The paper presents an areal classification of the phonological systems of 18 languages of Central and South-East Europe. By means of cluster analysis of phonological specifications of these languages two areal types are obtained: Balto-Balkan type, presented by two areas, and the Central type which is situated between these areas.

After the discovery of phonological features by N.Trubetzkoy and R.Jakobson led to conclude that the distribution of phonological features in different languages was not random. The areal distribution of the phonological features resembled the isolates of lexemes studied in linguistic geography. At the 1st International Conference of Linguists N.Trubetzkoy outlined the task of studying the areal configurations of linguistic features /9/. R.Jakobson presented the first example of such an area (the European Sprachbund or linguistic area) modelled on the basis of compact territorial distribution of languages of different degrees of genealogical relation. The phonological systems of the corresponding languages were characterized by the two following dominant features: palatalization of consonants and monotonous of vowels /4/. Areal linguistics studies show that the phonological system of a language derives not only from its genealogical relations, but is likewise influenced by the neighbouring languages /2, 3, 5, 6, 7/. The development of the phonological systems of a language is conditioned by the language situation in the given region and the spatial structure of the linguistic area.

While the task of finding out the distribution of phonological features is all the more evident, the choice of the guidelines for setting up a basis for defining the areal type of phonological systems, the number of positions of connectedness of a linguistic area) presents serious difficulties /3, 6/. E.g., though the phonological system of Irish is characteristic of the two abovementioned dominant features: the opposition of palatalized/non-palatalized consonants and monotonous of vowels, the Irish language can hardly be recognized as belonging to the Euro-asian language area. Two coincidences of features are not enough to decide whether the language belongs to a definite type /against: 3/. The Balkan language area (like other currently defined language areas) does not hold such features that might be regarded as necessary and sufficient in the traditional sense. It is frequently noted in areal typology that every feature defined as specific for a linguistic area can also be found in a language beyond the area /3, 4/. On the other hand, some languages belonging to the area mean no less features than described as specific. Essentially it is a high degree of similarity characterizes the languages belonging to the same domain of connectedness of a linguistic area /1, 2, 3, 6, 7/. The setting up of abstract ideal types to serve as a basis for a quantitative evaluation of real language systems /3/ seems to be of little value: the problem of constructing ideal types is to be used for a further description of the language systems of a given area. A high degree of similarity characterizes the languages belonging to a given area however is frequently due to non-exclusive intersecting features, that do not form a definite areal type.

This paper presents an attempt of suggesting areal types of phonological systems on the basis of the languages of the Balto-Balkan area.

A transparent procedure of construction, controlled objectivity on all stages and unambiguity of results are essential for setting up areal types of phonological features. Due to the abovementioned peculiarities of areal language relations the traditional Aristotelian classification is believed to be of little value. In the present case a quantitative approach seems preferable. All techniques of cluster analysis employed for the purpose of obtaining groups of objects characterized by a maximum degree of similarity, we have considered the Linker algorithm as best suited for our needs /8/. The Linker algorithm gives hierarchically arranged object clusters and determines the relative degree of similarity by which the objects are clustered. Generally the algorithm proceeds as follows: Consequently if the data matrix has inherent structure clusters constituting a linkage will succeed in identifying it. The result of the Linker algorithm is unambiguous only if the degrees of similarity (of the mean degrees of distance).

The algorithm can be presented in the form of a dendrogram mapping the sequence of elements and groups thereby showing the minimal distances at which the clusterings take place.

The languages considered present the following picture of language groups marked by an increasing degree of similarity in paradigmatic phonology (Fig. 1).

The next cluster is formed by Check, Slovak, Serbo-Croatian and Slovenian, while the following step adds Hungarian to the cluster. Last the group including Latvian and Lithuanian is added. It should be noted however, that the distance between the initial language cluster (Bulgarian-Macedonian) and the Greek is much greater than the distance between the languages of the Balkan group, both language clusters are united into one which is globally opposed to another cluster formed by consequent joining of-

\[ d = 1 - \frac{k}{n} \]

\[ 0 < 1 - (\alpha + \beta + \gamma) \]

\[ 0 < 1 - (\alpha + \beta + \gamma) \]

\[ 0 < 1 - (\alpha + \beta + \gamma) \]

\[ 0 < 1 - (\alpha + \beta + \gamma) \]

\[ 0 < 1 - (\alpha + \beta + \gamma) \]

\[ 0 < 1 - (\alpha + \beta + \gamma) \]
and marked by different hatching. In the altitude of the corresponding territories similarity of languages can be shown as an similarity of languages can be represented on the map as a structure of the correspond-
ging area. Thus the languages belonging to the 2nd. group
demonstrate a high degree of Overall similarity, while they do not necessarily contain specific features. As specific features characterizing areaal types tend to be absent, the procedure of establishing the types is believed to be rather objective.

REFERENCES

The features of syntagmatic phonology as well as features of other language levels should be likewise included in the data matrix. The interpretation of areaal types thus obtained will yield the best results when an increased degree of approximation of the global characteristics of the language does not change the output of the algorithm.

CONCLUSIONS

1. From the point of view of paradigmatic phonology the Balto-Balkanic region is a single relatively connected linguistic area with two language areas: a marginal Balto- Balkanic area proper and a central, compact area, separating the two territories of the Balto-Balkanic language area.

2. A quantitative analysis of the degree of similarity characterizing the phonological systems of a given area helps to establish the systems of maximum similarity, the hierarchy of their clustering and the optimum of the borderlines between the obtained areaal types.

3. The areaal types obtained as a result of the cluster analysis described in the present paper show a high degree of overall similarity, while they do not necessarily contain specific features. As specific features characterizing areaal types tend to be absent, the procedure of establishing the types is believed to be rather objective.