CHAPTER 2

Non-nominative and depersonalized subjects in the Balkans

Areality vs. genealogy

Victor A. Friedman and Brian D. Joseph
The University of Chicago / The Ohio State University

The languages of the Balkan sprachbund are surveyed here with regard to their constructions that show non-nominative subjects, typically in impersonal constructions. The issue of origins is considered, specifically as to whether these constructions represent inheritances from some earlier stage of the relevant languages or instead reflect the effects of contact. In the end, it is argued that a mix of areality, i.e. contact, and genealogy, i.e. inheritance, is needed to explain these constructions, with a nod required as well to typologically common patternings.

1. Introduction[1]

Masica (1976, 2001) claims that the construction he calls a 'dative subject experiencer' is an areal characteristic of South Asia and the Caucasus, and notes further that it can be found to some extent in Europe too. Although subject-type has

1. Throughout, in providing glosses for our examples, we follow the Leipzig Glossing Rules (LGR), though we augment the Leipzig list with: "IMPF" for imperfect (past tense); "IPFV" for imperfective (aspect); "MP" for medio-passive, in regard to certain nonactive verb forms in Albanian and Greek; "PC" for the Albanian so-called particle of concord; and "SP" for the subordinating particle found in all the Balkan languages. We also deviate from the LGR in the Balkan Romance examples (Aromanian, Daco-Romanian, and Megleno-Romanian) in using a hyphen after the weak object pronouns not to indicate a morpheme boundary but to follow the conventional orthography for those languages; also, "VOC" here refers to a vocative particle, not a vocative per se. For the element with the form *se* in Balkan Slavic and Balkan Romance and *pe* in Romani, we variably give the gloss "INTR" for intransitive or "REFL" for reflexive, or "MP", depending on the particular usage in a given example. We use the following abbreviations for languages discussed here: Alb, for Albanian; Armn, for Aromanian; BCSM, for Bosnian-Croatian-Serbian-Montenegrin; Blg, for Bulgarian; Grk, for Greek; Mac, for Macedonian; MR, for Meglenoromanian; Rmn, for Romanian; Rmi, for Romani; and, Trk, for Turkish; note also BR, for Balkan Romance, taking in Armn, MR, and Rmn; and BS, for Balkan Slavic, taking in Blg and Mac.

https://doi.org/10.1075/slcs.200.02fri

received considerable attention for individual branches of Indo-European as such (e.g. Barðdal et al. 2012), those subsequently experiencing contact in the Balkans (see Map 1) – Albanian, Greek, Indic, Romance, and Slavic – have not been examined areally for such features. Here, we look both to and beyond dative experiencers to constructions where marking of person for ostensible subjects is not to be found on the verb, the otherwise usual locus for such marking in the languages of the Balkans. We can identify seven Balkan construction-types with non-nominative or non-canonical subjects; of these, six are various impersonals (see Guentchéva 2010 on the Balkan situation with impersonals and Malchukov and Ogawa 2011 for some general typological considerations), while the seventh involves the neutralization of person marking, so that all seven may be subsumed under a label of "non-personal":

1. atmospheric and natural phenomena
2. experiencer constructions
3. impersonal constructions with corresponding personal forms
4. impersonal passives
    i. gnomic expressions (communicative)
    ii. generalized activity
    iii. potential
5. impersonal modals
    i. internal disposition
    ii. modalities of possibility and necessity
6. 'have' existential and 'be' possession
7. narrative imperatives

A tendency towards generally emotive non-nominative and impersonal subject usage is characteristic of the Balkan contact zone, and, as with several other Balkan developments, reveals BS-BR-Alb as central and Greek as distinctly marginal, with Romani usually closer to the BS-BR-Alb core, especially in co-territorial dialects. While some of these constructions are explainable as inherited (genealogical) or universal (typological) tendencies, the distributions of some modals and experiencers, as well as the narrative imperative, do appear to be areal. This statement reveals a typical methodological dilemma in dealing with the Balkan Sprachbund. Just as one has to be aware of the possibility of genealogical relatedness as a basis for a convergent feature, one also has to take into consideration the possibility of independent emergence of a construction or feature in the languages in contact, an independence that can be suggested by the occurrence of a feature outside of the contact zone. We argue that even when one has such considerations in mind, geography can often carry the day and point to contact as the basis for convergence. The goal of this paper, therefore, is to tease out, to the extent possible, the areal features involved in non-canonical subject marking in the Balkans. To do so, we

survey the various nonpersonal constructions, starting in § 2 with impersonals and then moving in § 3 to the imperative construction; in our conclusion in § 4, we assess the origins of the constructions in the languages of the Balkans as to areality, genealogy, and universality.

Romani Balkan I south II North Vlax south North			Turkic Turkish Gagauz Rumelian East West
Albanian Geg Tosk Lab Çam Arvanitika Arbëresh	Balkan Romance Daco-Romanian Wallachia Moldavia Banat Crişana Maramureş Istro- Romanian Vlah Aromanian North/West South/Eest Megleno-Romanian Tsărnareca Other	Hellenic Greek North South Tsakonian Judezmo East West	Balkan Slavic Torlak (SE BCS) Timok-NiŠava Prizren-south Morava Macedonian East West Bulgarian East West

Map 1. Languages and dialects of the Balkan Sprachbund

Many of the Balkan languages are overlapping and co-territorial, and it would therefore be misleading or inaccurate to label a specific territory with a specific language. While nation-state languages dominate in most of the territories of the respective nation-states, there are regions of various sizes in all of them where such is not the case. Moreover, at the beginning of the twentieth century all seven groups were represented on all the territories that would become today's nation-states, and in most states this is still the case today. The Aromanian isogloss bundle roughly follows the Macedonian one after intersecting with it. The distribution of Romani dialects is too complex to show on the map and so is given here. The Balkan I (South) dialects are spoken in Albania, southern Montenegro, Kosovo, southern Serbia, Wallachia and southern Moldavia, Bulgaria, Turkey, Greece, and Macedonia as well as Crimea. The Balkan II (North) are spoken in Bulgaria north of the Balkan range, southern Serbia, southern Kosovo, and northern Macedonia. The South Vlax dialects are approximately co-territorial with Balkan I (South), but exclude Crimea and Moldavia and include the rest of Montenegro and Serbia as well as Bosnia-Hercegovina, Dalmatia, and the Romanian Banat. The North Vlax dialects are spoken in Northern Serbia (Vojvodina), adjacent parts of southern Hungary and eastern Croatia across the Banat, Crişana, Transylvania, central Moldavia, and adjacent parts of southern Ukraine.

Source for Map 1: Revised from p. 203 in Friedman, Victor A. (2007). Balkanizing the Balkan Sprachbund: A Closer Look at Grammatical Permeability and Feature Distribution, in A. Y. Aikhenvald & R.M.W. Dixon (Eds.), *Grammars in Contact: A Cross-Linguistic Typology*, pp. 201–219, Oxford: Oxford University Press. © Victor A. Friedman 2007.

2. Impersonals

So-called impersonal constructions are traditionally taken to be those that, in a certain way, lack an overt subject or morphological specification of a subject, or that lack a fully referential subject. Referentiality is important because there are constructions with overt but nonreferential subjects – usually termed *expletive subjects* – that typically are subsumed under the rubric of impersonals. For example, the *it* in English expressions such as *It is raining* is overt but does not refer to any sort of specific entity, and yet, is a subject, behaving like a subject in that it is copied in tag questions (*It is raining, isn't it*) and can be raised with verbs like *seem* (*It seems to be raining*).

The qualification "in a certain way", given above, is necessary because impersonals typically appear to be subjectless, lacking even an expletive, as in Macedonian *Vrne* '[it] rains' but they differ systematically from pro-drop (null-subject) sentences. In pro-drop constructions, in languages that distinguish such categories as person and number of the subject via verbal inflection, the subject is specified and identified by that marking on the verb; by contrast, in an impersonal construction,

the verb is typically limited to the third person, usually the singular, and thus without any overt specification of the subject.[2] We, thus, take a somewhat broad view of "impersonals" here, deliberately so as not to arbitrarily exclude any potentially interesting sentence-type from consideration; we thus follow many linguists, e.g. Siewierska (2008), in identifying impersonals in this way. In what follows, the types identified in § 1 are surveyed.

2.1 Atmospheric and natural phenomena

Expressions for atmospheric and natural phenomena in the Balkans show considerable diversity within genealogical Balkan language groups, as well as commonalities that cross genealogical lines (see also Eriksen et al. 2010 on the typology of "weather constructions"). Whether or not the commonalities are contact induced turns out to be difficult to judge, since it can be argued that any given verb of weather is inspired by nature rather than nurture. Moreover, a given language can have more than one expression, especially for degrees of intensity. Still, the developments seen in the Balkans, if one focuses on the most common, intensity-neutral expressions, are striking, and suggest a number of interesting Balkan specificities.

For verbs of raining and snowing, Modern Greek *vréxi* and *xionízi* and Romanian *plouă* and *ninge* are noteworthy in that they preserve inherited specialized verbs translatable by the English verbs 'rain' and 'snow', respectively. South Danubian Balkan Romance, i.e. Aromanian and Meglenoromanian, however, have each innovated independently: Aromanian uses *da* 'gives' and Meglenoromanian uses *meardzi* 'goes' plus respective nouns for 'rain' and 'snow' (Armn *ploai, neauùă*, MR *ploaiă, neauă*).[3] In the Balkans, 'go' is not otherwise commonly used (although it can be encountered), and the only other language using 'give' is Romani (*del)*, where, interestingly enough, the subject is understood as *devel* 'god', as was the case in Sanskrit and Ancient Greek (admittedly with different verbs from 'give'). The Balkan Slavic languages are striking in their innovation and diversity in this regard. Except in a few dialects, none of them preserve the Common Slavic verb for 'rain' attested in Old Church Slavonic: *dъ̀ždi̯ti*. Both Bulgarian and Macedonian have innovated weather-specific impersonal

2. Suggestions such as Perlmutter and Moore's (2002) that there exists an expletive that controls agreement but has no phonological content is a theoretical issue that is beyond the scope of our presentation here. Narrative imperatives discussed below resemble impersonals in that there is no overt person and number marking of the understood subject.

3. Capidan (1925: 176) also reports reflexes of the inherited weather verbs: *ploaiă* and *neao* for MR. He also considers the use of 'go' in Meglenoromanian to be a typological parallel with East Slavic (cf. Russian *dožd' idët* 'rain goes'), since there is no evidence for a Balkan Slavic intermediary.

verbs that could be translated by English 'precipitate' since they can refer equally to rain and snow. In Bulgarian, the verb is *vali* from an earlier meaning 'roll', while Macedonian has a different verb, *vrne*. Georgiev et al. (1979: 211), speculate that *vrne*, which is attested in some so-called secret languages in Bulgaria, is a combination of *vali* and *rămi* 'drizzle', but given the choice of 'roll' in Bulgarian, it seems reasonable to speculate that a perfectization of *vrti* 'turn' with *-ne* could have produced **vrtne* whence *vrne*. In such a case, the Balkan Slavic speech area arguably shared a semantic shift, but with different lexicalizations. We can also note here that like Bulgarian and Macedonian, Turkish has a single specialized verb for raining and snowing, Standard Turkish *yağ-*, which, however, derives from 'pour, saturate' and also is the basis for the Turkish noun *yağmur* 'rain' (cf. *kar* 'snow').

Further to the north and west where Slavic is spoken, rain and snow 'fall' (e.g., Bosnian-Croatian-Serbian-Montenegrin (BCSM) *kiša pada* 'rain falls'), which is what they do in Albanian as well (e.g., *bie shi/borë* 'falls rain/snow'). While it is true that 'fall' is a natural verb to choose to describe what rain and snow do, it is nonetheless striking that in South Slavic, it developed precisely in that part where Albanian was spoken, keeping in mind that prior to 1878 mixed Slavic-Albanian populations extended east past the South Morava and as far north as Aleksinac. Then again, the Romanian borrowing of Slavic noun *zăpadă* 'snow' (itself containing the root *pad-* 'fall') alongside inherited *nea*, for which the verb is native *cădea* 'falls', points to a more widespread usage.

For expressions of cold weather of the type that in English is rendered *it is cold*, in the Balkans, the two principal verbs used are 'be' and 'make, do'.[4] Both are well attested outside the Balkans as well, but their distribution within the Balkans is nonetheless suggestive. In Modern Greek and Albanian, the verb of choice is 'make' (e.g. *káni krío* and *bën ftohtë*, respectively, for 'it is cold'), while in Bulgarian, Macedonian, Meglenoromanian, Romanian, and Romani, the verb 'be' is used (i.e., *studeno e*, *ladno e, iasti frig, este frig, šudro i[si]*, respectively, for 'it is cold').[5] Aromanian, however, straddles the two zones linguistically as well as geographically in this respect, having both types of expressions, i.e. both *fatsi arãstimi* and *easti arãstimi* or *arãstimi-i* for 'it is cold'.[6] Given the use of 'make' elsewhere in Romance,

4. Other possibilities, such as the Macedonian specialized verb *studi* 'it's [freezing] cold', are outside the scope of our consideration here, since the semantic parallels of auxiliary expressions are more readily comparable.

5. Turkish is also a 'be' language in this respect.

6. Capidan (1925: 177) identifies the Aromanian usage as a calque on Greek and notes that Weigand claims that *faţi* 'it makes' can also be used in Meglenoromanian, although Capidan did not encounter it.

it would appear that we have here a west-east divide, with Greek and Albanian representing the west, Slavic the east, and Balkan Romance influenced by Slavic, except Aromanian, which gives evidence of contact with both.

There is also an interesting difference between Macedonian and Bulgarian (cf. Guentchéva 2010 on Bulgarian). Although Bulgarian expressions of the type *vali* 'it's raining' or *studeno e* 'it's cold' are generally viewed as impersonal and therefore subjectless, for expressive purposes the expletive neuter pronoun *to* can be used, as in (1) and (2):

(1) Blg *Abe to naistina valjalo!*
VOC it truly rain.PRF
'Wow, it really [has] rained / (is raining)'[7]

(2) Blg *To naistina e studeno!*
it truly is cold(N).ADJ
'It really is cold'

In Macedonian, however, the equivalent sentences **Abe toa navistina vrnelo* and **Toa navistina e studeno* are not acceptable. There is thus an interesting difference here in a structural detail, with Macedonian being stricter about the possible presence of expletive subjects than Bulgarian. The Bulgarian here arguably represents an innovation, shared with East Slavic.[8]

The comparative results of the 'rain, snow, cold' impersonals are summarized in Tables 1 and 2. While Balkan expressions of weather do not show the across-the-board commonalities seen in some other areas of idiomatic expression in the languages, as documented by Papahagi (1908), Ikonomov (1968), and Thomaj et al. (1999), nonetheless the differential distributions are instructive and sub-areal connections that occur in other domains are to be found here, too. For instance, Bulgarian and Macedonian share an innovative semantic development, but with distinct realizations. Greek and Albanian converge, but differ in these expressions from Slavic and most of Balkan Romance, while Aromanian straddles the dividing line between east and west in some idioms of temperature. The languages are clearly differentiated, especially in the occurrence of impersonal expletive subjects

7. See Friedman (2012) for a discussion of the past tense in *-l*; the present reading is an admirative.

8. As Guentchéva (2010) notes, Skorniakova (2008) documents the occasional use of the ostensible neuter singular pronoun *ono* as an expletive subject with 'weather' verbs (and other impersonals) in both colloquial and literary Russian. Skorniakova also notes a corresponding use of the cognate *vono* in Ukrainian and *wono* in Lower Sorbian, this latter possibly under German influence.

and the lexical items used for precipitation, but they are also linked in terms of the semantics of innovations in describing precipitation.

Table 1. Expressions meaning 'it is raining/snowing'

	Greek	Romanian	Macedonian	Bulgarian	Albanian	Meglenoromanian	Aromanian	Romani
rain	vréxi	plouă	vrne dožd	vali dăžd	bie shi	meardzi ploaiă	da ploai [e]	del biršim
snow	xionízi	ninge	vrne sneg	vali snjag	bie bore	meardzi neauă	da neauă	del iv

Table 2. Expressions meaning 'it is cold'

	Greek	Albanian	Aromanian	Romanian	Meglenoromanian	Macedonian	Bulgarian	Romani
is cold	káni krío	bën ftohtë	fatsi/easti arătsimi~ arătsimi-i	este frig	iasti frig	ladno e/studi	studeno e	šudro i[si]

2.2 Experiencer constructions

The expression of internal experience, including emotions and feelings such as regret, shame, being cold, and the like, involves an impersonal construction in many languages of the Balkans, and as with the weather verbs of § 2.1, there is diversity in form but also some parallels suggestive of contact-induced convergence. The parameters of diversity are whether the expression is primarily noun-centered or verb-centered, which verb is used, how the experiencer is encoded, i.e. as a subject or an object, and, if an object, what case-marking the experiencer receives.

By way of illustrating the situation with experiencers and the variation these constructions show across the languages, Balkan forms for (3) 'I am sorry', (4) 'I am ashamed', and (5) 'I am cold' serve as representative examples of the class of constructions; note that unless marked as MP or INTR all verbs are active in these examples. They are then followed by some observations on the import of the structures evident here and the groupings that emerge from the data.

(3) Grk *lipáme*
regret.PRS.1SG.MP

Alb *më* *vjen* *keq*
me.DAT comes.PRS.3SG bad'

Rmi *pharo si mange*
heavy is me.DAT

Rmn *mi- este* *milă*
ime.DAT-is pity

Armn *njilă nj- easti*[9]
ipity me.DAT-is
Mac *žal mi e*
sorry me.DAT is
Blg *žalno mi e*
sorry me.DAT is
'I'm sorry'

(4)[10] Grk *drépome*
shame.PRS.1SG.MP
Rmi *ladžava*
shame.PRS.1SG[11]
Alb *më vjen turp*
me.DAT comes.PRS.3SG shame
Rmn *mi- e ruşine*[12]
me.DAT-is shame
Armn *nj- easti arshini*
me.DAT-is shame
Mac *mi e sram*
me.DAT is shame
Blg *sram me e*
shame me.ACC is
'I'm ashamed'

(5) Grk *krióno*
cold.PRS.1SG
Alb *kam ftohtë*
have.PRS.1SG cold
Rmi *šudro i mange / pahol man*
cold is me.DAT / cold.INTR.3SG me.ACC

9. The Balkan Romance use of *milă/njilă* (with *m* > *nj* /__*i* via regular sound change) is an early borrowing from Slavic, and reflects the OCS *milъ* 'pity, compassion', not the modern South Slavic meaning of *mil* 'dear'.

10. There are variant formulations in some of the languages, e.g. Alb *turpërohem* 'be.ashamed. MP.1SG' or Mac *se srami* 'REFL shame.3SG', often with slightly nuanced differences in meaning.

11. The Romani verb is derived from *ladž[avo]* 'shame' and shows a variety of conjugations in various dialects. As a regular *a*-stem verb it is unmarked for voice, 3SG *ladžal* has the meaning 'be ashamed'. In some dialects, the verb is derived with the suffix *-ov-* 'become'.

12. Romanian also has the possibility of a full verb with an accusative here: *mă ruşinează* 'I am ashamed'. We thank an anonymous reader for this example, from the noncancase database associated with the EVALISA project <www.evalisa.ugent.be>.

Rmn	*mi- e*	*frig*					
	me.DAT-is	cold					
Mac	*ladno*	*mi*	*e*	/	*mi*	*se*	*studi*
	cold(N).ADJ	me.DAT	is	/	me.DAT	INTR	cold.PRES.3SG
Blg	*studeno*	*mi*	*e*				
	cold(N).ADJ	me.DAT	is				
	'I'm cold'						

In each of these data sets, Greek and Albanian stand out from the other languages, though in different ways. Greek shows verb-based expressions in each case, with the experiencer encoded as the subject of the verb, marked via the person/number endings (active or nonactive (mediopassive), depending on the verb itself); an overt subject nominal could of course occur, e.g. *eyó lipáme*, but since Greek is a pro-drop language, generally no such nominal is found.

For 'shame', Romani also has a fully verbal expression; in this case, though, as with all the Greek examples, it is a matter of a retention of a verbal construction from earlier stages of the languages: New Testament Greek had a mediopassive verb, *splankhnízomai*, for 'I feel sorry', and Sanskrit similarly has the mediopassive verb *lajje* for 'I feel ashamed'. Albanian typologically seems to be transitional in that it has a blend of a verbal and a nominal construction: the noun in (3) through (5) carries the primary semantic weight – *keq* '(something) bad', *turp* 'shame', and *ftohtë* 'cold', respectively – but the verbal part is not a typical "light" verb; rather *vjen* 'come' and *kam* 'have' occur, each of which carries some semantic weight of its own.[13] The encoding of the experiencer in Albanian 'cold' is via the subject of the verb 'have', but in the other two examples, the experiencer is a dative nominal (here the 1SG.DAT weak pronoun *më*). Such constructional variation between a verbal and a nominal construction, representing a different type of transitional state, occurs also in Romani for 'cold'. The other languages have noun-based constructions throughout, and encode the experiencer through an oblique case-marked nominal; like Albanian, both Romanian and Macedonian have a dative experiencer.[14]

Although Bulgarian patterns with most of the other Balkan languages (including Macedonian) in using dative experiencers, it also has accusative experiencers for a small, closed class of ten nouns denoting difficult feelings (shame, fear, disgust, worry, disquiet) plus the noun (not necessarily denoting a negative but definitely

13. By contrast, a typical "light" verb would be 'be' or 'make', as found in various periphrases cross-linguistically.

14. Macedonian has a single accusative experiencer in the fixed expression *sram [da] te bilo!* 'Shame on you!' (lit. shame [SP] you.ACC be.PST), where the verb form preserves an archaic optative usage in addition to the archaic use of the accusative.

something potentially irritating) *gădel* 'sensation of being tickled'.[15] Such nouns take datives elsewhere in the Balkans, as do most other Bulgarian nouns of negative feeling. The specific distinctions are consistent with those in Barðdal (2011, 2015), but Bulgarian use of the accusative requires further investigation.

As is often the case, Greek is the most archaic outlier in terms of these experiencer constructions, whereas Albanian and Romani show some parallelism with what is found elsewhere in the Balkans. The greatest convergence, however, is between Balkan Slavic and Balkan Romance, in regard to oblique marking of experiencers and the use of a noun-based construction. Here, the geographical and structural facts thus support the possibility of a contact-induced change. In this regard it is significant that Balkan Romance diverges in this respect from its Latin source. The corresponding Latin constructions involved verbs, whether expressed personally, as in (6a, b), or impersonally, as in (6c, d):

(6) a. *misereor*
1SG.PRS.MP
'I feel compassion'
b. *frīgeō*
1SG.PRS
'I am cold'
c. *me* *miseret*
me.ACC sorry.3SG.PRS
'I feel sorry for'
d. *me* *pudet*
me.ACC shames.3SG
'I'm ashamed'

Thus, the experiencer constructions show some convergence that is likely to be contact-related, specifically involving Balkan Romance assimilating to Balkan Slavic nominal constructions, but overall the patterns of relationships that emerge within the Balkans for experiencer impersonals is different from that seen with the weather verbs; Albanian and Greek, for instance, do not match up here in the way they do with certain weather expressions.

15. The ten break into five pairs: (1) *jad* = *gnjav* (archaic) 'upset'; (2) *gnus* = *gad* (dialectal) 'disgust'; (3) *griža* = *enja* (< Grk *énja*; southeast dialect, used only in negative and ironic clauses) 'care'; (4) *sram* = *grjah* (archaic) 'shame'; (5) *strah* = *băz* (colloquial) 'fear'. In each pair, the item on the left is standard Bulgarian while the item on the right is stylistically marked as indicated in the parentheses. With the exception of *enja*, these words are all inherited.

2.3 Impersonal constructions with corresponding personal forms

The third type of impersonal subsumes those that, unlike the others discussed here, show a systematic correspondence, or alternation, with a personal construction with the same verb. The phenomenon in question is somewhat like verb lability, except that instead of pairings of intransitive and transitive uses for the same labile verb form (as in Grk *pijéni* 's/he.goes' / *pijéni kápjon* 's/he.takes someone (somewhere)' or Mac *zaspiva* 's/he.sleeps' / *go zaspiva dete-to* 'it.ACC s/he.puts.to.sleep child-the', i.e. 's/he puts the child to sleep'), the alternation with the impersonals here is between an impersonal expression with a source adjunct and personal counterpart with the source as subject. A paradigm case is the verb for 'drip', e.g. impersonal '(it) drips from the faucet' and personal 'the faucet drips', illustrated here with the verb in bold in each example (cf. Guentchéva 2010: 38):

(7)

Alb	*robineti*	***pikon***	=	***pikon***	*nga*	*robineti*					
	the.faucet	drips	=	drips	from	the.faucet					
Mac	*tapata*	***kape***	=	***kape***	*od*	*tapata*					
	the.faucet	drips	=	drips	from	the.faucet					
Blg	*trăbata*	***kape***	=	*ot*	*trăbata*	***kape***					
	the.faucet	drips	=	from	the.faucet	drips					
Rmi	*i*	*češma*	***thavdela***	=	***thavdela***	*e*	*češmastar*	/	*tar-i*	*češma*	
	the	faucet	drips	=	drips	the	faucet.ABL	/	from-the	faucet	
Grk	*i*	*vrísi*	***stázi***	=	***stázi***	*apó*	*ti*	*vrísi*			
	the	faucet	drips	=	drips	from	the	faucet			
Rmn	*țeava*	*se*	***scurge***	=	*se*	***scurge***	*din*	*țeava*			
	the.faucet	INTR	drips	=	INTR	drips	from	the.faucet			

In none of the right hand equivalents is an overt expletive subject possible, so that, for instance, **aftó stázi apó ti vrísi* 'it drips from the faucet' with neuter singular *aftó* 'it', is impossible in Greek. Only Romanian marks the verb overtly as intransitive, through the use of the so-called reflexive (i.e., intransitive/nonactive) marker *se*. One can see superficial structural similarities in collocations such as those illustrated in (7), specifically the existence of the alternation itself and the fact that the same form of the verb is used in each of the two expressions. However, other than those similarities, there are no geographic distributional facts or striking or unusual convergent details that would warrant a contact explanation for the facts. Nonetheless, in an inventory of Balkan impersonals, the convergence of dual patterning for the same verb form seen in (7) is noteworthy.

2.4 Impersonal passives, real and potential

While the prototypical passive in a nominative-accusative language shifts focus to a direct object by making it a subject while the active subject becomes an omissible agent, the impersonal passive focuses on the event itself, with no overt expression of the agent. This event can be gnomic, generalized, or potential.

2.4.1 *Gnomic impersonal passives: Verbs of speaking/communicating*

The gnomic impersonal passive is exemplified by verbs of speaking and communicating more generally, e.g. 'say' or 'write'. Such impersonals have a complement as their object. In English, *It is said that …* functions as the impersonal passive of *They say that …*, where the active subject *They*, rather like the passive expletive subject *It*, has no specific or definite referent. Such phrases can introduce gnomic generalizations (e.g., proverbs) as well as specific clauses (e.g., quotations from the Bible). The unspecified nature of the subject warrants considering them to be impersonals, as in Guentchéva (2010),[16] and their passive form warrants placing them here.

All of the Balkan languages have passives with verbs of communicating, formed in the usual way for each language, thus synthetically via passive voice morphology on the verb in Greek and Albanian and periphrastically via the use of the reflexive/intransitive marker, *se* in Balkan Romance and Balkan Slavic, and *pe* in the Romani of the Balkans, with the active verb form.;[17] The *se*/*pe(s)* morpheme of, respectively, Balkan Romance and Romani shows person agreement with the subject, varying in form according to the person of the subject, whereas the corresponding element in BS (and also in some Romani dialects influenced by Slavic) is invariant, always in the *s-* form in Slavic and *pe(s)* in Romani. Relevant examples are given in (8), along with a complementizer that would introduce the clausal complement.

(8) Grk *léjete* *óti* …
is.said.PRS.3SG.MP that …
Alb *thuhet* *se* …
is.said.PRS.3SG.MP that …
Rmn *se* *spune* *că* …
MP say.PRS.3SG that …

16. See also Siewierska (2011) on the use of third-person forms in impersonals. A reviewer astutely pointed out that such third persons do have a thematic role in the argument structure, despite their indefinite and nonspecific referentiality, but their inclusion here is in keeping with the broad view of impersonals we take in this study.

17. In the Albanian aorist and related paradigms built on that stem, the mediopassive is periphrastic and marked by *u*, which corresponds etymologically, in ways too complicated to allow for presentation here, to the *se* in BR and BS.

Blg	*kazva*	*se*	*če*	…
	say.PRS.3SG	MP	that	…
Mac	*se kažuva*		*deka*	
	MP say.PRS.3SG		that…	
Rmi	*vakerela*	*pe*	*kaj*	…
	say.PRS.3SG	MP	that	…

'It is said/They say that …'

In Greek, one finds also the indefinite-subject 3rd person plural active form *léne* 'they say', a reduced form of the fuller form *léyun(e)*, which is less common in this context, much like the English *they say* noted above. This usage also occurs in Latin, e.g. *legunt* 'they say', so that the Romanian use of a nonactive singular form here seems to represent a favoring, possibly contact-related given what is found in neighboring languages, of one alternate means of expression over another. Still, in addition to the nonactive type in (8), Păna Dindelegan (2013:106) notes that a few third person singular active verbs, e.g. *zice* 'it calls [=s/he is called]', can be used as impersonals of the 'they VERB …' type, in what is apparently an innovation away from the impersonal third plural of Old Romanian.

In Slavic generally (e.g. Russian *govorjat*), the indefinite active 3rd person plural usage is possible, or even obligatory, although for Balkan Slavic it not as usual as the nonactive usage illustrated by (8) (cf. Guentchéva 2010: 41). The preference of 3SG gnomic impersonals is seen in Old Church Slavonic (OCS), where the form, however, is active, e.g. *pišetъ* 'write.3sg.IPFV.PRS' meaning 'it is written / they write'. This form occurs four times in Matthew (4:4, 4:6, 4:7, 4:10), and each time, importantly, it corresponds to a Greek passive verb, and a perfect passive at that (*gégraptai* 'it has been written', i.e. 'it is in a state of having been written', thus 'it is written'). This possibility has been given up in modern Balkan Slavic, and the passive construction in (8), which was a possibility in OCS as well, has been selectively adopted and favored. The basis for the parallel favoring of a nonactive expression in Balkan Romance and Balkan Slavic is not clear, but it represents another way in which these languages show convergence in the domain of impersonal expressions. Similarly, the parallel Romani construction shows a movement away from earlier Indic syntax, both the Sanskrit use of a quotative particle (*iti* 'thus'), with or without an overt verb of communicating, and the Middle Indo-Aryan (cf. Bubenik 1998: 150–154) use of a synthetic passive or an ergative construction; Balkan Romani here has calqued from Slavic the use of an invariant form of the reflexive pronoun *pe[s]* as a passive/intransitive marker.[18]

18. Romani in general has derived, synthetic passive marking (see Boretzky & Igla 2004 I:179–181), so the use of the reflexive is noteworthy here. Moreover, as noted above, most Romani dialects, like Balkan Romance, have an agreeing reflexive marker (showing agreement in person), the invariant marking being a calque on Balkan Slavic.

2.4.2 *Impersonal passives for generalized activity*

Impersonal passives with unspecified agents to express a generalized activity can also be formed by means of mediopassive or intransitivized verbs in all the Balkan languages except Greek. These forms thus focus entirely on the action expressed in the verb, inasmuch as there is no specified subject and no logical object, and they represent an activity as going on in a general way. A typical translation for such a passive with a verb in English would be an impersonal 3rd plural active. This impersonal passive use is found in most of the Balkan languages.[19] A representative sampling is given in (9) for 'There was fighting over there; fighting was taking place there; they were fighting over there':

(9) Alb *andej luftohej*
over.there fought.IMPF.3SG.MP

Mac *tamu se boreše*
there MP fought.IMPF.3SG

Rmn *atunci se lupta*
there MP fought.IMPF.3SG

Rmi *okothe marela pe sine*
there fight.PRS.3SG MP be.IMPF

Blg *Tam* se *boreše*
there MP fought.IMPF.3SG

vs. Grk *polemúsan*
fight.IMPF.3PL

'They were fighting/Fighting was going on over there'

In (9), Albanian uses an overt mediopassive verb and Balkan Slavic, Balkan Romance, and Romani use their equivalents of a passive, consisting of an (etymological) reflexive marker with an active verb. Greek is exceptional here in that it uses a 3rd person plural active form with an understood indefinite subject, much like the alternant in § 2.4.1 concerning verbs of speaking. The convergence that the other languages show is striking, and the Greek exceptionality parallels what is found with some of the other impersonals. The historical interpretation of all this is difficult, but some important observations can be made. In the case of Balkan Slavic, this usage occurs in North Slavic as well, as Kibort (2001) and Gawelko (2005) demonstrate for Polish, the latter also giving the Indo-European background. Moreover, Example (10), which is comparable to (11a) below, is cited by Večerka (1996: 241) from a manuscript outside the canonical OCS corpus but of similar antiquity:

19. This construction occurs outside of the Balkans, for instance in Polish, as a reviewer observed. As noted in the introduction, this sort of issue arises repeatedly, and quite rightly, when one is trying to establish an areal basis for a particular phenomenon. However, occurrence outside of the geographic area of interest does not in itself vitiate a claim of areal convergence through contact.

(10) *pridetъ sę*
comes.PRS.3SG INTR
'someone comes'

Thus, there were probably antecedents within early Slavic for the Bulgarian and Macedonian usage in (9); indeed, the occurrence of impersonal passives in Baltic (e.g. in Lithuanian), admittedly with a different structure but impersonal and passive nonetheless, and elsewhere in Indo-European, e.g. Sanskrit and Old Norse, would suggest as much. Moreover, Latin also had this usage, as shown in (11):

(11) a. *itur*
'go.PRS.3SG.MP'
'someone goes' (literally: "it.is.gone (by someone)")
b. *pugnatum est acriter*
'fought(N).PASS.PTCP.SG is.PRS.3SG fiercely'
(literally "it-was-fought fiercely (by someone)"
'there was a fierce fight' (Caesar, *Gallic Wars* 3.21.1)

This Latin usage is presumably the basis for the Balkan Romance, updated to the current morphosyntactic analogue to the Latin passive. The Balkan Slavic usage also appears to have its source in Common Slavic. Therefore, the striking convergence seen in (9) probably reflects a superficial parallel involving independent lines of descent into the modern languages rather than a contact-related convergence, although with insufficient evidence for Albanian, we cannot be sure of the construction's status in that language.[20]

2.4.3 *Impersonal passives expressing potential*

The impersonal passive that expresses generalized activity can also be used to indicate the potential for an activity to occur as illustrated by Example (12):

(12) Alb *Këtu flihet mirë*
here sleep.PRS.3SG.MP well
Mac *Ovde se spie dobro*
here INTR sleep.PRS.3SG well
Blg *Tuk se spi dobre*
here INTR sleep.PRS.3SG well

20. Pană Dindelegan (2013: 173) notes that a reflexive as non-agentive construction is attested in Vulgar Latin, and that such constructions are found in all the Romance languages except French. However, she also notes that the construction is absent from earliest Romanian texts (16th century) and is rare in Old Romanian (17th-18th century). To the extent that the convergence is not merely typological, therefore, contact rather than inheritance seems a more likely explanation.

Rmi	*akathe*	*sovela*		*pe*	*šukar*		
	here	sleep.PRS.3SG		INTR	well		
Rmn	*aici*	*se*	*doarme*		*bine*		
	here	INTR	sleep.PRS.3SG		well		
vs. Grk	*borí*	*na*	*kimiθí*		*kanís*	*kalá*	*eðó*
	can.3SG	SP	sleep.PRS.3SG		someone	well	here

'One can sleep well here'

By extension, such impersonal passives, when negated, are used to indicate impossibility, essentially an interdiction, as in (13):

(13)

Alb	*këtej*	*s'*	*kalohet*			
	here	NEG	pass.PRS.3SG.MP			
Blg	*ottuk*	*ne*	*se*	*minava*		
	from.here	NEG	INTR	pass.PRS.3SG		
Mac	*otdovde*	*ne*	*se*	*minuva*		
	from.here	NEG	INTR	pass.PRS.3SG		
Rmn	*de*	*aici*	*nu*	*se*	*trece*	
	from	here	NEG	INTR	pass.PRS.3SG	
Rmi	*akatar*	*na*	*nakhela*		*pe*	
	from.here	NEG	pass.PRS.3SG		INTR	
vs. Grk	*ðen*	*borí*	*na*	*perási*	*kanís*	*eðó*
	NEG	can	SP	pass.3SG	someone	here

'One cannot pass here' = 'It is impossible to pass here' (literally, in all but Grk, "it is not passed here")

Once again, there is convergence involving all of the languages except Greek, which uses an overt modal verb, here *borí* 'can'. The pattern of superficial structural convergence is, thus, the same as with other impersonal passives, especially of the generalized activity type. Here, however, it may be that the earlier stages of at least Balkan Slavic and Balkan Romance did not have this nuance associated with the generalized activity impersonal construction, again judging from the evidence of OCS and Latin, so that the possibility that the specifics of the convergence in (13) are contact-induced must be seriously entertained. Thus the convergence is clear, even if the exact causation and the directionality of external influence, if any, are not.

2.5 Impersonal modals

As the potential impersonal passive of § 2.4.3 indicates, impersonal constructions can be used in the expression of various types of modality. Going beyond the potentiality of the impersonal passives, and looking further at the semantics of other modal expressions, one can distinguish impersonal constructions which, somewhat akin semantically to the internal experience of feelings and emotion (see § 2.2), highlight nuances of speaker disposition and intent, as well as impersonals that express the modality of possibility and necessity. Both show interesting patterns of convergence in the Balkans.

2.5.1 *Internal disposition*

The expression of internal disposition, positive or negative, on the part of a speaker can be expressed in most, and to some extent all, Balkan languages through the use of an impersonal passive construction with a dative (or equivalent) personal pronoun indexing the experiencer. These forms translate as 'X feels like …' or 'X doesn't feel like…', though literally they are "to-X VERBs/is-VERBed …", as illustrated in (14):

(14)
Mac	*mi*	*se*	*jade*		*(burek)*
Armn	*nji*	*-si*	*mãcã*		*(burec)*
MR	*ăń*	*-ți*	*máncă*		*(burec)*
	me.DAT	REFL	eat.PRS.3SG		(burek)
Alb	*më*	*hahet*			*(byrek)*
	me.DAT	eat.PRS.3SG.MP			(burek)
Blg	*jade*	*mi*	*se*		*(bjurek)*
	eats.PRS.3SG	me.DAT	REFL		(burek)
Rmi	*hala*	*pe*	*mange*		(*bureko*)[21]
	eat.PRS.3SG	REFL	me.DAT		(burek)
	'I feel like eating (burek)'				

The parallelism among these languages is striking, in both the form and the functional nuance associated with it.

---

21. This particular example is based on a judgment given by an informant (VAF, field notes); in a naturally occurring example, the *feel like* construction was not chosen, even though the Macedonian that the speaker code-switched into (bolded here) does use it:

shukar	*i,*	*na*	*mangav*	*te*	*hav*	***ne***	***znam***	***zoshto,***	***ne***	***mi***	***se***	***jade***
good	is	not	I.want	SP	I.eat	not	I.know	why,	not	me.DAT	REFL	eat.PRS.3SG

'It is good; I don't want to eat – **I don't know why, I don't feel like eating**'.

This type of construction is absent from Romanian and Greek, which for (14) would use lexical constructions as in (15) and (16), respectively:

(15) *vreau să mănânc (burec)*
I.want SP I.eat
'I want to eat (burek)'

(16) *éxo ðjáθesi na fáo (buréki)*
I.have mood SP I.eat (burek)
'I feel like eating (burek)'

There is, however, an interesting dialectal point of contact involving Greek: Papanastassiou and Papadamou (2012) report that in the Greek of the Kastoria region – where, we can add, Macedonian was the majority language until after the Balkan Wars (1912–1913) – there are villages that have the following constructions:

(17) a. *mi trójiti*
me.ACC eat.PRS.3SG.MP
'I feel like eating' (literally: "(to-)me (it-)is-eaten")
b. *mi píniti*
me.ACC drink.PRS.3SG.MP
'I feel like drinking' (literally: "(to-)me (it-)is-drunk")

which exactly calque the type illustrated in (14), with mediopassive voice forms and with an indirect object weak pronoun.[22] Given the geographical limitation of this type within Greek to just those northern varieties in contact with Macedonian (and also Albanian and Aromanian), it can be safely assumed that this represents a contact-induced innovation in Greek. Moreover, since the construction is found throughout Slavic, and is absent from Balkan Romance north of the Danube as well as from Romani dialects not in contact with Slavic, it is safe to conclude that the usage is of Slavic origin in the other Balkan languages, although, as usual, the lack of ancient evidence for Albanian leaves a degree of uncertainty.

2.5.2 *Modality of possibility and necessity*

Possibility and necessity are expressed similarly in the Balkan languages, with an impersonal verb in a 3rd person singular form followed by a complement with a personal verb headed by the subordinating subjunctive particle. In that way these impersonals are parallel to auxiliaries in other languages in terms of their function,

22. The verbs in (17) show the vowel raising characteristic of a northern dialect, so that they are from an earlier *trójete / pínete*, the forms found in southern dialects. Similarly, the experiencer object is accusative here, reflecting the northern use of the accusative for indirect objects, parallel to the genitive in southern dialects of Greek and datives in other languages.

but they behave like regular complement-taking main verbs. Examples are given for possibility in (18) and for necessity in (19); in (18) the sentences mean 'Perhaps/It-might-be-that I will eat X (X = burek)' and literally are "(it-)can that I-eat X"), while in (19) the meaning is 'I must eat' and the forms literally are "(it-)must that I-eat":

(18)

Mac	*može*	*da*	*jadam*	*burek*
Blg	*može*	*da*	*jadem*	*bjurek*
Alb	*mund*	*të*	*ha*	*byrek*
Grk	*borí*	*na*	*fáo*	*buréki*
Rmi	*šaj*	*te*	*hav*	*bureko*
Armn	*poati*	*s-*	*mãc*	*burec*
MR	*pọti*	*să*	*mănânc*	*burec*
Rmn	*poate*	*să*	*mănânc*	*burec*
	it.can	SP	I.eat	burek

'Perhaps I can eat burek'

(19)

Mac	*treba*	*da*	*jadam*	*burek*
Blg	*trjabva*	*da*	*jam*	*bjurek*
Alb	*duhet*	*të*	*ha*	*byrek*
Grk	*prépi*	*na*	*fáo*	*buréki*
Rmi	*valjani/trebul*	*te*	*hav*	*bureko*
Rmn	*trebuie*	*să*	*mănânc*	*burec*
Armn	*lipseashti*	*s-*	*mã*	*burec*
	it.need.	SP	I.eat	burek

'I need to eat burek'

These verbs can also have personal forms in all of the languages except Romani. Regarding complementation, it can be noted too that 'can' is one of the last verbs in Romanian that still allows an infinitival complement, e.g. *eu pot merge* 'I can go' (alongside *eu pot să merg*).[23]

An important diachronic development with these paticular impersonals is that in some instances they have developed, or can be assumed to have developed, out of personal constructions. For instance, inflected 'can' (root *mog-*) is found in all Slavic languages, and impersonals with this root are, outside of South Slavic, more likely to involve a predicative adverb *možno* (< **mogĭno-*), as in Russian, than a verb, so that the impersonal *može* of South Slavic is likely to be the innovation, although

23. In Bulgarian, the remnant of the Common Slavic infinitive (without final *-ti*) can also occur after *moga* 'can', e.g. *moga izjade edna kruša* 'I can devour a pear', but Nicolova (2008: 444) writes that it is no longer used except in extremely rare cases with *nedej(te)* 'don't!', and, we can add, in some individual speech styles or some rural dialects.

quite old, as it is attested in OCS. In Albanian, inasmuch as *duhe-* is formally the mediopassive of *dua* 'want; love', it can be assumed that the 3sg *duhet* 'it is necessary, it is needed' is a specialization of a personal *duhet* 'she/he/it is wanted'.[24] Similarly, in Ancient Greek, the root *prep-* could occur in a personal construction meaning 'be visible; be like' as well as, more rarely, 'be fitting', while the 3sg, impersonal, form meant 'it is fitting' (with dative of the affected nominal plus and infinitive complement), and it is a safe assumption that the impersonal form developed from the personal. The depersonalization of finite verbs meaning 'want' is also the source of the future markers in all the Balkan languages.

The source of this depersonalization is not clear, and it has been speculated, at least regarding the future (Joseph 1983), that such simplifications are due to a drive to reduce redundancy. While there may be some validity to such a notion, other morphosyntactic developments in the Balkans, in particular the replacement of the infinitive by finite forms, add redundancy via personal marking on subordinate clause verbs. However, while infinitive replacement does increase redundancy, it also reduces the inventory of morphological forms, and thus, together with the generalizations of 3rd person forms in the modal constructions under consideration here, consitutes a type of simplification, itself typical of many contact situations. Thus, depersonalized modals contribute to the superficial parallelism in Balkan syntax, and even if directionality is difficult to determine, it is nonetheless a convergent phenomenon consistent with general tendencies in language contact.[25]

2.6 'Have' existential vs. 'be' possession

The final type of impersonal to be discussed here has to do with the expression of existence and possession. It is well known that these two notions show relationships in terms of their formal expression cross-linguistically, with what might seem like the quintessential verb of possession, 'have', often figuring in existential constructions, and the quintessential verb of existence, 'be', in possessive constructions. Both constructions are found in the Balkans, with some patterns reflecting convergent inheritances and others sprachbund-related contact.

---

24. On the relationship between 'want' and 'need', note the use of active *θélei* in Greek, which generally means 'wants', in the sense of 'need' as in *to stifáðo θélei líyo aláti* 'the stew wants/needs a-little salt'. Such usage also occurs elsewhere in the Balkans, e.g. Macedonian *saka* 'it wants', in similar constructions, and other languages outside the Balkans.

25. We can add that depersonalizing in the future shows that futurity in the Balkans behaves like a mood.

The basic facts concerning existential constructions with 'have' in the Balkans are that Greek, Albanian, Balkan Slavic, and Balkan Romance all signal existence and nonexistence via an impersonal 3sg form of the verb 'have', negated if appropriate, with the nominal for which existence is predicated expressed in the accusative case, for those languages that distinguish nominal case; some examples are given in (20):

(20) Mac *ima ženi tuka*
has.3SG women here
'There are women here'
Blg *njama ženi tuk*
not.has.3SG women here
'There aren't any women here'
Alb *ka dikë këtu*
has.3SG someone.ACC here
'There is someone here'
Grk *éxi axinús s ti θálasa*
has.3SG urchins.ACC in the sea
'There are sea-urchins in the sea'
Armn *aoatsi-ari multsã mileti*
here-has.3SG much people
'There are a lot of people here'
MR *ari lucru*
has.3SG work
'There is work'

The Romanian situation is somewhat more complicated. Caragiu-Marioțeanu (1958) states that Romanian existential 'have' occurs in only in two contexts – with a negative plus indefinite object as in (21a) and the fixed phrase in (21b):

(21) a. *n-are cine aici să ajute*
NEG-has.3SG whom.ACC here SP helps.SBJV.3SG
'There is no one here to help'
b. *n-are nimic*
NEG-has.3SG nothing
'Never mind!'

The fixed expression in (21b) resembles Albanian *s'ka gjë* 'it's nothing; it's all right' (literally 'not has thing'), though the provenance of this parallel is unclear as to whether it involves borrowing in some way or is simply an independent development in each language.

The history involved seems fairly clear. Greek has had an existential construction with 'have' like that in (20) since Roman times in the post-Classical period, although it is not found in Classical Greek; given the chronology of the emergence of this construction in Greek, it is most likely a calque on a late Latin model. The

Latin construction, with the form *habet* '(it-)has' plus accusative – and for Western Romance the adverb *ibi* 'there' as well[26] – is quite likely the source, via calquing, of the Albanian construction, and is definitely the ancestor of the Balkan Romance construction. The restricted nature of the 'have' existential in Romanian, however, needs to be explained. Influence from early Slavic would give a possible explanation, since the Balkan Slavic formation of (20) with 'have' represents an innovation away from the early Slavic type with 'be'; that is, the 'be'-type is seen in OCS (present tense *jes-*, future *bǫd-*, past *bě-*, cf. Lunt 2001: 164) and is found still elsewhere in Slavic, e.g. Russian *jest'* 'there is/are', even though the 'have'-type occurs to some extent outside of South Slavic.[27] A reasonable scenario, therefore, is that the Aromanian and Meglenoromanian unrestricted use of 'have' in existentials reflects an inheritance from Late Latin *habet*, as elsewhere in Romance, while the Romanian restriction and spread of 'be' in existentials reflects Slavic influence, either learned influence from OCS or influence at a more colloquial level before Balkan Slavic innovated the use of 'have'. The South Slavic innovative generalization of a 'have'-existential is perhaps best seen as the result of the South Danubian Balkan Romance substratum (cf. Gołąb 1976). The spread into BCSM can be compared to the spread of the 'will' future, which affects all of South Slavic except Slovene and the Kajkavian dialects of Croatian. Both contact with old populations of Balkan Romance speakers and the influence of Greek in the Orthodox Church may have been factors.

A further point of convergence with the 'have'-existential, but also an area of some divergence among the languages, is in the possibility of the use of an object pronoun as the entity for which existence is being predicated. Thus Macedonian (and, *mutatis mutandis* with regard to word order, Bulgarian, and BCSM), Albanian, and Aromanian have this structure, while Greek does not, as the examples in (22) show:

(22)	Mac	*gi*	*ima*
	Alb	*i*	*ka*
	Armn	*l-*	*are*
	Grk	**ta*	*éxi*
		them.ACC	has.3SG
		there are …' (= 'they exist')	

26. The presence of *ibi* is posited to account for the *y* in the Spanish existential *hay* and the French existential *il y a*.

27. Both Polish and Ukrainian show a limited use of impersonal 'have' in existentials, specifically in negative constructions in the present tense (*nie ma* / *nemaje*, respectively), as discussed in Twardzisz (2012). Such facts suggest that the 'have' existential may have begun in late Common Slavic, since it is found in all branches of Slavic. Still, the complete generalization of the 'have'-type in South Slavic can be taken as significant, differentiating it from East and West Slavic.

This is potentially of typological interest since the predicated entity in existentials is typically indefinite when full noun phrases are involved, as in (20), and generally excludes definites and definite pronouns (cf. English **There is the man* / **There is him*).[28] However, there are non-Balkan languages that allow such object pronouns; in Spanish, for instance, *Lo hay*, literally 'it.ACC has', is possible for 'There is one (e.g. book)'. And while this could support the argument that the facts of (22) may reflect independent developments, the areal distribution (Balkan Slavic – Albanian – South Danubian Balkan Romance), reflecting an areal clustering found with other features, combined with the reconstructible chronology points to an areal explanation.

Thus there is a convergence in form in the Balkan languages regarding the use of an impersonal 'have' in existential expressions, and contact is relevant for some aspects of the convergence, although the details are complex and involve considerable time depth and some differentiation.[29]

As for possessives, the verb 'be' is found with a dative of possessor in earlier stages Classical Greek, Latin, Old Church Slavonic, and Sanskrit.[30] This most likely is a Proto-Indo-European construction inherited into each of those languages. In all except Sanskrit, a verb 'have' developed in ancient times and occurs alongside the 'be' possessive.[31] The 'be' construction is not technically an impersonal since the possessum is the subject of 'be', controlling verb agreement, for instance, but it is relevant here since it gives way in all but Indic to the use of 'have', so that the later instantiations of these languages in the Balkans use a verb 'have' for possession; this construction therefore provides an interesting counterpart to the existential 'have'. More importantly, the Romani construction is a true impersonal, with the verb 'be', as was the situation in Sanskrit, but the possessor is in the accusative, unlike Sanskrit, and the possessum is also in the accusative and does not govern verb agreement (note the singular verb in (23)), again unlike Sanskrit; examples of this

28. We are aware that definites can occur in existentials, as in English *If you want a solution to your loneliness, well, there is the man I've been dropping hints to you about for months ….* Nonetheless, such uses have a marked and highly contextually dependent status, and in the unmarked instance, existentials seem to occur more readily predicated of indefinites.

29. Note that both Banfi (1985) and Demiraj (1994) consider the 'have' existential to be a Balkanism, as does Asenova (2002: 248–249). Note also, however that Asenova's reference to "Bulgarian dialects in Southwest Macedonia" in fact refers to Macedonian, not to Bulgarian.

30. Actually, in Sanskrit it is a genitive possessor with 'be', but the genitive subsumes many dative functions in Sanskrit so this is equivalent, in a sense, to a dative possessor; see Danesi (this volume), for relevant discussion.

31. See, most recently, McAnallen (2011) on possession in OCS, with relevant literature, where she notes the marginal occurrence too of a construction with the locative preposition *u* with the genitive for the possessor.

double accusative construction are given in (23), where the reduplicated pronoun in (23a) is emphatic while (23b) has a nominative as a dislocated topic along with the double accusative:

(23) a. *(man) si man duj čhaven*
(me.ACC) is me.ACC two children.ACC
'I have two children'
b. *jekh daj si la duj čhaven*
one mother is her.ACC two children.ACC
'A mother has two children'

Relevant here too is the use of negated 'be' with accusative in the future construction in some Romani dialects of the Balkans, as it represents a calque on the negative future with 'have' in Balkan Slavic as in Romani in Example (24), which is calqued on the Macedonian example:

(24) Rmi *nae man te hav*
NEG.is me.ACC SP eat.1SG
Mac *nema da jadam*
NEG.have.3SG SP eat.1SG
'I will not eat'

3. Non-imperative imperatives

The use of an imperative as a narrative device, i.e. an element that enhances the narration and moves the action in a story along or elaborates on it, occurs for the rendering of first and third person actions more vivid in South and East Slavic. The East Slavic usage is limited to perfective verbs and denotes semelfactive actions, whereas in South Slavic it can be either perfective or imperfective and therefore not so limited. Insofar as all the Balkan languages normally mark the subject on the verb, the use of an imperative for first and third person narratives is quirky. Example (25) is based on the Macedonian in Koneski (1967: 418), translated into Albanian, Romani (Skopje Arli dialect), and Standard Turkish, respectively:

(25) Mac *Se vrakjavme pijani: toj padni, jas stani, jas padni, toj stani.*
REFL return.IMP.1PL drunk he fall.IMP I rise.IMP I fall.IMP he rise.IMP
Alb *U kthyem të pirë: ai bjer, unë çohu, unë bjer, ai çohu.*
REFL return.AOR.1PL PC drunk he fall.IMP I rise.IMP I fall.IMP he rise.IMP

Rmi *Irin sine amen mate: ov per, me ušti,*
return.PRS.1PL be.IMP.3SG we drunk he fall.IMP I rise.IMP
me per, ov ušti.
I fall.IMP he rise.IMP

Trk *Sarhoş dönüyorduk: o düş, ben kalk, ben düş,*
drunk we.were.returning he fall.IMP I rise.IMP I fall.IMP
o kalk.[32]
he rise.IMP
'We were coming back drunk: he falls down, I get up. I fall down, he gets up.'

We should note here that in Romani, the usage does not appear to occur in Greece (Igla 1996), nor in the North Vlax dialects outside the Balkans (Matras 1994), and for Turkic it does not occur in Central Asia. Such usage also does not occur in Greek, and in the case of Aromanian, the distribution is dependent on the contact environment. The dialects spoken in Greece do not have the usage, while the Aromanian of Macedonia does, as illustrated in Example (26):[33]

(26) *shi mini fudz shi năs fudz*
and I run.IMP and he run.IMP
'I run/ran and he runs/ran'

While the construction occurs in both Bulgarian and Romanian, it has a very different status in those languages from that in Macedonian and the Aromanian of Macedonia. In Bulgarian, the construction was already considered dialectal in the late nineteenth century, and the last standard grammar to even make mention of it was Teodorov-Balan (1940: 195). Thus, while it still occurs in some rural dialects, it has been excluded from the standard language. A similar situation prevails in Romania, where educated speakers find the construction to be archaic or dialectal. Although not so specified in the academy grammars (e.g. Ramolo 2008: 383), the examples given are all from nineteenth century literature, and Păna Dindelegan (2013: 54–55, 546–47) makes no mention of the usage. We can add here that in BCSM, this usage is perceived or prescribed as more characteristic of the south, although it also occurs as far north as the Croatian border with Slovenia (Maretić 1963: 625; Stevanović 1986: 708–709). Given the areal and genetic distribution of this construction, it is fair to conclude that in Balkan Romance, Romani, and possibly Albanian (again, in the absence of ancient sources we cannot be certain) it represents the influence of Balkan Slavic. As mentioned above, the construction does not occur in Central Asian Turkic, thus raising the possibility that this is a Rumelianism of Slavic origin that spread into Anatolia.

32. Standard Turkish would favor the gerundive constructions *düşe kalka* 'falling down, getting up', but narrative imperatives are acceptable, and they are quite normal in Macedonia.

33. We do not have data for the Aromanian dialects of Albania.

4. Conclusion

In the foregoing, we have explored various constructions in the languages of the Balkans in which there is no overt person marking that reflects the subject of a verb.[34] These nonpersonal constructions cover several types traditionally subsumed under the rubric of "impersonals" as well as one that involves an extended use of ostensibly second person imperative verbs to serve for any person. These constructions, thus, show different aspects of the nonrealization of canonical subjects on verbs, and they show as well a mix of convergences and divergences as to both form and function. As is well known, the Balkans are linguistically notable for multilingualism, for language contact, and for contact-induced structural and lexical convergence, yet for all of these constructions, we document the need to be sensitive not only to the possibility of language contact being responsible for the nonpersonal convergences but also to the possibility of language contact not being a factor.

Several observations sum up our findings about these nonpersonal constructions in the Balkans. Focusing first on the impersonals, per se, it is clear above all else that the languages that pattern together most closely are those on the territory of the Macedonian and Albanian Ottoman vilayets. The vilayet was the largest administrative unit of the late Ottoman Empire. The Albanian and Macedonian vilayets were Işkodra (modern Shkodër in Albania), Yanya (modern Ioannina in Greece), Manastir (modern Bitola in the Republic of Macedonia), Selânik (modern Thessaloniki in Greece), and Üsküp (modern Skopje in the Republic of Macedonia; this vilayet was also called Kosova). In modern-day terms, the territory included all of modern Albania, Kosovo, and the Republic of Macedonia as well as adjacent parts of Montenegro, Serbia, Bulgaria, and Greece. This patterning suggests a relevance for contact as an explanation for convergences there.

Also, there is a general tendency to move from the marking of person to the non-marking of person in constructions with modal senses, which may reflect universal tendencies with some contact involved as well. Finally, for those languages with ancient attestations and with both Balkan and non-Balkan offspring, we find that both Romance and Slavic appear as significant sources of shared convergences. To be sure, we cannot know given our current state of knowledge, what the situation was in the ancestor of Albanian.

Nonetheless, the complex picture given by the consideration of Balkan impersonals *sensu largo*, as considered here, demonstrates precisely the fact that the sprachbund is not the result of a single language dominating the others but rather a

34. We use this formulation here because imperatives could be argued to show person marking for second person; what is relevant here, though, is that there is no overt marking for the subject referred to by the verb, e.g. 'I' in the case of *stani* in the Macedonian version of (25).

complex interplay of factors in which different languages contribute to the creation of a distinctive shared linguistic space.

A consideration of the non- (or multi-)personal imperatives augments these findings in interesting ways. For one thing, the pan-Balkan character of this construction shows, *pace* Asenova (2002: 193), that the true narrative imperative occurs not only in Slavic, but also in Albanian, and Balkan Romance as well as in Romani and Turkish. Moreover, as with impersonals, geography is crucial, here in two ways. First, the use of the Balkan narrative imperative is more frequent as one moves south for both Balkan Slavic and Balkan Romance, but the construction is absent from Greek and dialects of other languages in contact with Greek (cf. Friedman 2008 on similar phenomena with respect to object reduplication). Moreover, with the exception of Slavic, it does not occur in related languages/dialects outside the Balkans, though data for Arvanitika (Albanian in Greece) or Arbëresh (Albanian in Italy) might help in dating the phenomenon, as would data from Greek outside of the Balkans, e.g. Pontic, Cappadocian, and Cypriot Greek. Second, it is most vital in languages/regions in contact with Macedonian (Albanian, Balkan Romani, northern Aromanian). Moreover, the East Slavic narrative imperative is quite distinct from the South Slavic, but it is not unreasonable to speculate that both have their origins in a colloquial Common Slavic usage that developed differently in the two regions, and in its Balkan manifestation developed most robustly precisely in the regions of most complex contact. The use of a 2sg imperative – which is cross-linguistically often the bare root – to render a narrative more vivid makes sense typologically. However, given the distribution in the Balkans and absence beyond the Balkans, the narrative imperative, which is better preserved in Macedonian than in Bulgarian, appears to be a Balkanism, i.e., an areal rather than a typological phenomenon.

Thus there are areas of contact-induced change in nonpersonals alongside independent inheritances that give parallelism in some of the languages. The overall situation with the wide range of Balkan nonpersonal constructions shows that both inheritance and contact must be taken into account in understanding convergent structures in the Balkans,[35] but at the same time, the areas of convergence to contribute to our understanding of the Balkans as a sprachbund.

35. See Barðdal et al. (2016) for similar conclusions regarding dative-subject constructions in Germanic.

References

Asenova, Petja. 2002. *Balkansko ezikoznanie*. Veliko Tărnovo: Faber.

Banfi, Emanuele. 1985. *Linguistica balcanica*. Bologna: Zanichelli.

Barðdal, Jóhanna. 2011. Lexical vs structural case: A false dichotomy. *Morphology* 21: 619–654. https://doi.org/10.1007/s11525-010-9174-1

Barðdal, Jóhanna. 2015. Icelandic valency classes: Oblique subjects, oblique anticausatives and the actional passive. In *Valency Classes: A Comparative Handbook*, Bernard Comrie & Andrej Malchukov (eds). Berlin: Mouton de Gruyter.

Barðdal, Jóhanna, Smitherman, Thomas, Bjarnadóttir, Valgerður, Danesi, Serena, Jenset, Gard B. & McGillivray, Barbara. 2012. The dative subject construction in Old Norse-Icelandic, Latin, Ancient Greek, Old Russian and Old Lithuanian. *Studies in Language* 36(3): 511–547. https://doi.org/10.1075/sl.36.3.03bar

Barðdal, Jóhanna, Arnett, Carlee, Carey, Stephen Mark, Eythórsson, Thórhallur, Jenset, Gard B., Kroonen, Guus & Oberlin, Adam. 2016. Dative subjects in Germanic: A computational analysis of lexical semantic verb classes across time and space. *STUF: Language Typology and Universals* 69(1): 49–84.

Boretzky, Norbert & Igla, Birgit. 2004. *Komentierter Dialektatlas des Romani*, 2 Vols. Wiesbaden: Harrassowitz.

Bubenik, Vit. 1998. *Historical Syntax of Late Middle Indo-Aryan (Apabhraṃśa)* [Current Issues in Linguistic Theory 165]. Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.165

Capidan, Theodr. 1925. *Meglenoromânii*, I. Bucharest: Academia Română.

Caragiu–Marioțeanu, Matilda. 1958. *Habēre* impersonal în Aromînă și meglenoromînă. In *Omagiu lui Iorgu Iordan*, 139–141. Bucharest: Academia Republicii Popolare Romîne.

Danesi, Serena. 2018. The expression of predicate possession in Vedic. In *Non-canonically Case-Marked Subjects: The Reykjavík-Eyjafjallajökull Papers*, Jóhanna Barðdal, Na'ama Pat-El & Stephen Mark Carey (eds). Amsterdam: John Benjamins. (this volume)

Demiraj, Shaban. 1994. *Gjuhësi ballkanike*. Skopje: Logos-A.

Eriksen, Pål, Kittilä, Seppo & Kolehmainenm, Leena. 2010. The linguistics of weather: Cross-linguistic patterns of meteorological expressions. *Studies in Language* 34(3): 565–601. https://doi.org/10.1075/sl.34.3.03eri

Friedman, Victor A. 2012. Perhaps mirativity is phlogiston, but admirativity is perfect: On Balkan evidential strategies. *Linguistic Typology* 16(3): 502–527. https://doi.org/10.1515/lity-2012-0019

Gawelko, Marek. 2005. La perspective indo-européenne de l'impersonel Polonais. *Wiener Slawistisches Jahrbuch* 51.21–30.

Georgiev, Vladimir I. (chief ed.) 1979. *Bălgarski etimologičen rečnik* II. Sofia: BAN.

Gołąb, Zbigniew. 1976. On the mechanism of Slavic–Rumanian linguistic interference in the Balkans. In *Bulgaria, Past and Present: Studies in History, Literature, Economics, Music, Sociology, Folklore & Linguistics*, Thomas Butler (ed.), 296–309. Columbus OH: AAASS.

Guentchéva, Zlatka. 2010. l'Impersonnel dans les languges balkaniques: Convergences et divergences. In *Glagolata sistema na balkanskite ezici – nasledstvo i neologija / The Verbal System of the Balkan Languages – Heritage and Neology*, Petya Assenova, Anastasiya Petrova & Tsenka Ivanova (eds), 24–50. Sofia: Faber (for the Commission on Balkan Linguistics, International Committee of Slavists).

Hale, William Gardner & Darling Buck, Carl. 1903[1966]. *A Latin Grammar*. Tuscaloosa AL: University of Alabama.

Igla, Birgit. 1996. *Das Romani von Ajia Varvara*. Wiesbaden: Harrassowitz.
Ikonomov, Nikolai. 1968. *Balkanska narodna mûdrost*. Sofia: BAN.
Joseph, Brian D. 1983[2009]. *The Synchrony and Diachrony of the Balkan Infinitive. A Study in Areal, General, and Historical Linguistics*. Cambridge: CUP.
Kibort, Anna. 2001. The Polish passive and impersonal in lexical mapping theory. In *Proceedings of the LFG01 Conference, University of Hong Kong, Hong Kong*, Miriam Butt & Tracy Holloway King (eds). Stanford CA: CSLI. <http://csli-publications.stanford.edu/>
Koneski, Blaže. 1967. *Gramatika na* makedonskiot literaturen jazik. Skopje: Kultura.
Lunt, Horace G. 2001. *Old Church Slavonic Grammar*, 7th edn. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110876888
Malchukov, Andrej L. & Ogawa, Akio. 2011. Towards a typology of impersonal constructions: A semantic map approach. In *Impersonal Constructions: A Cross-Linguistic Perspective* [Studies in Language Compaion Series 124], Andrej L. Malchukov & Anna Siewierska (eds), 19–56. Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.124.02mal
Maretić, Tomo. 1963. Gramatika hrvatskoga ili srpskoga jezika. Zagreb: Matica Hrvatska.
Masica, Colin. 1976. *South Asia as a Linguistic Area*. Chicago IL: University of Chicago Press.
Masica, Colin. 2001. The definition and significance of linguistic areas. In *The Yearbook of South Asian Languages and Linguistics*, Peri Bhaskararao & Karumuri Venkata Subbarao (eds), 205–268. New Delhi: Sage.
McAnallen, Julia. 2011. Predicative possession in Old Church Slavic Bible translations. In *Indo-European Syntax and Pragmatics: Contrastive Approaches*, Eirik Welo (ed.), *Oslo Studies in Language* 3(3):155–172. <http://www.journals.uio.no/osla>
Matras, Yaron. 1994. *Untersuchungen zu Grammatik und Diskurs des Romanes: Dialekt der Kelderaša/Lovara*. Wiesbaden: Harrassowitz.
Nicolova, Ruselina. 2008. *Bălgarska gramatika: Morfologija*. Sofia: Universitet "Sv. Kliment Ohridski."
Pană Dindelegan, Gabriela. 2013. *The Grammar of Romanian*. Oxford: OUP.
Papahagi, Pericle N. 1908. *Parallele ausdrücke und redensarten im rumänischen, albanesischen, neugriechischen und bulgarischen*. Leipzig: J.A. Barth.
Papanastassiou, George & Papadamou, Eleni. 2012. The position of the northern dialects of the Prefecture of Kastoria. Paper presented at Fifth International Conference on Modern Greek Dialects and Linguistic Theory (MGDLT 5). University of Ghent, 21 September.
Perlmutter, David M. & Moore, John. 2002. Language-internal explanation: The distribution of Russian impersonals. *Language* 78(4): 619–650. https://doi.org/10.1353/lan.2003.0049
Romalo, Valeria Guțu (chief ed.). 2008. *Gramatica limbii române: Cuvîntul*. Bucharest: Academia Română.
Siewierska, Anna. 2008. Introduction: Impersonalization from a subject-centred vs. agent-centred perspective. Special guest-edited issue on *Impersonal Constructions in Grammatical Theory of Transactions of the Philological Society* 106(2): 1–23.
Siewierska, Anna. 2011. Overlap and complementarity in reference impersonals: Man-constructions vs. third person plural-impersonals in the languages of Europe. In *Impersonal Constructions: A Cross-linguistic Perspective* [Studies in Language Companion Series 124], Andrej L. Malchukov & Anna Siewierska (eds), 57–90. Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.124.03sie
Skorniakova, Oxana. 2008. The existence of expletive pronominal subjects in Russian impersonal constructions. In *Issues in Slavic Syntax and Semantics*, Anastasia Smirnova & Matthew Curtis (eds), 35–53. Newcastel upon Tyne: Cambridge Scholars.

Stevanović, Mihailo. 1986. Savremeni srpskohrvatski jezik, I. Beograd: Naučna Kniga.
Teodorov–Balan, Alexandar S. 1940. *Nova bălgarska gramatika*. Sofia: Čipev.
Thomaj, Jani et al. 1999. *Fjalor Frazeologjik Ballkanik*. Tirana: Dituria.
Twardzisz, Piotr. 2012. To be and not to have in Polish locationals. *Rice Working Papers in Linguistics* 3:1–23.
Večerka, Radoslav. 1996. *Altkirchenslavische (Altbulgarische) Syntax* [Monumenta Linguae Slavicae Dialecti Veteris 27], Vol. 3. Frieburg: U.W. Weiher.

Irregular morphology in regular syntactic patterns: A case of constructional re-alignment

Mirjam Fried
Charles University, Prague

Table 1. Morphological patterns of OCz PAs and pseudo-PAs

gloss	present stem		act. ppl	CNG	pseudo-PAs	
'carry-ing'	*nes*	**– *ú***	– *c*	– …		
'pass-ing'	*min*	**– *ú***	– *c*	– …		*min-**ujúc-***
'pour-ing'	*lej*	**– *ú***	– *c*	– …		
'do-ing'	*čin*	**– *ie***	– *c*	– …	*čiň-**úc-***	*čiň-**ujúc-***
'desire-ing'	*žád-aj*	**– *ú***	– *c*	– …	*žád-**úc-***	
'buy-ing'	*kup-**uj***	**– *ú***	– *c*	– …		

Table 2. The morphology of OCz 3pl present tense

gloss	present stem (= 3pl)	
'they carry'	*nes*	**– *ú***
'they pass'	*min*	**– *ú***
'they pour'	*lej*	**– *ú***
'they do'	*čin*	**– *ie***
'they desire'	*žád-aj*	**– *ú***
'they buy'	*kup-**uj***	**– *ú***

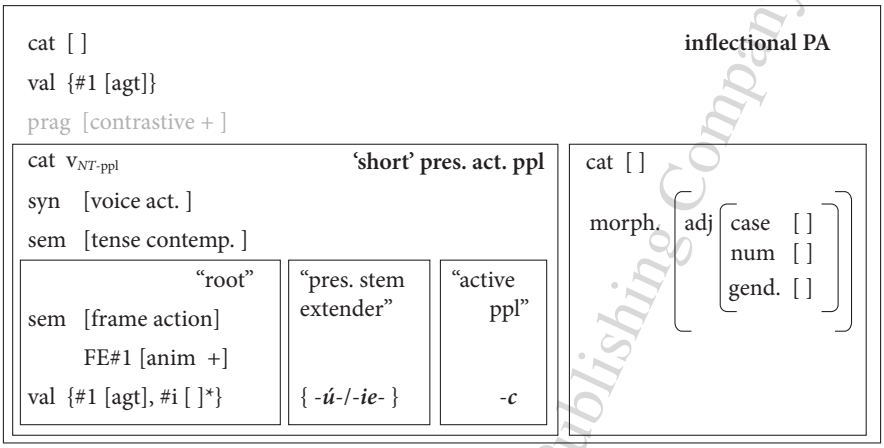

Diagram 1. Regular PA construction, morphological classes I-V

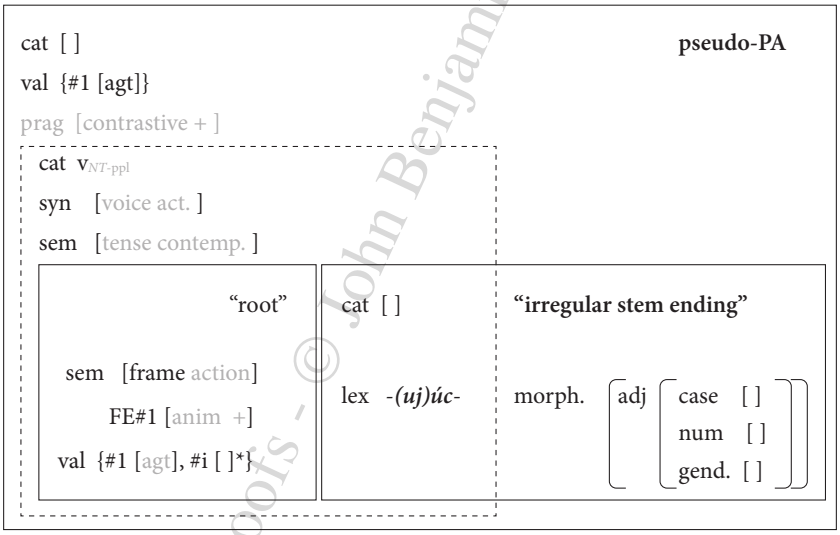

Diagram 2. Pseudo-PA construction

Table 3. Overall functional distribution of PAs and pseudo-PAs

	(in %) Pred.	Attrib.	Ambig.
PAs	25	60	15
pseudo-PAs	10	79	11

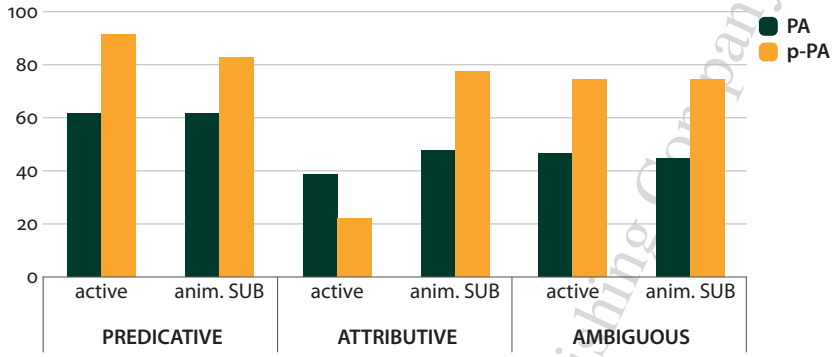

Figure 1. PAs and pseudo-PAs in the linearization pattern A

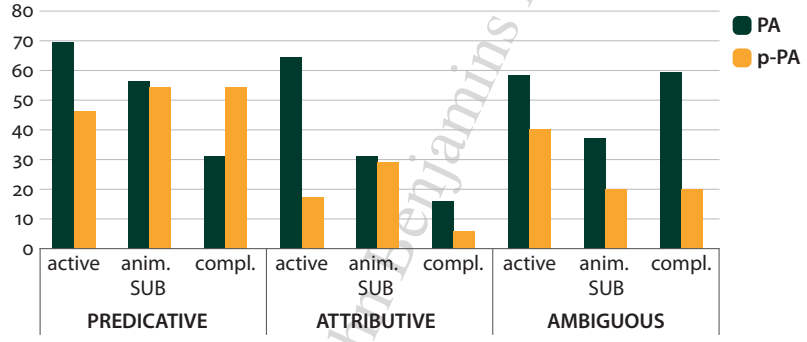

Figure 2. PAs and pseudo-PAs in the linearization pattern B

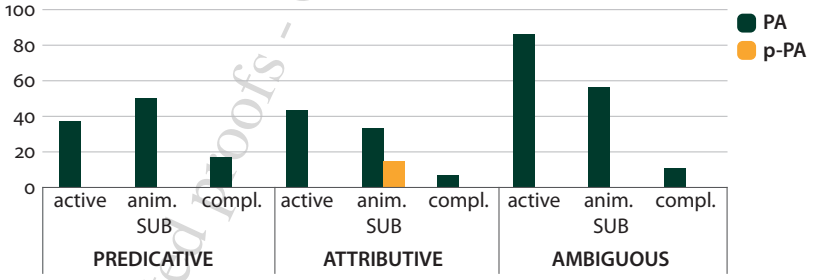

Figure 3. PAs and pseudo-PAs in the linearization pattern C

Table 4. Roots with shifted meaning, patterns B and C (relative frequencies)

	pseudo-PAs		regular PAs	
meaning:	**B**	**C**	**B**	**C**
non-PA	76%	73%	46%	80%
metaphoric	24%	53%	9%	23%
non-lexicalized metaphor	6%	20%	–	–

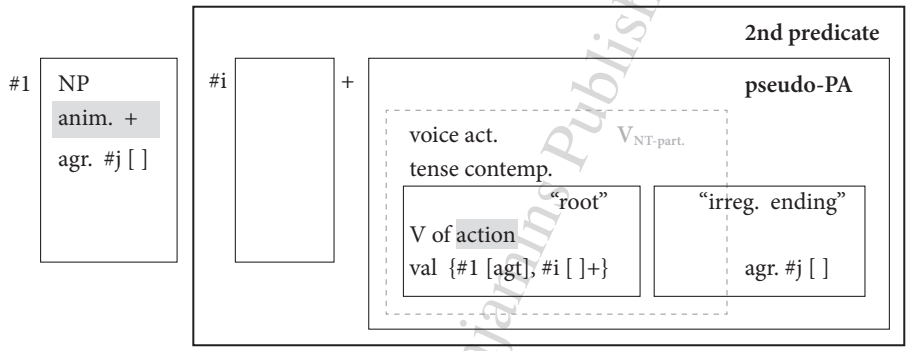

Diagram 3. Pseudo-PAs in linearization A – a verbal construction

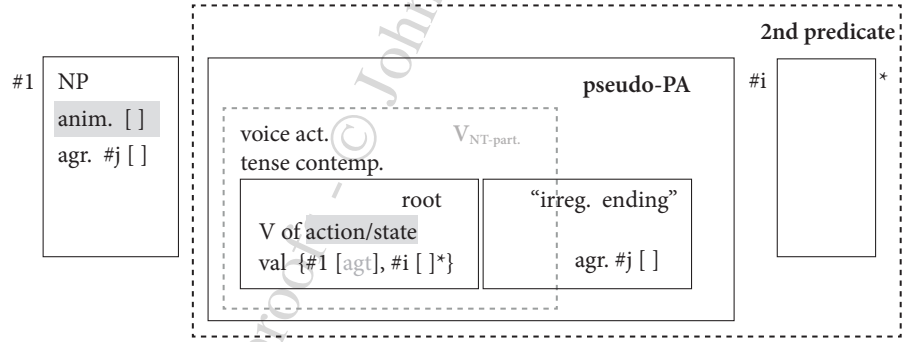

Diagram 4. Pseudo-PAs in linearization B, with non-subject complements

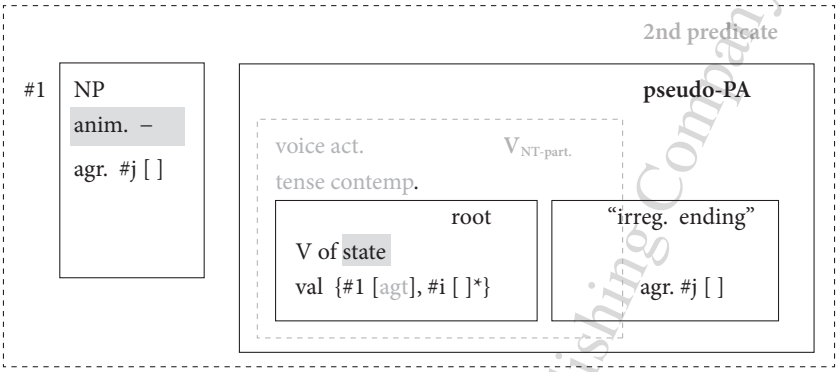

Diagram 5. Pseudo-PAs in linearization B, without non-subject complements

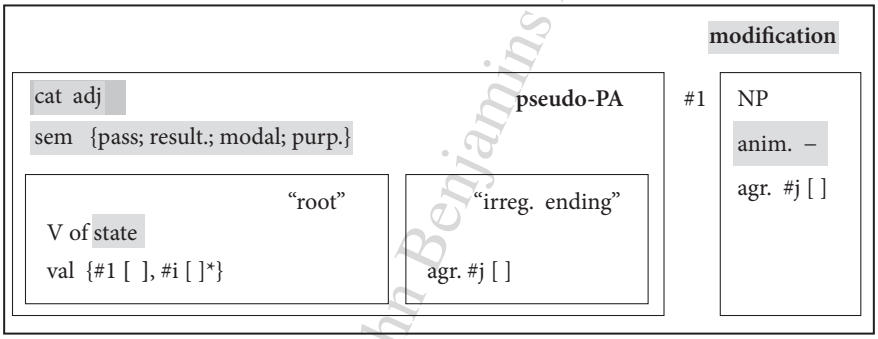

Diagram 6. Pseudo-PAs as adjectives in linearization C

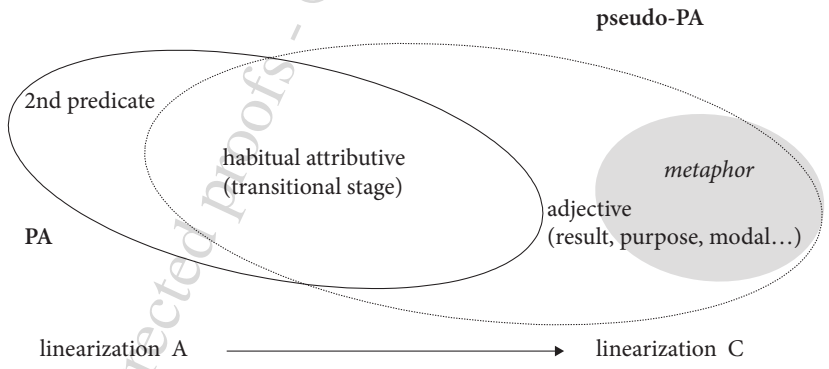

Diagram 7. Constructions with PAs and pseudo-PAs on the predicative/attributive continuum (a simplified constructional map)

Author queries

Please provide a citation for the reference id "CIT0101 (Danesi, Serena. 2018), CIT0109 (Hale, William Gardner & Darling Buck, Carl. 1903[1966])" since citation is missing in the article.
Do the tables and figs in c2-app1 belong to the article?