Proc. 5th int. Congr. phon. Sci., Münster 1964, pp. 374–377
(S. Karger, Basel/New York 1965).

From the Instytut Podstawowych Problemów Techniki, Poznań

Measurements on the Distinctive Features of Polish Phonemes

By Wiktor Jassem, Poznań

Investigations into the nature of the phonemic distinctive features in Polish are being carried out mainly with a view to obtain data for the design of speech compression systems and mechanical speech recognition possibly leading to real-time translation of spoken language. The distinctive features as here defined are based on results of acoustic (mainly spectrographic) analysis of undistorted speech and experiments with frequency and time filtered speech. In the following definitions it is to be understood that if the feature is present the phoneme is classified as one to which the first of the two-term label applies (e.g. if the phoneme has more than 3 formants, with amplitude levels within a range of 15 db beginning from the second formant, below 3 kc, and/or any antiformants, it is consonantal, otherwise it is vocalic, etc.).

1. *Consonantal : vocalic.* More than 3 formants, with amplitude levels within 15 db beginning from the second formant, below 3 kc, and/or any antiformants.

2. *Supraglottal : glottal.* Noise-like aperiodic excitation (alone or superimposed on quasi-periodic excitation).

3. *Nasal : oral.* More than 3 formants below 3 kc, with amplitudes within a range of 15 db beginning from the second formant.

4. *Smooth-abrupt.* No appreciable pulse-like rapid event.

5. *Compact-diffuse.* [General description: concentration of energy in a middle range of frequencies.] Specific definitions: (A) vocalic phonemes: $F_1 > 450$ cps, (B) consonantal phonemes: (a) nasal: no formants in the range 1.2 to 2 kc, (b) other: most, or all, energy concentrated in the range 1 to 4 kc.

6. *Acute : grave.* [General description: relatively high level of higher frequencies]. Specific definitions: (A) vocalic: (a) compact: $\frac{F_2 + F_3}{2} - F_1 > 1.45$ kc, (b) diffuse: $\frac{F_2 + F_3}{2} - F_1 > 2.2$ kc, (B) consonantal phonemes: (a) nasal: (α) compact: the level of the formant just above 2 kc more than 10 db higher than level of other formants except the first, (β) diffuse: a formant between 1.4 and 1.8 kc, (b) other: almost all energy above 3 kc.

7. *Low-tone : high-tone.* [General description: Low frequency of all formants.] Specific definitions: (A) vocalic phonemes: $F_1 + F_2 < 1.7$ kc, (B) consonantal phonemes: $F_2 + F_3 + F_4 < 8$ kc.

8. *Short : long.* (A) vocalic phonemes: Position next to another vocalic phoneme and duration of steady F_1 segment less than 100 msec, (B) consonantal phonemes: duration of the noise segment less than 100 msec.

9. *Voiced : voiceless.* Quasi-periodic glottal source excitation.

It will be seen that some of the features have complementary variants. E.g., the feature acute:grave has one definition for vowels and another for consonants. Even among the consonants the definition of the acute:grave contrast depends on whether the contrasting phonemes are diffuse or compact. The phoneme /ʒ/, to take an example, is defined according to (1), (2), (4), (5Bb), (6Bb), (7B), (8B), (9).

Wherever, in the above definitions, the distribution of energy in wide frequency bands is mentioned, it is to be understood that a band below 0.7 kc (in which energy, if present, is due mainly to the glottal source) is not considered.

Using two elementary signs (e.g. + and —, or 0 and 1, etc.) each phoneme is uniquely coded by a definite combination of such signs. A binary distinctive feature code is naturally suitable for transmission over bistable systems. Any sequence of phonemes, coded according to the system here proposed, with the signs in temporal succession, may be uniquely decoded from a text message without any additional signs being necessary for the indication of phoneme boundaries*.

The order of the distinctive features is such that the redundancy of the code is reduced to a minimum. The code is constructed in such a way that of the two elementary signs one is always attached

* The present system is a revision of a first version, cf. QPSR, STL 1/1962, Stockholm, pp. 7–14.

Table I

	ī	u	w	i	j	a	o	e	r	l	m	n	ŋ	ɲ	p	b	t	d	k
Consonantal	−	−	−	−	−	−	−	−	+	+	+	+	+	+	+	+	+	+	+
Supraglottal									−	−	−	−	−	−	+	+	+	+	+
Nasal									−	−	+	+	+	+					
Smooth									−	+					−	−	−	−	−
Compact	−	−	−	−	−	+	+	+			−	−	+	+	−	−	−	−	+
Acute	−	−	−	+	+	−	−	+			−	+	−	+	−	−	+	+	−
Low-tone	−	+	+			−	+												
Short		−	+	−	+														
Voiced															−	+	−	+	−
N	4	5	5	4	4	4	4	3	4	4	5	5	5	5	6	6	6	6	6
10^3 p	53	35	20	43	45	90	92	91	36	27	34	44	9	18	27	17	48	24	22

	g	c	ɟ	f	v	s	z	ts	dz	x	c,	z,	tc,	dz,	ʃ	ʒ	tʃ	dʒ
Consonantal	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Supraglottal	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Nasal																		
Smooth	−	−	−	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Compact	+	+	+	−	−	−	−	−	−	+	+	+	+	+	+	+	+	+
Acute	−	+	+	−	−	+	+	+	+	−	+	+	+	+	+	+	+	+
Low-tone											−	−	−	−	+	+	+	+
Short						−	−	+	+		−	−	+	+	−	−	+	+
Voiced	+	−	+	−	+	−	+	−	+		−	+	−	+	−	+	−	+
N	6	6	6	6	6	7	7	7	7	5	8	8	8	8	8	8	8	8
10^3 p	14	8	1	20	32	34	19	14	3	11	15	1	9	5	19	13	13	0*

* $p < 0.0005$.

to the group with more stages of subdivision (thus compact phonemes receive a + because some of them are divided again at 4 stages whereas no diffuse phonemes are divided again more than 3 times. Similarly, acute phonemes may be divided again 3 times, but grave phonemes only twice, so acute is a «plus» feature, etc.). This results in the phonemes tending strongly to order themselves according to the number of features. The table shows this to be certainly true of the consonantal phonemes. Strangely enough, the tendency is almost reversed in the small group of vocalic phonemes.

Measurements leading to the most accurate definitions of the individual features are still in progress at the Acoustic Phonetics Laboratory of the Polish Academy of Sciences, and some details concerning female voices (the frequencies given above in the definitions refer to male voices) have not yet been solved.

The entropy of the information source of the Polish phonemes being $H_s = -\sum_{i=1}^{i=37} p_i \log_2 p_i = 4.744$ bits/phoneme, and the entropy of our natural code being $H_c = \sum_{i=1}^{i=37} p_i N_i = 5.046$ distinctive feat./phoneme (N – number of distinctive features for phoneme with occurrence frequency p), the redundancy of our code is 6%. This redundancy is chiefly due to the fact that N does not always tend to increase with decreasing p. Considering, however, that our features are entirely natural the code redundancy is considered very low. Redundancy due to transitional probabilities may be used for human or mechanical error correction.

Author's address: W. Jassem, Instytut Podstawowych Problemów Techniki, ul. Matejki 48/49, *Poznan* (Poland).

Discussion

Daneš (Praha): I am afraid that the surprising agreement between both rows of members (viz. the numbers of bits and of distinctive features, respectively) is only an incident. Without doubt, one can imagine such a vowel change by which the frequency of a phoneme increases and that of an other one proportionally decreases (e.g. in the case of an incomplete merging of two phonemes), while the phonemic structure of the given language, the structure of the distinctive features, does not change at all.

Hill (Harlow): I should like to raise two points (which concern the discussion we have just heard). First, when constructing a Shannon code the most frequently occurring items for transmission should be represented by the least number of binary choices (in this case distinctive features) for greatest efficiency. Thus the point at issue is resolved by saying that, if a certain phoneme, requiring a certain number of binary choices (distinctive features) for identification, becomes more frequently used, then the speakers of that language are using it less efficiently than they were.

I will ask Mr. *Jassem* to elucidate further the second point, which is how, precisely, he arrived at the series of divisions of the total phonemes to produce the tree illustrating the grouping of phonemes by frequency of occurrence.

Bluhme (Amsterdam): The F_1/F_2 diagram exposed by Mr. *Jassem* shows quite considerable overlapping of the Polish vowel phonemes. However, as we know from phonemic analyses that they are distinctively different from each other, we have serious doubts in the validity of this acoustic approach. Mr. *Mol* has been able to show that the significant differences come out much better if the formants of vowels of one and the same speaker are compared with each other. So it would have been more appropriate if Mr. *Jassem* had used the formular $\sum\left(\frac{F_1}{F_2}\right)$ instead of $\frac{\sum F_1}{\sum F_2}$.