and troughs in the perceptual spectrum, poor frequency resolution, and relatively high ability to identify vowels (TURNER and VAN TASELL, [32]; LEEK *et al.*, [16]).

The identification of vowels may be facilitated by spectral enhancement leading to an increase of their spectral contrast. The spectral enhancement, causing a greater than normal concentration of the spectrum energy around the formant frequencies, may improve recognition of vowels because increasing of spectral contrast partly compensates for the poorer than normal frequency resolution accompanying a sensorineural hearing impairment. Some researchers have attempted to change the spectral contrast of speech sounds by changing their formants bandwidths (BOERS, [5]; VAN VEEN and HOUTGAST, [34]; SUMMERFIELD *et al.*, [31]). Increasing formant bandwidths reduces the formant peak-to-trough differences and consequently reduces the spectral contrast, while decreasing formant bandwidths causes opposite effect. BOERS [5] found only an insignificant effect of narrowed bandwidths on sentence intelligibility. In certain experimental conditions it was even established that for the normal-hearing and hearing-impaired subjects, increased contrast resulted in poorer speech reception threshold scores. VAN VEEN and HOUTGAST [34] found that decreasing formant bandwidths had a small effect on the judged similarity of vowels. SUMMERFIELD *et al.*, [31] varied spectral contrast by varying formant bandwidths in the synthetic CVC syllables. They found little improvement in identification of stop consonants resulting from narrowed bandwidths. LEEK *et al.*, [16] examined a minimum spectral contrast for identification of synthesized vowels by normal-hearing and hearing-impaired subjects. They measured identification ability in noise as a function of peak-to-trough differences in the spectrum of those vowels and found that the peak-to-trough amplitude differences required for 75% identification accuracy amounted to 1–2 dB for normal-hearing subjects and 6–7 dB for hearing-impaired subjects. Hearing-impaired subjects were able to score even above 90% correct identification when the peak-to-trough difference was increased to 14 dB. FRANCK *et al.*, [11] investigated the effects of the reduced dynamic range on the vowel perception by compression and compensation of reduced frequency resolution by spectral enhancement. They found in some measurable conditions that spectral enhancement produced improvements of vowel scores, but this was counteracted by deterioration of the consonant scores for almost all subjects.

The study reported here extends ongoing works examining the impact of hearing impairment on a subject's ability to identify vowels. Since vowels have relatively higher amplitude and longer duration than consonants, they are quite important for persons with hearing impairment. The basic purpose of this study was to examine the influence of spectral enhancement on the identification of natural Polish vowels by hearing-impaired subjects. The internal representations of vowels were modified by means of the spectral enhancement algorithm which increases the peak-to-trough differences in the spectral envelope. Natural (unprocessed) vowels (Experiment 1) and spectrally enhanced vowels (Experiment 2) were used in the studies.

2. Changes in the spectral contrast of vowels

If a vowel sound passes through a set of auditory filters corresponding to normal-hearing, the structure of the vowel formants is well reproduced at the output of those filters that is in the course of the excitation pattern mapping the envelope of the vibrations of the basilar membrane. However, if the filters are considerably broadened, which is common in the case of sensorineural hearing loss, the reproduction of particular formants of the vowel in the excitation pattern is much worse, which leads to difficulties in vowel identification. In order to reduce the influence of the broadened auditory filters on a speech intelligibility, the acoustic signal supplied to an impaired-hearing system should be transformed in such a way that it would produce a similar excitation as a non-transformed signal in the normal hearing system. If P_N and F_N stand for the matrices representing the excitation patterns and auditory filters in a normal hearing system, P_S and F_S in the impaired-hearing system, and X is the vector representing the spectrum of a signal, then

$$P_N = F_N X \tag{1}$$

and

$$P_S = F_S X\,. \tag{2}$$

The signal the spectrum of which would be represented by the vector $\Theta(X)$, that is the signal which having passed through the broadened filters would stimulate the same activity of the hearing system as the signal X in the normal hearing system, should satisfy the equation:

$$P_N = F_S \Theta(X) \tag{3}$$

and finally

$$\Theta(X) = F_S^{-1} F_N X\,. \tag{4}$$

The above described concept of signal transformation has been implemented in the study reported here on the basis of the algorithm developed at the Cambridge University by Baer and Moore (BAER and MOORE, [4]; BAER *et al.*, [3]). The principle of the algorithm is as follows. The vowels were stored on a hard drive of a PC host at a sampling rate of 24 kHz with 24-bit resolution. Each vowel, was divided by the Hanning window into short time segments (256 points, 10 ms long, 50% overlapping) and converted to the frequency domain (X vector), by the Fast Fourier Transform (FFT). The short-term spectra obtained in this way were processed by the enhancement procedure in which only the magnitudes (not phases) of the spectra were processed. In the enhancement process, the spectra were subjected to a procedure which enhanced their peaks and valleys. Similarly like in Baer and Moore's studies, the solution of the Eq. (3) was based on the non-negative, least-square algorithm. In the inverse FFT, the adapted magnitude data were combined with the original phase data and retransformed into the time domain. The time segments obtained were added using an overlap-add procedure to get the output signal. Finally, the signals were low-pass filtered at 4.5 kHz. This algorithm was implemented in Matlab 6.1 environment (Mathworks, [17]). In the enhancement procedure, the spectral contrast was introduced accordingly to the auditory filter parameters, individually for each subject. It allowed changes in the contrast in these frequency bands in which broadening in the auditory filters was previously measured. The spectral contrast procedure affected mainly the higher frequency region, as the auditory filters tended to be much broader in this region. This procedure affected the low frequency band only in a limited range. The effect of spectral enhancement for chosen vowels is illustrated in Fig. 5.

3. Method

3.1. Subjects

Four normal-hearing subjects and three hearing-impaired subjects with roughly similar shape of audiograms took part in the experiments. Normal hearing was defined in terms of pure-tone thresholds of less than 15 dB HL (ANSI, [2]) at octave frequencies from 0.25 to 4 kHz. The paid hearing-impaired subjects (age range 24–40 years) had longstanding, bilaterally symmetrical hearing impairment within the range from about 35 to 75 dB HL. No air gaps were observed for any of the subjects. Basic audiologic tests (air and bone conductions, speech audiometry and SISI test) indicated sensorineural impairment of cochlear origin. The air conduction audiograms for hearing-impaired subjects who participated in the study are shown in Fig. 1.

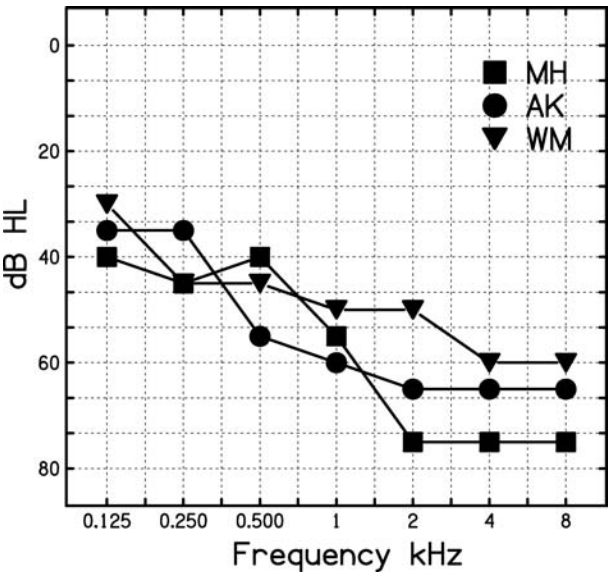


Fig. 1. Audiograms of the three hearing-impaired subjects at octave intervals. The subjects are identified as AK, WM, MH.

As seen in Fig. 1 subjects AK and WM showed hearing losses which gradually increased from 30–35 dB at 125 Hz to about 60–65 dB at 4 kHz. The hearing loss of subject MH increased from 40 dB at 125 Hz to approximately 75 dB at 4 kHz.

For the subjects with sensorineural hearing loss, the characteristics of the auditory filters centered at 1, 2 and 4 kHz were determined using the notched-noise method (PATTERSON, [24]; PATTERSON and MOORE, [25]). The relative bandwidth of notched-noise (the bands relative to center frequencies) was constant and equal 0.4. The equivalent rectangular bandwidth (ERB) was calculated from the steepness parameters of the filter slopes using the following formula: $\mathrm{ERB} = 2f_c/p_l + 2f_c/p_u$ (f_c – center frequency in kHz, p_l – slope of the lower skirt, p_u – slope of the upper skirt (GLASBERG and MOORE, [12]). The auditory filters determined for the tested subjects are indicated by solid lines in Fig. 2 (for details see SKRODZKA *et al.*, [30]). The shallower slopes of the auditory filters reflect their broad bandwidths (large ERB values). The broken lines correspond to the shapes of the auditory filters transmittance of the subjects with normal-hearing for whom the auditory filters have been well established and their equivalent bandwidth varied from 11% to 17% of the central frequency (GLASBERG and MOORE, [12]; MOORE and PETERS, [19]). In this study it was assumed that the equivalent bandwidth of the auditory filters of subjects with normal-hearing was 16% of the central frequency of the filter.

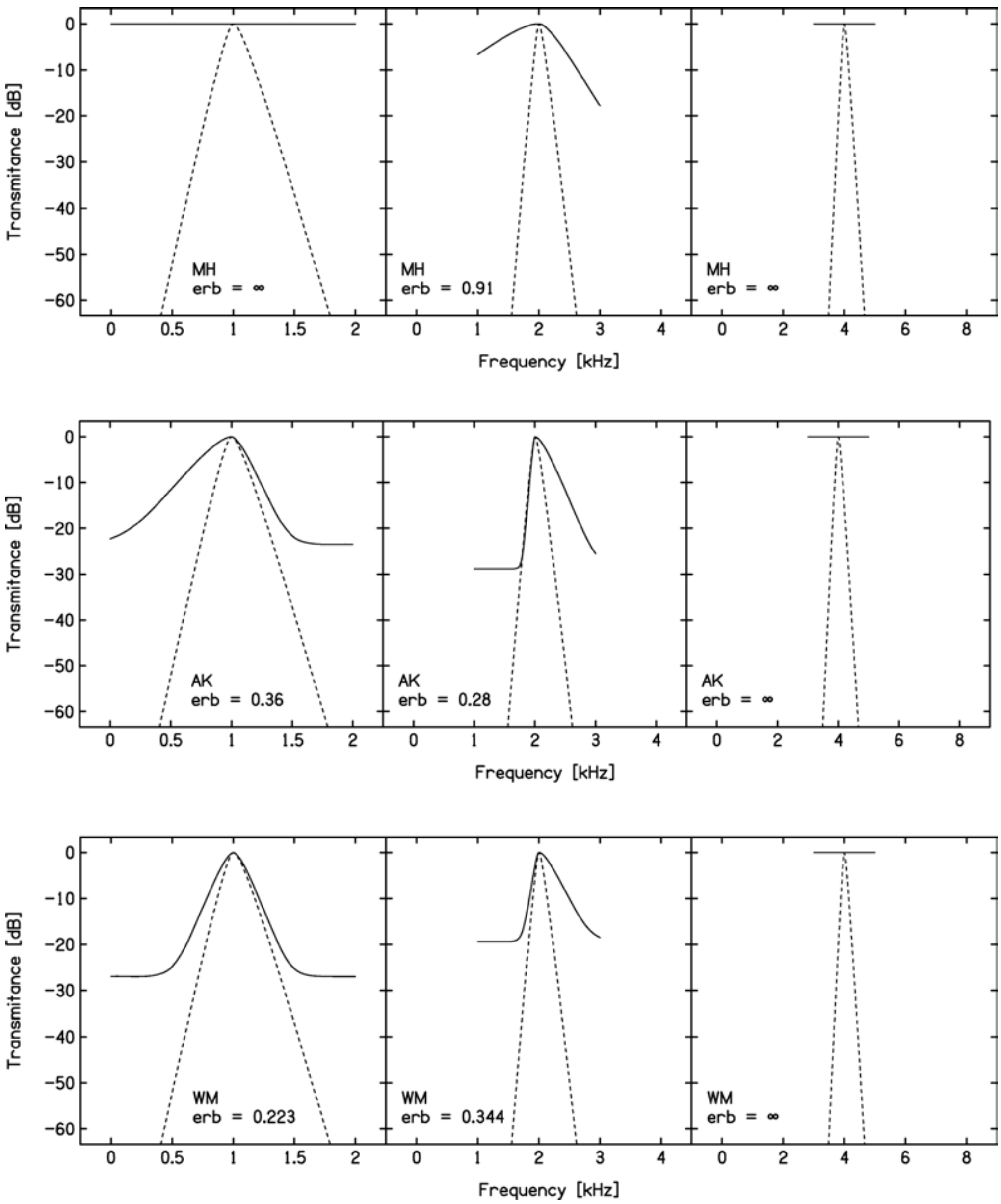


Fig. 2. The auditory filters for subjects AK, WM, MH with sensorineural hearing losses (solid lines) and normal hearing subjects (broken lines).

Figure 2 shows that the auditory filters determined for the hearing-impaired subjects, at the frequencies for which significant hearing loss was observed, are asymmetric, have lower dynamic range and are much broader than the corresponding filters of normal-hearing subjects. At 4 kHz the auditory filter bandwidths were undefined.

As mentioned above, the broadening of the auditory filters, having a deteriorating effect on the frequency selectivity, can impair vowel identification, especially when the vowels are presented in masking noise. Amplification of the signal reaching the ear of a subject with a broadened filter cannot improve this identification, because the signal to noise ratio remains the same. In the present study it was assumed that improvement in vowel identification could be achieved by spectral enhancement of the signal reaching the ear.

3.2. Stimuli and procedure

Six vowels (/o/, /a/, /i/, /e/, /u/, /y/) produced by two talkers (female and male) were used as stimuli. Each vowel lasted approximately 400 ms for female voice and 500 ms for male voice. The vowels were stored on a computer disk and presented to subjects in random sequence. The stimuli were presented in 8 blocks of 4 trials to each group of subjects. Each block of trials consisted of eight randomly ordered presentations of six vowels. Subjects were asked to listen to each of the blocks of stimuli a minimum of four times on different days. After each stimulus presentation, the subject wrote on a list which vowel was heard. Correct answer feedback was not given. All testing was completed in four sessions, each lasting three hours. The spectral enhancement was done off-line. Subjects were seated in a sound-treated booth. They were listening to stimuli via Sennheiser HD580 headphones, coupled to the right ear of the normal-hearing subjects and to the preferred ear for the hearing-impaired subjects. Presentation level for normal-hearing subjects was 75 dB SPL. For hearing-impaired subjects, the level for each individual was adjusted, so that the subject could hear the stimuli at the most comfortable level (MCL) which fell within a range from 85 to 105 dB SPL. For the hearing-impaired subjects experimental procedure was identical to that for the normal-hearing subjects. To avoid distortions, a high-pass (cutoff frequency 50 Hz) and a low-pass filter (cutoff frequency 4500 Hz) were implemented.

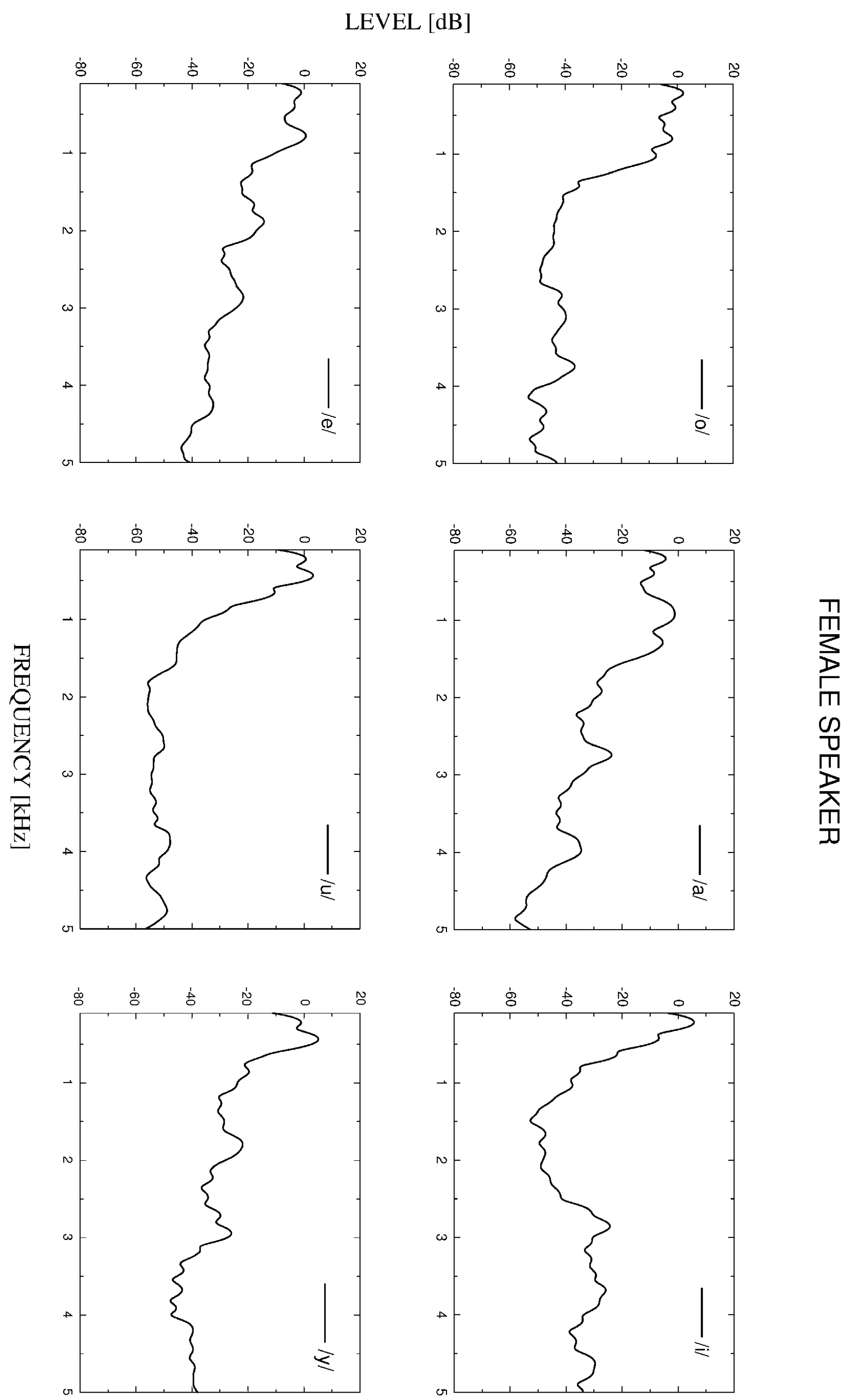
4. Results

4.1. Experiment 1. Identification of natural (unprocessed) vowels

In experiment 1, identification scores were determined for six natural Polish vowels perceived by four normal-hearing and three hearing-impaired subjects. Figures 3a, b show the spectra of six natural Polish vowels produced by the female and the male talker.

The formant frequencies and the level differences in spectral peaks and troughs of tested vowels are given in Table 1.

As follows from Table 1, the frequencies of the formants F_1, F_2, F_3 and F_4 for the female talker are higher than the corresponding ones for the male talker. However, as far as the parameter $\Delta L = L_{\text{peak}} - L_{\text{through}}$ is concerned, no clear relation between the male and female speech sounds has been found.


FEMALE SPEAKER
LEVEL [dB]
FREQUENCY [kHz]
/o/
/a/
/i/
/e/
/u/
/y/

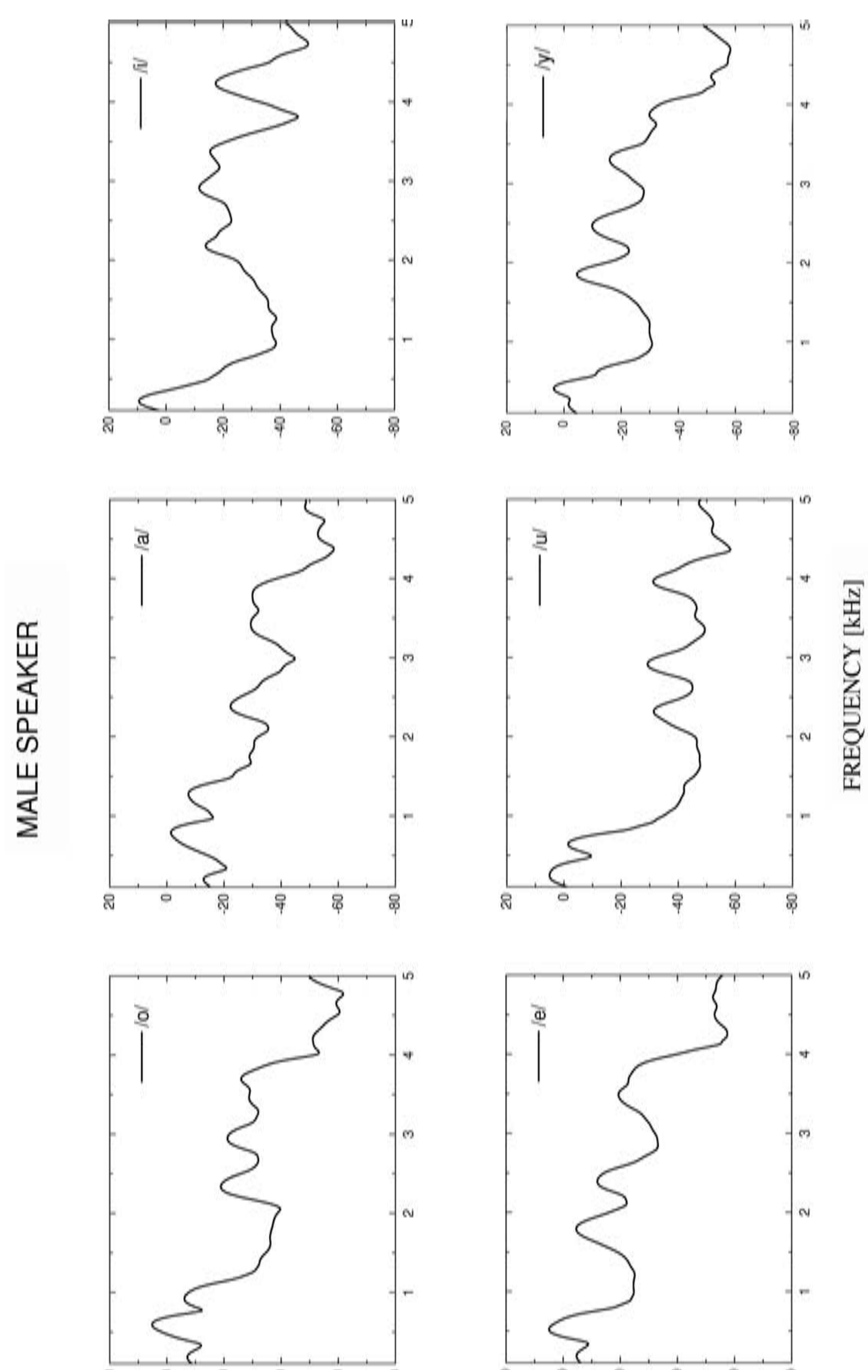


Fig. 3. Spectra of six natural Polish vowels produced by a female and a male talker.

Table 1. The values of the formant frequencies and the level differences in spectral peaks and troughs for natural vowels produced by female and male talkers.

Talker	Vowel	Formant frequency [Hz]				$\Delta L = L_{\mathrm{peak}} - L_{\mathrm{through}}$ [dB]			
		F_1	F_2	F_3	F_4	F_1	F_2	F_3	F_4
Female	/o/	747	1061	3065	3769	6.7	41.6	6.2	11.6
	/a/	888	1321	2805	3964	12	28.2	17.4	22.6
	/i/	227	2816	3715	4625	55.3	6.9	11.1	43.4
	/e/	823	1863	2870	3942	22.5	14.8	14.5	12.8
	/u/	465	714	2534	3866	60	41	3.3	12.7
	/t/	468	1808	2902	4863	38.4	14.4	19.9	4.6
Male	/o/	606	985	2361	3011	18.5	34.6	17.6	15.5
	/a/	790	1301	2469	3476	16.3	20	17.1	3.3
	/i/	220	2198	2924	3379	51.7	13.4	10.9	21.3
	/e/	584	1798	2465	3509	28.8	17	23.7	35.2
	/u/	270	649	2274	2967	12.5	46.5	12.3	27.9
	/t/	433	1741	2491	3336	31.8	19.5	25.5	40.3

Individual data on vowel identification for normal-hearing subjects (NHS) and hearing-impaired subjects (HIS) are shown by hatched bars in Fig. 4. The mean correct identification of tested vowels for NHS and HIS groups is given by black bars.

The data show that the identification scores for all natural vowels, for the NHS group, amount to almost 100%. Such a high score for the NHS group is similar to that obtained by VAN TASSEL *et al.*, [33] who found that natural vowels were identifiable by untrained normal subjects with an average identification score of 98.2%, and by HILLENBRAND and GAYVERT [13] who obtained identification score of 95%.

For the HIS group, lower identification scores and significant variability in performance were observed. The accuracy of identification is indicated by standard deviation. In this group the percent of correct identification depended on the talker (female or male) and the type of vowel. For the male talker better vowel identification was observed than for the female one. For example the subject AK scored 97% correct identification for the male talker and 74% correct identification for the female one. For the female talker, the HIS group showed the best identification for the vowel /y/ and a specific decrease in identification of the vowel /o/. Overall, percent-correct vowel identification performance for the HIS group was 84%, and specifically for the individual participants: 81% for AK, 83% for MH and 88% for WM.

VAN TASELL *et al.* [33], who tested hearing-impaired subjects, found that one of their three subjects identified the seven synthetic vowel stimuli well (93%) while the other two performed with only 70% accuracy. NÁBĚLEK *et al.*, [20] reported a range of vowel-identification performance of 68% to 93% for subjects with a mild-to-moderate

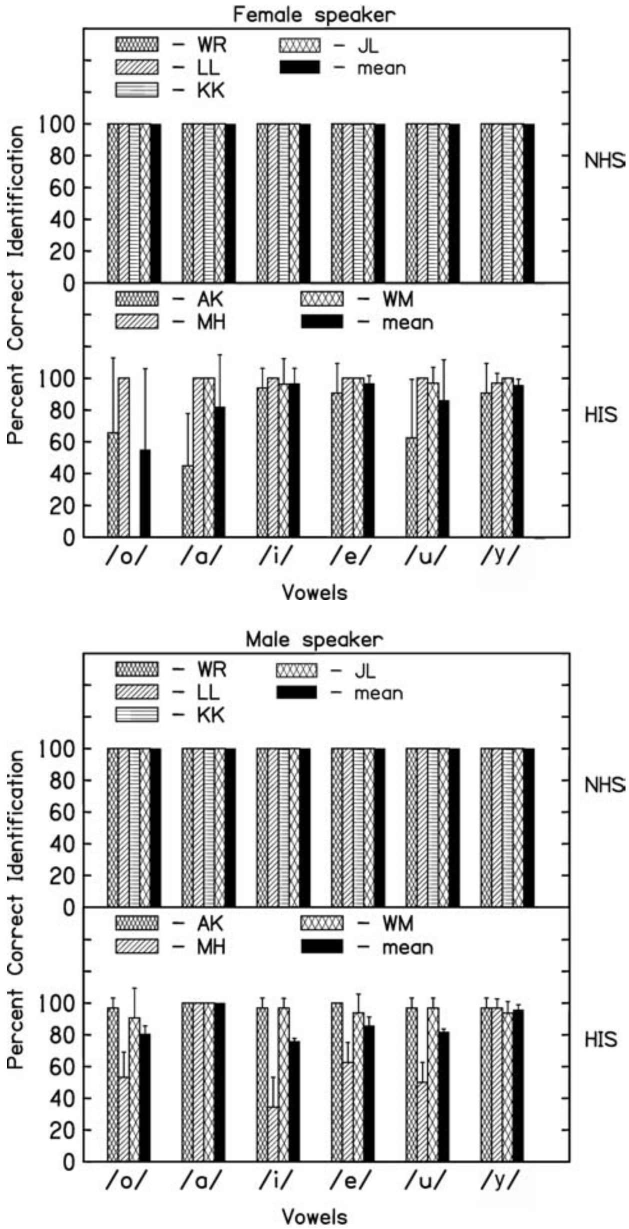


Fig. 4. Individual and mean correct identification (in percent) of natural vowels obtained by the NHS and HIS group, for the female and the male talker.

sensorineural hearing loss. According to LEEK *et al.*, [16], hearing-impaired subjects were able to score 75% identification accuracy for the peak-to-trough amplitude differences of 6–7 dB and 90% correct identification for the peak-to-trough differences of 14 dB. Table 1 indicates that for natural Polish vowels, the peak-to-trough differences are quite significant. This may explain why the hearing-impaired subjects with poor spectral resolution show a high identification ability for some natural vowels.

The results shown in Fig. 4 were subjected to the variance analysis (ANOVA) in order to check the statistical significance of the effect of the talker on vowel identification. It was found that the effect of the talker was statistically significant [$F(1, 30) = 12.13$, $p < 0.05$].

4.2. Experiment 2. Identification of spectrally enhanced vowels

In experiment 2, the vowel identification scores were determined for spectrally enhanced vowels perceived by the same hearing-impaired subjects who participated in experiment 1. Apparatus and stimulus presentation paradigms were identical to those used in experiment 1. Figure 5 (broken line) shows example spectra of vowels processed by the spectral enhancement procedure applied.

The spectral contrast used for stimuli in this experiment was chosen individually for each subject with hearing impairment, depending on his (her) auditory filter shape. The enhanced vowels are characterised by a larger dynamic range, better peak resolution, and better preservation of the vowel formant structure relative to the natural (unprocessed) vowels.

Individual data (hatched bars) and data averaged across subjects (black bars) on vowel identification for hearing-impaired subjects are shown in Fig. 6. For comparison, an average score for unprocessed vowels is also presented by unfilled bars. Accuracy of identification is indicated by standard deviation. Generally, the subject performance for the spectrally enhanced vowels was better than for the unprocessed vowels. The degree of improvement depended on the type of the vowel and the talker.

For the male talker, the highest identification improvement due to the spectral enhancement was observed for the vowels /o/, /i/ and /u/. For this talker, the hearing-impaired subjects scored in this experiment about 96% correct vowel identification (86% was for unprocessed vowels) and for the female talker they scored on average 93% (84% was obtained for unprocessed vowels).

The results shown in Fig. 6 were subjected to variance analysis (ANOVA) in order to check the statistical significance of the improvement of the vowel identification due to spectral enhancement. It was found that the effect of spectral enhancement was statistically significant [$F(1, 60) = 8.03$, $p < 0.06$]. The effects of the speaker was also statistically significant [$F(1, 60) = 10.59$, $p < 0.05$]. No statistical significance was found for the dependence between the identification improvement due to spectral enhancement and the shape of the auditory filters.

5. Discussion

The data from experiment 1 show that the mean percent of correct identification for the HIS group averaged across vowels was 84% with SD $= 8\%$, while for the NHS group it was 100%. The easiest identifiable vowels usually showed a large fre-

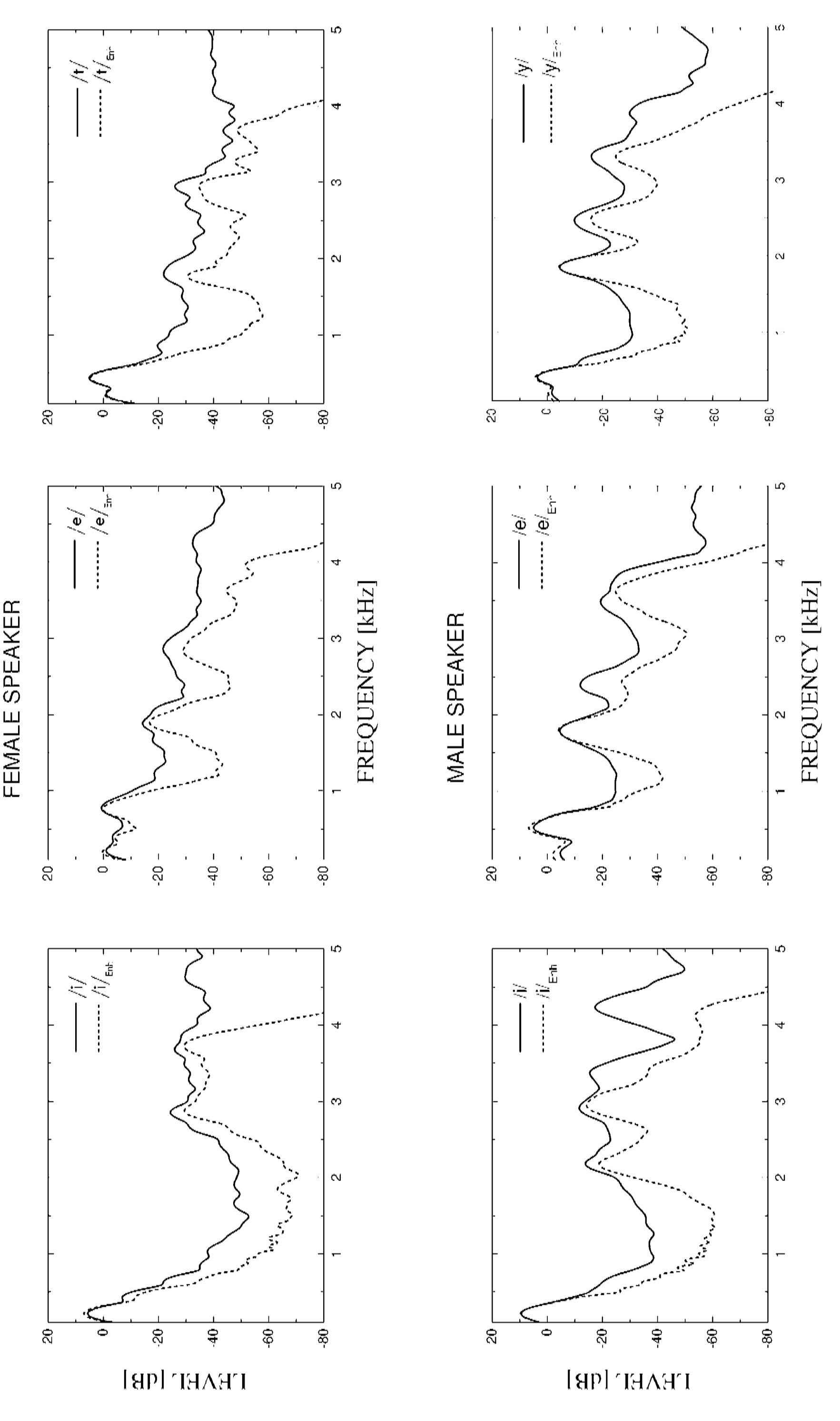


Fig. 5. Spectra of three vowels processed by the enhancement procedure (broken lines). For comparison, the spectra of the unprocessed vowels are also presented (solid lines).

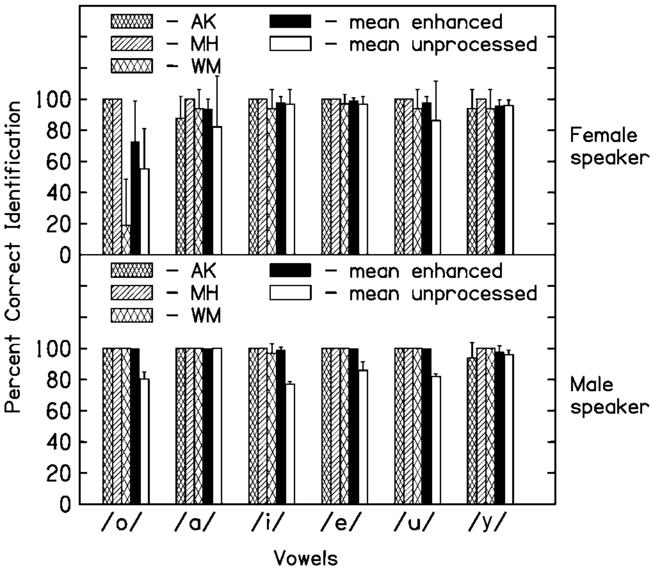


Fig. 6. Individual and mean correct identification (in percent) of spectrally enhanced vowels obtained by the HIS group. Unfilled bars refer to the unprocessed vowels.

quency difference between the formants $F1$ and $F2$, whereas the worst identifiable vowels represented a small frequency difference between those formants. Worsening of the performance of the HIS group relative to that of the NHS group generally supports the assumption that broadening of the auditory filters bandwidth, associated with sensorineural hearing impairment, reduces the peak-to-trough amplitude differences in the internal auditory representation of the vowel spectra. The decrease in the identification scores for these subjects is probably due to the fact that they have some difficulty in discrimination of closely spaced formant peaks, particularly those with relatively low amplitudes. Such difficulties are a consequence of a smoothing of the internal representation of vowels by broadened auditory filters.

There is no consensus in the literature on the effect of vowel formant structure on the identification of vowels. It is usually assumed that to get good vowel identification, the internal representations of vowels should clearly exhibit formant peaks. DUBNO and DORMAN [8] stressed a special role of the first formant in vowel identification. COUGHLIN's *et al.*, [7] data revealed a significant correlation between the ability to identify correctly the vowel and the difference limens for formant discrimination. They concluded, however, that vowel identification was partly predicted by a reduced ability to discriminate spectral differences in the $F2$ region. Thus, a mild-to-moderate high-frequency hearing loss may have for some subjects a significant effect on vowel identification. However, CHISTOVICH [6] has found that poorly resolved formant peaks would not necessarily imply poor vowel identification performance. Other researchers have suggested that accurate identification of vowels is dependent on combined information on such acoustic properties as spectral cues, vowel duration, and formant dynamics

(ANDRUSKI and NEAREY, [1]; JENKINS *et al.*, [14]). Our data indicate that the frequency information provided by the spectrum in the region of the first two formants is particularly important for recognition of vowels for subjects with moderate hearing loss.

Results of experiment 2 have shown that the spectral contrast enhancement improved the vowel identification in the hearing-impaired subjects. Contrast changes in the formant structure seem to account for the clear differences in vowel identification. The positive effect of spectral enhancement found in this experiment is similar to that reported by LEEK *et al.* [16], who found that spectral contrast in vowels provided a useful cue to vowel identification for persons with moderate hearing impairment. However, it is not an obvious outcome since it is not in agreement with the KLATT [15] findings who established that of overriding importance in vowel identification were the formant frequencies and the formant amplitudes, whereas peak-to-valley differences and overall spectral slope were not particularly important. Our data suggest that the vowel identification requires only gross estimation of formant peaks rather than resolution of details across the spectrum. It seems that peaks in the internal representation do not need to coincide with vowel formant frequencies, as long as a general pattern is preserved.

The increase in the spectral contrast of vowels might partly compensate for the reduced frequency resolution of the hearing impaired subjects, so as to produce an internal representation similar to that produced in the normal-hearing subjects. Although contrast enhancement was not equally successful in improving the vowel identification for the hearing-impaired subjects, there was a distinct improvement in the average performance of this group. The improvement is clearly seen for subjects MH (by about 17%) and for AK (by about 12%), and little less for WM (2%). Several authors have examined the effect of spectral contrast on the identification of speech sounds and some of them failed to demonstrate its beneficial effects for hearing-impaired subjects (BOERS, [5]; SUMMERFIELD *et al.*, [31]; DUBNO and DORMAN, [8]). BOERS [5] found only insignificant effect of narrowed bandwidths on the sentence intelligibility. But for some experimental conditions it was established that an increased spectral contrast resulted in poorer speech-reception threshold scores of hearing-impaired subjects. SUMMERFIELD *et al.*, [31] varied spectral contrast by varying formant bandwidth in the synthetic CVC syllables and found little improvement in identification of stop consonants resulting from narrowed bandwidths. FRANCK *et al.*, [11] reported the positive effect of spectral enhancement on the vowel identification, however, it was counteracted by the negative effect on the consonants. Perhaps the difference between the results of the present study and those mentioned above is caused by the fact that in our study spectral contrast increased the amplitude of formant peaks in the midfrequencies relative the low frequency peak which was almost unchanged. In this way the upward spread of masking from F1 was reduced.

The inconsistencies in the literature on vowel identification may stem from several possible reasons. The spectral contrast in the internal representation may be significantly reduced, but it is still sufficient for vowel identification. LEEK *et al.*, [16] showed that normal-hearing subjects required only a 1–2 dB difference in the amplitude of harmonics at spectral peaks and troughs to achieve greater than 75% accuracy in identification of some synthetic vowels. Subjects with moderate hearing losses required only a 6–7 dB peak-to-trough differences for identification. Besides, hearing-impaired subjects may use some additional cues for vowel identification such as vowel duration and pitch (PETERSON and LEHISTE, [26]), formant transition (VERBRUGGE *et al.*, [35]) or some linguistic factors which may improve their identification performance. One may also assume that some vowels are characterised by unique patterns in their internal representations. Such an assumption would help explain why vowel identification is not seriously impaired in cases of mild and moderate hearing losses.

The problem with interpretation of the role of spectral enhancement in vowel identification refers among others to the confounding of the amplitude and envelope of the formants. Spectral enhancement changes these two parameters, thus the improved vowel identification resulting from this procedure may be due to changes in the formant amplitude, the formant spectral envelope, or combination of both effects. It would be interesting to determine independently the contributions of these two parameters to perceptual effects associated with vowel identification.

The results obtained in this paper suggest that the spectral enhancement in the vowels compensates for the limited frequency resolution of the hearing-impaired subjects. The enhancement produces some advantages in vowel identification for those subjects. This study has been performed for a very limited number of speech stimuli and hearing-impaired subjects. Besides, the enhancement procedure produces also some distortions that can be not easily acceptable by some people with hearing impairment. Thus further investigation, for a much wider range of stimuli and a greater number of hearing-impaired subjects is needed before the spectral enhancement procedure finds any possible application in hearing aids.

6. Conclusions

The conclusions drawn from the results of the experiments are as follows:

- Normal-hearing subjects demonstrated excellent, 100% identification score for unprocessed (natural) Polish vowels. For the same vowels, hearing-impaired subjects showed slight-to-moderate difficulty with identification. The identification scores for those subjects ranged from 83 to 88%.
- For the hearing-impaired subjects, spectral enhancement improved the vowel identification by about 10%. Identification scores fall within the range from 90 to 100%. Analysis of the results of individual subjects has shown that this improvement clearly depended on the type of vowel and the talker.

Acknowledgments

This research was supported by Grant 2 P05D 122 26 from State Committee for Scientific Research. Permission from the Cambridge group (B. Moore and T. Baer) for the use of their spectral enhancement algorithm is gratefully acknowledged.

References

[1] ANDRUSKI J., NEAREY T., *On the sufficiency of compound target specification of isolated vowel and vowels in /bVb/ syllables*, J. Acoust. Soc. Am., **91**, 390–410 (1992).

[2] ANSI S3.6 – 1969. *Specification for audiometers*, American National Standards Institute, New York 1969.

[3] BAER T., MOORE B. C. J., GATEHOUSE S., *Spectral contrast enhancement of speech in noise for listeners with sensorineural hearing impairment: Effects on intelligibility, quality and response times*, J. Rehabil. Res. Dev., **30**, 49–72 (1993).

[4] BAER T., MOORE B. C. J., *Spectral enhancement to compensate for reduced frequency selectivity*, J. Acoust. Soc. Am., **95**, 2992 (1994).

[5] BOERS P. M., *Formant enhancement of speech for listeners with sensorineural hearing loss*, I. P. O. Ann. Prog. Rep., **15**, 21–28 (1980).

[6] CHISTOVICH L. A., *Central auditory processing of peripheral vowel spectra*, J. Acoust. Soc. Am., **77**, 789–805 (1985).

[7] COUGHLIN M., KEWLEY-PORT D., HUMES L. E., *The relation between identification and discrimination of vowels in young and elderly listeners*, J. Acoust. Soc. Am., **104**, 3597–3607 (1998).

[8] DUBNO J. R. DORMAN F. M., *Effects of spectral flattening on vowel identification*, J. Acoust. Soc. Am., **82**, 1503–1511 (1987).

[9] DELGUTTE B., KIANG N. Y. S., *Speech coding in the auditory nerve: I. Vowel-like sounds*, J. Acoust. Soc. Am., **75**, 866–878 (1984).

[10] EVANS E. F., *Peripheral auditory processing in normal and abnormal ears: Physiological consideration for attempts to compensate for auditory deficits by acoustic and electrical prostheses*, [in:] *Sensorineural Hearing Impairment and hearing Aids*, edited by C. Ludvigsen and J. Barford. Scand. Audiol. Suppl. 6, 9–44 (1978).

[11] FRANCK B. A. M., SIDONNE C., VAN KREVELD-BOS G. M., DRESCHLER W. A., *Evaluation of spectral enhancement in hearing aids, combined with phonemic compression*, J. Acoust. Soc. Am., **106**, 1452–1464 (1999).

[12] GLASBERG B. R., MOORE B. C. J., *Derivation of auditory filter shapes from notched-noise data*, Hearing Res., **47**, 103–138 (1990).

[13] HILLENBRAND J., GAYVERT R. T., *Identification of steady-state vowels synthesized from the Peterson and Barney measurement*, J. Acoust. Soc. Am., **94**, 668–674 (1993).

[14] JENKINS J. J., STRANGE W., MIRANDA S., *Vowel identification in mixed-speaker silent-center syllables*, J. Acoust. Soc. Am., **95**, 1030–1043 (1994).

[15] KLATT D. H., *Prediction of perceived phonetic distance from critical-band spectra: A first step*, Proc. IEEE Int. Conf. Speech Acoust. Signal Process. 129, 1278–1281 (1982).

[16] LEEK M. R., DORMAN M. F., SUMMERFIELD Q., *Minimum spectral contrast for vowel identification by normal-hearing and hearing-impaired listeners*, J. Acoust. Soc. Am., **81**, 148–154 (1987).

[17] Mathworks. (1996). Matlab 5. Signal processing toolbox.

[18] MOORE B. C. J., GLASBERG B. R., *Masking patterns for synthetic vowels in simultaneous and forward masking*, J. Acoust. Soc. Am., **73**, 906–917 (1983).

[19] MOORE B. C. J., PETERS R. W., *Auditory filter shapes at low center frequencies*, J. Acoust. Soc. Am., **88**, 132–140 (1990).

[20] NÁBĚLEK A. K., CZYŻEWSKI Z., KRISHNAN L. A., *The influence of talker differences on vowel identification by normal-hearing and hearing-impaired listeners*, J. Acoust. Soc. Am., **92**, 1228–1246 (1992).

[21] NÁBĚLEK A. K., DAGENAIS P. A., *Vowel errors in noise and in reverberation by hearing-impaired listeners*, J. Acoust. Soc. Am., **80**, 741–748 (1986).

[22] OWENS E., BENEDICT M., SCHUBERT E. D., *Further identification of vowel items in multiple-choice speech discrimination testing*, J. Speech Hear. Res., **14**, 841–847 (1971).

[23] OWENS E., TALBOTT C., SCHUBERT E., *Vowel discrimination of hearing-impaired listeners*, J. Speech Hear. Res., **11**, 648–655 (1986).

[24] PATTERSON R. D., *Auditory filter shapes derived with noise stimuli*, J. Acoust. Soc. Am., **59**, 640–654 (1976).

[25] PATTERSON R. D., MOORE B. C. J., *Auditory filters and excitation patterns as representations of frequency resolution*, [in:] Frequency Selectivity in Hearing, edited by B. C. J. Moore (Academic, London).(1986).

[26] PETERSON G. E., LEHISTE I., *Duration of syllable nuclei in English*, J. Acoust. Soc. Am., **32**, 693–703 (1960).

[27] PICK G. F., EVANS E. F., WILSON J. P., *Frequency resolution in patients with hearing loss of cochlear origin*, [in:] Psychophysics and Physiology of Hearing, ed. by E. F. Evans and J. P. Wilson (Academic, London).(1977).

[28] SACHS M. B., YOUNG E. D., *Encoding of steady-state vowels in the auditory nerve: Representations in terms of discharge rate*, J. Acoust. Soc. Am., **66**, 470–479 (1979).

[29] SIDWELL A., SUMMERFIELD Q., *The effect of enhanced spectral contrast on the internal representation of vowel-shaped noise*, J. Acoust. Soc. Am., **78**, 495–506 (1985).

[30] SKRODZKA E., WICHER A., OZIMEK E., SĘK A., *Auditory filters in sensorineural hearing-impaired subjects*, Archives of Acoustics, **27**, 3, 159–174 (2002).

[31] SUMMERFIELD Q., FOSTER J., TYLER R., BAILEY P., *Influences of formant bandwidths and auditory frequency selectivity on identification of place of articulation in stop consonants*, Speech Comm., **4**, 213–229 (1985).

[32] TURNER C., VAN TASELL D. J., *Sensorineural hearing loss and the discrimination of vowel-like stimuli*, J. Acoust. Soc. Am., **75**, 562–565 (1984).

[33] VAN TASELL D., FABRY D. A., THIBODEAU L. M., *Vowel identification and vowel masking patterns of hearing-impaired subjects*, J. Acoust. Soc. Am., **81**, 1586–1597 (1987).

[34] VAN VEEN T. M., HOUTGAST T., *Spectral sharpness and vowel dissimilarity*, J. Acoust. Soc. Am., **77**, 628–634 (1985).

[35] VERBRUGGE R., STRANGE W., SHANKWEILER D., EDMAN T., *What information enables a listener to map a talker's vowel space?*, J. Acoust. Soc. Am., **60**, 198–212 (1976).

[36] WIGHTMAN F., MCGEE T., KRAMER M., *Factors influencing frequency selectivity in normal and hearing-impaired listeners*, [in:] Psychophysics and Physiology of hearing, edited by E. F. Evans and J. P. Wilson, Academic, London.(1977). ?

[37] YOUNG E. D., SACHS M. B., *Representation of steady-state vowel in the temporal aspects of the discharge patterns of population of auditory-nerve-fibers*, J. Acoust. Soc. Am., **66**, 1381–1403 (1979).