Anna Korytowska
Nikolai Kopernicus University, Institute of Slavic Philology, Toruń
Poland

VOCALIC CLUSTERS IN THE BALKANIC SLAVIC LANGUAGES

SUMMARY

The paper deals with the typological classification of Slavic languages with respect to the so-called vocalic and consonantal types. Special attention is devoted to vocalic clusters. The vocalic clusters have been analyzed with respect to their position within the word (initial, medial, final; intramorphemic or intermorphemic), with respect to the presence or absence of stress, as well as with respect to word origin (original, loanword). The paper also addresses the notion of intervocalic context and deviations from reduction rules in vocalic clusters (in Bulgarian). The comparison of languages (Macedonian, Serbian, Bulgarian and Polish) is based primarily on text analysis, rather than on the analysis of phoneme inventories. Preliminary results of a larger research that is currently under way are presented.

*Key words: phonetic typology, typology, Slavic languages*

When speaking about the phonetic typology of languages, first of all we speak about the division of languages into the so-called vocalic languages and consonantal ones. The first such classification of the Slavic languages was done by Issačenko (1939).

According to Majewicz (1989:185-186), the author of the well known classification of the world languages, all Slavic languages should be classified as consonantal ones. Majewicz's classification is based on the frequency of vocalic and consonantal segments in the sufficient body of long texts of particular languages. In vocalic languages the ratio between the number of vowels and consonants in texts is higher than 1, in consonantal languages it is smaller than 1, and it is equal to 1 in languages of the balance. Thus, according to this criterion none of the Slavic languages belong to the type of the balance, nor to the vocalic type.

Majewicz rejects the classification based on the number of vowels and consonants in the phonemic inventory. In such a classification those languages which have at least 30 % of vowels in the phonemic inventory are considered vocalic. Milewski divides languages in this way (1962). If we take into account this criterion, the differences among Slavic languages will be quite big, though the number of the vocalic and consonantal phonemes in the system does not correlate always with their textual frequency.

Even if according to the Majewicz's criteria all Slavic languages belong to the consonantal type, there exist differences between particular languages also with regard to the textual frequency of vowels and consonants. One border line goes along Carpates, between Southern Slavic and Northern Slavic languages. There is certain parallelism to the number of vocalic phonemes in phonemic inventories – south from Carpates vowel quantity is preserved as a phonemic feature. Thus in Croatian, Serbian, Slovene, as well as in Czech and Slovak vocalic phonemes are doubled. On the contrary, in Northern Slavic languages, and in Bulgarian, the number of consonants increases thanks to the rich correlation of palatalization and the occurrence of rich consonantal clusters. For example, we have in Polish about 730 of initial and final consonant clusters, whereas in Serbian their number is around 250 (Sawicka 1974, 1996:105-129). However, the differences in the number of vowels and consonants in texts between the Northern and Southern Slavic are not big and, in any case, they are not comparable to those in the phonemic inventories (Figure 1; data from other Slavic languages see Sawicka, to be published).

Figure 1. Frequency of vowels and consonants on the basis of 3 pages of literary texts (white colour indicates vowels).

Slika 1. Učestalost samoglasnika i suglasnika na temelju 3 stranice književnih tekstova (samoglasnici su označeni bijelom bojom)
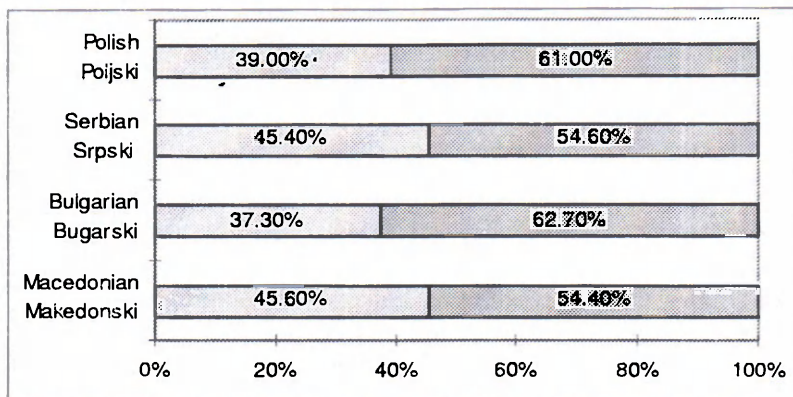
Frequency of vocalic clusters is another feature discriminating Slavic languages. (By vocalic clusters I understand here clusters of syllabic segments). The „Balkanic" languages are distinguished by this factor. First of all Macedonian, and, to the lesser extent, Serbian (and probably Croatian too) belong to this group which is characterized by an extremely high frequency of vocalic clusters, even several times higher than in the rest of Slavic languages. Equally high frequency of VV clusters is typical for Albanian and, probably Greek and Italian dialects. Figure 2 presents the frequency of two-segmental vocalic clusters in Balkanic languages in comparison with Polish.

Figure 2. Frequency of VV clusters on the basis of 3 pages of literary texts

Slika 2. Učestalost samoglasničkih skupina (VV) na temelju 3 stranice književnih tekstova
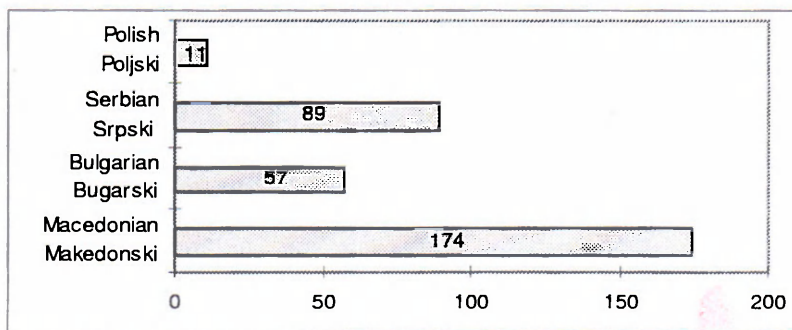
As it can be seen, vocalic clusters occur very often in the Balkanic languages. In this small group of languages certain subgroups can be distinguished with respect to the frequency and quality of vowels. The Macedonian language is certainly the richest one. Subsequent tables show the details concerning vocalic clusters in Balkanic languages. Besides some historical phonetic changes (such as loss of [x], or loss of iota before front vowels), it is morphology that favors the creation of vocalic sequences in Macedonian (see Sawicka & Spasov 1991:47-60). Consequently, all possible combinations of vowels, including the clusters [Vr] occur in Macedonian.

As in the rest of the Slavic languages, foreign lexicon is the rich source of vocalic clusters occurring within morphemes. In the domestic words vocalic sequences are found mainly at morphemic boundaries. Vocalic geminates are frequent in Macedonian. Combination of two identical vowels may result in contraction into one long vowel, but such a realization is not accepted by literary standard. All possible combinations of vowels may occur in the middle of the word in the domestic words, as well as in loanwords. Thanks to the morphology of verbs the inventory of final VV clusters in Macedonian words is quite rich too (from 30 possible combinations of vowels as many as 12 occur word finally). On the other hand, initial VV clusters are not so numerous (only 4 combinations were found from 30 possible). All initial clusters found in the analyzed text were divided by morphemic boundaries occurring between the preposition or prefix and the orthotonic word beginning with a vowel. The [Vr] clusters were found only in the middle of the word, at the same boundaries between the preposition or prefix and the orthotonic word with an initial vowel. Situation in loanwords is a little bit different – initial, as well as final clusters are rare in borrowed words (in *Słownik polsko-macedoński i macedońsko-polski* 1990 I have found only 2 types of final anf 4 types of initial clusters). Besides, the classification according to the place of stress in the VV clusters can be done. The cluster may bear the stress on its first or on its second segment, or it may be unstressed – the last type is the most frequent in Macedonian.

Macedonian clusters containing [i] or [e] are the most difficult to describe. Two facultative types of their realization can be found: with or without more or less distinct intervocalic iota. The problem is the phonemic value of this inserted iota. Neither phonemic nor the morphemic system give an unequivocal solution. Instability appears not only in pronunciation, but even in spelling (cf. *сеалка* but *сејач* with identical realization of the context in question in both cases). Worth mentioning is that in loanwords the insertion of iota before [e] and [i] usually does not take place (cf. [moe], [moʲe] and [moje], but only [poeta] without intervocalic iota) and the process itself appears with various intensity (even in the [Vi] context inserted iota may be very weak).

Table 1.   Vocalic clusters in Macedonian
                 (on the basis of 3 pages of literary texts)
Tablica 1.   Samoglasničke skupine (VV) u makedonskom
                 (na temelju 3 stranice književnih tekstova)

Total number of VV **Ukupan broj VV**	174
Number of VV in domestic words Broj VV u domaćim riječima	154
Number of VV in foreign words Broj VV u stranim riječima	20
Number of VV at morphemic boundaries* Broj VV na granicama morfema*	72
Number of VV within morphemes Broj VV unutar morfema	102
Number of initial VV Broj inicijalnih VV	1
Number of medial VV Broj medijalnih VV	120
Number of final VV Broj finalnih VV	53
Number of vocalic geminates Broj samoglasničkih geminata	47

* I am taking into consideration only such morphemic boundaries which are somehow marked phonetically; which can be therefore treated as phonemic units (junctures). Boundaries after prepositions, prefixes, before articles and boundaries in compound words have such a value in Slavic.

* Uzete su ubzir samo one morfemske granice koje su na neki način obilježene fonetski, pa ih se stoga može tretirati kao fonemske jedinice. Granice poslije prijedloga, predmetaka, prije članova kao i granice unutar složenih riječi imaju takvu vrijednost u slavenskim jezicima.

Serbian is in second place with regard to the frequency of vocalic clusters – it has Balkanic as well as non-Balkanic features. Rich data on vocalic sequences in Serbian are shown in Ivić's article from 1968. The frequency numbers checked on 3 pages of literary text are shown in Table 2. Whereas in the Macedonian text I have found 174 VV clusters – the Serbian text has only 89, but still many more than any of the Northern Slavic languages. The Serbian vowel system is richer than the Macedonian one. Serbian preserves quantity as a phonemic feature – this doubles the number of vocalic units. Besides, among possible combinations of syllabic segments we have also [V r] and [rV] – the former occurs only at morphemic boundaries, cf. [zarzati]; the latter is also possible as optional within morphemes, cf. [grocе] or [groce].

Generally, all combinations of vowels occur in the domestic, as well as in foreign words in Serbian. Restrictions apply only to combinations with the syllabic [r] – such sequences do not occur at all in loanwords; in the domestic words only [ar], [or], [er], [ur] and [ro] occur. All together there occur as many as 30 types of combination of vowels in Serbian. Nearly all of them are found at morphemic boundaries. Neither [Vr] clusters nor vocalic geminates occur within morphemes ([oo] and [uu] are, as it seems, the only exceptions, cf. *zoološki, vakuum*). Also

the [iV] clusters are not found within morphemes. According to the orthoepic rules letters *iV* are realized as threese-gmental, with distinct iota inserted between the two vowels. As in Macedonian all combinations of vowels may appear in word medial position. Initial and final clusters are relatively numerous – the last ones emerge mainly in the inflection of foreign nouns, cf. *bifea, bifeu, bifee*, and also as the result of the shift of syllable final [l] into [o] (*hteo, video, gledao*, etc.).

Table 2.   Vocalic clusters in Serbian
                   (on the basis of 3 pages of literary texts)
Tablica 2.   Samoglasničke skupine (VV) u srpskom
                   (na temelju 3 stranice književnih tekstova)

Total number of VV **Ukupan broj VV**	**89**
Number of VV in domestic words Broj VV u domaćim riječima	82
Number of VV in foreign words Broj VV u stranim riječima	7
Number of VV at morphemic boundaries Broj VV na granicama morfema	40
Number of VV within morphemes Broj VV unutar morfema	49
Number of initial VV Broj inicijalnih VV	16
Number of medial VV Broj medijalnih VV	35
Number of final VV Broj finalnih VV	38
Number of vocalic geminates Broj samoglasničkih geminata	0

Against the non-Balkanic background the Bulgarian language has also quite high frequency of vocalic clusters. I have found 57 clusters in the 3 pages of the Bulgarian literary text. Nearly all possible combinations of vowels (with very few exceptions) occur in the domestic, as well as in foreign words. VV in loanwords never contain [ə]. [ə] is rare also in Bulgarian words where it appears only at morphemic boundaries. All types of vocalic combinations occur in the middle of the words. Final clusters are also numerous – 18 (of the 31 possible) were found in the word final position. Initial clusters are rare. As far as the stress is concerned, clusters with the second vowel stressed and unstressed clusters are the most frequent in Bulgarian. The third pattern with the first vowel stressed is very rare (see Table 3).

According to the Bulgarian standard orthoepic norm, the insertion of iota in all types of vocalic clusters with the second front vowel is not acceptable – it is considered dialectal. However, in the colloquial speech the insertion of iota is quite a frequent phenomenon, cf. *пиано* [pijáno], *вариант* [vʌrijánt], *авиатор* [ʌvijátor], *идиот* [idij'ɔt] (see *Граматика...* 1982: 133).

Table 3.   Vocalic clusters in Bulgarian
               (on the basis of 3 pages of literary texts)
Tablica 3.   Samoglasničke skupine (VV) u bugarskom
               (na temelju 3 stranice književnih tekstova)

Total number of VV **Ukupan broj VV**	**57**
Number of VV in domestic words Broj VV u domaćim riječima	53
Number of VV in foreign words Broj VV u stranim riječima	4
Number of VV at morphemic boundaries Broj VV na granicama morfema	7
Number of VV within morphemes Broj VV unutar morfema	50
Number of initial VV Broj inicijalnih VV	4
Number of medial VV Broj medijalnih VV	39
Number of final VV Broj finalnih VV	14
Number of vocalic geminates Broj samoglasničkih geminata	6

The loss of one element of the cluster is another phenomenon present in the colloquial speech. It happens in foreign words, only within morphemes. More often it is [i] which is lost after palatalization of the preceding consonant, cf. *материал* [mʌterial] > [mʌterjál] > [mʌter‚ál]. Contraction of geminates is also possible in loans, cf. *вакуум* may be pronounced as [vakum], besides [vakuum]. Simplification of any cluster is not possible when the cluster is divided by morphemic boundary, regardless of the origin of the word (foreign or domestic).

The presence of a high vowel in the second position in the cluster favors the loss of syllabicity by this vowel. This regards [i] and [u] in clusters: [ai], [oi], [ei] and [au]. This process can vary in intensity – syllabicity may be only weakened or the vowel may be fully changed into the glide. The nonsyllabic realization of [i] is characteristic of the colloquial speech. It happens in the middle of the word as well as in word final position, cf. [bʌsėjn], [ger'ɔjte], [krájʃte], instead of [bʌsćin], [ger'ɔite], [kráiʃte]. Loss of syllabicity results in the occurrence of double forms even in the standard form, cf. *воин* [vɔin] and [vɔjn] *таинствен* [t'ʌinstven] and [t'ʌjnstven].

De-syllabification of [u] takes place in a number of foreign words, cf. [daxaʋ], [aʋtɔpsja], or in words such as [m‚aʋ], [aʋ].

Another problem connected with the realization of vocalic clusters in Bulgarian is the application of rules of reduction of unstressed vowels into vowels in clusters. The same problem occurs in Russian, where deviations from general rules of vowel reduction were observed (see Zinovieva 1996).

Bulgarian vowel reduction differs from the Russian one. In the Standard form unstressed /a/, /ɔ/ and /ə/ are noticeably reduced, whereas reduction of /e/ is rather weak or even absent. The reduced /ɔ/ does not become centralized as in

Russian – it becomes higher and sometimes even changes into [u]. These rules are often broken in the pronunciation of loanwords containing vocalic clusters, for example in words such as *боа, оазис* all vowels have usually full pronunciation. On the other hand, there are foreign words containing the same vocalic clusters occurring in the same accentual position where vowels in clusters undergo dinstinct reduction, cf. *воал* [voál], *експлоатация* [eksploʌtátsiʌ], *боржоа* [borȝoá], *боржоазен* [borȝoázen] or even [burȝuázen], etc. Double forms also exist even in the realizations of the same speaker, cf. *какаов* [kʌkáof], and [kʌkáuf]. Besides such phenomena as de-syllabification of high vowels, insertion of iota or contraction were observed in realization of vocalic clusters in loans. The described situation has been established on the basis of an experiment: the text containing vowel clusters was read by several Bulgarian speakers.

The situation in the Northern Slavic languages is different than that in the „Balkanic" Slavic. For example, in Polish the frequency of vocalic clusters is very low: in the comparable portion of text only 8 vocalic clusters occurred (which is more than 10 times less than in Serbian and more than 20 times less than in Macedonian). Foreign words and morphemic boundaries are the main source of vocalic clusters in Polish. In Polish, all types of vowel combinations may occur at the morphemic boundaries as well as within morphemes.

Table 4. Vocalic clusters in Polish
(on the basis of 3 pages of literary texts)
Tablica 4. Samoglasničke skupine (VV) u poljskom
(na temelju 3 stranice književnih tekstova)

Total number of VV **Ukupan broj VV**	**8**
Number of VV in domestic words Broj VV u domaćim riječima	8
Number of VV in foreign words Broj VV u stranim riječima	0
Number of VV at morphemic boundaries Broj VV na granicama morfema	8
Number of VV within morphemes Broj VV unutar morfema	0
Number of initial VV Broj inicijalnih VV	0
Number of medial VV Broj medijalnih VV	8
Number of final VV Broj finalnih VV	0
Number of vocalic geminates Broj samoglasničkih geminata	3

The exceptions are very few; they regard certain combinations with /i/ or /y/. Nearly all existing sequences of vowels may appear in the middle of the word. Final and initial clusters are not so numerous (17 final sequences and 15 initial ones against 31 possible combinations). Polish data are presented in the Table 4.

Thus, when we compare the situation in the Balkanic languages with that in the Northern Slavic we can see not only quantitative differences in the occurrence of vowels and consonants, but, first of all, extreme

differences in the frequency of vocalic clusters. Whereas vocalic clusters are more frequent in the South, consonant clusters are more numerous, more complicated and more frequent in the North. The high frequency of vocalic clusters in Serbian, Croatian, Bulgarian and Macedonian is a result of certain domestic historical changes ([l] > [o] in Serbian and Croatian, loss of iota before front vowel in Macedonian and Bulgarian), as well as a result of foreign interference (loss of intervocalic iota in Greek, see Sawicka 1997).

The results obtained constitute one more factor in the evaluation of the phonetic type of a language as more or less vocalic/consonantal. They also allow us to recognize the high frequency of vocalic clusters as a Balkanic feature.

If we take into consideration the extension of the *Balkan Sprachbund*, the bigger difference in the vocalic clusters frequency between Bulgarian and Macedonian than between Serbian and Macedonian seems to be a little bit strange. However other phonetic features go together with this difference. Among others, differently than in Serbian and Macedonian, the Bulgarian language has developed correlation of palatalization – this feature is characteristic of North-Eastern Slavic (see more details in Grzybowski 1997).

Besides the phonetics divisions of the Balkanic languages and dialects do not overlap with those resulting from morphosyntax (see Sawicka 1997).

REFERENCES

Граматика на съвременния български книжовен език, том 1. Фонетика. ред. Д. Тилков (1982). София: Издателство на българската академня науките.

Grzybowski, S. (1997). "Status fonematyczny palatalności w językach słowiańskich". In: I. Sawicka & S. Grzybowski (eds.). *Studia z palatalności w językach słowiańskich*, Toruń: Wydawnictwo Uniwersytetu Mikolaja Kopernika.

Isačenko, A. (1939). Versuch einer Typologie der slavischen Sprachen. *III Medjunarodni Kongres Slavista*, Beograd. 76-80.

Ivić, P. (1968). Vokalske grupe u srpskohrvatskom jeziku. *Slavia Orientalis, Rocznik XVII*, Nr 3, 315-319. Warszawa: PWN.

Majewicz, A. F. (1989). *Języki świata i ich klasyfikowanie.* Warszawa: PWN.

Milewski, T. (1962). *Wstęp do językoznawstwa.* Warszawa: PWN.

Sawicka, I. & Spasov, L. (1991). *Фонологија на современиот македонски јазик.* Скопје: Македонска Академија на Науките и уметностите.

Sawicka, I. (1974). *Struktura grup spółgłoskowych w językach słowiańskich.* Ossolineum.

Sawicka, I. (1996). *Podstawowe zróżnicowanie fonetyczne Słowiańszczyzny. Govor XII*, 1-2, 105-129.

Sawicka, I. (1997). *The Balkan Sprachbund in the Light of Phonetic Features.* Warszawa: Wydawnictwo "Energeia".

Sawicka, I. Z typologii fonetycznej języków słowiańskich – to be published

Zinovieva, T. (1996). Grupy samogłoskowe w języku rosyjskim. *Acta Universitatis Nicolai Copernici, Studia Slavica I.* Nauki Humanistyczno-Spoleczne, zeszyt 311, Toruń.

Sources

1. Dictionaries:

Pianka, W., Topolińska, Z. & Vidoeski, B. (1990). *Słownik macedońsko-polski i polsko-macedoński.* Warszawa-Skopie: PWN, Wydawnictwo „Makedonska Kniga"

Sabol, E. (1997). *Mały słownik języka polskiego.* Warszawa: Wydawnictwo Naukowe PWN.

Frančić, V. (1987). *Słownik serbsko-chorwacko-polski*, t. I-II. Warszawa: PW Wiedza Powszechna.

Попов, Д., Симов, К. & Виднмска, С. (1998) *Речник за правоговор, правопис, пунктуация.* София: Атлантис. КЛ.

2. Literary texts:

Bajac, V. (1992). "Podmetači za snove", Beograd, wyd. Svetovi, Novi Sad, 69-73

Йордан Йоков (1989). По жицата в Деспина Илнева. Български език за студенти, изучаващи български език вън България, Slovenské Pedagogické Nakladatel'stvo, Bratislava, 343-346

Ортаков, Д. (1974). Развојот на македонската соло-песна, [in:] *Предавања на VII семинар за македонски јазик, литература и култура, Скопје и Охрид, 15.VIII-31.VIII.1974.* 172-178, Скопје: Унпверзитетска печатница «Кирил и Методиј».

Żeromski, S. (1995). Wierna rzeka, Kraków, Wydawnictwo Literackie, 5-9

Anna Korytowska
Sveučilište Nikole Kopernika, Institut za slavensku filologiju, Toruń
Poljska

SAMOGLASNIČKE SKUPINE U BALKANSKIM SLAVENSKIM JEZICIMA

SAŽETAK

U članku se govori o samoglasničkim skupinama u slavenskim balkanskim jezicima. Kao što je poznato jedan je od načina fonetske tipološke klasifikacije jezika njihova je podjela na takozvane samoglasničke i suglasničke jezike. Autor prve takve klasifikacije bio je A. Isačenko (1939).

Prema mišljenju Alfreda Majeviča (1989) sve slavenske jezike valja svrstiti u skupinu suglasničkih jezika u kojima je odnos samoglasnika i suglasnika u tekstovima nekog jezika manji od jedan (ako je V/C 1, onda je to suglasnički tip jezika, ako je V/C ~ 1, onda je to uravnotežen tip, a ako je V/C 1, onda je to samoglasnički tip). Ipak, čini se da među slavenskim jezicima postoje razlike koje se odnose na količinu suglasničkih i samoglasničkih elemenata u književnim tekstovima, a i u vezi s brojem samoglasničkih skupina. Prvi kriterij dijeli slavenski svijet na dva dijela - sjeverni i južni, prema drugom iz skupine svih slavenskih jezika izdvajaju se slavenski balkanski jezici u kojima je broj skupina sastavljenih od dva samoglasnika veći nego u ostalim slavenskim jezicima.

U članku se iznose rezultati istraživanja samoglasničkih skupina u makedonskom, bugarskom i srpskom jeziku te rezultati analize književnih tekstova u tim jezicima (ti se rezultati uspoređuju s rezultatima istraživanja poljskog jezika). U opisu su uzete u obzir sljedeće karaktersitike:
- *položaj samoglasničke skupine unutar morfema ili na spoju morfema,*
- *položaj samoglasničke skupine na početku, u sredini ili na kraju riječi,*
- *podrijetlo riječi,*
- *mjesto naglaska.*

Osim toga, u članku se iznose podaci o odnosu suglasničkih i samoglasničkih elemenata u tekstovima danog jezika.

Veliki broj samoglasničkih skupina pojavljuje se prvenstveno u makedonskom jeziku. Njihovom broju pogoduje morfologija i fonološki sustav jezika. Veliki problem su realizacije skupina u kojima se nalaze segmenti /i/ ili /e/, budući da se susreću i fakultativne realizacije sa /j/ ili bez /j/. Pitanje se odnosi i na kvalitetu realizacije glasnika /j/.

Na drugom je mjestu po broju samoglasničkih skupina srpski jezik. U tom jeziku, kao i u svim ostalim slavenskim jezicima, mnoge riječi sadrže skupine

dvaju samoglasnika stranog podrijetla. U srpskom je jeziku relativno mnogo skupina samoglasnika koje se nalaze na kraju riječi. Njihovo pojavljivanje pogoduje deklinaciji posuđenih imenica i prijelazu -l→ -o.

  Znatan je broj glasničkih skupina u bugarskom jeziku. U skladu s književnom izgovornom normom pojavljivanje /j/ između dvaju elemenata u nekim tipovima samoglasničkih skupina predstavlja dijalektalnu crtu tog jezika. Često je gubljenje jednog samoglasnika u skupini ili nastaje prijelaz jednog od samoglasnika u polusamoglasnik. U bugarskom jeziku, kao i u ruskom, primjećuju se odstupanja od načela realizacije nenaglašenih samoglasnika. To je očito u izgovoru stranih riječi. Ponekad se susreću dvije varijante izgovora u jednog govornika: s redukcijom nenaglašenog samoglasnika ili bez njegove redukcije.

  U poljskom jeziku u usporedbi s balkanskim slavenskim jezicima samoglasničkih je skupina znatno manje. One se pojavljuju ili na morfemskim spojevima ili unutar morfema u posuđenicama. I u poljskom je jeziku problematično svajstvo glasnika /j/ koji se pojavljuje između elemenata samoglasničke skupine.

Ključne riječi: fonetska tipologija, tipologija, slavenski jezici