Ekaterina Tarpomanova
Sofia University Saint Kliment Ohridski

Mapping the *Sprachbund* and the Grammaticalization Theory: the Case of the Balkan Languages

The paper aims at analyzing the tense system of the Balkan Sprachbund in light of the grammaticalization theory. Similarly to other Indo-European languages, in the Balkan languages new analytical tenses emerged and changed the overall temporal system. Furthermore, new verbal categories developed based on reanalysis of perfect and future: evidentiality in Bulgarian and Albanian and presumptive in all languages. The process of grammaticalization of future and perfect, on one hand, and evidentiality and presumptive, on the other, is traced by applying semantic, morphological, phonetic and functional criteria of grammaticalization.

Keywords: Balkan *Sprachbund*, grammaticalization, perfect, future, evidentiality, presumptive

Sprachbund

The similarities between the Balkan languages were first noticed by Jernej Kopitar (1829) and Franc Mikloshic (1861), but it was Nikolai Trubetzkoy who theoretically motivated the term *Sprachbund*. Trubetzkoy used the term *языковой союз* in a 1923 paper in Russian and in 1928, at the International Congress of Linguistics, he introduced the German calque *Sprachbund*. According to him, languages in a *Sprachbund* share similarities in phonological systems, cultural vocabulary, morphological structure and syntax, as opposed to *Sprachraum* 'language space', or the space of genealogically related languages, characterized by systematic phonological and morphological correspondences and a common basic vocabulary (Trubetzkoy 1930). In modern linguistics *Sprachbunds* are defined as groups of languages whose similarities are acquired through a convergence process due to language contact. Language families, by contrast, inherited their similarities from proto-languages and further underwent divergence process.

The Balkan languages are considered representative for the *Sprachbund* theory: they developed common features in all language levels due to language contact and bilingualism. Phonological correspondences include the central vowel *ъ/ă/ë* in Bulgarian, Romanian and Albanian; the vowel reduction in Bulgarian and Albanian and in dialects of Romanian and Greek; *'a/e* metaphony in eastern dialects of Bulgarian and in Romanian; iotation in Bulgarian and Romanian; rhotacism in Romanian, Albanian and Greek; sonorisation of *p, t, k* after a nasal consonant (*mp>mb*, *nt>nd*, *nk>ng*) in Albanian, Greek and Aromanian. Similarities in the Balkan lexical system are found in the cultural vocabulary, phraseology and semantic calques. Balkan features in morphosyntax are the following: analytic development of the noun system, syncretism of dative and genitive, postponed definite article, clitic doubling, analytic degrees of comparison, loss of infinitive, future tense, conditional mood, perfect tense, opposition perfect/aorist (cf. Feuillet 1986; Асенова 1989/2002; Demiraj 1994; Feuillet 2012, among others). Morphosyntactic similarities are believed to be central for the Balkan *Sprachbund* as they affect the language structure (Асенова 2002: 75). Many of them can be viewed in terms of grammaticalization.

Grammaticalization

The concept of grammaticalization is present in a number of studies since the beginning of the 19th century, mostly in the field of historical linguistics, but it was the French linguist Antoine Meillet who first used and motivated the term in a 1912 paper, *L'évolution des formes grammaticales*. Meillet stated that there are certain types of language change different from sound change, analogy and borrowing. Unlike analogy, that does not affect the language system, grammaticalization creates new forms and introduces categories which had no linguistic expression, thus transforming the overall system. He defined grammaticalization as "attribution of grammatical character to an erstwhile autonomous word" (Meillet 1912: 385). One of the examples that Meillet used to illustrate the process is the grammaticalization of the verb *être* 'be' as auxiliary in *passé composé*, the most used past tense in French, pointing out the difference between *je suis qui je suis* 'I am who I am', where the verb is autonomous, and *je suis parti* 'I have departed', where it is a grammatical element (Meillet 1912: 385).

With the growing influence of the structuralism in linguistics and the domination of the synchronic approaches to language, the theory of grammaticalization became somewhat unfashionable. Later on, in the 1980-1990-ies the theory became popular again when several influential works were written: Christian Lehmann's *Thoughts on Grammaticalization* (1982), Heine

and Reh's *Grammaticalization and Reanalysis in African Languages* (1984), Heine, Claudi and Hünnemeyer's *Grammaticalization: a conceptual framework* (1991), Hopper and Traugott's *Grammaticalization* (1993), etc. Grammaticalization is regarded both as a process of language change: "... a process whereby a lexical item assumes a grammatical function, or a grammatical morpheme assumes a even more grammatical function" (Heine, Claudi and Hünnemeyer 1991); "… a linguistic process, both through time and synchronically, of organization of categories and of coding" (Traugott and Heine 1991: 1), and as a research framework: "... refers to that part of study of language change that is concerned with such questions as how lexical items and constructions come in certain linguistic context to serve grammatical functions, or how grammatical items develop new grammatical functions" (Hopper and Traugott 2003: 1). A basic notion in the theory is the grammaticalization pathway (cline, chain, continuum, channel), i.e. the stages of transformation of a content word into a grammatical marker. The grammaticalization pathway may be presented with the famous scheme of Hopper and Traugott: *content word > grammatical word > clitic > inflectional affix*, in which every right item is less lexical and more grammatical (Hopper & Traugott 1993/2003: 7).

In a number of studies grammaticalization parameters are defined, i.e. processes that are related to the grammaticalization of forms. Heine describes four consecutive parameters: semantic shift (desemanticization): loss of semantic content; morpho-syntactic shift (decategorization): loss of morphological and syntactic characteristics; morpho-phonological shift (cliticization): loss of autonomous word status; phonological shift (phonological erosion): loss of phonetic material (Heine 1993). An important grammaticalization parameter is the obligatorification: a process whereby a language structure becomes more obligatory (Heine and Kuteva 2007: 34). Lehmann describes two groups of parameters. Paradigmatic parameters include: i. integrity: possession of a certain substance which allows the sign to maintain its identity, its distinctness from other signs; ii. paradigmaticity: the formal and semantic integration both of a paradigm as a whole and of a single subcategory into the paradigm of its generic category; iii. paradigmatic variability: the freedom with which the language user chooses a sign (similar to obligatorification in Heine and Kuteva 2007). Syntagmatic parameters are the following: i. structural scope: the structural size of the construction which the grammatical means helps to form (as a constituent); ii. boundedness: the intimacy with which a sign is connected with another sign to which it bears a syntagmatic relation; iii. syntagmatic variability: the ease with which a sign can be shifted around in its context (Lehmann 2002). Grammaticalization is usually seen as a unidirectional process, although there are examples of degrammaticalization or back lexicalization (Heine and Kuteva 2002: 4).

The theory of grammaticalization has been criticized for relying on broad and unclear definitions, for not clarifying the relation with analogy as a means of language change, for the concept of unidirectionality given that there are counterexamples (cf. Joseph 2001). Nevertheless, it is one of the most popular frameworks in moderns linguistics and it is applied both in synchronous and diachronous approaches to language analysis.

Mapping two theories

In what follows, I will analyze the process of grammaticalization in the temporal system of the Balkan languages (Bulgarian, Greek, Albanian and Romanian) taking into account the "new" tenses: future, which is one of the core contact induced Balkan features, and perfect, considered as a marginal one. These two tenses not only changed the configuration in the Balkan way to express temporality, but in addition they gave rise to new verbal categories, presumptive in all the Balkan languages and evidentiality in Bulgarian and Albanian.

In order to evaluate the grammaticalization level of the perfect and the future tense, on the one hand, and of the evidentiality and the presumptive, on the other, I will apply some of the criteria of grammaticalization defined by different authors, not following a certain scheme, but rather using a set of parameters that best describe the processes in the Balkan *Sprachbund*. The parameters may be labeled as semantic, morphosyntactic, phonetic and functional.

The semantic parameter applies differently to the grammaticalization process of the two tenses and the subsequent grammaticalization of the presumptive and the evidentials. **Desemanticization** in Heine's sense (Heine, 1993) is suitable when analyzing the grammaticalization of future and perfect as it is related to the transformation of a lexical verb into a grammatical item (auxiliary or particle). On the other hand, when a form is grammaticalized based on another grammatical form, the process may be viewed as a **semantic reanalysis**.

The morphosyntactic parameters may be defined as a development to morphologization, which can include several factors. **Form fixation** is the existence of a distinct form that expresses the grammatical meaning. **Paradigmatization** is the creation of a uniform paradigm that subsumes systematic relations between its elements. **Integrity** may be called the strong connection between the constituents of a form.

The phonetic parameters include the **loss of stress** of a function word and the **phonetic erosion** as a loss of phonetic material.

As a functional parameter I labeled the **obligatorification** of the gram-maticalized form, i.e. the limitation of freedom of the language users with regard to the grammatical paradigm (cf. Lehmann 2002: 124).

Tense system of the Balkan *Sprachbund*

As already mentioned, the grammaticalization of future and perfect was crucial for the establishment of the now existing temporal system of the Balkan languages. Table 1 presents a general picture of the tense system in Bulgarian, Greek, Albanian and Romanian.

	Bulgarian	**Greek**	**Albanian**	**Romanian**
Present	*ходя*	*γράφω*	*shkoj*	*fac*
Imperfect	*ходех*	*έγραφα*	*shkoja*	*făceam*
Aorist	*ходих*	*έγραψα*	*shkova*	*(făcui?)*
Perfect	*съм ходил*	*έχω γράψει*	*kam shkuar*	*am făcut*
Pluperfect	*бях ходил*	*είχα γράψει*	*kisha / pata shkuar*	*făcusem*
Future	*ще ходя*	*θα γράφ/ψω*	*do të shkoj*	*voi face* *o să fac* *am să fac*
Future in the past	*щях да ходя*	*θα έγραφα*	*do të shkoja*	*(aveam să fac?)*
Future anterior	*ще съм ходил*	*θα έχω γράψει*	*do të kem shkuar*	*voi fi făcut*
Future anterior in the past	*щях да съм ходил*	*θα είχα γράψει*	*do të kisha shkuar*	*-*

Table 1. Tense system of the Balkan languages.

The first three rows present the old synthetic tenses in all four languages, i.e. present, imperfect and aorist (the form of the imperfect in the modern Slavic languages is a Slavic innovation (Vaillant 1966: 61), it does not inherit the old Indo-European imperfect, on the contrary, this is a common Slavic process, not related to the language contacts in the Balkan area).

In the diachronous development of Romanian and Greek new analytical tenses replaced old Latin and Greek perfect and future, with correspondences in Albanian and Bulgarian. Future and perfect are innovations that took place on the Balkans, except for perfect in Bulgarian which is a Slavic feature already existing in Old Bulgarian. Subsequently new tenses were formed, based

on future and perfect: future in the past with a morphological marking of future and imperfect; future anterior using the marking of future and perfect; future anterior in the past with a complex marking for future, perfect and imperfect.

As a result of common processes in the Balkans, tenses in Albanian, Greek and Bulgarian are almost identical. In Romanian, though, there are quite a lot exceptions that do not follow the Balkan tense model. Similarly to the common Romance development, aorist in Romanian is almost lost, its functions are taken by the perfect tense which indicates a (simple) event in the past. Romanian is one of the Romance languages (together with Ibero-Romance) and the only Balkan language that preserved the synthetic pluperfect. There are three models of future tense formed with different auxiliaries. The future perfect is based on the *voi* + infinitive model, and the future in the past based on the 'have'+ conjunctive model appeared only recently and is found in colloquial speech only, it is not even mentioned in grammars. And finally, there is not a future perfect in the past.

Perfect

The synthetic Indo-European perfect was replaced by periphrastic constructions in all Indo-European languages. The new analytic perfect appeared in Balto-Slavic, formed with the auxiliary 'be' and the past active participle of the lexical verb. Slavic used the participle ending in -*l* that originates from a verbal adjective (Vaillant 1966: 83). The analytic forms of perfect and pluperfect are already attested in the earliest written documents of Old Bulgarian (Maslov 1988: 76) and both tenses are considered a Slavic feature. In modern Bulgarian among the numerous dialectal varieties of the perfect, the Slavic type 'be' + *l*-participle (aorist active participle of both perfective and imperfective verbs) has been generalized in the standard language: *съм писал* 'I have written, imperfective' *съм написал* 'I have written, perfective'. However, in the south and south-east dialects (Strandzha, Trace, Southwest Macedonia) and in Macedonian[1], perfect appears in its Balkanized version, formed from 'have" auxiliary and the past passive participle: *имам гледано* 'I have watched'. What is commonly accepted as a Balkan feature in Bulgarian with respect to the perfect is the fact that this is the only Slavic language that preserved the opposition between perfect and aorist, as in all the other Slavic languages perfect fully replaced aorist.

---

[1] In Macedonian two models coexist: the Slavic one, 'be' + aorist active participle (*сум видел* 'I have seen'), and the Balkan one, 'have' + past passive participle (*имам видено* 'I have seen').

In Koine Greek, a functional fusion between perfect and aorist took place that in the 4th century AD ended up with a complete uniformity of their functions. In the early Byzantine papyri perfect was replaced by aorist in most of the occurrences and the fusion between the two tenses lead to the disappearance of the synthetic perfect. In the same period different periphrastic constructions started to express the perfect semantics. Syntactic constructions based on two auxiliaries, 'have' and 'be' + participle, were used in Medieval Greek initially to express voice opposition: *ἔχω γεγραμένον* 'have written' for active and *εἰμί γεγραμένος* 'be written' for passive. The construction of the type *ἔχω γράψει*, formed by the auxiliary 'have' and aorist infinitive, appeared only in the 16th century, but in Modern Greek this is the regular perfect form (cf. Асенова 2002: 241-242 and the authors cited there).

In classical Latin, analytical passive perfect was already formed from the auxiliary 'be' and the past participle: *factus sum* 'be done'. Vulgar Latin developed analytical active perfect constructions based on the auxiliaries 'have' and 'hold' (Maslov 1988: 72). The modern Romance languages use the auxiliaries 'have' and 'be' to express voice opposition in perfect. Romanian follows the Romance model of the perfect using the auxiliary 'have' and the past participle: *am făcut* 'I have made'. Unlike French, perfect in Romanian is formed with the auxiliary 'have' both from transitive and intransitive verbs: Fr. *j'ai donné* 'I gave' / *je suis venu* 'I came', Rom. *am dat* 'I gave' / *am venit* 'I came'. Similarly to other Romance languages (French, Italian), aorist has been superseded by perfect, the latter being generalized to express events in the past.

It is difficult to say when the perfect in Albanian emerged, but this process must have been finished centuries before the 16th century, as the analytical perfect and pluperfect forms were already established in the works of the old Albanian authors (Demiraj 1976: 101). Perfect in Albanian is formed following the 'have' + participle model, voice opposition is expressed by the use of different auxiliaries, 'have' for active and 'be' for non-active: *kam parë* 'I have seen' / *jam parë* 'I am seen'.

The Balkan languages started using analytical perfect forms following the overall Indo-European trend. The beginning of grammaticalization of the new forms in the Indo-European languages can be seen in the syntactic constructions containing different verbs further developed into auxiliaries and different types of deverbatives (participles, adjectives, infinitives, etc.). If there is a Balkan model of perfect, that would be the 'have' + passive participle model, a feature that can be found in Macedonian, Bulgarian dialects, Albanian and Romanian:

Bulg./Mac. *имам видено* 'I have seen'
Alb. *kam parë* 'I have seen'

Rom. *am văzut* 'I have seen'

The Modern Greek perfect displays similarities and differences with respect to the Balkan model: it uses 'have' as auxiliary, on one hand, but on the other, the deverbal form of the lexical verb is not a participle, but an aorist infinitive, and additionally, the voice opposition is based on morphological alternation in the lexical verb form and not on the auxiliaries: *έχω γράψει* 'I have written' / *έχω γραφτεί* 'I am written'.

The historical development of many languages shows the potential of the verb 'have' to be grammaticalized as an auxiliary. As argued by Benveniste, 'have' has different nature as compared with 'be'. Both verbs may express existence, but in the case of 'be' there is a relation of identity between two items, while the items connected through 'have' are in relation of possessor and possessed (Benveniste 1982). This relation is preserved in many languages, as for example in French and it can be seen in the gender and number agreement of the participle with the object. In the Balkan model the connection is lost and the participle is invariable:

Fr. *C'est peut-être la première **oeuvre** d'art que j'ai **vue** dans ma vie.* 'It's probably the first piece of art that I have seen in my life.'

Мас. *Ова е најлошата **повреда** што ја имам **видено** во мојата тренерска кариера.* 'This is the worst damage that I have seen in my trainer career.'

In Bulgarian and Greek similar means to express result are used, but they are not grammaticalized. In the resultative syntactic constructions of the type 'have' + passive participle the concept of possession is still preserved and it is expressed by the agreement in gender of the participle and the possessed. In addition, the level of morphologization is very low which can be proven by the fact that we can easily insert items between the two parts of the construction.

*Τα **έχω γραμμένα** όλα για να τα βρει η κόρη μου. (= τα **έχω** όλα **γραμμένα**)* 'I have written all that so that one day my doughter could find it.'
***Имам написани** четири теми за Свищовския университет. (= **имам** четири теми **написани**)* 'I have written four items for the exams at the University of Svishtov.'

Grammaticalization parameters of perfect

The Balkan languages follow the common grammaticalization path of the Indo-European languages. The semantic bleaching of some frequent verbs, such as 'be'[2] and 'have', is a common process that resulted in the loss of

---

[2] According to Dik, the copula does not have independent meaning and must be treated as a supportive device both as a copula and as auxiliary (Dik 1987: 56).

their semantic content and their auxiliarization. Considering the Hopper and Traugott's scheme, only the first step to grammaticalization is completed, i.e. the transition from a content to a function word: *content word > grammatical word > clitic > inflectional affix*.

In all Balkan languages perfect has a fixed form that can be easily distinguished from other temporal forms. As it is an analytical form, its integrity is an important indication of grammaticality. With this respect, the Balkan type of 'have'-perfect shows a stronger integrity than the 'be'-perfect in the Balkan Slavic. A sign of morphologization is the strong connection between the two elements of the form: no other lexical or functional item can be inserted between the elements of the 'have'-perfect and the elements themselves have a strict position that cannot be changed, while in the 'be'-perfect pronominal clitics are allowed and the auxiliary and the participle may change their position depending on the word order of the clause (the copula takes the second position in the clause structure):

Alb. ** kam vajzën parë, * kam e parë, * kam nuk parë, * parë kam*;

Bulg. *съм му го дал* 'I have given it to him', *аз съм дал / дал съм* 'I have given'

The paradigmatization of the perfect comprises its development as a subcategory and further reorganization of the category of tense by the creation of other subcategories using the perfect morphological markers, such as pluperfect and future perfect:

Alb. *kam parë, ke parë, ka parë ...* 'I have seen, you have seen, he has seen …'

kisha parë, kishe parë, kishte parë ... 'I had seen, you had seen, he had seen'

do të kem parë, do të kesh parë, do të ket parë ... 'I will have seen, you will have seen, he will have seen'

The obligatorification consists in the regular use of the perfect in a typical context. As opposed to other tenses, perfect is required in a context where an action occurred in the past with a focus on its resulting state in the present:

Alb. *Do t'i themi gjërat që deri tani nuk i **kemi thënë**.* 'We will tell him the things that we haven't told him so far.'

From perfect to evidentiality

Based on perfect, a new category arose in two of the languages of the Balkan *Sprachbund*, Bulgarian and Albanian. Evidentiality is defined as a grammatical means to express the source of information in a given statement (Aikhenvald 2004: xi). Different opinions have been expressed about the period of grammaticalization and with this respect the question of the Turkish

influence has been widely discussed. No matter if we accept the hypothesis of external influence on Bulgarian and Albanian, the internal predisposition of the two languages to create a new category is of great importance. A major role in this process has the preservation of the opposition between perfect and aorist. In that opposition the perfect expresses an event that is non-localized in the past and actual in the moment, and the aorist, on the contrary, denotes an event that is localized in the past and non-actual in the present (Asenova 1996). Subsequently, the event perceived by its result in the present is reinterpreted as a non-witness position, and the concrete event in the past is related to a witness position. In Romanian, where the perfect took over the function of aorist, this semantic reinterpretation was impossible. So far, the question why an evidentiality system was not developed in Greek, although the opposition between aorist and perfect is preserved and the perfect shares the functional and semantic characteristics of the Balkan temporal model, has no satisfactory answer. Brian Joseph offers an extralinguistic explanation: Greek did not develop evidentiality due to social and cultural factors, i.e. the attitude of Greeks toward the preservation of their language, the literary tradition of Greek, the identification of the language with religion, etc. (Joseph 2003: 317).

In Bulgarian the evidential system includes perfect-like forms marked for indirect source of information, opposed to the indicative forms which are marked for direct source. The emergence of the evidential system produced an important morphological change: the formation of an imperfect active participle, which is unique for the Slavic morphology. The process was accompanied by the change of temporal reference of the new participle that started to indicate present. The evidential paradigm contains three morphologically marked indirect evidantials: renarrative marked with lack of auxiliary in the 3rd person (*пишел* 'he writes. ren.'), inferential whose marker is the presence of the auxiliary in the 3rd person (*е пишел* ' he writes, inf.'), and dubitative marked with an additional past participle of the verb 'be' – (*бил пишел* 'he writes, dub.'). Additionally, there is an admirative whose forms are homonimical with the renarrative. All indirect evidentials developed a reduced temporal paradigm with tenses grouped in pairs, i.e. one form combines the meaning of two tenses: present and imperfect, perfect and pluperfect, future and future in the past, future anterior ant future anterior in the past, aorist.

In Albanian, there is no formal difference between all indirect evidentials and admirative. It is difficult to estimate which is the primary meaning of the form, taking into account the fact that in the modern language the admiratve usage prevails considerably over the evidential, but on the other hand occurrences in old texts such as epic poems and folklore tales have a clear evidential character. The evidential/admirative is formed through inversion of the two items of perfect, the auxiliary and the participle, with the respective change

of temporal reference, but the indicative is unmarked for evidentiality. Four tenses constitute the temporal system of evidentiality: present and imperfect are considered synthetic tenses, formed by inversion and further fusion of the indicative perfect and pluperfect, respectively (*kam shkuar* > *shkuakam*, *kisha shkuar* > *shkuakësha*); perfect and pluperfect are analytical tenses formed following the model of the respective indicative tenses, but using the evidential form of the auxiliary (*paskam shkuar*, *paskësha shkuar*).

Grammaticalization parameters of evidentiality

In her typological study of evidentiality Aikhenvald makes a difference between grammaticalized evidentiality and evidential strategy, i.e. the use of other categories (moods, tenses, modal constructions, etc.) to express evidential meaning (Aikhenvald 2004: 105). Bulgarian and Albanian are considered languages with grammaticalized evidentiality, as the evidential forms differentiated from the initial perfect forms.

The semantic parameter consists in the reanalysis of the grammatical meaning of the perfect, which represents a prior event whose results are visible in the present. As opposed to the aorist, the perfect meaning was reinterpreted as representing a non-observed event and gave rise to the indirect evidentiality (Asenova 1996). In Albanian the main and the most frequent meaning is the admirative one expressing surprise of an unexpected event. The prevalence of the admirative over the evidential meaning is due to expressive force of the inversion (Gerdzhikov 2003: 120). The way Albanian evidential is formed may also explain its emphatic functions.

Paradigmatization of evidentiality consists in the development of a temporal system that covers all the temporal values in Bulgarian, except for the inferential, which due to semantic restrictions cannot be used in the non-past, and the non-future plan in Albanian. The form fixation of the evidentials and the admirative should be regarded as a level of differentiation from the initial perfect form. This parameter is fully completed in Albanian, where all of the forms of the evidential/admirative paradigm are unique, i.e. none of them coincides with the perfect. In Bulgarian the homonymy[3] between inferential aorist and indicative perfect has been widely discussed (cf. Алексова 2003). Additionally, renarrative aorist differs in the 3rd person only, by the omission of the auxiliary: *ходил е* 'he has gone, indicative perfect' / 'he went, inferential aorist' vs. *ходил* 'he went, renarrative aorist'. The low level of differenciation

---

[3] According to Gerdzhikov, the forms under discussion are not homonymous, which would mean they coincide just accidentally, but bipartite, i.e. one form takes part in two grammatical paradigms (Герджиков 2003: 230).

of the Bulgarian evidentials from the initial perfect form is often exaggerated (probably because these are the most frequent evidential forms) and interpreted as a low level of grammaticalization. However, all the other tenses of the renarrative and the inferential, as well as the whole paradigm of the dubitative, have unique forms.

A complete integrity of the form is achieved in Albanian by the fusion of the two elements of the synthetic present and imperfect. The inverted participle and auxiliary formed an accentual (and graphic) unity and the auxiliary assumed the function of affixoid. Additionally, the form moved to a phonetic reduction by losing its ending. In Bulgarian, there is no difference between evidentials and perfect in terms of integrity.

Obligatoriness also displays peculiarities in each of the two languages. In Bulgarian, the use of evidentials is obligatory in a certain context subsuming indirect information, but the admirative may be freely replaced by an indicative form in contexts related to surprising events. In Albanian, due to the grammaticalization pattern that includes inversion, only the most expressive forms, i.e. admirative and dubitative, are to a great extent obligatory, while for the neutral evidential meanings of reporting or inferring a nonwitnessed event the speaker is free to choose between an evidential or an indicative form.

Future as a Balkan feature

Indo-European languages use several auxiliaries to grammatically express futurity, such as 'have' (most of the Romance languages), 'want' (English), 'become' (German), 'go' (English, French), 'be' (Slavic languages), etc. Auxiliaries are used in different formats or sometimes are further transformed into affixes, as in Fr. *je viendrai* 'I will come'. In the Balkan languages the 'want'-future is considered one of the most important features that prove the convergence process and the formation of a *Sprachbund*. What is remarkable in the Balkan future type is not the choice of the same auxiliary, but the fact that all the languages had a different initial point to express futurity and, going through the same stages of grammaticalization, achieved almost identical final point.

Future in Ancient Greek is expressed by the present of the aorist stem, similarly to Old Bulgarian where the present of the perfective stem is used to express futurity (the so-called simple future). Latin future is synthetic, marked with inflectional suffix (*portabo* 'I will carry'). Later on, periphrastic constructions coocur and tend to replace the old future models: 'have', 'want', 'must, be obliged', 'intend', etc. + infinitive in late Koine Greek, 'be', 'want', 'have', etc. + infinitive in Old Bulgarian, 'have', 'want', 'must', etc. + infinitive in late Latin. The development of the Balkan future model is summarized

by Petya Asenova in the following way: 1) reduction of the verbs that take part in the periphrastic constructions to two, 'want' and 'have'; 2) desemanticization and auxiliarization of the verbs; 3) replacement of the infinitive of the lexical verb with conjunctive constructions; 4) transformation of the auxiliary into invariable particle which is induced by the replacement of the infinitive with a finite form (Асенова 2002: 213-215).

As a result of the grammaticalization process in Bulgarian and Greek future is expressed with the same model, 'want'-particle + present finite verb form (Gr. *θα κάνω*, Bulg. *ще правя* 'I will make'), but in Bulgarian 'have'-future is kept in the negative form (*няма да правя* 'I will not make'). Standard Albanian (based on Tosk dialect) differs from that model by preserving the conjunctive of the lexical verb inflectionally marked only in 2nd and 3rd p. sg., and additionally the preposed conjunctive marker is often omitted in the colloquial speech: *do (të) bëjë* 'he/she will make'. The Gheg dialect uses the auxiliary 'have'. There are three future formats in Romanian: 'want'-auxiliary + infinitive (*voi face* 'I will make'), 'have'-auxiliary + conjunctive (*am să fac* 'I will make'), and 'will'-particle + conjunctive (*o să fac* 'I will make'). Only the latter among them adheres to the Balkan model. Interestingly, Serbo-Croatian partly shares this Balkan feature using the 'want'-auxiliary combined with infinitive (*ja ću raditi* 'I will work') and an inverted form with a shortened infinitive in which the auxiliary is in fact transformed into affixoid (*radiću* 'I will work').

Grammaticalization parameters of future

The semantic bleaching as a grammaticalization parameter includes the transition from the notion of volition expressed lexically to the more abstract notion of futurity.

The morphologization consists in the transformation of a syntactic structure into a grammatical form and comprises both elements: 1) the functional element follows the pathway: lexical verb > auxiliary > invariable particle; 2) the lexical element evolved from infinitive to subjunctive, and further to present form in Bulgarian and Greek. The change is less advanced in Albanian, where the preposed conjunctive marker (the conjunction *të* 'to') is often omitted, but the inflectional marker is preserved. In Romanian the omission of the conjunctive marker is impossible. Future has a fixed form in all the Balkan languages and does not differ from perfect in terms of integrity, i.e. only pronominal clitics may be inserted between its elements: Bulg. *ще му го дам*, Gr. *θα του το δώσω*, Alb. *do t'ia jap*, Rom. *o să i-l dau* 'I will give it to him'. Similarly to perfect again, paradigmatization comprises the subcategorial level and further the restructuring of the category of tense.

Phonetic erosion is attested in the grammaticalization process of the functional element. Its development towards invariable particle is accompanied with a phonetic simplification, except for Albanian where the 3rd person sg. has been established as a future particle: Bulg. *хощѫ > ще*, Gr. *θέλω > θα*, Rom. *voiu > o*; Alb. *do = do*.

Future is obligatory in contexts requiring future reference, aside from tense transpositions, such as using present forms with future reference.

From future to presumptive

In all Balkan languages future tenses gave rise to the category of presumptive, based on the epistemic interpretation of futurity. Presumptivity as a probabilistic category is expressed in the Indo-European languages by different means: other verb categories (such as conditional in French), modal verbs, adverbials, etc., but none of them may be accepted as grammaticalized. However, in the languages of the Balkan *Sprachbund* there is a trend towards grammaticalization of the presumptive.

Romanian is widely accepted as a language with a grammaticalized presumptive mood. Among the three future formats, only one served as a basis for the new category – the 'want'-auxiliary + infinitive format (*voi face* 'I will make'). This format has been further specialized to express presumptivity with a phonetically reduced auxiliary (with apheresis of the initial *v*): *oi face* 'I must be makeing', *oi fi făcând* 'I must be making', *oi fi făcut* 'I must have made', etc. On the other hand, there are three other formats with presumptive reading, based on conditional, conjunctive and infinitive forms (cf. Mihoc 2013). The discussion which forms must be included in the presumptive paradigm is still ongoing. A review of different opinions is made by Victor Friedman, who outlines two opposite trends: a minimalist approach that includes one form based on the future format (future auxiliary + *be*-infinitive + gerund of the lexical verb: *va fi cântând* 'he/she must be singing'), and a maximalist approach that combines different forms of three formats – future, conjunctive and conditional (Friedman 1998: 396–397).

In Bulgarian presumptive emerged as a result of reanalysis of two tenses, future and future anterior. Unlike future tenses, where the omission of the conjunction *да* 'to' is considered a final stage of grammaticalization, it is preserved in the presumptive form, although its usage is optional: *ще (да) ходи* 'he/she must walk', *ще (да) е ходил* 'he must have walked'. The presumptive paradigm is further extended with two other tenses, imperfect and pluperfect, in which the entire formation *ще да* is interpreted as a presumptive marker and *да* 'to' cannot be omitted: impf. *ще да ходеше*, pluperf. *ще да беше ходил*. Additionally, an alternative paradigm emerged which seems to be a

fusion between presumptive and inferential: impf. *ще да е ходел*, *ще да е бил ходил*.

In Greek one form is labeled as presumptive, it is a combination of the future particle and the verb in aorist: *θα πήγε* (Holton et al. 2000: 229). However, the future imperfective, the future in the past and the future anterior may also have presumptive reading (Giannikidou and Mari 2016). In Albanian several temporal forms may have presumptive interpetation: future, future in the past, future anterior, and future anterior in the past (GGjSh 1: 316 – 317). A combination of the future particle and a periphrastic present progressive form can be used both with a temporal (future progressive) and a presumptive value.

Table 2 summarizes the presumptive forms *per se* and the forms that may have presumptive reading in the Balkan languages.

	Romanian	Albanian	Bulgarian	Greek
Present (simple)	*va/o merge*	*do të shkojë*	*ще (да) ходи*	*θα πηγαίνει*
Present progressive	*va/o fi mergând*	*do të jetë duke shkuar*	-	-
Imperfect	-	*do të shkonte*	*ще да ходеше* / *ще да е ходел*	*θα πήγαινε*
Perfect / Aorist	*va/o fi mers*	*do të ketë shkuar*	*ще (да) е ходил*	*θα πήγε*
Past progressive	*? va/o fi fost mergând*	*? do të ishte duke shkuar*	-	-
Pluperfect	*? va/o fi fost mers*	*do të kishte shkuar*	*? ще да беше ходил* / *? ще да е бил ходил*	*? θα είχε πάει*

Table 2. Presumptive in the Balkan languages.

In the table, the underlined forms are unique, the rest are temporal forms that may have a presumptive reading in a certain context. The forms with question mark occur only seldom, although some of them are unique.

Grammaticalization parameters of presumptive

The grammaticalization of the presumptive relies on the modal reinterpretation of an already existing grammatical meaning, i.e. futurity. The new category may be defined as probabilistic (Gerdzhikov 2003: 129). Simultaneously, the semantic reanalysis involves change of temporal reference from future to non-future, i.e. present or past. In Romanian, the semantic reanalysis comprises several initial grammatical categories and meanings, respectively.

The morphologization process is not entirely uniform in the Balkan languages. The paradigmatization trend is advanced in all of them, although there are only a few forms that are exclusively presumptive, none in Albanian. The preservation of the conjunction *да* 'to' in Bulgarian eventually differentiates the presumptive and the indicative use of the homonymous forms (indicative future / presumptive present, indicative future anterior / presumptive perfect). In all the other forms *да* cannot be ommited, which shows that the entire formation *ще да* is interpreted as presumptive marker that can be added to certain temporal forms: *ще да + ходеше*, *ще да + беше ходил*. Phonetic reduction takes place only in the future format of Romanian, both phonetically changed and non-changed forms co-occurring in speech. The reduced forms have exclusively presumptive reading.

While the morphologization is advanced to some extent, the obligatorification of the presumptive is a very week process due to the possibility to express presumption by lexical means, mainly modal verbs and adverbials:

Alb. *do të ketë ardhur* 'he/she must have come'

duhet të ketë ardhur 'he/she must have come'

ndoshta ka ardhur 'he/she has probably come'

Recent studies show decreasing frequency of the presumptive forms (see for Romanian Zafiu 2009, for Bulgarian Търпоманова 2017).

Conclusions

The paper discusses some processes in the verbal system of the Balkan *Sprachbund* in light of the grammaticalization theory, tracing the development of two tenses and their further reanalysis as new categories. The grammaticalization of the perfect and the future took place in the period of intensive language contacts and situation of bilingualism and multilingualism in the Balkans, except the Slavic 'be-'perfect. However, the development of 'have'-perfect in Bulgarian dialects and Macedonian is a strong proof for the convergence process between the Balkan languages. The grammaticalization of the two tenses rearranged the whole temporal system of the languages of the Balkan *Sprachbund*, which became almost identical in Bulgarian, Albanian and Greek. Tenses display a strong obligatorification, except for some cases of transposition. Furthermore, perfect and future served as a basis for grammaticalization of evidentiality and presumptive, respectively.

Evidantiality is a stable category in Bulgarian and Albanian, with advanced morphologization and differentiation from the initial perfect forms. Peculiarities of each language are observed in obligatorification. The formation of presumptive in all Balkan languages can be regarded as a trend to temporal paragmatization with only a few unique forms. On the other hand,

the week obligatorification and the decreasing frequency make it a marginal category in the Balkan languages.

References

Aikhenvald, A. (2004). *Evidentiality*. Oxford: Oxford University Press.

Asenova, P. (1996). Habitorja e shqipes dhe mënyra e ritregimit e bullgarishtes. – In: *Seminari XVII ndërkombëtar për gjuhën, letërsinë dhe kulturën shqiptare*. Ti-ranë, 18-31 gusht 1996, 565 – 579. Tirana.

Benveniste, E. (1982). "Être" et "avoir" dans leur fonctions linguistques. – In: *Problèmes de linguistique générale*, 186 – 207. Paris: Gallimard.

Demiraj, Sh. (1976). *Morfologjia historike e gjuhës shqipe.* Pjesa II. Tirana: Mihail Duri.

Demiraj, Sh. (1994). *Gjuhësi ballkanike.* Skopje: Logos-A.

Dik, S. (1987). Copula Auxiliarization: How and Why? – In: *Historical development of auxiliaries.* Harris, M. & Ramat, P. (Eds.). Berlin, New York, Amsterdam: Mouton de Gruyter, 53 – 84.

Feuillet, J. (1986). Linguistique balkanique. *Cahiers balkaniques*, No 10, 1986.

Feuillet, J. (2012). *Linguistique comparée des langues balkaniques.* Paris: Institut d'Études Slaves.

Friedman, V. (1998). The Grammatical Expression of Presumption and Related Concepts in Balkan Slavic and Balkan Romance. – In: *American Contributions to the 12th International Congress of Slavists*. (Flier, M., Al.Timberlake, Eds.), pp. 390 – 405. Bloomington: Slavica.

GGjSh 1 = Demiraj, Sh. (red.). (2002). *Gramatika e gjuhës shqipe, Vëll. 1, Morfologjia*. Tirana: Akademia e Shkencave.

Giannakidou, A. and A. Mari. (2016). A unified analysis of the future as epistemic modality: the view from Greek and Italian. http://home.uchicago.edu/~giannaki/pubs/GiannaMari.Epistemic.future.NLLT.final.pdf (1 April 2019)

Heine, B. and M. Reh (1984). *Grammaticalization and Reanalysis in African Languages.* Hamburg: H. Buske.

Heine B., U. Claudi, and F. Hünnemeyer (1991). *Grammaticalization: a conceptual framework.* Chicago: University of Chicago Press.

Hopper, P. and E. C. Traugott (1993/2003). *Grammaticalization*. Cambridge: Cambridge University Press.

Heine, B. (1993). *Auxiliaries: Cognitive Forces and Grammaticalization*. New York & Oxford: Oxford University Press.

Heine, B. and T. Kuteva (2002). *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.

Heine, B. and T. Kuteva (2007). *The Genesis of Grammar*. Oxford: Oxford University Press.

Holton, D., P. Mackridge and I. Philippaki-Warburton. (2000). *Γραμματική της Ελληνικής Γλώσσας*. Αθήνα: Πατάκη.

Joseph, B. D. (2001). Is there such a thing as 'grammaticalization'? *Language Sciences*, 2001, pp. 163 – 186.

Joseph, B. D. (2003). Evidentials: summation, questins, prospects. – In: *Studies in Evidentiality* (*Typological Studies in Language* 54) (A. Aikhenvald and R. M. W. Dixon, Eds.), pp. 307-327. Amsterdam: Benjamins.

Kopitar, B. (1829). Albanische, walachische u. bulgarische Sprache. *Jahrbücher der Literatur*, Band 46, 59 – 106. Wien.

Lehmann, C. (1982/2002). *Thoughts on Grammaticalization: A Programmatic Sketch. Vol 1. Arbeiten des Kölner Universalien-Projekt.* Köln: Universitat Köln. [Revised edition *Thoughts on Grammaticalization. Arbeitspapiere des Seminars für Sprachwissenschaft der Universität Erfurt,* 2002, Nr. 9. Available at: https://www.christianlehmann.eu/publ/ASSidUE09.pdf]

Maslov, J. S. (1988). Resultative, Perfect and Aspect. *The Typology of Resultative Constructions.* (Nedjalkov, V. P., Ed.), 63 – 65. Amsterdam; John Benjamins.

Meillet, A. (1912). L'évolution des formes grammaticales. *Scientia: rivista internazionale di sintesi scientifica*, vol. 12, p. 384 – 400.

Mihoc, T. (2013). The Romanian presumptive mood: The key to the Romanian will-future. *Ottawa Papers in Linguistics* (online).

http://artsites.uottawa.ca/clo-opl/doc/Mihoc2013.final_.pdf (1 April 2019)

Miklosich, F. (1861). Die slavischen Elemente im Rumänischen. *Denkschriften der Wiener Akademie der Wissenschaft*, Band 12. Wien.

Traugott E. C. and B. Heine 1991. Introduction. In: *Approaches to Grammaticalization.* (Traugott E. C. and B. Heine, Eds.), 1 – 14. Amsterdam/Priladelphia: John Benjamins Publishing Company.

Trubetzkoy, N. S. (1930). Proposition 16. Über den Sprachbund. *Actes du premier congrès international des linguistes à la Haye, du 10-15 avril 1928*, Leiden: A. W. Sijthoff, pp. 17–18.

Vaillant, A. (1966). *Grammaire comparée des langues slaves.* Tome III, Le verbe. Paris: Klincksieck.

Zafiu, R.-I. (2009). Interpretări gramaticale ale prezumtivului. – In: *Studii de gramatică. Omagiu Doamnei Profesoare Valeria Guţu Romalo.* (R. Zafiu, B. Croitor, A.-M. Mihail, Eds.), 289 – 305. Bucharest: Editura Universităţii din Bucureşti.

Алексова 2003: Красимира Алексова. Удостоверителен перфект или умозаключителен аорист – от теоретичните основи към преподаването на чужденци. *Публикации от Юбилейната научнопрактическа сесия „40 години ИЧС"*, София, 60 – 66.

Асенова, П. (1989/2002). *Балканско езикознание.* Велико Търново: Фабер.

Герджиков, Г. (2003). *Преизказването на глаголното действие в българския език.* София: Университетско издателство "Св. Климент Охридски"

Търпоманова, Е. (2017). Лексикални, граматични и синтактични средства за изразяване на предположение в българския език. – В: *Надмощие и приспособяване. Сборник доклади от Международната научна конференция на Факултета по славянски филологии. Том II.* 52 – 59. София: Факултет по славянски филологии, СУ „Св. Климент Охридски".

Ekaterina Tarpomanova

Mapimi i *Sprachbund-it* edhe teoria e gramatikalizimit: rasti i gjuhëve ballkanike

Rezyme

Që nga fillimi i teorisë së Sprachbund-it, gjuhët ballkanike konsiderohen për-faqësuese për proceset e konvergjencës. Diakronikisht greqishtja, shqipja, rumanisht-ja dhe bullgarishtja (në një shkallë më të vogël serbo-kroatishtja) zhvilluan tipare të përbashkëta në të gjitha nivelet gjuhësore, d.m.th. fonologjike, leksikore, morfolog-jike dhe sintaksore. Sidoqoftë, ngjashmëritë morfo-sintaksore janë thelbësore për konceptin e Sprachbund-it, meqë ato janë rezistente ndaj ndryshimeve dhe kërkojnë një kontakt të zgjatur dhe të thellë gjuhësor.

Termi gramatikalizim u prezantua nga gjuhëtari francez Antoine Meillet në fillim të shekullit të 20-të, por korniza teorike u zhvillua disa dekada më vonë. Gramati-kalizimi i përcaktuar si "atributim i karakterit gramatikor një fjale autonome të di-kurshme" (Meillet 1912) konsiderohet në ditët e sotme si një nga proceset kryesore të ndryshimit të gjuhës.

Karakteristikat gramatikore të Sprachbund-it Ballkanik mund të analizohen për sa i përket gramatikalizimit. Në sistemin foljor të gjuhëve ballkanike mund të dallo-hen disa ndryshime të rëndësishme: foljet 'dua' dhe 'kam' janë gramatikalizuar dhe inkorporuar përkatësisht në format kohore të së ardhmes dhe të së kryeres. Për më tepër, bazuar në kohën e ardhme, forma të reja u specializuan për të shprehur supozim, ndërsa koha e kryer në shqip dhe bullgarisht e zgjeroi kuptimin e saj për të shprehur habitoren dhe evidencialitetin. Duke përdorur kritere të rrepta, ne mund të vlerësojmë nivelin e gramatikalizimit të tipareve ballkanizmave parësorë, si dhe të ndryshimeve më të vona.

Fjalë kyçe: Sprachbund Ballkanik, gramatikalizimi, e kryera, e ardhmja, eviden-cialiteti, habitorja