Syntactic Methods in Word Formation − on the Basis of a Contrastive Analysis of Word Formation in Balkan Languages

Artur KARASIŃSKI*

Key-words: *Balkan languages, word formation, semantic derivation*

This short article refers to an international project which is carried out in the Institute of Slavistics of the Polish Academy of Sciences. The project named *Contrastive word formation of the Balkan languages* embraces all major Balkan languages. The coordinator of the project is Professor Viara Maldjieva from Torun. Participants responsible for various languages are: Professor Irena Sawicka, PhD. Anna Korytowska, PhD. Tomasz Cychnerski. Professor Marjan Markovik`, Professor Eleni Buzarovska, PhD., Angelina Pancevska, Sonia Milenkovska, Artur Karasiński.

Scientific target of the project

The main aim is to identify the specific features of the Balkan word formation in its semantic as well as formal aspect. To achieve this, first we need to describe the formation systems using a common model. Thus, the starting point for comparison is the mediator-language, comprising semantic units (categories) and rules of their connections (relations among them). The comparison of the linguistic data on the basis of a semantic mediator-language will allow us to obtain an inventory of word formation types, of their activity and specificity in particular Balkan languages. In this sense it will be a contrastive, typological, as well as an areal study.

Significance of the project

The Balkan Language League is nowadays a frequent subject of systematic linguistic investigation. We consider that the Balkan languages are the languages belonging to the Balkan *Sprachbund* (a unit of the areal classification of languages) and not all languages simply occurring in the Balkans. Consequently, the subject of our research will be: Bulgarian, Macedonian, Serbian, Albanian, Greek, Romanian, Aromanian and, possibly, Romani. Whereas in the case of morphosyntax we are already able to formulate a set of specific Balkan features, the Balkan word formation has never been studied. There is neither a corpus material of particular languages prepared (with the exception of Serbian and Bulgarian), nor do we know

* University of Nicolas Copernicus, Torun, Poland.

BDD-A1048 © 2012 Institutul de Filologie Română „A. Philippide"
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 13:10:29 UTC)

anything about the Balkan connections. Apart from this, the existing scarce descriptions should be reformulated according to the common model. Accordingly, the description which will be made in the frames of the project, in certain cases will be only a description of word formation in these languages, and it can be expected that it will trigger subsequent investigation of particular Balkan languages word formation, as well as the investigation of the connections among them. Apart from this, thanks to the methodology applied in the project, the achieved results may serve as a basis for further cognitive investigations regarding the so-called "profiling". That is because the relation between referential values of derived names (lexical units) and their word-formation categories in particular languages constitutes the relation between the segmentation of the world by the language and linguistic performance of fragments of the world, distinguished on the basis of this segmentation.

The knowledge of common and different word-formation phenomena is also very important for lexicography and glottodidactics. The results of the project may be helpful in the future in the preparation of one- or two-language dictionaries (including the type of Polish - Balkan language dictionary and vice versa). They may also serve as didactic materials in teaching Balkan languages.

It should also be stressed that the achievements of the Polish linguistics regarding the elaboration of theoretical models, as well as practice in word formation description are a good basis for realization of the project.

What is known about the research topic

Most of the existing descriptions of the Balkan word formation are usually short and very traditional chapters in general grammars of a given language, simply listing formants, or, very rarely, articles on particular problems. Separate descriptions of word formation are published only for Serbian and Bulgarian; they are also traditional, as their starting point is always the form and the criteria for description are intuitive and heterogeneous. Whenever they deal with convergence, there are always diachronic observations concerning the borrowing of, for example, Slavic suffixes into Greek or Albanian, or Turkish or Greek formants into remaining Balkan languages. For Aromanian or Romani we did not find anything on word formation at all.

In our project we are trying to construct a model in which all compared languages should have equal position – this would highlight the differences and common phenomena.

Theory and methodology in a nutshell

In the proposed research word formation is understood as a synchronic investigation of the internal relations within a language, as an investigation of the relations between the morphemes composing a derivative. Consequently, word formation may be understood as the investigation of the word-internal syntax.

The methodology which will be applied in the project concerns the development and refinement of a semantic model of a comparative word formation research.

BDD-A1048 © 2012 Institutul de Filologie Română „A. Philippide"
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 13:10:29 UTC)

The basis and the starting point for comparison is the mediator-language, comprising semantic units (categories) and the rules of their connections (relations between them).

The basic postulate is assumed to be the presence of a structural isomorphism between the language units, sentence and lexeme, i.e. the same combinatorial rules license both the combinations of lexemes within the sentence and the combinations of morphemes within the lexeme.

Unlike in the sentence, there is a constant correlation between the semantic and the formal structure of derivatives: the formant always serves as an exponent of a general meaning, whereas the theme indicates the detailed meaning. The only exceptions are the compound and the derived words with affixes, composed of two detailed notions.

Another postulate, which is basic for the research, together with the isomorphism of the sentence and morpheme structures, is the relative autonomy of the semantic and syntactic structures of the language units. Regarding word formation, it means that in this case semantic and formal derivatives should be distinguished in analogy to sentence syntax, that distinguishes the concept structure (sentence) and the formal structure (sentential expression).

The consistent application of this distinction leads to the definition of word formation as morpheme syntax (similar to sentence syntax) and of the derivative word as a structure, within which the rules for formal and semantic morpheme combination are realized. This allows us to seek and find morpheme combinatorial rules, i.e. the relations between the morphemes, characterized by regularity.

The establishment of an inventory of word formation categories is the result of a multi-stage logical division based on the following (functional-semantic) dichotomy criteria:

1st stage: +/- general predicative notions expressed in the morphemic semantic structure by the formant: 1. Predicate – (material) object.

2nd stage: +/- essential predicate: 1. nucleus predicate (constituting a superior structure), 2. added predicate.

3rd stage: content of the relation of implication of constituent notions: 2.1. ... 2.n. (implied) types of argument positions.

The categories distinguished on the basis of the above described procedure constitute separate classes. Consequently, that means that the semantic component belonging to a given category may appear only once in a mono-predicative predicate-argument structure.

It is supposed that the inventory of word formation categories will consist of two groups: predicate categories (distinguished in the second stage of division) and argument categories (distinguished in the third stage of division).

The inventory of the word formation categories constituting elements of the mediator language (*tertium comparationis*) for the description of word formation is the following:

BDD-A1048 © 2012 Institutul de Filologie Română „A. Philippide”
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 13:10:29 UTC)

a. argument categories	b. predicate categories
agent	state
patient	event
result	place
experiencer	manner
disposer	predisposition
beneficiary	negation
object	causation
element	similarity
instrument	relation
materia	size
	pole
	youth
	emotion and/or appreciation
	quantity
	intensity

The function of a formant depends on its participation in the semantic and formal (morphological and syntactic) nature of the newly created derived lexical unit. In that way we identify two possible functions – *semantic* and *syntactic*. That means that these functions may be realized, though not obligatorily, in each predicate. The semantic function of a formant consists in the fact that it is an exponent (marker) of an component of the morphemic semantic structure of the derivative – predicate (nucleus or added) or argument. The syntactic function consists in the fact that the formant determines categorial properties (morphological and syntactic) of the derivative, and consequently decides its distributive characteristics (inflectional and selective). By combining the occurrence (+) and absence (-) of these two possible functions we obtain four functional classes of derivatives:

Class	Derivatives:	Formant function:	
		Semantic	Syntactic
1)	Mutation	+	+
2)	Modification	+	-
3)	Transposition	-	+
4)	Tautology	-	-

The formants with a semantic function are exponents of word formation categories, which were distinguished as units of the mediator language by the modeling of the structures of semantic derivatives. A constitutive component of the formal structure of the derivative is always the theme (themes) connecting various types of formants. The types of formants serve as a basis for distinguishing the formal classes of derived structures.

BDD-A1048 © 2012 Institutul de Filologie Română „A. Philippide”
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 13:10:29 UTC)

Examples of Albanian semantic derivation analysis

I. Adjectival derivatives derived from verbs:

The kind of relation between the defined noun and the derivational base of the adjective is a criterion of the division of Albanian adjective derivatives derived from verbs. These main types of relations are:

1. A relation in which the base of an adjective is the basic predicate.

i.e.
P` – quantity
makinë e kushtueshme – 'makinë që kushton shumë' –*ueshme*
'expensive car – the car that costs a lot of money'

tokë pjellore – 'tokë që pjell shumë' –*or*
'fertile ground – ground that gives high yields'

This is a polypredicative structure. The predicate that creates this relation is a simple action/state predicate. The nucleus predicate is consists of the theme of the adjectival derivative. The noun is the representation of the word formation – argument categories. All quantifiable things can belong to this category. Formants such as *-ueshme, -or* have the semantic function and are the exponents of word formation predicate category – quantity.

2. A relation in which the base of an adjective is the additional predicate. Some examples of predisposition:

Relation of predisposition Ob – P`

shkrim i lexueshëm – 'shkrim që mund të lexohet' –*ueshëm*
legible handwriting – 'handwriting that can be read'

zëri i dëgjueshëm – 'zëri që mund të degjohet' –*ueshëm*
audible voice – 'voice that can be heard'

dokument i botueshëm – 'dokument që mund të botohet' –*ueshëm*
printable document – 'document that can be printed or published'

The nucleus predicate consists of the word formation formant *–ueshëm.* The added predicate consists of the theme of the verb. The noun represents argument category – object –implicated by the added predicate P`.

II. Adjectival derivatives derived from nouns:

Relations Ag_x – Pat_y

togë ekzekutues – 'togë [x] që ekzekuton dikë (y)'
firing squad – 'squad [x] that carries out the execution of someone (y)'

BDD-A1048 © 2012 Institutul de Filologie Română „A. Philippide"
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 13:10:29 UTC)

The predicate that constitutes this predicate-argument structure is the word formation predicate category of causation. It is consists of the theme of the adjective. The noun represents the argument category – the agent. The patient is not expressed. The function of the formant is syntactic. The syntactic function consists in the fact that the formant determines the categorial properties (morphological and syntactic) of the derivative.

Relations: Ag Res – Instr
aparat fotografik – 'mjet që përdoret për të fotografuar'
photographic device – 'equipment for taking photographs'

vegël ndërtimi – 'mjet që përdoret për të ndërtuar'
construction tool – 'an implement used to build sth.'

The predicate that constitutes this predicate-argument structure represents the word formation predicate category of *causation*. The category of *agent* is necessary when there is a predicate-argument structure which also constitutes the category of *instrument*. Then the *agent* category is blocked. The *result* consists ofthe theme of the adjective and the *instrument* consists of the theme of the noun.

III. Adjectival derivatives derived from adjectives:

The adjectives derived from other adjectives are very interesting. This is the modification type of derivatives which means that the formants that create new lexical unit change the meaning of the derivative base. These formants with a semantic function are the exponents of word formation categories, which were distinguished as units of the mediator language by the modeling of the structures of semantic derivatives.

i.e.
Negation
i parëndësishëm– 'që nuk është i rëndesishëm' pa-
unimportant/insignificant – 'which is not important'

i panevojshëm – 'që nuk është i nevojshëm'
unnecessery – 'which is not necessary'

Spatial relation
ndërplanetar – 'i tillë P që behet ndër planetet / i tillë x që vëpron, punon / gjendet ndër planetet'
interplanetary – 'such P that occurs between planets / such x that acts / takes place between planets'

BDD-A1048 © 2012 Institutul de Filologie Română „A. Philippide"
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 13:10:29 UTC)

References

Baltova et al. 2003: J. Baltova, S. Karolak, V. Maldjieva, *Projekt VIII tomu GKBP: Słowotórstwo. Wybrane zagadnienia*, in: *Studia gramatyczne bułgarsko-polskie*, t. VII: *Przewodnik po akademickiej Gramatyce konfrontatywnej bułgarsko-polskiej*, Warszawa, p. 185−192.

Demiraj 1995: Sh. Demiraj, *Gramatika e gjuhës shqipe, vëllimi I morfologjia*, Tirana.

Kallas 1984: K. Kallas, *Słowotwórstwo. Przymiotnik*, in: *Gramatyka współczesnego języka polskiego, Morfologia*, ed. R. Grzegorczykowej, R. Laskowskiego, H. Wróbla, Warszawa, p. 408−455.

Kostallari (ed.), 1984: A. Kostallari, *Fjalor i shqipës së sotmë*, Tirana.

Maldjieva 2007: V. Maldjieva, *Funkcje derywatów w tekście (propozycja modelu analizy i opisu)*, in: *Słowotwórstwo i tekst*, ed. V. Maldjieva, Z. Rudnik Karwatowa, Warszawa, p. 120−129.

Maldjieva 2009: V. Maldjieva, *Gramatyka konfrontatywna bułgarsko-polska, Słowotwórstwo*, tome 9, Warszawa.

Nagórko 1987: A. Nagórko, *Zagadnienia derywacji przymiotników*, Warszawa.

Surdulli (ed.) 1985: M. Surdulli, *Gramatika e gjuhës së sotme letrare shqipe, Morfologja*, Prishtina.

Abstract

The author of this paper concisely presents a project which is carried out in the Institute of Slavistics of the Polish Academy of Sciences. The project embraces all major Balkan languages. Most participants in the project come from Poland and from the Republic of Macedonia. Its main task is to identify the most common patterns of word formation in the Balkan languages. These patterns are regarded as "mirrors" or as transformations of syntactic structures (predicate–argument structures).

BDD-A1048 © 2012 Institutul de Filologie Română „A. Philippide"
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 13:10:29 UTC)