The distinctive features of Polish phonemes

Jassem, W.

journal: STL-QPSR  
volume: 3  
number: 1  
year: 1962  
pages: 007-016

http://www.speech.kth.se/qpsr
The significance of the distinctive feature principle not only to linguistics but also to certain aspects of information theory has long become obvious. But the definition of the individual features as originally stated by Jakobson, Fant and Halle (7) and then reformulated by Jakobson and Halle (8) cannot be regarded as an unalterable and exhaustive set of tokens by which all that is phonematically relevant in the acoustic realization of any language can be described. In the only publication so far available that is based on an amount of data large enough to carry the description beyond the hypothesis stage (5) the author, one of the founders of the theory, was obliged to modify the original definitions on several points. G. Fant and B. Lindblom recently restated the definitions introducing further modifications (4).

An attempt is here being made to describe the system of the phonematically relevant acoustic features of Polish speech according to the general principles of the Jakobson-Fant-Halle theory. In matters of detail this description will be found to contain some formulations which differ from those hitherto suggested, the main reasons being that either a current definition has been found not to be satisfactorily operative when applied to the material under investigation or that a different statement has been considered to cover the data more consistently. Thus, the consonantal-nonconsonantal contrast has not been applied here because when overall amplitude differences are eliminated by means of gain control systems, no confusion between vocoids on the one hand and laterals and nasals on the other results. Nor has the vocalic-nonvocalic contrast been utilized because no adequate statement permits a clear-cut differentiation between energy concentrations that are formant-like and such as are not. The system here submitted is partly based on speech material which has been analyzed in connection with other projects and has previously been described (1)(10)(11)(13). Further data have been and are still being collected and analyzed.

(1) Phonemes are described as supraglottal if they show a continuous energy distribution with frequency due to an energy source placed above the glottis. All fricatives, affricates and stops are supraglottal whereas the vocoids, the nasals, the lateral /l/ and the

---

x) The following survey undertaken by W. Jassem is provisional only and a few modifications can be expected in future elaborations on this subject. Thus the distinctive features No. 3 and No. 4 in Fig. 1-4 will be labeled "nasal" and "lateral" replacing "shunted" and "pole-zero", respectively.
flapped or rolled /r/ are non-supraglottal. In some of the supraglottals there is also a quasi-periodic excitation at the glottis, which structurally operates at a different level; see point (10) below.

(2) The airflow may be stopped for a relatively short moment with a subsequent more or less rapid release. In the case of /r/, the only non-supraglottal interrupted phoneme, this articulation results in a brief reduction of overall level. In a flapped [r] there is only one such reduction, in a rolled [r] there is a succession of two or more. These "phones" are free variants of the /r/ phoneme. The supraglottal interrupted phonemes differ from the uninterrupted by the shortness of the non-periodic segment which does not appear to exceed 20 msec. The stops may therefore probably be described as representing the result of single-pulse or shock excitation whilst the fricatives and affricates result from noise excitation.

(3) It is well known from the current descriptions (e.g. (3)(4)) that nasal contoids represent acoustically the effect of shunting the airflow through the nose by a smaller or larger mouth cavity. This introduces both an anti-resonance - a reduction of level in the spectrum envelope in the vicinity of 1 kc/s - and additional energy concentrations, particularly in a region just above 2 kc/s. The shunted-non-shunted distinction is not applicable to the supraglottal phonemes.

(4) The /l/ phoneme appears to differ consistently from all the other non-supraglottal non-interrupted non-shunted phonemes by a "dip" in the spectrum envelope just above 1 kc/s without a distinct additional energy concentration (formant). There is thus an anti-resonance, but there are no significant additional resonances as in the case of shunted phonemes. This description is largely in agreement with that given by Fant in ref. (3) but may require modification in view of the more recent studies (14). Second and third formant transitions (ibid.) cannot be regarded as being distinctive in Polish since /l/ may occur without a neighbouring vocoid as in utterance-final "cykl" = /tsɛkl/.

(5) The two phonemes /j/ and /w/ always occur next to /i u e a/ or /o/. The most regular feature that has been found to appear in /j/ and /w/ is a peculiar pattern of F3. Whilst F1 and F2 form a smooth
transition between the position typical for the syllabic and that typi-
al for the non-syllabic (i.e. between the two "targets"), \( F_3 \) forms a
kind of "angle" or "dip". This has been described in some more detail
in ref. (10).

(6) The classification of phonemes into compact vs non-compact
is less satisfactory than the others. Only /i e a o u/ have so far
been clearly divided into two groups: compact /e a o/ and non-compact
/i e u/. This distinction is based on the position of \( F_1 \) (cf.(3)(4)(9)).
A demarcation line at 450 c/s works very satisfactorily, neatly dividing
the vast majority of our specimens into compact (\( F_1 > 450 \) c/s) and non-
compact (\( F_1 < 450 \) c/s) vocoids. Also non-interrupted supraglottal pho-
nemes can sufficiently well be divided into such as have at least one
very distinct energy concentration below 4.5 kc/s, these being compact,
and such as have no such distinct concentration (non-compact). But a-
mong the nasals and stops the compact-non-compact distinction requires
further elaboration.

(7) The flat-nonflat distinction is relevant in all vocoids.
It is based on the sum of the \( F_1 \) and \( F_2 \) frequencies (cf. (4)).
The dividing line lies at approximately 1.8 kc/s. In /o u/ and /w/, the
flat vocoids, \( F_1 + F_2 < 1.8 \) kc/s. In /i e a/ \( F_1 + F_2 > 1.8 \) kc/s. In the
supraglottal phonemes the flat-nonflat distinction only operates within
the non-interrupted compact group, /x/ being flat in contrast to all the
remaining phonemes. \( F_1 \) is heavily damped in all supraglottal phonemes,
and the flat /x/ is distinguished from /œ ɔ tʃ dʒ ʃ ʒ tʃ dʒ/ by having
only one distinct energy concentration below approx. 4 kc/s, whilst the
non-flat phonemes have two at least.

(8) The grave-nongrave contrast is, at present, a common label
for at least three different but mutually exclusive features. In the
vocoids it refers to the distance between \( F_1 \) and \( F_2 \). There are two
vocoid pairs representing this contrast, viz. /i/ vs /i/ and /a/ vs /o/.

\[ x \) It now appears that the position of \( F_3 \) is probably irrelevant
to the flat-nonflat distinction in Polish (cf. ref.(9)).
The grave-nongrave feature is redundant in the flat vocoids and in the gliding (non-syllabic) vocoids. Although in all voices hitherto analyzed $F_2 - F_1$ was greater in /i/ than in /ɪ/ and in /e/ greater than in /a/, more exact numerical values which would be generally representative for the language are not yet available. Also /f v/ have been considered to be in contrast to /s z ŋ z/ as grave vs non-grave. Since all these contoids are non-compact, their characteristics should be looked for in the energy distribution above 4.5 kc/s. Most of the energy is usually contained in a frequency band from approx. 4.5 kc/s to 6.0 kc/s in the non-grave /s z ŋ z/ whilst in /f v/ there is more energy between approx. 6.0 and 10 kc/s than between 4.5 and 6.0 kc/s. Since, however, /f/ and /v/ are fairly well recognized even if frequencies above 5 kc/s are filtered out (ref. (1)), it would appear that the distinction might primarily be based on differences in overall level, /f v/ being significantly lower in level than /s z ŋ z/ and the difference being of the order of 15 to 20 dB (cf. (6)). If this is the preferable description, there might be some doubt as to whether this contrast should receive the same label as that for /ɪ/ vs /i/ and /a/ vs /o/. The contrasts /m/:/n/, /ŋ:/ŋ/, /p/:/t/, and /k/:/c3/ as grave vs non-grave must for the moment be considered hypothetical. On the other hand /ʃ ʃ ʃ ʃ/ appear clearly to contrast with /ɡ ɡ ɡ ɡ/ as grave:non-grave by virtue of having all important energy concentrations (formants) lower by approx. 0.5 kc/s.

(9) All non-flat non-interrupted supraglottal contoids with the exception of /f v/ are either mellow or non-mellow. The distinction is based primarily on the duration of the noise segment. If this is less than approx. 80-100 msec the phoneme is mellow. Otherwise it is non-mellow.

(10) The voiced-non-voiced contrast applies to all supraglottal phonemes with the exception of /x/. The syllabic vocoids are, in ordinary speech, always voiced. The other phonemes in which the distinction is not relevant may be voiced or voiceless, partly according to context (conditioned variants) and partly irrespective (free variants)\(^{x}\).

\(x\) The energy distribution in the noise spectrum of voiceless supraglottal phonemes differs from that in the non-supraglottal phonemes, frequencies below approx. 1 kc/s being heavily damped in the former category as the effect of the source being placed above the glottis (cf. (3)(4)).
All the contrasts /s/:/z/, /tɛː/:/dɔ/, /f/:/v/, etc. depend uniquely on the presence or absence of $F_0$. The voiced phonemes are fully voiced and the non-voiced are completely voiceless. A lenis-fortis contrast does not exist in Polish.

The terminology here employed with reference to distinctive features is not entirely satisfactory. Some of it is quite provisional, such as gliding:non-gliding. No appropriate term could in this case be found to specify the acoustic phenomenon described above. It may also be found objectionable to refer to /l/ specifically as a "pole-zero" phoneme, seeing that an anti-resonance is clearly significant in the nasals also. The term "interrupted" may be objected to on the ground that, at least from the articulatory point of view, there is an "interruption" in the affricates. In general, it would be desirable to have all the terms refer directly to the characteristics of the speech wave irrespective of the articulations. We would, however, not like to make any final terminological suggestions before this investigation has progressed much further. Also, some of the contrasts, as will have been deduced from the above, are less final than others.

The terms "vowel" and "consonant" have not been used because we feel that they should refer to the function of the phonemes in the syllable (cf. (17)).

Although, then, our present results are provisional only, it may be of interest to evaluate the informational content of the Polish phonemes as coded by the binary distinctive features, on assumption that our structural analysis is correct. Fig. 1-4 represents a "branching" diagram of the entire phonemic system. A right-hand branch may be denoted by a plus and a left-hand branch by a minus. The individual signs (phonemes) are represented by different numbers of binary signals (12). Whilst, for instance, /r/ is represented by 2 ($\rightarrow$), /i/ is represented by as many as 8 (-----). Table I.D-1 shows a binary code of the polish phonemes based on distinctive features. The code is strictly binary, it employs no $\pm$ signals and does not require any additional signals for phoneme boundaries in a current text. This may be shown by an example chosen at random. The work "język" /jeːʐɨk/ will be coded:

--------+------------
Fig. 1-4. A branching diagram of the distinctive features of Dutch phonemes.
| Phoneme      | i | t | u | e | a | o | j | w | l | n | m | p | n | r | s | z | ts | dz | f | v | z | ts | dz | s | 3 | tf | dz | x | t | d | p | b | c | j | k | g |
| Distinctive features |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| Supraglottal | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| Interrupted  | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| Shunted      | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| Pole-zero    | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| Gliding      | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| Compact      | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| Flat         | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| Grave        | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| Mellow       | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| Voiced       | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
Since no phoneme is coded by either one or two, three or four minuses, or by four minuses and one plus, the sixth signal, a minus, is taken. By reference to Table 1.2.1 or the diagram in Fig. 1.4 it will be found that ---+ corresponds to /i/. No other phoneme beginning with this particular combination, there is a phoneme boundary after the first 6 signals. With this simple procedure, which can easily be automated, any text may be unambiguously decoded.

M. Steffen calculated the relative frequency of Polish phonemes in 1957 (16). It has since appeared that a modification of her phonemicization is desirable. \([\#] \) and \([\&] \) should be interpreted as /æ\]/ and /æ\/] (ref. (9)). Steffen's figures have been altered accordingly and col. 7 in Table 1.2.2 gives, under \(p_k\), the relative frequencies of the phonemes after the modification (in pro mille). In this Table the phonemes have been arranged accordingly to the number of signals in the distinctive feature code, with indices in col. 5 (under i). Col. 2 gives the index number of each phoneme according to its relative frequency (thus, /ɛ/, with the greatest relative frequency 0.392, has \(k=1\), /ɛ/ with relative frequency 0.091 has \(k=2\), etc. \(N_i\) indicates the number of binary signals in the natural (i.e. the distinctive feature) code (see Table 1.2.1). It can immediately be seen that natural coding disagrees with the information content of the individual signs (phonemes). If \(i=k\) for any phoneme (as in the case of /ɛ/), this is mere coincidence. With equal probability the entropy of 37 signs is

\[
H = \log_2 37 = 5.207 \text{ bits/signal}.
\]

The actual entropy of the distinctive feature code, as assumed here, is

\[
\hat{H}_d = \sum_{i=1}^{N_i=37} p_i N_i = 6.254 \text{ bits/phoneme}.
\]

With equal frequency the entropy of our code would be

\[
H_d = \frac{1}{37} \sum_{i=1}^{N_i=37} N_i = 5.892 \text{ bits/phoneme}.
\]

Thus, natural coding seems partly to run counter to the individual information content of the phonemes. This is mainly due to
the fact that most of the phonemes with highest probability, viz. the syllabic vocoids, have, in our code, the highest number of signals.

An optimal binary code, constructed according to the classical halving method (see e.g. ref. (15), p. 105-108) and based purely on the relative frequency of signs, gives, for each phoneme a number of binary signals \( N_k \), see Table I.D-2. The entropy of the optimal code is

\[
H_c = \sum_{k=1}^{K=37} p_k N_k = 4.797 \text{ bits/phoneme.}
\]

The entropy of the Polish phoneme inventory measured at the source is

\[
H_s = - \sum_{k=1}^{K=37} p_k \log_2 p_k = 4.744 \text{ bits/phoneme.}
\]

The absolute redundancy of our distinctive feature code is therefore

\[
R_d = H_d - H_s = 1.511 \text{ bits/phoneme}
\]

whilst the absolute redundancy of the optimal code is only

\[
R_c = H_c - H_s = 0.0053 \text{ bit/phoneme.}
\]

The relative redundancies are, respectively,

\[
R_d = (1 - \frac{H_s}{H_d}) \times 100 \% \approx 34 \% \text{ for the distinctive feature code and}
\]

\[
R_c = (1 - \frac{H_s}{H_d}) \times 100 \% \approx 1 \% \text{ for the optimal code.}
\]

If similar relations obtain in other languages (which we consider probable, though reliable data are scarce), the following conclusions can be drawn:

It is not unrealistic to consider the possibility of analysis-synthesis speech transmission systems based on automatic detection of

\[x\) The above calculations disregard Markoff chain relations.\]
distinctive features (cf., for instance ref. (18)). If maximal reduction of redundancy in the message is the primary consideration, then it would be preferable first to identify phonemes on the basis of the distinctive features, and transmit them in an optimal binary code rather than transmit binary signals corresponding directly to distinctive features.

The author wishes gratefully to acknowledge the technical assistance of Mr. A. Møller, who produced displays of over 300 one-word Polish utterances with a 48-channel spectrograph.

W. Jassem


NOTE for TABLE I.D-1: The relative frequency of /\d^3/ is less than 0.0005.
<table>
<thead>
<tr>
<th>i</th>
<th>k</th>
<th>distinctive feature code</th>
<th>(N_1)</th>
<th>(N_k)</th>
<th>(10^3 p_k)</th>
<th>(10^3 p_1 N_1)</th>
<th>(10^3 p_k N_k)</th>
<th>(-10^3 p_k \log_2 p_k)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9</td>
<td>r</td>
<td>2</td>
<td>5</td>
<td>36</td>
<td>72</td>
<td>180</td>
<td>173</td>
</tr>
<tr>
<td>2</td>
<td>14</td>
<td>l</td>
<td>4</td>
<td>5</td>
<td>27</td>
<td>108</td>
<td>135</td>
<td>141</td>
</tr>
<tr>
<td>3</td>
<td>29</td>
<td>x</td>
<td>4</td>
<td>7</td>
<td>11</td>
<td>44</td>
<td>77</td>
<td>72</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>n</td>
<td>5</td>
<td>5</td>
<td>44</td>
<td>220</td>
<td>220</td>
<td>198</td>
</tr>
<tr>
<td>5</td>
<td>12</td>
<td>m</td>
<td>5</td>
<td>5</td>
<td>34</td>
<td>170</td>
<td>170</td>
<td>165</td>
</tr>
<tr>
<td>6</td>
<td>22</td>
<td>n</td>
<td>5</td>
<td>6</td>
<td>18</td>
<td>90</td>
<td>108</td>
<td>104</td>
</tr>
<tr>
<td>7</td>
<td>31</td>
<td>f</td>
<td>5</td>
<td>7</td>
<td>9</td>
<td>45</td>
<td>63</td>
<td>61</td>
</tr>
<tr>
<td>8</td>
<td>18</td>
<td>v</td>
<td>5</td>
<td>6</td>
<td>20</td>
<td>100</td>
<td>120</td>
<td>113</td>
</tr>
<tr>
<td>9</td>
<td>13</td>
<td>v</td>
<td>5</td>
<td>5</td>
<td>32</td>
<td>160</td>
<td>160</td>
<td>159</td>
</tr>
<tr>
<td>10</td>
<td>5</td>
<td>t</td>
<td>5</td>
<td>4</td>
<td>48</td>
<td>240</td>
<td>192</td>
<td>210</td>
</tr>
<tr>
<td>11</td>
<td>16</td>
<td>p</td>
<td>5</td>
<td>5</td>
<td>24</td>
<td>120</td>
<td>120</td>
<td>129</td>
</tr>
<tr>
<td>12</td>
<td>15</td>
<td>b</td>
<td>5</td>
<td>5</td>
<td>27</td>
<td>135</td>
<td>135</td>
<td>141</td>
</tr>
<tr>
<td>13</td>
<td>23</td>
<td>c</td>
<td>5</td>
<td>6</td>
<td>17</td>
<td>85</td>
<td>102</td>
<td>100</td>
</tr>
<tr>
<td>14</td>
<td>32</td>
<td>j</td>
<td>5</td>
<td>7</td>
<td>8</td>
<td>40</td>
<td>56</td>
<td>56</td>
</tr>
<tr>
<td>15</td>
<td>35</td>
<td>k</td>
<td>5</td>
<td>9</td>
<td>1</td>
<td>5</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>16</td>
<td>17</td>
<td>l</td>
<td>5</td>
<td>5</td>
<td>22</td>
<td>110</td>
<td>110</td>
<td>121</td>
</tr>
<tr>
<td>17</td>
<td>25</td>
<td>g</td>
<td>5</td>
<td>6</td>
<td>14</td>
<td>70</td>
<td>84</td>
<td>86</td>
</tr>
<tr>
<td>18</td>
<td>6</td>
<td>j</td>
<td>6</td>
<td>4</td>
<td>45</td>
<td>270</td>
<td>180</td>
<td>201</td>
</tr>
<tr>
<td>19</td>
<td>19</td>
<td>w</td>
<td>6</td>
<td>6</td>
<td>20</td>
<td>120</td>
<td>120</td>
<td>113</td>
</tr>
<tr>
<td>20</td>
<td>11</td>
<td>s</td>
<td>6</td>
<td>5</td>
<td>34</td>
<td>204</td>
<td>170</td>
<td>166</td>
</tr>
<tr>
<td>21</td>
<td>20</td>
<td>z</td>
<td>6</td>
<td>5</td>
<td>19</td>
<td>114</td>
<td>95</td>
<td>109</td>
</tr>
<tr>
<td>22</td>
<td>26</td>
<td>t</td>
<td>6</td>
<td>6</td>
<td>14</td>
<td>84</td>
<td>84</td>
<td>86</td>
</tr>
<tr>
<td>23</td>
<td>34</td>
<td>d</td>
<td>6</td>
<td>8</td>
<td>3</td>
<td>18</td>
<td>24</td>
<td>25</td>
</tr>
<tr>
<td>24</td>
<td>10</td>
<td>u</td>
<td>7</td>
<td>5</td>
<td>35</td>
<td>245</td>
<td>175</td>
<td>169</td>
</tr>
<tr>
<td>25</td>
<td>5</td>
<td>o</td>
<td>7</td>
<td>3</td>
<td>92</td>
<td>644</td>
<td>216</td>
<td>317</td>
</tr>
<tr>
<td>26</td>
<td>24</td>
<td>c</td>
<td>7</td>
<td>6</td>
<td>15</td>
<td>105</td>
<td>90</td>
<td>91</td>
</tr>
<tr>
<td>27</td>
<td>36</td>
<td>z</td>
<td>7</td>
<td>10</td>
<td>1</td>
<td>7</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>28</td>
<td>30</td>
<td>w</td>
<td>7</td>
<td>7</td>
<td>9</td>
<td>63</td>
<td>63</td>
<td>61</td>
</tr>
<tr>
<td>29</td>
<td>33</td>
<td>d</td>
<td>7</td>
<td>7</td>
<td>5</td>
<td>35</td>
<td>35</td>
<td>38</td>
</tr>
<tr>
<td>30</td>
<td>21</td>
<td>s</td>
<td>7</td>
<td>6</td>
<td>19</td>
<td>133</td>
<td>144</td>
<td>109</td>
</tr>
<tr>
<td>31</td>
<td>27</td>
<td>t</td>
<td>7</td>
<td>6</td>
<td>13</td>
<td>91</td>
<td>78</td>
<td>81</td>
</tr>
<tr>
<td>32</td>
<td>28</td>
<td>f</td>
<td>7</td>
<td>7</td>
<td>13</td>
<td>91</td>
<td>91</td>
<td>81</td>
</tr>
<tr>
<td>33</td>
<td>37</td>
<td>d</td>
<td>7</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>34</td>
<td>8</td>
<td>l</td>
<td>8</td>
<td>5</td>
<td>43</td>
<td>344</td>
<td>215</td>
<td>195</td>
</tr>
<tr>
<td>35</td>
<td>4</td>
<td>t</td>
<td>8</td>
<td>4</td>
<td>53</td>
<td>424</td>
<td>212</td>
<td>225</td>
</tr>
<tr>
<td>36</td>
<td>2</td>
<td>e</td>
<td>8</td>
<td>4</td>
<td>91</td>
<td>728</td>
<td>364</td>
<td>315</td>
</tr>
<tr>
<td>37</td>
<td>3</td>
<td>a</td>
<td>8</td>
<td>4</td>
<td>90</td>
<td>720</td>
<td>360</td>
<td>312</td>
</tr>
</tbody>
</table>