0.0 Phonological typologies, statistical counts and mathematical models

The high structuring of phonology, the obvious classes of sounds, and the classes of their classes, have made phonological typologies a not too rare proposal. And even where typologies were not claimed as such, they were often implicit in the statements made. Both phonetic and phonemic, acoustic and articulatory, structural and non-structural, have all been proposed and have evoked discussions, critiques and applications.

One can mention works by the Prague linguists culminating in the writings of, among others, Skalička, Krámsky, and notably Trubetzkoy and Jakobson. In America, we have work by Greenberg, Hockett, Saporta and Voegelin as well as numerous followers and critics. Among other European contributions the acoustic typologies by Menzerath and Meyer-Eppler.

0.1 Mathematical models and mathematical (more precisely statistical) techniques of analysis have also been elaborated. Classification, distribution and frequency characteristics of various sound patterns have been a particular concern and represent the bulk of numerical phonological typologies, especially in the U.S.[1] Often, again, the subject is classification of inventories and particular types (articulatory) of phonemes. Such is, for instance, Pierce's

"A Statistical Study of New World Consonants", with counts of from the most to the least common consonants and classes of consonants in a great number of Amerindian languages. A critique and evaluation is found in Saporta 1957. We have probably the least structural end of the typological spectrum here.

2.0 Distinctive feature typologies

There is another series of phonological typologies based on Jakobson's distinctive feature analysis and of course their Praguian and particularly their Trubetzkoyan background.[2]

2.0 The distinctive feature indices of Andrej Avram

Typological indices derived from Jakobsonian features were proposed by Andrej Avram (1961), to study the distribution of distinctive features in the phonological system.

Avram proposes the following indices (if I understand him well):

P = number of phonemes

T = number of features

D = number of features with which each feature combines

D_m = average distribution of distinctive features

R = coverage of each feature (how many phonemes it characterizes)

Rm = coverage of the system as such (:R/T)

C = complexity of a phoneme, i.e. how many features it is a bundle of

C_m = average complexity of the system (=ΣC/P)

E = efficiency of a system; E=P/T

Avram also includes maxima and minima for D_m, R_m, C_m.

2.1 His typology was a good first step for studying systems via their feature distribution. However, there are cases, such as the following two hypothetical systems, different (but of undecided distance), which Avram's typology fails to distinguish.[3]

System A

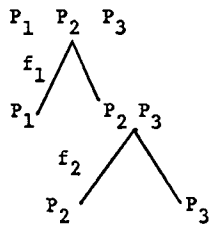
Phonemes: P_1 P_2 P_3

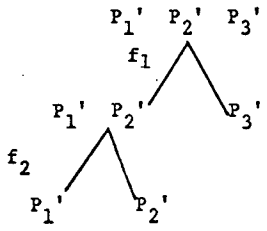
System A'

Phonemes: P_1' P_2' P_3'

Distinctive features

f_1 f_2





	P_1	P_2	P_3
f_1	+	−	−
f_2	0	+	−

	P_1'	P_2'	P_3'
f_1	+	+	−
f_2	+	−	0

Indices:

$P = 3$	$P' = 3$
$T = 4$	$T' = 4$
$D_m = 1$ $(Dm = \frac{D(f_1) + D(-f_1) + D(-f_2) + D(-f_2)}{4}$	$D'_m = 1$
$Rm = 1$ or 33.33%	$R'_m = 1$ or 33.33%
$E = P/T = 3/4 = .75$	$E' = P'/T' = 3/4 = .75$
$Cm = (1 + 2 + 2)/3 = 5/3 = 1.66$	$C'm = (2+2+1)/3 = 5/3 = 1.66$

It is obvious after inspection of the indices, that the two systems are not distinguished. Such a loss of information is characteristic of averaging.

3.0 Postovalova's valence and probability indices

More complicated yet much more adequate measures of distinctive feature distributions were proposed by the Soviet linguist Postovalova. Although they were first used for the study of just one system, typological applications were also suggested by the author.

Postovalova's paper in Problemy Linguističeskogo Analiza examines the subject of feature distribution in a phonological system. Several statistics are defined:

3.1 simple probability, i.e. frequency of utilization of a feature by the system for phoneme composition; the three possibilities, +, −, 0 are considered separately.

For feature a we have:

$$Pa = \frac{m}{n}$$

where m = number of rows with a
n = number of rows.

3.2 conditional probabilities, indicating how frequently different pairs of features characterize phonemes in the system, which is to say, given feature a for a phoneme what are the chances that feature b will combine with a in the same phoneme.

Given feature a the probability that b will combine with it for the same phoneme is:

$$P_a(b) = \frac{m'}{n'}$$

where m' = number of phonemes with a and b
n' = number of phonemes with a .

3.3 Finally, Valence is defined by Postovalova to disclose information on a feature's combinability with the other features and also information on the system as a whole (by including total number of features).

$$V_a(b) = \frac{Pa(b)}{n - 1}$$

where n = total number of distinctive features in the system.

(The probability of a feature appearing combined with another feature would be $\frac{1}{n-1}$ if all features were equiprobable.)

3.4 Illustration from the hypothetical case of 2.1 (The Valence proposed by Postovalova is modified by Afendras (1968), so that it appears as a feature by feature matrix: this step is very important as it makes comparison across languages a matter of comparison of features drawn from the "Universal" system, rather than comparison between vowels and their features.)

Valences

System A:

		f_1 +	f_1 −	f_1 0	f_2 +	f_2 −	f_2 0
f_1	+				0.	0.	1.
	−				.5	.5	0
	0				0.	0.	0.
f_2	+	0.	1.	0.			
	−	0.	1.	0.			
	0	1.	0.	0.			

System A':

		f_1 +	f_1 −	f_1 0	f_2 +	f_2 −	f_2 0
f_1	+				.5	.5	0.
	−				0.	0.	1.
	0				0.	0.	0.
f_2	+	1.	0.	0.			
	−	1.	0.	0.			
	0	0.	1.	0.			

Clearly, the two hypothetical systems are strongly distinguished.

4.0 Valence analysis of Balkan vocalic systems

And now an application of this quantitative typology to a specific problem: the Balkan linguistic convergence area. Non-phonological aspects have been thoroughly investigated, in the classic treatment by Sandfeld (1930) and most recently in some powerful typological

studies (Kazazis, Civ'jan, Birnbaum in several articles, Klagstadt, etc.) Balkan phonology has prompted many comments by Jakobson, Ivić and others, but to my knowledge only one systematic study (Havránek, 1933) which actually drew heavy criticism (Małecki Stankiewicz). Interesting results were obtained by applying the above method to the study of several Balkan idioms.[4]

4.1 But before discussing the results some of the basic problems encountered will be mentioned:

The systems were compared against a maximal matrix which included all the features occurring in the population of the systems analyzed.[5]

Any of the actual systems include a subset of this maximal set of features. In the final correlation each system was considered as having 0's throughout for the features which it did not utilize. But 0's were also indicative of impertinence of a feature for a given phoneme when the feature was distinctive for other phonemes in the system. Thus two kinds of concepts were collapsed as they both were represented by 0. However, this has probably been rectified by the fact that features not used in a system have a 0 throughout.

4.2 Another actual handicap is the non-availability of distinctive feature descriptions for the vast majority of the systems compared. And even when available, they were often tinted by both the author's views and his preferences (e.g. Petrovici on Rumanian) or were out of different periods of theoretical development of distinctive features.

In such cases, I took the liberty of normalizing the data by modifying the existing analyses (:the same method was followed throughout e.g. constructing branching-trees). In some other instances more than one solution were possible and for lack of data I kept the alternatives. Such systems appear in the figures as language X-1, 2, 3 etc. Other instances of numbered, multiple systems for one language refer to situations where such variety actually exists either stylistically or in social dialects (e.g. literary Makedonski).

4.3 Some features are very typical of vocalic systems either universally, or for the European languages, or, more specifically, for the Balkan languages. E.G. diffuseness, gravity, flatness, stress (simple occurrence is considered here, not combinations). Then, other features s.a. length, tone, nasality etc. are much less common. An ideal comparison should give different weights to such features. Sharing nasality, for instance, should be typologically very significant and two systems which do, should be classed as very similar. Conversely, if in a group of many languages which draw on 5-6 features to distinguish their vowels, but usually have 3 or 4, only one uses nasality this should be significant enough to set this particular language quite far apart. Now, in the correlation some factors take this into account but indirectly and not sufficiently. On the other hand since in reality (i.e. in the Balkan case) systems having "odd" features have also the "common" features, their typological distance is reflected in their having a higher number of features than the other languages,

a fact reflected in the Valence matrix (:lower values for each cell).[6]

4.4 The introduction of a new feature usually results in a whole series of new phonemes, and actually the more numerous these phonemes the more important the new feature to the system. This is expressed by the product of number of distinctive features X number of phonemes = total # of cells in the feature by phoneme matrix. An index incorporating this will reflect more qualities of the whole system. I propose therefore, tentatively, a modification of the valence formula to:

$$V'_a(b) = \frac{Pa(b)}{K(n-1)}$$ where K = # of phonemes

This weighting makes the index much more sensitive to variations in the number of features.

4.5 Higher order conditional probabilities can also be introduced, e.g.

$$P_{ab}(c) = \frac{K}{m}$$

where K = number of phonemes which have in common features a, b and c, and
m = number of phonemes with a and b in common.

A Valence $Vab(c) = \frac{Pab(c)}{(n-1)}$ can then be defined.

And so on until we have the $P_{1\ \ldots\ (n-2)(n-1)}$ and the resulting Valence.

The results presented here are based on Postovalova's original formula.

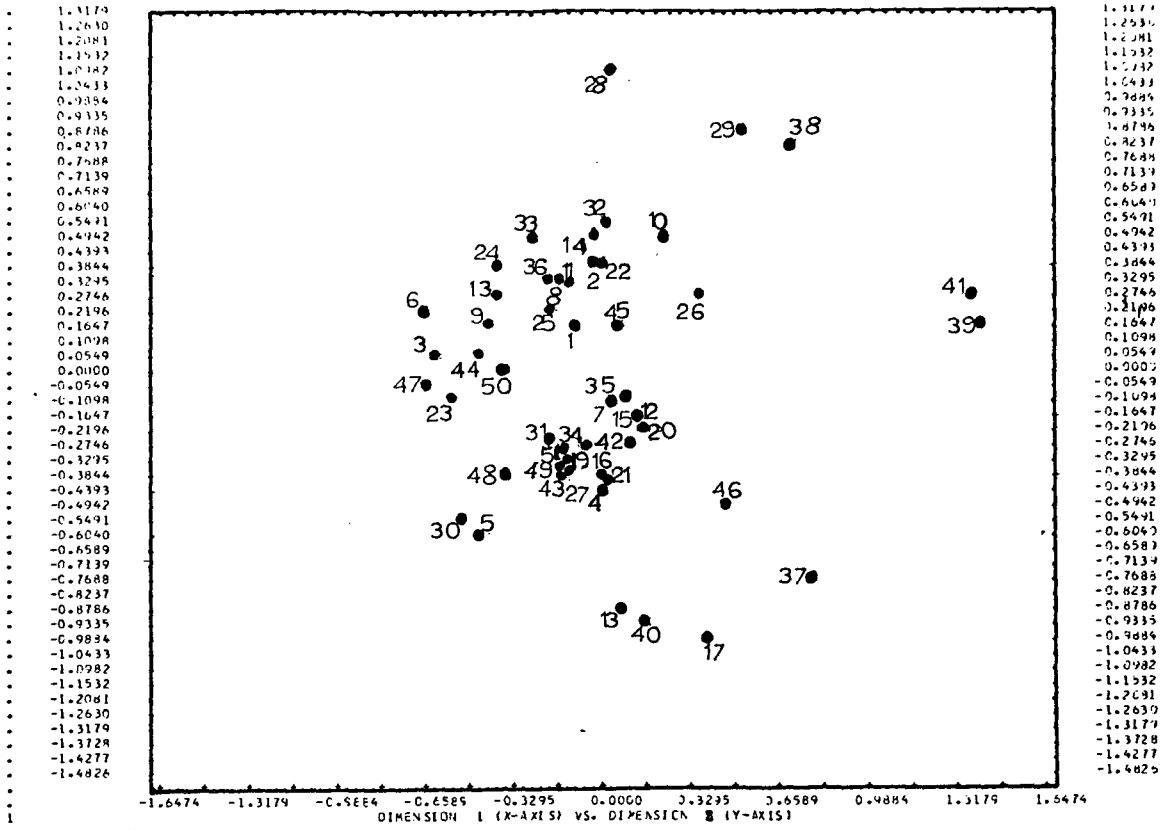


Fig. 1. KRUSKAL Multidimensional Scaling Program, Computer Plot: Clustering of Balkan Vocalic Systems based on Distinctive Feature Valence.

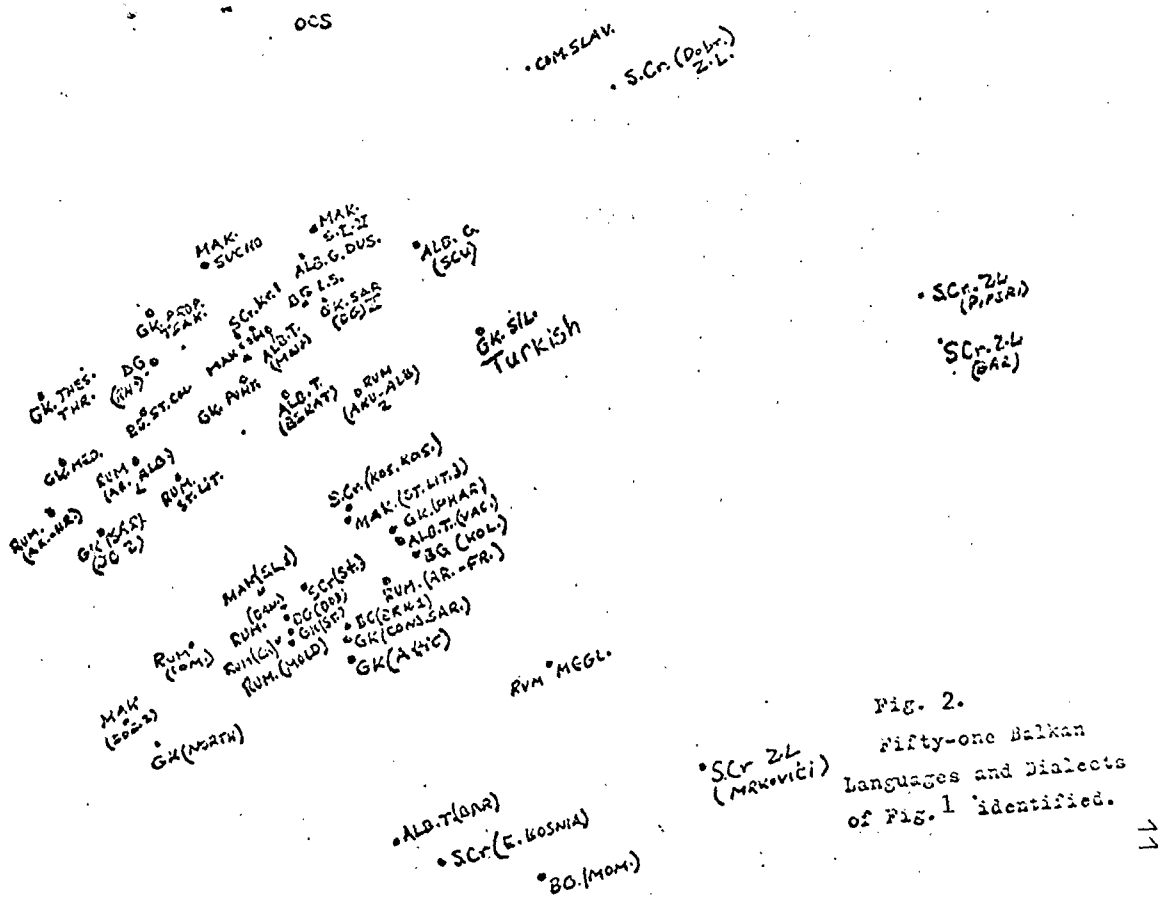


Fig. 2.
Fifty-one Balkan Languages and Dialects of Fig. 1 identified.

5.0 Statistical correlation of Balkan Valence matrices

51 vocalic system matrices were actually analyzed, their P_a, $P_a(b)$ and $V_a(b)$ matrices calculated, and these final matrices were correlated and plotted using two different methods according to distance from each other.[7]

Gammon (1967) used a similar statistical technique for finding the similarity (conversely, the distance) between several Polynesian languages. In his case, the information was not in the form of matrices but in the form of lists.

6.0 Results

In the final plotting (see figures 1, 2, 3 and 4) several groupings can be discerned.

First, in the multidimensional scaling analysis, we can speak roughly of 3 groups: two form a sort of a nucleus in the center and the third (distributed in two subgroups also) surrounds it. On the external group we have Old Church Slavic, and close by Common Slavic. Then spread around mostly Serbo-croatian dialects, with some other idioms (e.g. 7-vowel Momčilovci Bulgarian, (9+Nasal) vowels Barile Tosk Albanian, 8-vowel Meglenitic, 7-vowel E. Bosnia Serbo-croatian). This group seems to include only dialects with the feature of tenseness. Except for the Albanian dialect, all group members are dialects located

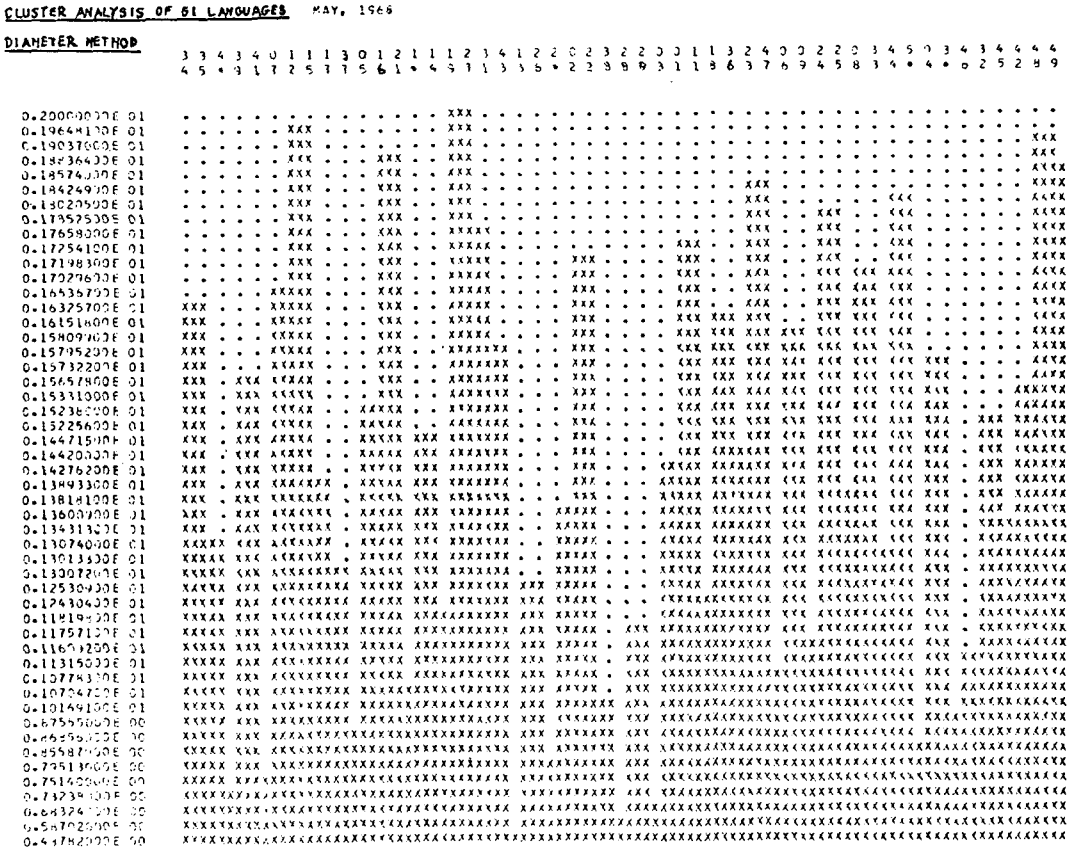
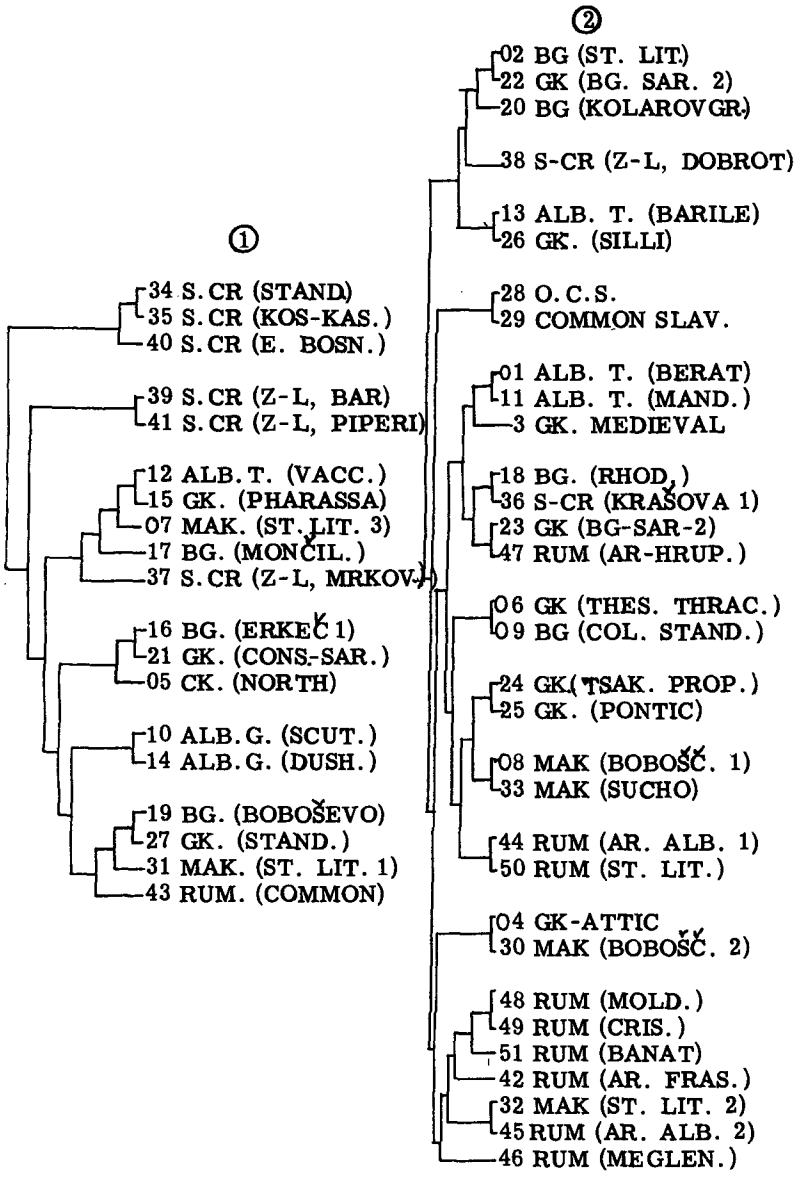


Fig.3 . Diameter Cluster Analysis of Valences: Computer Plot.

Fig. 4. Diameter Cluster Analysis:
Rearranged from Fig. 3.



See Appendix A for list of idioms on this table.

in the Central Balkans, i.e. we have here an areal grouping.

Four out of six Albanian systems (all of the dialects within the geographic area of the Balkans) fall in one group and are closer to Macedonian and Bulgarian than to Rumanian.

Greek dialects are quite diffused but stay within the two nuclear groups (this includes Classical Attic and Medieval Greek.)

In the diameter method (figure 4), the most interesting grouping is that of all seven systems with the 5-vowel pattern (irrespective of additional features such as length - nasality - tone) on the same side of the initial bifurcation.

Another subgrouping includes only systems which use flatness to distinguish a second series of back vowels (Rhodope Bulgarian, Krašova I Serbo-croatian, Bulgarian Sarakačan Greek, and Hrupišta Arumanian).

Among the other interesting results are the following: contrary to Old Church Slavic and Common Slavic, which in both analyses stand apart, Attic Greek is classed as very similar to the bulk of the systems analyzed: in the diameter method through early joining of the branching tree, and in the multidimensional sealing by being located in the middle of the one of the two central constellations. This is all the more intriguing in the light of the fact that in a similar statistical correlation (see figure 5) of some Balkan and other

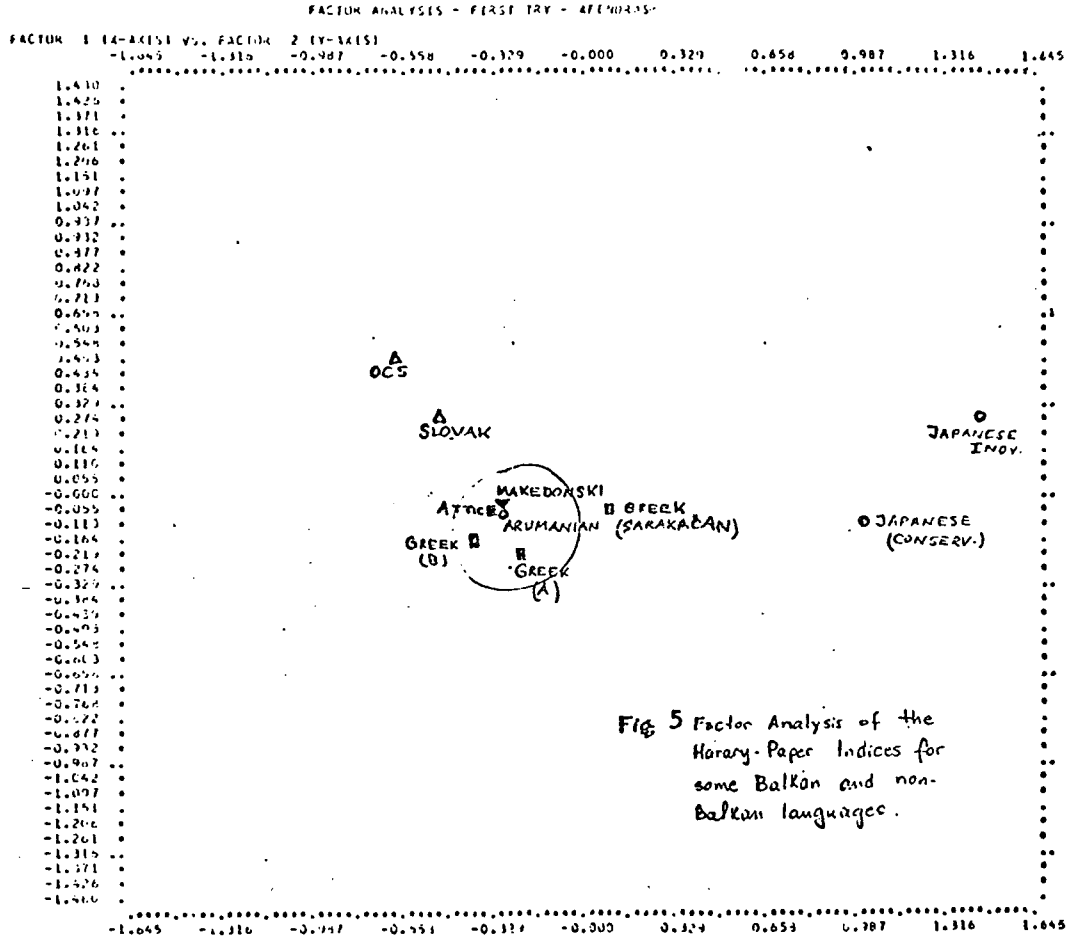


Fig 5 Factor Analysis of the Horary-Paper Indices for some Balkan and non-Balkan languages.

systems expressed in the Harary-Paper model for binary phoneme combinations. Attic was also in the center of the group and OCS (as well as Czech) quite distant from it. Is this not a strong indication of convergence of South Slavic towards a pattern characteristic of Greek since the time of Attic? If anything, one would intuitively class the old idioms (Common Slavic, OCS and Attic) as belonging to roughly the same types: large systems, length, tone, accent etc. as additional features. What "latent" structures are responsible for the outcome of the statistical analysis?

7.0 Some Conclusions

It is hoped that the advantages, possible uses and problems of this approach for typological classification as well as some uses of the latter, were sufficiently demonstrated. It is also hoped that criticism and suggestions on all aspects will further the utility and vigor of the approach.[9] Among the main objectives is a better anchoring of the model on mathematical theory and, parallel to this, better explanation of the linguistics behind the findings and in adapting the model. While this last is obvious, the search for a model mathematically well grounded is imperative if the approach is to profit more from the power of a fully developed mathematical theory.[10] As a concrete step in this direction, a Stochastic Process Model for the same problem is currently investigated, as well as other models applied to different aspects of phonologic structure.

Notes

1) See bibliography for particular works on typology.

On mathematical models especially pertinent is the work of Spang-Hanssen, Harary and Paper, Ungeheuer, and more recently a series of studies by Soviet Linguists e.g. Revzin; the work of other Soviet and East European linguists (Šaumjan, Marcus, Kulagina etc.) is very interesting and could be applied to typological investigations. For an overview see Kiefer, 1968.

On statistical techniques, Pierce (1957, 1962), Saporta (1957), Spang-Hanssen, Herdan; in Eastern Europe extensive work, of which one can mention Krámsky who typifies some of the Prague research and Andreyev's group in the Soviet Union.

For general discussions, see Birnbaum 1966, 1968, 1969, Edmundson 1967, Greenberg 1957, Horne 1966, Kučera and Monroe 1968, Plath 1963, Spang-Hanssen 1961-1964, Uspenskij 1966.

For a historical-phonetic typology see Grimes 1961.

2) Of course Jakobson's work as well as that of Trubetzkoy is of a typological nature. Such is for instance Trubetzkoy's *Grundzüge der Phonologie* as well as his articles on vocalic or consonantal systems. For a discussion of the typological nature of Jakobson, Fant and Halle's *Preliminaries to Speech Analysis*, see Voegelin 1956.

Notes (cont.)

3) It seems that no two "mirror" systems could be distinguished by this typology. A theoretical shortcoming in spite of the fact that not terribly many such cases exist.

4) Much of the data analysis used for the present paper was done at the Johns Hopkins University, as part of my doctoral research which culminated in a thesis (May 1968).

5) The idea of a maximal system thus defined can also be found in Voegelin 1963. Uspenskij (1965:63) defines it as follows, opposite to a minimal system: "Jazyk-étalon pervogo tipa (i.e. minimal) možno ponimat' kak teoretiko-množestvennoje proizvedenije vsex xarakterizujemyx (v opredelennom aspekte) jazykov (modelej), t.e. kak invariantnuju dlja vsex étix jazykov model'; jazyk-étalon vtorogo tipa (i.e. maximal) možno ponimat' kak teoretiko-množestvennuju summu usex priznakov opisyvaemyx jazykov (modelej). *Pri étom v kačestve točki otsčeta pri tipologičeskix sravnenijax dolžen ispol'zovat'sja jazyk-étalon minimal'nogo tipa; tem samym jazyk-étalon étogo tipa možet čitat'sja osnovnym.* Uspenskij points out that the minimal system in a sense "catches the essence" of the languages to be compared. See also the Ph.D. dissertation by Afendras, The Balkans as a Linguistic Area:

Notes (cont.)

A Study in Phonological Convergence. Baltimore, 1968, § 3.4, 3.5 (101-112), 4.9-4.10 (139-140), and Ch. 5 (141-152) for establishing and discussion of the maximal and minimal vocalic and consonantal systems in the Balkans.

Birnbaum (1966:20) in his discussion of Uspenskij above also expounds on the notions of maximal (:Boolean sum) and minimal (Boolean product) typological systems.

Actually, much the same is implicit in some of the American typologies (of course Voegelin's, mentioned above) for instance. Pierce's, with its "omnipresent" consonants, the basic core = minimum, and total collection of any consonants occurring in at least one language = maximum.

6) See, for instance, in the statistical correlation (KRUSKAL Multidimensional Scaling) OCS, Common Slavic stand quite apart from almost all of the other languages.

7) See appendix for a list of all the idioms analyzed.

8) Reproduced from Afendras (1968:145) fig. 10.

9) Parallel attempts, or rather converging attempts from other directions might also suggest improvements or better support

Notes (cont.)

our findings.

Grimes (1962), for instance, analyzes phonetic divergence (: "scatter") within Romance and finds French and Rumanian display "high scatter" from expected innovations, therefore distances from the rest. "It is tempting to guess that the scatter in Rumanian could reflect the influence of non Romance speech communities that have interacted with the Rumanian community (or communities)".

Now since in our analysis Rumanian is quite close to the other Balkan languages (unlike, for instance Serbocroatian) one could say that Grimes' study and this present complementary and mutually supporting results.

10) Edmundson's section on mathematical models in Borko 1967 provides a starting point with the must and must nots of the researcher, the does and the does nots of the model.

SUMMARY

In this study, the vocalic systems of a large number of Balkan idioms (past and present) were analyzed in terms of Jakobsonian distinctive features. Various methods for comparison and scaling for similarity, as well as the problems encountered, are discussed and evaluated. Some questions of typology, such as distinctive feature weights, are revealed; suggestions are made for their future incorporation into typologies of this nature.

It is a surprising fact in linguistic scholarship that no feasible, nor adequate manner for comparing phonological systems quantitatively has been devised. The notion of distinctive feature valence proposed by the Russian linguist, V. I. Postovalova, answers the need for such a feature distribution measure.

The valence matrices for the vocalic systems of fifty-one Balkan idioms, as well as simple and joint probabilities of distinctive feature occurrence are calculated. Finally, the results are correlated and submitted to computerized factor analysis (various programs).

APPENDIX A

A LIST OF ALL BALKAN IDIOMS ANALYZED

ALBANIAN

1. Dushmani Geg
2. Scutari Geg
3. Berat Tosc
4. Mandres Tosc
5. Barile (Italy) Tosc
6. Vaccarizzo (Italy) Tosc

BULGARIAN

7. Colloquial Standard
8. Literary Standard
9. Bobošev̆o
10. Erkeč̆ (2)
11. Kolarovgrad
12. Momč̆ilovci
13. Rhodope

GREEK

14. Standard; also most JUDEO-SPANISH
15. Conservative Sarakač̆an
16. North including Sarakač̆an; some JUDEO-SPANISH
17. Sarakač̆an of Bulgaria I

APPENDIX A (cont.)

GREEK (cont.)

18. Sarakačan of Bulgaria II
19. Propontis Tsakonian
20. Thessalian, Thracian
21. Micrasiatic - Pontus
22. Micrasiatic - Pharassa
23. Micrasiatic - Silli, also TURKISH
24. Early medieval
25. Attic Classical

MAKEDONSKI

26. Standard Literary I
27. Standard Literary II (Regional-Stylistic variant)
28. Standard Literary III (Regional-Stylistic variant)
29. Sucho
30. Boboščica I
31. Boboščica II (alternative phonemicization)

32. OLD CHURCH SLAVONIC

33. COMMON SLAVIC

34. COMMON RUMANIAN

APPENDIX A (cont.)

RUMANIAN

35. Standard
36. Banat
37. Crişean
38. Moldavian
39. Fraşeri Arumanian
40. Other Arumanian of Albania I
41. Other Arumanian of Albania II (alternative phonemicization)
42. Hrupişta Arumanian
43. Meglenitic

SERBOCROATIAN

44. Standard Literary
45. E. Bosnia Štokavian
46. Kosovo-Resava (Kasidol)
47. Zeta-Lovčen (Mrkovići)
48. Zeta-Lovčen (Piperi)
49. Zeta-Lovčen (Dobrota)
50. Zeta-Lovčen (Bar)
51. Ě-dialects (Banat) Krašova

APPENDIX B

1) SAMPLE DISTINCTIVE FEATURE MATRICES

(GREEK AND MAKEDONSKI DIALECTS)

2) SAMPLE VALENCE MATRICES FOR TWO OF THE ABOVE

(adapted from Afendras 1968)

ORIGINAL DISTINCTIVE FEATURE MATRIX OF GREEK (STANDARD)

	ˊ ˊ ˊ ˊ ˊ ieaouieaou	
DIFFUSE/NON. DIFF.	+---+---+	Np = 10 Nf = 4
COMPACT/NON. COMP.	0-+-00-+-0	
GRAVE/ACUTE	--0+-0+-	
STRESSED/UNSTRES.	+++++-----	

ORIGINAL DISTINCTIVE FEATURE MATRIX OF N. GREEK (CONS. SAR.)

	ˊ ˊ ˊ ˊ ˊ ˊ iueoæaiua	
DIFFUSE/NON DIFF.	+-----+-	Np = 9 Nf = 4
COMPACT/NON. COMP.	00---+00-	
GRAVE/ACUTE	-+-+-+-+-	
STRESSED/UNSTRES.	++000+---	

ORIGINAL DISTINCTIVE FEATURE MATRIX OF N. GREEK (GENERAL)

	ˊ ˊ ˊ ˊ ˊ ieaouiua	
DIFFUSE/NON. DIFF.	+------+-	Np = 8 Nf = 4
GRAVE/ACUTE	--0+--+0	
COMPACT/NON. COMP.	0-+-000+	
STRESSED/UNSTRES.	+0+0+---	

ORIGINAL DISTINCTIVE FEATURE MATRIX OF MAKEDONSKI (ST. LIT. I)

	ieaou	
DIFFUSE/NON. DIFF.	+---+	Np = 6 Nf = 3
COMPACT/NON. COMP.	0--+0	
GRAVE/ACUTE	--0++	

ORIGINAL DISTINCTIVE FEATURE MATRIX OF MAKEDONSKI (ST. LIT. II)

	iueoæa	
DIFFUSE/NON DIFF.	+++---	Np = 5 Nf = 3
GRAVE/ACUTE	-+-+-+	
FLAT/PLAIN	0-+0-+	

ORIGINAL DISTINCTIVE FEATURE MATRIX OF MAKEDONSKI (ST. LIT. III)

	iueoæa	
DIFFUSE/NON.DIFF.	+++---	Np = 6 Nf = 3
GRAVE/ACUTE	-+-+-+	
COMPACT/NON. COMP.	00--++	

VALENCE MATRIX, PAGE 1. GREEK (STANDARD)

		1 +	1 0	1 -	1 +
(1) DIFFUSE/NON. DIFF.	+	0.00	0.00	0.00	0.00
	0	0.00	0.00	0.00	0.00
	-	0.00	0.00	0.00	33.33
(2) COMPACT/NON. COMP.	+	0.00	0.00	11.11	0.00
	0	33.33	0.00	0.00	0.00
	-	0.00	0.00	22.22	0.00
(3) GRAVE/ACUTE	+	16.67	0.00	11.11	0.00
	0	0.00	0.00	11.11	33.33
	-	16.67	0.00	11.11	0.00
(4) STRESSED/UNSTRES.	+	16.67	0.00	16.67	16.67
	0	0.00	0.00	0.00	0.00
	-	16.67	0.00	16.67	16.67

2 0	2 -	2 +	3 0	3 -	3 +	4 0	4 -
0.00	0.00	16.67	0.00	16.67	13.33	0.00	13.33
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	16.67	33.33	16.67	20.00	0.00	20.00
0.00	0.00	0.00	33.33	0.00	6.67	0.00	6.67
33.33	0.00	16.67	0.00	16.67	13.33	0.00	13.33
0.00	33.33	16.67	0.00	16.67	13.33	0.00	13.33
16.67	16.67	0.00	0.00	0.00	13.33	0.00	13.33
0.00	0.00	0.00	0.00	0.00	6.67	0.00	6.67
16.67	16.67	0.00	0.00	0.00	13.33	0.00	13.33
16.67	16.67	16.67	16.67	16.67	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
16.67	16.67	16.67	16.67	16.67	0.00	0.00	0.00

VALENCE MATRIX, PAGE 1. MAKEDONSKI (ST. LIT. I)

		1			2			3	
		+	0	-	+	0	-	+	0	-
(1) DIFFUSE/NON. DIFF.	+	0.00	0.00	0.00	0.00	50.00	0.00	25.00	0.00	25.00
	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	-	0.00	0.00	0.00	50.00	0.00	50.00	25.00	50.00	25.00
(2) COMPACT/NON. COMP.	+	0.00	0.00	16.67	0.00	0.00	0.00	0.00	50.00	0.00
	0	50.00	0.00	0.00	0.00	0.00	0.00	25.00	0.00	25.00
	-	0.00	0.00	33.33	0.00	0.00	0.00	25.00	0.00	25.00
(3) GRAVE/ACUTE	+	25.00	0.00	16.67	0.00	25.00	25.00	0.00	0.00	0.00
	0	0.00	0.00	16.67	50.00	0.00	0.00	0.00	0.00	0.00
	-	25.00	0.00	16.67	0.00	25.00	25.00	0.00	0.00	0.00

BIBLIOGRAPHY

Note: Most (but not all) of the works listed in the Bibliography were accessible to the author. Very complete bibliographies of Typological Studies, Mathematical Linguistics, or, the overlap, Mathematical Typologies will be found in: 1, 4-6, 20, 25, 26 29, 31, 32, 45, 51, 54, 55, 65.
The bibliography of the primary sources used for the distinctive feature analysis as well as detailed discussion can be found in 1.

1. Afendras, Evangelos A. 1968. The Balkans as a Linguistic Area: A Study in Phonological Convergence. Unpublished Ph. D. Dissertation, The Johns Hopkins University, Baltimore.

2. Andreyev, N.D. (ed.) 1965. Statistiko-Kombinatornoje modelirovanije jazykov. Moscow.

3. Avram, Andrej. 1964. Sur la typologie phonologique quantitative. Revue Roumaine de Linguistique. IX. 131-134.

4. Birnbaum, Henrik. 1966. On Typology, Affinity, and Balkan Linguistics. Zb. za filologiju i lingvistiku. IX. 17-30. Novi Sad.

5. ———— . 1968. Rekonstrukcja wewnętrzna, kolejność synchonicznych reguł gramatyki syntetycznej i zagadnienie najdawniejszych stosunków między językami Bałtyckimi a Słowańskimi. IJOSLAP XI. 1-24.

6. ———— . 1969. Deep Structure and Typological Linguistics. Unpublished manuscript.

7. Borko, H. 1967. (Ed.) Automated Language Processing: The State of the Art. New York, Wiley.

8. Civ'jan, T.F. 1966. Imja suščestritel'noje v balkanskix jazykax (:k strukturno-tipologičeskij xarakteristike balkanskogo jazykovogo sojuza). Moscow, Akademija Nauk.

9. Edmundson, A.P. 1967. Mathematical Models in Linguistics and Language Processing. In Borko, (ed.).

10. Gammon, E. 1967. Quantitative Linguistic Typologies. 10ICL. Bucharest. (Preprint).

11. Greenberg, J. 1957. The Nature and Use of Linguistic Typologies. IJAL. Vol. 23. 68-77.

12. Grimes, J. 1962. Measures of Linguistic Divergence. P9ICL The Hague, Mouton. 44-50

13. Hamp, Eric . 1962. On the Interconnection of Sound Production, Perception and Phonemic Typology. P4ICPS. Mouton. 639-642.

14. Harary, F. and H. Paper. 1957. Toward a General Calculus of Phonemic Distribution. Lg. 33. 143-169.

15. Haugen, Einar. 1957. Language Contact. P8ICL. 771-785.

16. Havránek, B. 1933. Zur phonologische Geographie. (Das Vokalsystem des balkanischen Sprachbundes.) P1ICPS.

17. Herdan, G. 1960. Type-Token Mathematics. A Textbook of Mathematical Linguistics. (See also 7.5 which reviews Harary-Paper's Model.)

18. Hockett, G.F. 1958. A Manual of Phonology. Baltimore.

19. ———— . 1967. Language, Mathematics and Linguistics. The Hague, Mouton.

20. Horne, K.M. 1966. Language Typology. 19th and 20th Century Views. Georgetown University 10LAL.

21. Isačenko, A.F. 1939. Versuch einer Typologie der Slawischen Sprachen, LS 1, 64 ff.

22. Ivić, P. 1962. On the Structure of Dialect Differentiation. Word. 18. 53 ff. and numerous other works on Serbocroation, interference and general dialectology.

23. Jakobson, R. 1962. Sur la théorie des affinités phonologiques entre les langues. S.W. I. 234-246. Mouton.

24. ———— . 1962. Über die phonologischen Sprachbünde. S.W.I. 137-143.

25. ———— . 1962. K xarakteristike jevrazijskogo jazykovogo sojuza. Reprinted in: Selected Writings I. 117-136. Mouton.

26. ————, C.G. Fant and M. Halle. 1967. Preliminaries to Speech Analysis. 7th Printing. Cambridge. M.I.T. Press.

27. Kazaris, Kostas. 1964. Some Balkan Constructions Corresponding to W. European Infinitives. Unpublished Ph. D. Thesis, Indiana University, Bloomington.

28. ———— . 1967. On a Generative Grammar of the Balkan Languages. Foundations of Language. 3. 117-123. Dordrecht.

29. Kiefer, F. 1968. Mathematical Linguistics in Eastern Europe. American Elsevier, New York.

30. Krámský, J. 1959. Quantitative Typology of Languages. Language and Speech. II, no. 2.

31. Kučera, H. and G.K. Monroe. 1968. A Comparative Quantitative Phonology of Russian, Czech and German. New York. (A rather extensive bibliography with references to Kučera's other works is appended.)

32. Lehmann, W.P. 1962. Historical Linguistics. New York.

33. Lekomceva, M.I. 1963. Tipologija fonologičeskix sistem. In: Issledovanija po Strukturnoj Tipologii. Moscow.

34. Lyons, John. 1962. Phonemic and non-Phonemic Phonology; Typological Reflections. IJAL. 28. 127-133.

35. Mackey, W.F. 1953. Bilingualism and Linguistic Structure. Culture XIV. 143-149.

36. Makajev, É. A. (Ed.) 1966. Problemy Lingvističeskogo Analiza. Moscow. Nauka.

37. Mańecki, M. 1933. Systemy wokalne języków bałkanskich. Sprawozdania z Posiedzeń, Polska Akademia Umiejętności. 38. No. 8. p. 3 ff.

38. Marcus, S. 1963a. Typologie des langues et modèles logiques. Acta Mathematica Academiae Scientiarum Hungaricae. XVI. No. 3-4.

39. ———— . 1963b. Un model matematic al fonemului. Studii şi cercetări matematice, vol. 14, 405-421.

40. ———— . 1966. Le modelage mathématique en phonologie. Cahiers de Linguistique théorique et appliquée, 109-116.

41. ———— , and E. Vasiliu. 1960. Mathématiques et phonologie, Théorie des graphes et le consonantisme de la langue roumaine. Revue de mathématiques pures et appliquées, vol. 5, 519-540, 681-703.

42. Menzerath, P. 1950. Typology of Languages, JASA 22. 698-701.

43. ———— and W. Meyer-Eppler. 1950. Sprachtypologische Unter-Suchungen I. Lund.

44. Meyer-Eppler, W. Anwendungs der Kommunikationsforschung auf lautsprachliche und typographische Probleme. Sprachforum. 1: 70-77.

45. Papp, F. 1966. Mathematical Linguistics in the Soviet Union. Mouton, The Hague. (An extensive bibliography appended.)

46. Peterson, G. and F. Harary. 1961. Foundations of Phonemic Theory. Proceedings of the Symposia in Applied Mathematics. 12. R. Jakobson (ed.), Rhode Island.American Mathematical Society.

47. Petrovici, E. 1962. Les traits distinctifs des phonèmes roumains. P4ICPS. 723-728. Mouton, The Hague.

48. Pierce, J. E. 1957. A Statistical Study of Consonants in New World Languages. IJAL. 23. 36-45, 94-108.

49. _____. 1962. Possible Electronic Computation of Typological Indices for Linguistic Structures. IJAL. 28:215-226.

50. Pjotrovskij, R. G. 1969. Informacionnyje izmerenija jazyka. Leningrad. (Several earlier works can be found in his bibliographies as well as that in Papp (1966); some of his later work is also very pertinent to the statistical study of phonology.)

51. Plath, W. 1963. Mathematical Linguistics. In: Trends in European and American Linguistics 1930-1960. Mohrmann, Sommerfelt and Whatmough (Eds.). Spectrum, Utrecht-Antwerp.

52. Postovalova, V.I. 1966. O sočetajemosti differencial'nyx priznakov soglasnyx fonem sovremennogo russkogo jazyka. In Problemy Lingvističeskogo Analiza. Moscow.

53. Reichenkron, G. 1962-3. Der Typus der Balkansprachen. Zeitschrift für Balkanologie. E. 91-122. Wiesbaden.

54. Revzin, I.I. 1962. Modeli Jazyka. Moscow.

55. _____. 1967. Method modelirovanija i tipologija slavjanskix jazykov. Moscow.

56. Sandfeld, K. 1932. Linguistique Balkanique. Paris.

57. Saporta, S. 1957. Methodological Considerations Regarding a Statistical Approach to Typologies. IJAL. 23:109-113.

58. Skalička, V. 1958. Typologie slovanských jazyků, svláště ruštiny. ČR, 78 ff.

59. Spang-Hanssen, H. 1962. Mathematical Linguistics - Trend in Name or in Fact? P9ICL. The Hague, Mouton. 61-71 (incl. discussion).

60. Stankiewicz, E. 1958. Toward a Phonemic Typology of the Slavic Languages. AC4ICS The Hague, Mouton.

61. Trnka, B. 1936. General Laws of Phonemic Combinations. TCLP, 6. 57-62.

62. Trubetzkoy, N.S. 1929. Zur allgemeinen Theorie der phonologischen Vokalsysteme. TCLP. 1 39-67.

63. ———— . Principes de phonologie. (French translation of J. Cantineau). Paris.

64. Ungeheuer, G. 1959. Das logistische Fundament binärer phonemklassifikationen. Studia Linguistica. XIII: 69-97.

65. Uspenskij, B.A. 1962. Principy Strukturnoj Tipologii. Moscow.

66. Voegelin, C.F. 1956. The Scope of whole System (Distinctive Feature) and Subsystem Typologies. Word. XII. 444 ff.

67. ———— , F.M. Voegelin, S. Wurm, O'Grady and T. Matsuda. 1963. Obtaining an Index of Phonological differentiation from the construction of non-existent Minimax systems. IJAL. 29. No. 1 5-28.

68. Weinreich, U. 1953. Languages in Contact. New York.

69. ———— . 1957. Research Frontiers in Bilingualism Studies. P81CL. 786-810 (including discussion).

70. Wells, R. 1954. Archiving and Language Typology. IJAL. 20 101-107.

71. Wolff, H. 1959. Subsystem Typologies and Area Linguistics. AL 1, No. 7, 1-88.

Addenda

(#73 did not reach the author until after the completion of this paper. However, it does not alter the conclusions reached.)

72. Birnbaum, Henrik. 1968. Slavjanskije jazyki na Balkanax i ponjatije tak nazyvajemyx jazykovyx sojuzov. Glossa. vol. 2-1. 70-92.

73. Širokov, O.S. 1964. Voprosy balkanskogo vtoričnogo jazykovogo rodstva (primenenije Statistiki v diaxroničeskoj fonologii). In: Problemy sravnitel'noj grammatiki indojevropejskix jazykov. Naučnaja sessija. Tezisy dokladov. Moscow. 68-69

74. Uspenskij, B.A. 1965. Strukturnaja Tipologija Jazykov. Moscow.