THE RELEVANT REGIONS OF CZECH VOWELS FROM THE PERCEPTION POINT OF VIEW

BLANKA BOROVIČKOVÁ*

In the process of determination of the perceptional vowel cues we compare two kinds of experimental material. The first is derived from spectrograms, the second results from listening tests. The linguistic material used was the symmetrical Czech sound groups CVC and VCV. These groups contain all mathematically possible combinations of five Czech vowels and 25 consonants. Each of the 125-member group was uttered by 4 speakers (soprano, alto, tenor, bass).

In this manner we get a total group of one thousand utterances. These groups were recorded on magnetic tape. The record was used first for making spectrograms. We tried to find common or similar signs of the spectra for each phoneme. In our material each vowel phoneme exists in three hundred variants, each consonant phoneme exists in sixty variants. A detailed report on this work was published in 'Spectral Analysis of Czech Sound Combinations'.

The same record was also used for the listening test. The acoustic signal was filtered through half-octave band-pass filters. Fourteen of these filters cover the entire audible frequency range. The signal thus modified was listed by 16 listeners (students of the technical university). The $^3/_4$ of a million responses to the sounds were evaluated by means of a computer. From this step of our work we got, first the 3000 histograms of the correct responses to each of the sound variant related to the midle frequency band-passes, and second 3000 lists of sound interchanges. Then we divided the responses into three categories of relevancy: Up to 3 correct responses, irrelevant; from 4 to 12, middle relevancy and from 13 up to 16 responses, maximum relevancy.

In the next step we compared both the findings from the spectra and the findings from the listening test to obtain the perceptual cues of vowels and consonants respectively.

I will now try to convey to you our findings, which follow from this comparison of vowels. In addition to the well known facts, which we confirmed by our work, we found some fairly interesting new facts.

The necessity of the presence of two first formants for recognition of the vowel is

* Czechoslovak Academy of Sciences Praha, Czechoslovakia.

a well known assertion. But for the Czech vowel /*i*/ the presence of one formant only is sufficient, the first or the second formant respectively. Here it is also interesting that the second region of relevancy for /*i*/ comprises all higher formants (F2 — F5), which makes nearly a single compact region also on the spectrograms. The maximum value of correct responses is in the three bands, with the middle frequency 2,0; 3,0 and 4,5 kcps. The middle frequency of the first region of relavency is 0,2 kcps.

Vowel /*e*/ also has two regions of relevancy, which correspond to both, the first and the second formant frequency. The middle frequency of the first region is 0,4 kcps, of the second one 2,0 kcps. But the percentage of the correct responses in both regions is very low. This suggests a necessity of the joint presence of at least both first vowel formants.

The third vowel /*a*/ has only a single region of relevancy at the middle frequency of the band 1,1 kcps. It is caused by the fact that both formants of /*a*/ are close in frequency, only half an octave. So both of the formants fall, at least by their parts, into a single band-pass (1,1 kcps) with 100 % of correct responses.

The vowel /*o*/ has two regions of relevancy (0,6 and 1,1 kcps), which merge into a single one. The distance between the first and second formant is again very small, but a little larger than in the case of the preceding vowel /*a*/. It amounts to approximately one octave. From the fact that the percentage of correct responses is not always in the category of maximum relevancy, we deduce the necessity of the presence of both formants of /*o*/ for correct recognition.

The last vowel /*u*/, in spite of the spectral similarity with the preceding /*o*/, has a quite different characteritic of the relevancy region, developed from the listening test. There is no stable frequency band (region of relevancy) as was found with the preceding vowels. The frequency ranges of maximum of correct responses vary very much and the region of middle relevancy is very large (from the frequency band 0,2 kcps up to 2,0 kcps). The maximum is in the band of middle frequency 1,4 kcps. This maximum is shifted to a higher frequency in connection with the palatal articulated consonants /*ď*/, /*ť*/, /*ň*/, /*j*/ and with the consonants /*ž*/, /*š*/, /*č*/. But according to the spectrograms the band 1,4 kcps is neither the frequency of the second nor of the third formant.

Thus we assume, that the main cue for recognition of the vowel /*u*/ is the transient of its second formant. As the frequency of the first two formant of /*u*/ is very low, the formant frequencies are determined by a single harmonic frequency of the voice only (particularly by the high voices). But this harmonics varies with the variations of fundamental frequency. This uncertainty of formant determination causes a perceptual orientation towards another cue, the formant transient.

As a conclusion we can say that: these assumptions confirm also our hypothesis about minimal speech unit from the perception point of view. This unit is the consonant-vowel connection and the middle part of this unit, the transients of formants (particularly the second one), is the most important part for both, the consonant and even the vowel recognition.