

Phonostatistical study of Komi Zyryan vowels and consonants

Phonostatistical investigations deal with more and more Finno-Ugric languages. Komi Zyryan is one of the major Finno-Ugric languages, and that is why it too must be studied by phonostatistical methods.

The stories, plays and novels of famous Komi Zyryan writers were analysed statistically by computer. The list of authors included Ivan Toporov (Regyd das Kvajt. Syktyvkar 1975, 323 pages), Gennadij F'odorov (Novels. Syktyvkar 1979, 444 pages), F. V. Šerbakov, A. S. Tarabukin, Z. I. Rogova, I. I. Tarabukin, N. E. Nikulin (Tolyś Vylo. Stories and Plays, Syktyvkar 1979, 108 pages).

The material was transcribed by E. A. Igušev (the author expresses his gratitude for help received to the Dean of the Philological Department of Syktyvkar University, docent E. A. Igušev). The transcription was made according to the accepted Komi Zyryan system of phonemes established by V. I. Lytkin (Lytkin, 1955, 1966). The statistical analysis was made at the Computing Centre of Novosibirsk University (Computer EC-1033). The total volume of the sample was 80 168 vowel and consonant phonemes. This size of sample allowed us to produce statistically valid results (Tambovcev, 1980).

According to V. I. Lytkin's system, the following phonemes are to be found in Komi Zyryan:

1. Vowels *i, j, u, e, ē, o, a*
2. Consonants *p, t, t', k, b, d, d', g, f, s, š, ś, x, z, ž, ź, c, č, č̣, ẓ̌, ẓ̣̌, m, n, n̄, v, l, l', r, j*

Thus there are 36 phonemes: 7 vowels and 29 consonants.

After computing the sample the following ordered series of

Komi Zyryan phonemes was obtained (Table 1). This series was divided into three subseries which were called high, middle and low. Then the absolute mean probability was calculated, which in this case was found to be equal to 0,00278 or 2,78 %. The middle subseries comprised the phonemes whose frequency of occurrence was close to the value of the absolute mean probability: *e* (3,10 %) and *u* (2,76 %). The high subseries comprised the phonemes whose frequency of occurrence was greater than 2,88 % — *a, e, i, s, i, n, k, o, r, t, l, m, d, v, j*. These phonemes can be regarded as more characteristic of Komi Zyryan. The phonemes whose values were less than 2,56 % were found to be in the low subseries, thus — *ś, p, ċ, š, g, b, ŋ, z, l', ž, ž, d', t', ž, č, ž, f, c, x*.

The analysis of Table 1 showed the following regularities in the Komi Zyryan language:

1. The ratio of vowels and consonants is 2:3 (vowels 41,33 %, consonants 58,67 %)
2. The ratio of vowels of middle, high and low height was 8:7:5
3. The ratio of vowels of middle, front and back zone was 18:6:5 or 3,6:1,2:1
4. The ratio of prelingual, mediolingual, labial and postlingual consonants was 6:2:2:1
5. The ratio of the sonants, occlusives and fricatives was 1,6:1,1:1)
6. Komi Zyryan speech is melodic since about 70 % (i.e. 66,84 %) of it is vowels and sonants.

It should be mentioned that Komi Zyryan and Mansi (Tambovcev 1977, 1979, 1980, 1981) have some features in common: they have the same ratio of vowels and consonants (2:3); the order of the sonants, occlusives and fricatives is the same and the ratio of them is very similar also; both languages are very melodic and have a similar frequency of occurrence of vowels and sonants (Mansi 70 % and Komi Zyryan 66,84 %). The most frequent vowel in both languages is /a/. If one takes into account the fact that in Hungarian (Jékel, Papp, 1974), in Nenets (Popova, 1978) and in Selkup (Morev, 1973) the most frequent vowel is also /a/, while in languages of the Slavonic and Germanic families the most frequent vowel is not /a/, but /e/ or /i/ — in Russian (Jolkina, Judina, 1964) it is /i/, while in Czech (Ludvíková, Kōniková, 1967) it is /e/ as well as in Polish (Segal, 1972), German (Kučera,

Monroe, 1968) and Swedish (Fant, 1954) — then perhaps the coincidence of one and the same phoneme /a/ as the most frequent in Komi Zyryan, Mansi, Hungarian, Nenets and Selkup may not be by chance, and some similarity of vowel distribution can be seen in Finno-Ugric and Samoyed languages.

JURI A. TAMBOVCEV

Table 1

Absolute frequency of occurrence of Komi Zyryan vowels and consonants, %

| I | / | II | / | III |
|-------------------|---|-------------------|---|--------------------|
| 1. <i>a</i> 9,71 | | 16. <i>e</i> 3,10 | | 18. <i>ś</i> 2,73 |
| 2. <i>ę</i> 8,93 | | 17. <i>u</i> 2,76 | | 19. <i>p</i> 1,78 |
| 3. <i>j</i> 7,18 | | | | 20. <i>č</i> 1,60 |
| 4. <i>s</i> 6,82 | | | | 21. <i>š</i> 1,19 |
| 5. <i>i</i> 5,26 | | | | 22. <i>g</i> 1,15 |
| 6. <i>n</i> 5,18 | | | | 23. <i>b</i> 1,08 |
| 7. <i>k</i> 4,83 | | | | 24. <i>ń</i> 1,05 |
| 8. <i>o</i> 4,39 | | | | 25. <i>z</i> 0,98 |
| 9. <i>r</i> 4,02 | | | | 26. <i>l'</i> 0,75 |
| 10. <i>t</i> 4,00 | | | | 27. <i>ž</i> 0,73 |
| 11. <i>l</i> 3,79 | | | | 28. <i>ž</i> 0,60 |
| 12. <i>m</i> 3,72 | | | | 29. <i>d'</i> 0,52 |
| 13. <i>d</i> 3,70 | | | | 30. <i>r'</i> 0,43 |
| 14. <i>v</i> 3,65 | | | | 31. <i>ž</i> 0,42 |
| 15. <i>j</i> 3,35 | | | | 32. <i>č</i> 0,30 |
| | | | | 33. <i>ž</i> 0,21 |
| | | | | 34. <i>f</i> 0,04 |
| | | | | 35. <i>c</i> 0,03 |
| | | | | 36. <i>x</i> 0,02 |
| | | | | Σ -80 168 -100 % |

REFERENCES

- Елкина В. Н., Юдина Л. С., Статистика слогов русской речи. Вычислительные системы 10, Новосибирск 1964, с. 61.
- Fant G., *Relativa förekomsten av ord och talljud i svenska språket*. Radiovetenskaplig konferens. Stockholm 1954.
- Jékel P., Papp F., *Ady Endre összes költői műveinek fonémastatisztikája*. Budapest 1974.
- Kučera H., Monroe G. K., *A comparative Quantitative Phonology of Russian, Czech and German*. New York 1968.
- Ludvíková M., Königová M., *Quantitative Research of Graphemes and Phonemes in Czech*. The Prague Bulletin of Mathematical Linguistics, 1967, No 7, pp. 15 — 29.
- Лыткин В. И., Коми-зырянский язык. Языки народов СССР, Том 3, Финно-угорские и самодийские языки. Москва 1966, с. 283 — 284.
- Морев Ю. А., Звуковой строй среднеобского (ласкинского) говора селькупского языка. Томск 1973, с. 4 — 12.
- Попова Я. Н., Фонетические особенности лесного наречия ненецкого языка. Москва 1978, с. 128 — 147.
- Сегал Д. М., Основы фонологической статистики. Москва 1972, с. 137 — 138.
- Тамбовцев Ю. А., Некоторые характеристики распределения фонем мансийского языка. СФУ XIII, 1977.
- Тамбовцев Ю. А., Распределение гласных фонем в мансийской поэзии. СФУ XV, 1979, с. 164 — 167.
- Тамбовцев Ю. А., Частотные характеристики первого слога мансийского языка. Звуковой строй сибирских языков. Новосибирск 1980, с. 72 — 75.
- Тамбовцев Ю. А., Закономерности частотного функционирования долгих и кратких гласных в ударных и неударных слогах мансийского слова. СФУ XVII, 1981, с. 105 — 109.