Slavic morphosyntax is primarily determined by its geographic location and contact configuration

Ilja A. Seržant
Christian-Albrecht-University of Kiel

Abstract: This paper seeks to identify causal factors constraining the diachronic dynamics of particular morphosyntactic categories of Slavic. It is suggested that the modern inventory of Slavic languages is not a result of random accumulation of historically accidental changes and non-changes. Instead, it is argued that macro-areal pressures constrained by the geographic location and the particular language-contact configuration determine the selection of inherited properties for either retention or loss and, subsequently, innovation. I primarily provide evidence from two categories: verbal person-number indexes (subject agreement markers) and partitivity markers and I also briefly discuss some other fusional categories.

Keywords: macro-areal pressures, Slavic, person-number indexes, case, partitives, morphosyntactic change

1. Introduction

There is a trend in linguistic typology to shift the attention from finding universal properties of language to the understanding of the linguistic diversity around the globe (compare the programmatic "what's where why?" in Bickel 2007; see also Dixon 2003; Kusters 2003; Trudgill 2009, 2011). In the same vein, this paper pursues the question: Why are Slavic languages the way they are? While we are not in a position to provide an exhaustive account of the causal factors that have conditioned the development of modern Slavic languages, the modest goal of this paper is to provide evidence for the claim that macro-areal pressures are an important factor that shaped the development of Slavic after its split from Proto-Indo-European.

Despite considerable intragenealogical variation, Slavic languages share a large amount of linguistic material that comprises all domains of grammar and lexicon. Thus, Slavic languages share a number of morphosyntactic properties based on largely etymologically related morphological devices such as derivational aspect, past forms based on the suffix *-l* (originally participles), verbal person-number inflection, singular-plural distinctions, passive morphology based on the *-no-/-to-*resultatives, etc. (e.g. Seržant 2012; Hansen 2014; Wiemer 2019). This is, of course, not unexpected given the shallow time depth of this subfamily (ca. 1300 years, cf. Holman et al. 2007) and the intensive mutual contacts due to small geographical distances between Slavic languages.

However, the question of why these languages share precisely this specific set of properties is not trivial at all. For example, one may wonder why the old, Indo-European middle morphology was entirely abandoned in Slavic while it is still retained in Modern Greek. Or, why the old Indo-European perfect still found with *vědě* [know/see.1PERF.SG] 'I know' in the earliest Old Church Slavonic documents was entirely lost in modern Slavic while, for example, Germanic languages generalized it as the only past-tense form (cf. Gołąb 1992: 65). We do not understand the selectional mechanism which is responsible for the retention vs. loss of inherited properties. Moreover, this common set of properties does not only contain common inheritance (Section 2) but also has a number of common innovations (Section 3). Again, while languages innovate parts of their inventories continuously, one may wonder why Slavic introduced its particular set of innovations and not some other.

1

This paper explores the selectional mechanism for retention and innovation of particular properties. First, it claims that two factors, namely, the macro-areal pressures and the particular contact configuration – in addition to universal pressures – must be viewed as important components of this mechanism. Secondly, I provide evidence for the claim that retentions of inherited properties need not be historically accidental, inert processes (the null hypothesis) but may be constrained by areal pressures. I show that (macro)areal pressures support the selection of particular inherited properties for retention (Section 2) and the selection of other properties for loss (Section 3). Generally, a clear-cut separation between areal innovations and areal retentions of inherited structures and, thus, the traditional complementary division of labour between historical and areal linguistics (since Jakobson 1930, 1931; cf. Nichols 1992: 163) does not seem to be meaningful.

More specifically, I focus on two phenomena: the retention of the Proto-Indo-European verbal person-number indexes (Section 2), on the one hand, and the innovation of the partitivity markers (Section 3), on the other. Thus, I argue that the quite faithful preservation of person-number indexes from Proto-Indo-European into modern Slavic languages is due to the very conservative dynamics of the neighbouring languages in this domain. Slavic languages occupy a part of the transitional zone that is found between highly conservative languages in the East of Eurasia and the innovative languages in the Northwest of Europe. By contrast, the old, genitive partitivity marker has been innovated in Slavic due to a strong macroareal pressure.

Areal pressures themselves accumulate from a large number of different kinds of contact events involving different transfer mechanisms. In view of the macro-perspective of the paper, I do not discuss these mechanisms in detail here.

Thirdly, I argue that the effect of macro-areal pressures is constrained by the particular *contact configuration*. Thus, languages normally have mutual contacts with the languages that occupy geographically adjacent territory. However, different local factors such as the time depth of the contacts, the particular historical events, the physical shape of the territory (cf. Nichols 1992), and other factors – cumulatively referred to as *contact configuration* here – may considerably constrain the contact effects and thus the effect produced by the macro-areal pressures (cf. Thomason 2001). For example, the abruptness of contact – understood as a high proportion of adult L2 vs. L1 speakers – has been argued to have strong reduction effects (Trudgill 2009, 2011; Sinnemäkki 2020).

Causal explanations in terms of a macro-areal impact for certain – primarily phonetic and phonological patterns of Slavic (such as palatalization) – have already been ventured by a number of scholars in most cases by making reference to the extra-genealogical language contact with Turkic and/or Uralic languages (Kattein 1983; Galton 1997; Stankiewicz 1986; see also Gvozdanović 2015 on Celtic and Slavic). While joining this tradition, I focus primarily on morphosyntactic properties of Slavic.

The paper is structured as follows: Section 2 explores the dynamics of verbal person-number indexes in Eurasia and compares the areal trends with the evidence from Slavic. Section 3, by contrast, explores the common Slavic innovation of partitivity markers, replacing the old, inherited possessive marking strategy with the new spatial, separative strategy. Section 4 briefly provides evidence on other morphosyntactic phenomena of Slavic from such categories as fusional case and aspectual inflection. Finally, Section 5 summarizes the conclusions and discusses the results thus achieved and the method.

2. Verbal person-number indexes

In what follows, I avoid the more traditional terms like *bound pronouns* or *agreement markers* and follow Lazard (1998) and Haspelmath (2013) and refer to these as (*bound person-number*) *indexes* (the term was introduced in Boelaars 1950 or even earlier).

In general, Slavic verbal person-number indexes are surprisingly conservative both in terms of the number of segments retained and in terms of the small amount of sound change, as can be observed in Table 1:

	1SG	2SG	3SG	1PL[1]	2PL	3PL
PIE	*-oh$_2$*	*-e-s-i*	*-e-t-i*	*-o-m-os*	*-e-th$_2$-e*	*-o-nt-i*
Russian	*-u*	*-eš'*	*-et(')*	*-'om*	*-'ete*	*-ut(')*
Serbian	*-ēm*	*-ēš*	*-ē*	*-emo*	*-ēte*	*-ū*

Table 1: Modern Russian and Serbian vs. Proto-Indo-European (PIE) subject indexes (the *thematic conjugation*, cf. the discussion in Meier-Brügger 2010: 173-184)

Despite some minor differences in the phonetic form of the first singular and third plural (where a nasal has been added and the pre-nasal vowel has been raised from Proto-Slavic **on* to Common Slavic *ǫ* and then to East Slavic *u*), other forms are strikingly reminiscent of those of Proto-Indo-European, both as to their phonetic realization and as to their lengths (cf., *inter alia*, Olander 2015). Given that Proto-Indo-European was spoken around 5,000-3,000 BC (Anthony 1995: 558; Nichols & Warnow 2008: 781; Meier-Brügger 2010: 194; Holman et al. 2011), the approximate age of the Russian indexes is at least 5,000 years.

By contrast, in some other modern Indo-European languages such as Scandinavian languages, French, Hindi or Marathi, the person-number indexes are almost entirely lost. In other words, despite the same prerequisites in the common proto-language (Proto-Indo-European), these languages and language families underwent other type of developments. In order to explore the variation of different degrees of the decay across Eurasia, a database comprising person-number indexes of six families of Eurasia has been created.

2.1. The database
This study relies on the database of the six person-number indexes of the intransitive (masculine) subject – first singular (1SG), second singular (2SG), 3SG, 1PL, 2PL, 3PL – from 150 modern and 6 proto-languages from 6 families from Eurasia: Indo-European, Tibeto-Burman (Kiranti & Rgyalrongic only), Turkic, Uralic and Semitic. Each language is represented by only one paradigm used in the present tense except for Semitic for which the so-called imperfect paradigm was chosen. Only Dravidian and Semitic distinguish between genders here. For these languages, only the masculine/common forms were taken into account. Irregular verbs and rare forms were excluded.

The paradigms within one family are always historically interrelated so that one can meaningfully talk about the dynamics when comparing the paradigms across the modern languages of one family and with the proto-language. The reconstructed paradigms of the proto-languages have been taken from the authoritative literature: Meier-Brügger (2010: 311) for the thematic conjugation of Proto-Indo-European; Janhunen (1982: 35) for Proto-Uralic; Erdal (2004: 233) and Róna-Tas (1998: 74) for Proto-Turkic; Andronov (2009: 224-231) for Proto-Dravidian; Hasselbach (2004: 32), Huehnegard (2000) and Lipiński (2001: 378) for Proto-Semitic; Jacques (2016), DeLancey (2010: 15, 2011: 2, 2014) and LaPolla (2003: 30) for Proto-Gyalrongic and Proto-Kiranti (Tibeto-Burman).

In what follows, Tables 2-4 illustrate the database with some examples and provide the decay factor (discussed in detail below Section 2.2). The entire database has been published at zenodo.org (Seržant 2020b).

	1SG	2SG	3SG	1PL[2]	2PL	3PL	DECAY

---

[1] The alternative reconstructed form **o-me-s(i)* was not taken into account here.
[2] The alternative reconstructed form **o-me-s(i)* was not taken into account here.

PIE	oh₂	e-s-i	e-t-i	o-m-es	e-th₂-e	o-nt-i	
Persian	*am*	*i*	*ad*	*im*	*id*	*and*	0.13
Greek	*o*	*is*	*i*	*ume*	*ete*	*un*	0.14
Macedonian	*em*	*eš*	*e*	*eme*	*ete*	*at*	0.11
Upper Sorbian	*u*	*eš*	*e*	*emy*	*eće*	*u/ja*	0.15
Slovenian	*em*	*i*	-	*mo*	*te*	*so*	0.24
German	*e*	*st*	*t*	*en*	*en*	*en*	0.3
Dutch	*0*	*t*	*t*	*en*	*en*	*en*	0.41
English	-	-	*s*	-	-	-	0.91
French	-	-	-	*ǫ*	*e*	-	0.77

Table 2: Indo-European indexes (the "thematic conjugation" cf. Meier-Brügger 2010: 173-184)

	1SG	2SG	3SG	1PL[3]	2PL	3PL	DECAY
Proto-Turkic	*men*	*sen*	-	*biz*	*siz*	*-lar/ler*[4]	
Azeri	*am*	*san*	-	*yg*	*synyz*	*lar*	0.07
Bashkir	*myn*	*hyŋ*	-	*byz*	*hygyz*	*lar*	0.07
Dolgan	*bïn*	*gïn*	-	*bït*	*gït*	*lar*	0.07

Table 3: Turkic indexes (cf. Erdal 2004: 233; Róna-Tas 1998: 74)

	1SG	2SG	3SG	1PL[5]	2PL	3PL	DECAY
Proto-Uralic	*m*	*n / t*	-	*måt/mät/(mäk)*	*tåt/tät/(täk)*	*t*	
Southern Saami	*am*	*ah*	*a*	*ebe*	*ede*	*ieh*	0
Erzya	*an*	*at*	*y*	*otano*	*otado*	*yt'*	0
Moksha	*n*	*t*	*j*	*tama*	*tada*	*jxt'*	0
Mari	*am*	*at*	*eš*	*yna*	*yda*	*yt*	0
Selkup	*am*	*al*	*ətə*	*ej*	*əlī*	*ətī*	0

Table 4: The Proto-Uralic subject indexes and their lengths (Janhunen 1982: 35)[6]

In order to compare the dynamics of person-number indexes across the languages of the database, I have computed the decay factor for each language.

2.2. Measuring the decay factor
In this subsection, I present the method for measuring the degree of decay of an index paradigm (henceforth the *decay factor*). The decay factor has the maximal value of 1 if the erstwhile paradigm has totally disappeared in the course of development. However, in most languages of the database, different kinds of transitional stages are found. In order to measure this variation in the morphological functionality of verbal person-number indexes across Eurasia, I put forward the following metrics based on three indicators. An inflectional paradigm may be considered to have lost its functions if all person-number slots (i) have undergone total length reduction when compared to the proto-language, (ii) are mutually syncretic and (iii) are zeroes

---

[3] The alternative reconstructed form **o-me-s(i)* was not taken into account here.

[4] Note that most of the modern Turkic languages employ the nominal plural marker *-lar/-ler* here. This plural marker stems from the nominal domain (possibly originally only with an animate subject) in Old Turkic (the data presented in Erdal 2004: 232).

[5] The alternative reconstructed form **o-me-s(i)* was not taken into account here.

[6] There is some uncertainty about the quality of the second person singular since many Uralic languages point to the original **-t* while the Eastern periphery (Komi, Ob-Ugric, Samoyed) point to a nasal **-n* (Janhunen 1982: 34; Kulonen 2001; Honti 2010: 21). Furthermore, sometimes it is assumed that the final consonant of the first and second plural is the plural marker *-k* and not the ancient plural marker *-t* agglutinatively added to the stems of the first and second person pronouns (cf. Laanest 1982[1975]: 229-30 discussing this proposal; Honti 2010: 21). Note that these controversies do not affect the study since the quantity remains the same on both alternatives.

by form (phonetically unmarked). In this extreme case, these three factors are tautological, but not otherwise. For example, Scandinavian languages have almost lost the original, present-tense paradigm, all present-tense slots being marked by *-r*. Thus, while all slots are syncretic (ii) and all slots have been reduced when compared to Proto-Indo-European, these are not zeroes, cf. Norwegian *jeg komme-r* '1SG come-*r*' 'I am coming', *du komme-r* '2SG come-*r*' 'you are coming', *han komme-r* '3SG.M come-*r*' 'he is coming', *vi komme-r* '1PL come-*r*' 'we are coming'. By contrast, Rumanian has first-person singular zero (iii) but is quite conservative with respect to the other two indicators in that it has little reduction (i) and retains all the distinctions with no syncretism (ii).

A paradigm that did not undergo any reduction, syncretism and creation of zeroes counts as maximally preserving, regardless whether the forms have remained the same or have been replaced via analogy, reanalysis or phonetic changes. Such a paradigm has the minimal value of the decay factor, i.e. 0. I also adopt the convention that any kind of lengthening of the forms – in contrast to reductions – cannot decrease the functionality of the slot and of the entire set. For example, the analogical lengthening of the Proto-Slavic 1PL index *-mŭ* to Old Church Slavic (occasionally) or Polish (obligatorily) 1PL *-my* from *-mū* does not have any effect on the functional load of this index.

In what follows, I explain how these three indicators – (i) *the reduction of length,* (ii) *the degree of syncretism* and (iii) *the number of zeroes* – are measured. The higher the values of these indicators the more the paradigm is functionally impaired and the closer it is on its way to the full decay.

(i) *The reduction of the length* of the indexes is measured in phonetic segments as the difference between the length of the synchronic form and the length of the respective proto-form for each modern language. For example, a change such as the *replacement* of the Proto-Indo-European 2SG *-e-si* by the Proto-Slavic *-e-šī* does not anyhow affect the functionality of the morphological system. By contrast, reduction changes such as the entire loss of the dental consonant in the third-person indexes of many Slavic languages (PIE *-e-ti* > Polish *-ⁱe-Ø*) did affect the morphological system of the person-number inflection because it potentially leads to an unmarked third-singular index, i.e. zero. Thus, while a change containing a phonetic replacement does not affect the morphological structure and the functionality of the inflectional paradigm, a reduction change or even an entire loss may potentially represent traces of an ongoing decay of the indexing system. Accordingly, the loss of the final syllable in Polish (PIE *-e-ti* > Polish *-ⁱe-Ø*) is counted as the reduction from three segments to only one, i.e., as a reduction of 2 (segments). In order to level out the differences between the reduction of long proto-indexes and short proto-indexes, the reduction is divided by the total of the respective proto-index, i.e. 2/3 = 0.67 for the third-person index of Polish. Subsequently, the means of these resulting values is calculated as the reduction value of the language.

Note that since the focus is on the functional capacity of the paradigm, it plays no role whether there is a reduction by sound law (cf. the second singular in Polish *-sz* from Proto-Indo-European *-esi* via Proto-Slavic *-ešī*) or due to some other type of reduction (cf. Polish 3SG that dropped the final *-ć* < *-ti* not by a sound law, contrast it with nouns such as *sieć* < *s'eti* 'net' retaining the final *-ć*).

(ii) *The degree of syncretism* measured as the maximum number (six) minus the number of distinct morphological strings employed by the paradigm.[7] For example, German has two syncretic pairs: (a) 1PL & 3PL *geh-en* 'go-1PL/3PL' 'we/they go' and (b) 3SG & 2PL *geh-t* 'go-3SG/2PL' 'he/you go(-es)'. The total number of strings used in this paradigm is four. Six minus four yields the value of two. Thus, the degree of syncretism of the German paradigm is 2. The maximum value here is five found in Scandinavian. Scandinavian languages have generalized the affix *-r* indistinguishably for all person-number slots in the present tense.

---

[7] I thank an anonymous reviewer for helping me to better formulate this measure.

(iii) *The number of zeroes* in the paradigm. Only unequivocal zeroes are counted. For example, in English, all person-number slots except for the third singular are zeroes. By contrast, I did not count the Slavic third-singular forms with the dental loss (e.g. Polish *czyt-a* '(s)he reads') as zeroes because such an analysis is very much dependent on how one treats the stem-final vowel (-*a*- in this case). This is not a substantial problem for the approach since the higher degree of reduction in Polish as compared to, for example, Russian (*čitaj-et* '(s)he reads') is captured by the first indicator (i) above anyway.

I have filled in the values of these indicators for each index in every language of the database. Subsequently, I normalized the values, arranging all values between 0 and 1 along (1):

(1)   *Minimum-maximum feature scaling[8]*
        $X_{norm} = X / X_{max}$

Finally, I have computed the decay factor for each language by averaging the normalized values of all three indicators. The resulting decay factors are represented in Figure 1.

2.3 Results and discussion

The decay factors of 154 languages from 6 families (Indo-European, Uralic, Turkic, Semitic, Dravidian, Rgyalrongic+Kiranti) are mapped on Figure 1:
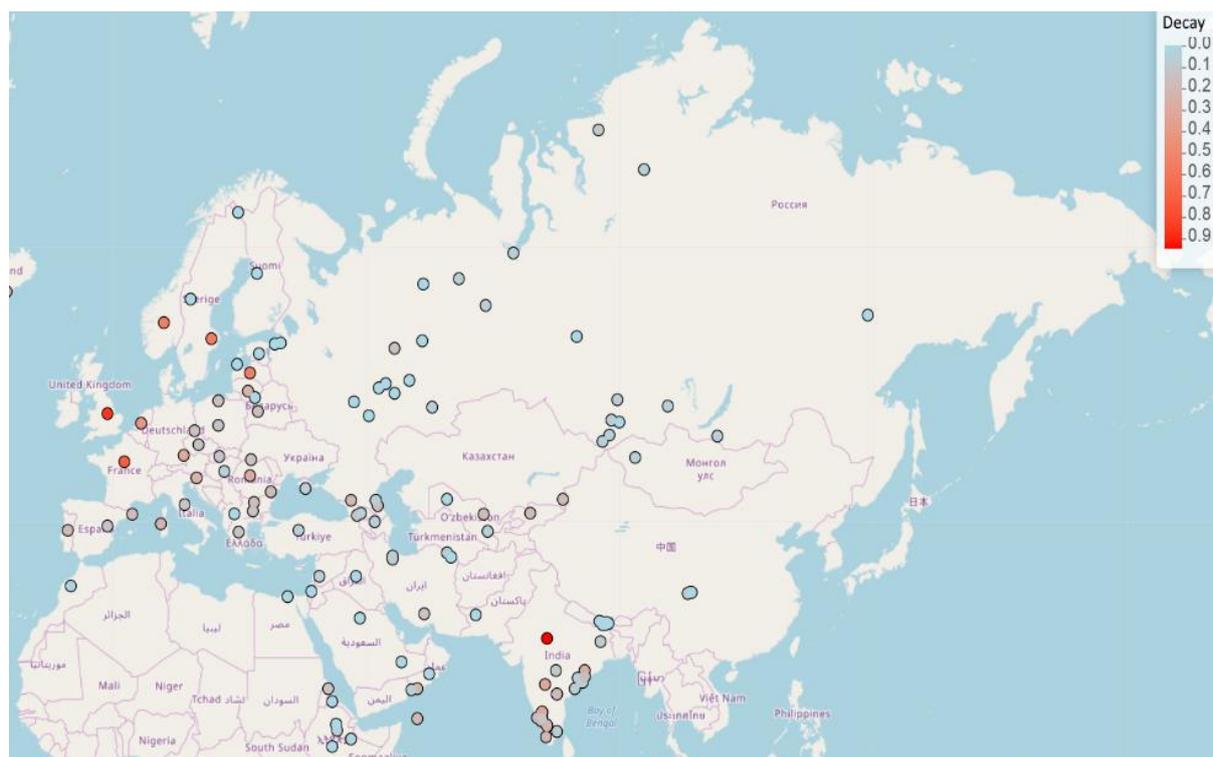

Figure 1: The degree of decay factor across different languages of Eurasia

As one can see in Figure 1,[9] there are two decay hotspots (marked by the red dots): (i) the subcontinent *Hindustan* (cf. Siewierska 2004: 278) where primarily Indo-European (Indo-Aryan), Dravidian and Austroasiatic (Munda) languages are spoken, and (ii) *Northwestern*

---

[8] This formula looks as follows: $X_{norm} = (X-X_{min}) / (X_{max}-X_{min})$. However, since X_{min} is always zero with all three factors I have simplified the formula above.

[9] Figure 1 has been created with the lingtypology package of R (Moroz 2020).

Europe with Indo-European (Indo-Aryan) languages. The high degree of decay in Hindustan is not of immediate relevance here and will not be discussed further (see Peterson 2017 for an account). Just note that this area is strongly affected by Southeast Asian languages (foremost Austroasiatic) which is the largest area strongly dispreferring indexing in the world (cf. the map in Donohue & Denahm 2020: 455 based on 2378 languages[10]), also diachronically (Bisang 2014).

By contrast, the Turkic and Uralic languages are the most conservative here (blue dots, decay factor < 0.1). Slavic languages, in turn, are somewhat intermediate here: these languages generally show quite a low degree of decay (grey dots, < 0.2), albeit not as low as most of the Turkic and Uralic languages.

The following three adjacent areas are of immediate relevance to Slavic: *Northwestern Europe, Northeastern Eurasia* and *the Transitional area*. I discuss these in detail in what follows.

First, there is *Northwestern Europe.* This area is the European hotspot of the indexes decay.[11] It encompasses languages such as French, English, Scandinavian but also to some extent German, Dutch or North Italian varieties. These languages abandon the indexes inherited from Proto-Indo-European to various degrees. For example, Dutch totally abandons person distinctions in the plural, cf. *lach-en* 'laugh-1/2/3PL' 'we/you/they are laughing' and French in the singular. English retains the inflection only with one verb (*to be*) and elsewhere only the 3SG marker -*s* (cf. *he go-es*), while Scandinavian does not have any person-number distinctions at all.[12]

Second, there is the most conservative, huge area primarily populated by Turkic and Uralic languages, stretching from the Baltic sea to Siberia (henceforth *Northeastern Eurasia*). A number of other languages of the database belong here as well, e.g. Iranian (Indo-European) Persian or Ossetic (decay factor <0.13 in both).

Finally, there is the *Transitional area* with conservative Indo-European languages of the Balkans such as Greek (decay factor < 0.14), Albanian (< 0.07), Romanian (< 0.3). Slavic languages (< 0.15) also belong here.

I summarize all three areas in Table 5:

Northwestern Europe	Ø 0.61	Germanic	Ø 0.49
		French	0.72
Transitional area	Ø 0.12	Greek	< 0.14
		Albanian	< 0.07
		Slavic	Ø < 0.15
Northeastern Eurasia	Ø 0.05	Turkic	Ø < 0.07
		Uralic	Ø 0.02

Table 5: Decay factors across the three areas (Ø – averaged across languages)[13]

A number of observations can be made on the basis of these figures.

First, the decay factors map perfectly on the geographical space despite the fact that the decay factors have been computed entirely independently of geography. In other words, the

---

[10] Unfortunately, there was no information on the database that Donohue & Denahm (2020) have used.

[11] Note that areas with a high degree of decay may also undergo developments towards new indexes, for example, from cliticization of personal pronouns. This is the case in Non-standard French (Givón 1976) or in some North-Italian dialects (Gerlach 2002) and in some languages of the Eastern part of Hindustan, e.g., in Munda languages (Peterson 2017: 241-242). I do not take these recent developments into account because recent developments are primarily constrained by their grammaticalization sources and paths and not by their target functions.

[12] Note that irregular verbs such as English *to be* were not taken into account.

[13] Since the decay factor shows only a small range of variation within these subfamilies, averages are a good approximation here.

geographical distribution of different degrees of decay in Europe suggests a strong correlation between the geographical position and the decay factor on what I shall refer to *the East-West axis* (in addition to the North-South axis). In what follows, I refer to this correlation as to the *East-West cline*. Slavic languages take an intermediate position on this cline: both with regard to their decay factor of ca. 0.15 (which is numerically between 0.61 of Northwestern Europe and 0.05 of Northeastern Eurasia), and with regard to their geographical position (Eastern Europe).

Now that the correlation between the geographical space and the decay factor has been established, it can be reasonably inferred that Slavic languages have retained the morphological functionality[14] of their inflectional person-number indexing system from Proto-Indo-European into Early and Modern Slavic due to their geographic position on the East-West cline.[15] What is more, the East-West cline may also be observed within Slavic as well. The only index that does undergo considerable reduction across Slavic languages is the third-person index. Yet, the reduction of this index increases from the more conservative Russian in the East to the more innovative Polish in the West, Ukrainian and Belarusian being in between when it comes to the dental segment:

	3rd singular	3rd plural
Russian	*-Vt, -Vt'*	*-Vt, -Vt'*
Ukrainian	*-V(t')*	*-Vt'*
Belarusian	*-V(t')*	*c' (< -t')*
Polish	*-V*	*-V*

Table 6: Reduction of the third person indexes from East to West
(Brackets mean differential marking)

Compare 3SG: Russian *nes-et* 'carry-3SG', dialectal *-et'*, Ukrainian *čyta-e* 'read-3SG', Belarusian *viedaj-e* 'know-3SG' (*-t'* is present in other conjugations), Polish *nies-ie* 'carry-3SG'; 3PL: Russian *nes-ut(')* 'carry-3PL', Ukrainian *čyta-jut'* 'read-3PL', Belarusian *vodz=jac'* 'drive=3PL', Polish *nios-ą* 'carry-3PL'. Thus, I conclude that even the internal Slavic variation is constrained by the East-West cline.

Secondly, in what follows, I argue that the method pursued here finds strong support in what we know about the history of Slavic ethnicities. Although the decay factor has been computed entirely independently of the historical facts about the migration and linguistic contacts of Slavic, it strongly correlates with what we know about the particular contact configuration of Slavic.

While the East-West cline roughly explains the intermediate position of the Slavic languages, it fails to explain why the Slavic decay factor ($< Ø$ 0.15) strongly gravitates towards the Transitional area ($Ø$ 0.12) as well as the languages of the conservative Northeastern Eurasia ($Ø$ 0.05) and is considerably more distant from the Northwestern Europe ($Ø$ 0.61). Thus, the East-West cline can only be a part of the explanation. I argue that the particular contact configuration of Slavic is responsible for this skewing: Slavic languages were in much more

---

[14] In this paper, I only focus on the *morphological functionality* of an indexing paradigm which is solely about the ability of a paradigm to unequivocally express all six person-number indexes but not, for example, whether these indexes may be used as referring expressions on their own or not (the so-called *pro*-drop), etc.

[15] I did not discuss recently emerged indexing structures such as the past-tense paradigm of Polish (and of some western Ukrainian dialects) (Andersen 1987) because the young age of these paradigms does not allow estimation of any diachronic trends here. However, I predict that also the new paradigms will follow the areal pressures established above. Finally, note that I do not make any predictions about whether or not such paradigms may emerge because grammaticalization sources and paths are subject to very different constraints from the resulting patterns.

intensive contact with languages of Northeastern Eurasia and the Transitional area than with the languages of Northwestern Europe.

Thus, different Slavic populations were in intensive contact with a number of Uralic and Turkic tribes, sometimes even leading to a partial or total assimilation of the latter by the Slavs:

(2)   Turkic (*inter alia*, Galton 1997: 21ff; Gołąb 1992: 310, 401ff; Nichols 1993; Stachowski 2014)
        Danube Bulgars, Khazars, Pechenegs, Coumans, Tatars, etc.

(3)   Uralic (*inter alia,* Andersen 2003: 47-48; Haarmann 2014; Rjabinin 1997)
        Ugric Hungarian, Finnic Merya, Muroma, Tver', partly Karelian (Old Novgorod)

Furthermore, intensive contact is documented with the conservative Indo-European languages of the Transitional area that have also retained their person-number indexes. Thus, Early Slavs must have had intensive contact with East Goths as many old loanwords demonstrate. Later, South Slavs had intensive contact with Greeks and Balkan Romance varieties (cf. Gołąb 1992: 64). These languages, too, preserve the Proto-Indo-European person-number inflection quite faithfully (cf. Table 2 above). The contact with these languages as well as with some of the Turkic and Uralic languages listed above must have been very intense, involving assimilation of these speakers by Slavs and thus considerable substrate effects. Substrate effects generally have a much stronger impact on all domains of language than occasional cultural contact. Substrates normally involve massive child bilingualism during the transitional stage before the total assimilation. In this way, they allow for the transfer of the properties of the substrate into the dominant language. Having said this, child bilingualism – in contrast to adult bilingualism (Trudgill 2009) – does not cause substantial reduction of the morphologically less transparent patterns (see a recent overview in Kempe & Brooks 2018), especially if the source and the target language share the category (e.g. indexing).

By contrast, languages of Northwestern Europe were not in such intense contact with Slavic. These were the languages associated with cultural and political prestige and may have had an impact only on particular social subgroups of Slavic population. For example, we know of different kinds of military, political and trade contacts with Normans from Scandinavia – especially with early East Slavs – but this contact did not involve any significant population migration from Scandinavia into the East Slavic territory with a subsequent assimilation or other kinds of situations that would licence intense language contact. Similarly, the latter impact of Low and, subsequently, High German of Hansa on East Slavic, still later, the impact of French and, finally, of English on all Slavic languages was certainly not as intense as to go very much beyond lexical cultural borrowings.

Having said this, Slavic languages (Ø 0.15) are not as conservative as Northeastern Eurasia (Ø 0.05) and do show a minor degree of decay in parallel to the entire Transitional area which, in turn, must have been under some influence of Northwestern Europe. More specifically, German has and has had intense contacts with the languages of the West Slavic branch.[16] Yet, German (0.3) is the least innovative language in Northwestern Europe. Hence, its effect on the decay factor of West Slavic is moderate.

To sum up, the historical data briefly summarized above evidently point to very different degrees of language contacts of Slavic. Crucially, the degrees of contact intensity strongly correlate with the decay factors. The stronger the contact between two languages the more similar their decay factors are. Generally, the strong numerical association of Slavic with Northeastern Eurasia and the Transitional area correlates with the intensity of linguistic contacts

---

[16] Germans assimilated some West Slavic branches, e.g. Sorbian (partly) or Polabian (totally). The lowest decay factor of German within Northwestern must be interpreted as due to the supporting effect of Slavic.

with these areas. To conclude, the pure geographic position on the East-West cline is not the only factor that contributes to the retention of the verbal inflection in Slavic. Another important factor is the particular *contact configuration* of Slavic.

The claim that the degree of decay of the person-number indexes paradigm is primarily conditioned by the geographic position and the contact configuration also finds support elsewhere. For example, the two languages of the East Baltic branch of IE that are closely related to Slavic show an unexpectedly high decay factor of 0.24 in Lithuanian and 0.55 in Latvian. This is despite the fact that elsewhere these two languages are demonstrably the most conservative among modern Indo-European languages. Both Baltic languages have generalized the singular-plural syncretism across all conjugational paradigms in the third person, cf. Lithuanian *neši-a* 'carry-3SG=PL' '(s)he carries / they carry'.[17] Furthermore, Latvian also shows second-third person zero syncretism in the singular of all verbs of the present tense of the first conjugation, cf. *ņem-Ø* 'take-2SG/3SG/3PL'.

This high decay factor finds its explanation once one looks closer into the contact configuration of these languages. Lithuanian and Latvian are situated geographically close to Northwestern Europe and, crucially, had much stronger contact with its languages than Slavic, as early as some runic inscriptions that were found in Latvia. Later Low (and subsequently High) German heavily dominated this area: Livonia (of which Latvia was a part) and later also the Duchy of Courland and Semigallia (Courland and Semigallia are parts of modern Latvia) as well as the so-called Minor Lithuania (*Mažoji Lietuva*) were primarily dominated by German-speaking elite (stemming from the Teutonic Order and, later, from Hansa) for centuries. The impact of Swedish must have been stronger than in Slavic at least during Swedish Livonia (1629-1721). Finally, Polish also played an important role in the East of Latvia (Polish Livonia). Recall that Polish is slightly less conservative in this respect than, for example, Russian.

3. The marker of the Partitive construction

Above I have explored the retention mechanism on the backdrop of the macro-areal East-West cline. In this section, I examine an innovation in Slavic that is also due to macro-areal pressures, ultimately resulting from language contact. I focus on the marking strategy of partitivity in Slavic. Partitivity is a semantic relation which codes a subset of a superset ("part-whole" relation) (Enç 1991; von Heusinger 2002: 261-62; Koptjevskaja-Tamm 2001, 2009; Seržant, forthc.-d):

(4)   *Some of the students of our university / Two of the students of our university*

Thus, both examples in (4) encode partitivity relations, featuring a subset of the superset consisting of all the students of our university.

Proto-Indo-European, Proto-Slavic and Early Slavic employed the possessive strategy to encode this relation, namely, the genitive (the partitive genitive, *genetivus partitivus*) (Miklosich 1883: 449, 473):

(5)   Old Russian (Kiev Chronicle, 108)
        komuždo   ixъ   razdajalъ   volosti
        each.DAT.SG   3PL.GEN grant.PERF.PARTC   territory.ACC.PL
        'he granted territories to each of them'

---

[17] This is also somewhat parallel to Finnish which occasionally employs the third-person singular form for the third-person plural (agreement suspension).

In (5), the recipient participant 'each' is a member (the *subset*) of the well-defined *superset* 'they' (scil. the sons of Vladimir) – a proportional part-whole relation that is coded by the genitive case on the superset NP 'of them'.[18]

Yet, the genitive can no longer be used in this function in modern Slavic in most instances. Instead, the new periphrastic construction with a separative preposition such as Modern Russian *iz* 'from' has to be used. Thus, (5) cannot be rendered in Modern Russian without the (new) partitive preposition *iz* 'from':

(6)   Modern Russian (p.k.)
        *každomu   *(iz   n)ix   razdal   volosti*
        each.DAT.SG from   3PL.GEN grant.PST   territory.ACC.PL
        'he granted territories to each of them'

Other Slavic languages are similar, cf. *od* 'from' in Serbian:

(7)   Serbian
        neki   od   nas
        some   from   1PL.GEN
        'some of us'

One might think that the separative strategy of marking the part-whole relation is just a default strategy to mark partitivity relations. However, partitive markers are very diverse cross-linguistically, ranging from various adpositions and cases (possessive, locative, instrumental, comitative) to verbal clitic partitive particles (cf. Budd 2014; Hoeksma, ed., 1996; Koptjevskaja-Tamm 2001, 2009; Seržant, forthc.-d) or just involving no marker at all. The following example from Malayalam features the locative marking (the locative strategy):

(8)   Malayalam (Dravidian; Asher & Kumari 1997: 218)
        aaŋkuṭṭikaḷil   raṇṭə peer   vannu
        male.child.PL.LOC   two   person come.PST
        'Two of the boys came.'

As a matter of fact, the separative strategy is typical only for Eurasia as I argue in Section 3.2 below.

3.1. The database
This study is based on a convenience sample of partitive expressions covering 138 languages, 171 entries (one language may have more than one strategy to encode partitives) from 46 families and all six macro-areas. The sample is biased toward Eurasia (48% of the entries, 82/171) but not for Indo-European. The database is published at zenodo.org (Seržant 2020a).

3.2. The separative strategy of partitives as a macro-areal pressure

---

[18] I distinguish between proportional part-whole relations (such as *seven of the students*) in which both the subset and the superset belong to the same kind and meronymic part-whole relations (*a leave of a tree*) (see Seržant, forthc.-d).

Coding strategies reveal areal biases as Figure 2 suggests.[19] The separative strategy (red dots) is typical only for Eurasia. Other areas prefer other strategies. For example, the locative strategy as in (8) above (orange dots) is more typical for Africa.[20]

Below, Table 7 presents the

Figure 2; Distribution of the major coding strategies of partitives (from Seržant, forthc.)

row counts and highlights the outliers. Unfortunately, the sample is not large enough to provide for a more precise picture. However, it is sufficient to argue that the null hypothesis of the coding strategies being normally distributed across all macro-areas can be rejected. To test this I computed the standardized residuals in the *chi*-square test. The separative strategy in Eurasia reveals itself as a statistically significant outlier (standardized residuals < 2.5).[21]

	Africa	Australia	Eurasia	North America	Oceania	South America
separative	3	2	**44 (57%)**	-	3	3
possessive	1	1	17 (22%)	1	4	-
locative	**15 (56%)**	-	8 (10%)	-	1	1
unmarked	8 (30%)	8	**8 (10%)**	**5**	10 (50%)	4
other	-	-	1	-	2	2
	27 (100%)	*11*	*77 (100%)*	*6*	*20*	*10*

Table 7: Coding strategies across the macro-areas, bold indicates statistical significance (from Seržant, forthc.)

Given that the underlying sample is a convenience sample, I have also computed Fisher's Exact test in a two-by-two table ($p < 0.001$) that puts fewer requirements on the distribution of the observed data (Janssen et al. 2006: 425):

	Eurasia	elsewhere
separative	44	11
other	33	63

Table 8: Contrasting the separative strategy in Eurasia with elsewhere

---

[19] Figure 2 was created with the lingtypology package of R (Moroz 2020). Unfortunately, languages with two different partitivity markers cannot be properly presented on the map for technical reasons. For example, Russian has two strategies – the recessive possessive (marked by the genitive case only) and the innovative separative one (marked by the separative preposition *iz* 'from') – and only the latter is indicated on the map.

[20] The locative strategy may also be found elsewhere, for example, in Samoan (Austronesian; not included in the database), cf. the locative preposition *i* being used as a generalized partitive (cf. Mosel & Hovdhaugen 1992: 108).

[21] The standardized residuals for the outlier cells in bold were above/below 2.5 in a *chi*-square test. Despite the fact that the underlying dataset is a convenience sample, the test is a legitimate method because the availability of a good grammatical description does not interfere with particular strategies; the latter can be considered to be picked out randomly.

The areal bias towards the separative strategy in Eurasia is not only supported statistically by genealogically unrelated languages of this area that select this pattern (e.g. Turkic, Basque, etc.) but also by those languages of the sample that are related. For example, many modern European languages have partitive prepositions that all stem from separative prepositions such as French *de,* Dutch *van* or Latvian *no* but which are nevertheless unrelated among each other etymologically. Thus, French *de* is not related to German *von* or Scandinavian *av/af.* What is more, even within subfamilies we find unrelated prepositions such as German (West Germanic) *von* vs. Scandinavian (North Germanic) *av/af* or Lithuanian (East Baltic) *iš* 'from' and Latvian (East Baltic) *no* 'from'. This is also true for the Slavic languages, cf. Modern Russian (East Slavic) *iz* 'from' in (6) vs. *od* 'from' in (7) in Modern Serbian (South Slavic). This means that the partitive function of these prepositions is recent and not due to common inheritance. At the same time, the fact that all these languages innovate their coding strategies of partitives on the basis of the same locational metaphor, i.e. on the metaphor of separation from the Ground, strongly suggests that this pattern is not selected accidentally by the langauges of Eurasia. Once common inheritence is excluded as an explanation for correlations (even for the genealogically closely related languages), this correlation must be due to an areal pressure. The latter is ultimately grounded in multiple pattern borrowing across Eurasia.[22]

To conclude, modern Slavic languages are on their way to abandoning the inherited possessive strategy of encoding part-whole relations which was crucially based on the possessive metaphor. Instead, all Slavic languages tend to adopt a coding strategy that is based on the spatial metaphor of separation. This process is not accidental but is motivated by a strong macro-areal pressure.[23]

The emergence of this common Slavic morphosyntactic propertie is essentially different from person-number indexes discussed above (Section 2): while the former was selected for retention due to the macro-areal pressures, the latter, by contrast, was replaced by the separative strategy that is the dominant strategy in Eurasia.

4. Other morphosyntactic phenomena and macro-areal pressures

In Section 2, I have argued that macro-areal pressures in combination with the specific contact configuration constrain the retention and innovation of morphosyntactic properties. I have argued that the retention of person-number indexes in the languages of the Balkans can be accounted for in terms of the East-West cline as well as the particular contact configuration of the Transitional area. In this section, I argue that the areal method pursued above may account for other phenomena selected for retention or loss as well. Thus, the macroareal East-West cline in combination with the local contact configuration has a supporting effect in other domains of verbal inflection in the languages of the Balkan sprachbund.

For example, the ancient inflectional aspectual distinctions between the aorist (the perfective past) and the imperfect (the imperfective past) have been well-preserved (*mutatis mutandis*) in South Slavic. There is a cline within Slavic that stretches from East Slavic with no traces of these categories (since at latest the 18[th]C. when the artificial use of the aorist was lost even in the written language), via West Slavic (with some traces of these, e.g., in Upper Sorbian) to South Slavic which has quite faithfully preserved these categories (e.g. Bulgarian). Other Indo-European languages of the Balkans such as Greek and Albanian have also preserved this distinction. More generally, the Balkan Indo-European languages are unique in this respect:

---

[22] Copying of the separative strategy is also found far beyond Europe, cf., for example, Pakendorf (2010: 727-729) on pattern borrowing of the separative partitive coding in Evenki (Tungusic), Dolgan and Sakha (Turkic).

[23] I have no explanation on the origin of the separative bias in Eurasia that so heavily affected Slavic languages.

all other modern branches of this huge family have lost this distinction entirely (except for modern Armenian).

At the same time, Balkan languages underwent a considerable loss of fusional complexity in another domain: very early on, the original case inflection was lost. This loss is also not unexpected given the macro-areal cline from East (many case distinctions) to West (with no cases in languages such as Scandinavian), cf. Figure 3. This cline is supported by non-Slavic languages. In the South, Greek has lost most of its cases already by the Roman period (cf. Rafiyenko & Seržant 2020 on case in Postclassical Greek). Neither Albanian nor Rumanian have preserved much of the Proto-Indo-European or even, in the case of Rumanian, Proto-Romance case system (as observed in Latin).

Figure 3: Number of cases in languages of Eurasia (Iggesen 2013)

By contrast, Finnic languages of the Uralic stock have not only preserved the Proto-Uralic cases but also developed a number of new cases: the so-called secondary spatial cases and the partitive. Standard and North Russian not only quite faithfully preserved the case system of Proto-Slavic but they even underwent an increase of cases to include the defective partitive and locative cases (Breu 1994; Seržant 2014, 2015, forthc.-a).

To conclude, selection of inherited properties for retention and loss is subject to macroareal properties as well as to specific contact configuration.

5. Discussion and Conclusions

In this paper, I have presented an approach towards causal explanations of the properties of Slavic languages. This approach crucially combines two factors: the particular contact configuration and the macro-areal pressures and clines.

More specifically, I have argued that the preservation of the inherited person-number paradigm in modern Slavic languages is due to the areal pressures constrained by (i) the position of Slavic on the East-West cline and (ii) the contact configuration (Section 2). I have argued that Slavic languages occupy the Transitional area with a minor exposure towards decay. I have also claimed that West and South Slavic languages align with the languages of this area in the degree of retention, while East Slavic, especially Russian is somewhat closer to the languages of Northeastern Eurasia. I have also argued that this match in the degree of retention can hardly be accidental and must be due to the preserving effect of language contact. L2-speakers have less difficulty learning the pattern that they already have in their L1. Therefore, they do not have a strong need to override the L2-system. Since any reduction produced by L2 speakers has also to be adopted by native speakers in order to be established in the language, retention appears to be a more efficient option on the population level when there is a functional and formal match between the relevant L1 and L2 categories.

In parallel, the innovation of the partitivity marker is equally non-accidental and must be explained as the result of the macro-areal pressure to encode partitivity with a separative (ablatival) marker (Section 3). I have also briefly discussed other morphosyntactic phenomena such as inflectional aspect (Section 4). This evidence also points to strong macroareal

constraints and to the effect of the particular contact configuration. Thus, the genealogical preconditions and areal pressures conspire to shape the morphosyntax of the Slavic family.

More generally, this study provides evidence against the mutually exclusive separation of genealogical and areal effects (as, for example, in Nichols 1992: 163) and claims that inherited patterns too are subject to areal pressures. The effect of areal pressures on inherited patterns may be either loss – optionally with a subsequent innovation (Section 3) – or retention (Section 2). I have claimed that the preservation of person-number indexes in Slavic is a case in point. Similarly, the preservation of inflectional aspectual distinctions between aorist and imperfect (South Slavic) as well as the preservation of inflectional case may also be accounted for by the supporting effect of language contact.[24]

Having said this, the studies carried out here rely on different degrees of resolution of the areal variation. In the first study (Section 2), the distribution of indexing systems across Eurasia is treated "under a higher zoom-in factor". Here, I have taken into account different indicators of decay and worked with averaged measures for each language instead of just relying on the binary distinction of whether or not the languages of the sample have preserved the inherited indexing system. By contrast, with partitives (Section 3), the binary distinction of whether or not the languages code their partitives with the separative strategy was in focus. With respect to both phenomena, I have argued for a strong areal pressure on Slavic. However, only in the first study, I had enough data at disposal to argue that a macro-areal pressure (the East-West cline) may in addition be constrained by the particular contact configuration. More specifically, I have shown that Slavic languages align with the Transitional area (Southeastern Europe) and Northeastern Eurasia and less so with Northwestern Europe. Unfortunately, I do not have comparable representative data on partitives at my disposal that would point in favour (or against) the effects of the particular contact configuration of Slavic. Nevertheless, given the somewhat sporadic evidence at my disposal, contact configuration does seem to play a role here as well. Thus, the functional properties of partitives in East Slavic, especially in North Russian, align with those of Finnic languages (Uralic) much stronger than with those of the languages of Northwestern Europe (Seržant 2014, 2015). Moreover, there are morphosyntactic properties as well that are primarily determined by the contact configuration. Thus, Russian is the only Slavic language that was on the way to developing a dedicated partitive case similar to the partitive case in the Finnic languages (Breu 1994; Seržant 2014). Finally, both families align here in another respect. They both have two partitive markers: the recent, separative marker (*iz* 'from' in Russian and the elative case in Finnic) which encodes exclusively part-whole meanings and the old, non-transparent marker (the genitive case in Russian and the partitive case in Finnic), which only rarely denotes part-whole relations and is primarily used for other functions related to pseudo-partitivity.

An additional factor constraining the development towards loss (and innovation) or retention might be the stability of the category itself. Thus, it has been argued that person-number indexes are a stable category: languages do not tend to lose this category unless they come into one of the areas where indexing is dispreferred (Nichols 1995: 343; Wichmann & Holman 2009; Seržant, forthc.-b). By contrast, partitives very often develop into pseudo-partitives and thus lose their original meaning which is why new partitives emerge again ("Partitivity cycle" in Seržant, forthc.-c).

It is the task of future research to examine how high-contact situations affect macro-areal clines discussed in this paper. Trudgill argues for a strong correlation between the abruptness of contact and the decay of fusional morphology (2009). According to Trudgill

---

[24] The supporting effect of language contact may also lead to an increase in fusional morphological complexity. For example, the preservation of the suffixal and prefixal actionality derivations due to language contact allowed the Slavic languages to develop an aspectual system that is typologically unusual with regard to its morphological properties (Wiemer & Seržant 2017).

(2009, 2011), socio-historical factors contribute to the decay vs. retention of fusional morphology (a proxy of complexity adopted in this research) and thereby they affect areal clines. High-contact situations which crucially involve massive adult bilingualism resist the preservation of fusional morphology such as verbal person-number indexes (Trudgill 2011: 34ff). This is because adult L2-learners typically do not achieve the level of proficiency that would guarantee faithful transmission of non-transparent fusional morphological structures across generations (cf., *inter alia*, Clahsen & Muysken 1996; Lupyan & Dale 2010; Trudgill 2011: 39; Kempe & Brooks 2018). It is possible that high-contact situations affect areal clines. Thus, we have observed that the Balkan sprachbund is very much conservative when it comes to the verbal morphology while it is innovative when it comes to fusional case. By contrast, Northwestern Europe is innovative in both respects. It is possible that Northwestern Europe underwent even more intensive and abrupt contact situations involving a high proportion of adult L2 speakers than the Balkans. This would explain why not only fusional case but also the cross-linguistically stable category of person-number indexing has been lost here.[25]

Finally, I have not discussed universal pressures, which must have also affected Slavic morphosyntax. Thus, the third person singular tends to be shorter and is shortened more drastically across Slavic (cf. Table 6 above). This tendency is in line with the universal pressure of the third person to be shorter than other persons (Benveniste 1971; Koch 1995; Bybee 1985: 53; Cysouw 2003: 61-2; Siewierska 2010; Bickel et al. 2015). I summarize:
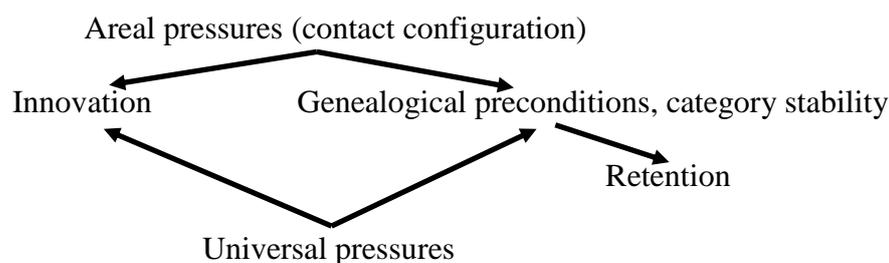

Figure 4: Schematic ontology of factors constraining language

Figure 4 illustrates the relation between retention and innovation, which are both constrained by the areal pressures emerging from contact configuration (and universal pressures).

---

[25] Another potential explanation for the stronger reduction in Northwestern Europe than in the Balkans might have been an assumption of an unknown isolating substrate language. However, this assumption is ruled out. Thus, the autochthonous population of Scandinavia spoke Proto-Saami, which was certainly not an isolating language. Similarly, the autochthonous population of England spoke early Celtic, which was not isolating either.

References

Andersen, Henning. 1987. From auxiliary to desinence. In Martin Harris & Paolo Ramat (eds.), Historical development of auxiliaries. Berlin: Mouton de Gruyter. 21–51.

Andersen, Henning 2003: Slavic and the Indo-European migrations. In: Andersen, Henning (ed.), *Language Contacts in Prehistory. Studies in Stratigraphy.* Amsterdam/Philadelphia: John Benjamins. 45-76.

Andronov, Michail S. 2009: *A Comparative Grammar of the Dravidian Languages*. [Beiträge zur Kenntnis südasiatischer Sprachen und Literaturen 7]. Wiesbaden: Otto Harrassowitz.

Anthony, D. W. 1995. Horse, wagon & chariot: Indo-European languages and archaeology, *Antiquity* 69, 554–65.

Asher, R. E. & T. C. Kumari. 1997. *Malayalam.* Descriptive Grammars Series, Descriptive Grammars. London & New York: Routledge.

Benveniste, Émil 1971: Problems in general Linguistics. Translated by Mary Elisabeth Meek. [Miami Linguistics Series 8]. Coral Gables (Florida): University of Miami Press.

Bickel, Balthasar 2007: Typology in the 21st century: Major current developments, *Linguistic Typology* 11, 239–251.

Bickel, Balthasar; Witzlack-Makarevich, Alena; Zakharko, Taras; Iemmolo, Giorgio 2015: Exploring diachronic universals of agreement: alignment patterns and zero marking across person categories. In: Fleischer, Jürg; Rieken, Elisabeth; Widmer, Paul. Agreement from a diachronic perspective. Berlin: De Gruyter Mouton, 29 - 51.

Bisang, Walter. 2014. On the strength of morphological paradigms. In Martine Robbeets & Walter Bisang (eds.), *Paradigm change: In the Transeurasian languages and beyond*. 23–60. Amsterdam: John Benjamins.

Boelaars, J. H. M. C. 1950. *The Linguistic Position of South-Western New Guinea.* (Orientalia Rheno-Traiectina, 3.) Leiden: E. J. Brill.

Breu, Walter 1994: Der Faktor Sprachkontakt in einer dynamischen Typologie des Slavischen. In: H.-R. Mehlig (ed.), *Slavistische Linguistik* 1993. München. 41–64.

Budd, Peter. 2014. Partitives in Oceanic languages. In Luraghi, Silvia & Tuomas Huumo (eds.), *Partitive Cases and Related Categories.* (Empirical Approaches to Language Typology, 54). Berlin & Boston: Mouton DeGruyter. 523-562.

Bybee, Joan 1985: *Morphology: A Study of the Relations between Meaning and Form.* Amsterdam/Philadelphia: John Benjamins.

Clahsen, Harald & Pieter Muysken 1996: How adult second language learning differs from child first language development, *Behavioural and Brain Studies* 19, 721-723.

Cysouw, Michael 2003: *The paradigmatic structure of person marking*. Oxford: Oxford University Press.

DeLancey, Scott 2010: Towards a history of verb agreement in Tibeto-Burman, *Himalayan Linguistics* 9(1), 1–38.

DeLancey, Scott: 2011: Agreement prefixes in Tibeto-Burman, *Himalayan Linguistics* 10(1), 1–29.

Dixon, R.M.W. 2003: A program for linguistics, *Turkic languages* 7, 157-180.

Enç, Mürvet 1991: The semantics of specificity, *Linguistic Inquiry* 22, 1–25.

Erdal, Marcel 2004: *A Grammar of Old-Turkic.* [Handbook of Oriental Studies. Handbuch der Orientalistik. Section eight. Central Asia. Volume three.]. Leiden/Boston: Brill.

Galton, Herbert 1997: *Der Einfluss des Altaischen auf die Entstehung des Slavischen.* Wiesbaden.

Gerlach, Birgit 2002. *Clitics between syntax and lexicon*. Amsterdam: John Benjamins.

Givón, Talmy 1976. Topic, pronoun and grammatical agreement. In: Charles Li (ed.), *Subject and topic*. London/New York: Academic Press, 149–88.

Gołąb, Zbignew 1992: *The Origins of the Slavs: A Linguist's View.* Columbus, Ohio: Slavica Publishers.

Gvozdanović, Jadranka. 2015. Evaluating prehistoric and early historic linguistic contact. In Dag T. T. Haug (ed.), *Historical linguistics 2013 (Selected papers from the 21st International Conference on Historical Linguistics, Oslo, 5–9 August 2013)*,89–108. Amsterdam, Philadelphia: Benjamins.

Haarmann, Harald 2014: Finnougrisch-slavische Sprachkontakte. In: Karl Gutschmidt, Sebastian Kempgen, Tilman Berger & Peter Kosta (eds.), *Die slavischen Sprachen: Ein internationales Handbuch zu ihrer Struktur, ihrer Geschichte und ihrer Erforschung.* Vol. 2. Berlin: De Gruyter Mouton.

Hansen, Björn 2014: Partial typologies: Grammatical categories. In: Karl Gutschmidt, Sebastian Kempgen, Tilman Berger & Peter Kosta (eds.), *Die slavischen Sprachen: Ein internationales Handbuch zu ihrer Struktur, ihrer Geschichte und ihrer Erforschung.* Vol. 2. 2202-2221. Berlin: De Gruyter Mouton.

Haspelmath, Martin 2013: Argument indexing: A conceptual framework for the syntax of bound person forms. In: Bakker, Dik & Haspelmath, Martin (eds.) 2013. *Languages across boundaries: Studies in memory of Anna Siewierska*, 197-226. Berlin: De Gruyter Mouton.

Hasselbach, Rebecca 2004: The Markers of Person, Gender, and Number in the Prefixes of G-Preformative Conjugations in Semitic, *Journal of the American Oriental Society* 124(1), 23-35.

von Heusinger, Klaus 2002: Specificity and Definiteness in Sentence and Discourse Structure, Journal of Semantics 19(3), 245–274.

Hoeksema, Jacob, ed., 1996. *Partitives. Studies on the Syntax and Semantics of Partitive and Related Constructions.* (Groningen-Amsterdam Studies in Semantics 14). Berlin: De Gruyter.

Holman, Eric W., Christian Schulze, Dietrich Stauffer & Søren Wichmann 2007: On the relation between structural diversity and geographical distance among languages: Observations and computer simulations, *Linguistic Typology* 11, 393-421.

Holman, Eric W., Cecil H. Brown, Søren Wichmann, André Müller, Viveka Velupillai, Harald Hammarström, Sebastian Sauppe, Hagen Jung, Dik Bakker, Pamela Brown, Oleg Belyaev, Matthias Urban, Robert Mailhammer, Johann-Mattis List, Dmitry Egorov 2011: Automated Dating of the World's Language Families Based on Lexical Similarity, *Current Anthropology* 52(6), 841-875.

Honti, László 2010: Personae ingratissimae? A 2. személyek jelölése az uráliban, *Nyelvtudományi Közlemények* 107, 7-57.

Huehnergard, John. 2000. Comparative Semitic Linguistics. Unpublished. Cambridge, Mass.

Iggesen, Oliver A. 2013. Number of Cases. In: Dryer, Matthew S. & Haspelmath, Martin (eds.) *The World Atlas of Language Structures Online.* Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info/chapter/49, Accessed on 2020-11-16.)

Jacques, Guillaume 2016: Le sino-tibétain: polysynthétique ou isolant? *Faits de langues* 47(1), 61-74.

Jakobson, Roman 1930: *K charakteristike evrazijskogo jazykovogo sojuza.* In: *R. Jakobson, Selected Writings I, Phonological Studies.* The Hague. 2nd Reprint 1971. 144-201.

Jakobson, Roman 1931: Über die phonologischen Sprachbünde, *Travaux du cercle linguistique de Prague* 4, 234–240.

Janhunen, Juha 1982: On the structure of Proto-Uralic, *Finnisch-ugrische Forschungen* 44, 23-42.

Janssen, Dirk, Balthasar Bickel, Fernando Zúñiga. 2006. Randomization tests in language typology, *Linguistic Typology*, 10(3). 419-440.

Kattein, Rudolf 1983: *Diachrone Typologie der slavischen Sprachen.* Frankfurt-am-Main/Bern/New York: Peter Lang.

Kempe, Vera & Patricia J. Brooks 2018: Linking Adult Second Language Learning and Diachronic Change: A Cautionary Note, *Frontiers in psychology* 9, 480.

Koch, Harold 1995: The creation of morphological zeros. In Geert Booij (ed.), *Yearbook of morphology*, 31–71. Dordrecht: Kluwer.

Koptjevskaja-Tamm, Maria 2001: A Piece of the Cake and a Cup of Tea: Partitive and Pseudo-Partitive Nominal Constructions in the Circum-Baltic Languages. In: Östen Dahl & Maria Koptjevskaja-Tamm (eds.), *The Circum-Baltic Languages. Typology and Contact*, vol. 2: *Grammar and Typology* (Studies in Language Companion Series). Amsterdam & Philadelphia, 523–568.

Koptjevskaja-Tamm, Maria 2009: A lot of grammar with a good portion of lexicon": towards a typology of partitive and pseudopartitive nominal constructions. In Helmbrecht, Johannes and Nishina, Yoko and Shin, Yong-Min and Skopeteas, Stavros and Verhoeven, Elisabeth (eds.), *Form and Function in Language Research. Papers in Honour of Christian Lehmann*, 329-346. Berlin, New York: Mouton de Gruyter.

Kulonen, Ulla-Maija 2001: Zum n-Element der zweiten Personen besonders im Obugrischen. *Finnisch-Ugrische Forschungen* 56, 151-174.

Kusters, Wouter 2003: *Linguistic complexity: The influence of social change on verbal inflection.* Leiden: Leiden University Press.

Laanest, Arvo 1982: Einführung in die ostseefinnischen Sprachen. Autorisierte Übertragung aus dem Estnischen von Hans-Hermann Bartens. Hamburg: Buske.

LaPolla, Randy 2003: Overview of Sino-Tibetan morphosyntax. In: Thurgood, Graham, Matisoff, James A. & Bradley, David (eds.), *Linguistics of the Sino-Tibetan area: The state of the art*. Canberra: Department of Linguistics, Australian National University [Pacific Linguistics Series C, 87]. 22-42.

Lazard, Gilbert. 1998. *Actancy*. Berlin & New York: De Gruyter Mouton.

Lipiński, Edward 2001: *Semitic Languages: outline of a comparative grammar*. 2ed. Leuven: Peeters.

Gary Lupyan & Rick Dale 2010: Language Structure Is Partly Determined by Social Structure, *PLOS One* 5(1), e8559.

Meier-Brügger, Michael 2010: *Indogermanische Sprachwissenschaft*. 9., durchgesehene und ergänzte Auflage. Unter Mitarbeit von Matthias Fritz und Manfred Mayrhofer. Berlin: De Gruyter.

Miklosich, Franz 1883: *Vergleichende Grammatik der slavischen Sprachen*. Band 4: Syntax. Wien: Wilhelm Braumüller.

Moroz, George 2020: R package 'lingtypology'. Full title: Linguistic Typology and Mapping. Version 1.1.3. URL https://CRAN.R-project.org/package=lingtypology, https://github.com/ropensci/lingtypology/, https://ropensci.github.io/lingtypology/

Mosel, Ulrike & Even Hovdhaugen. 1992. *Samoan Reference Grammar*. Oslo: Scandinavian University Press.

Nichols, Johanna. 1992. *Linguistic Diversity in Space and Time*. London / Chicago: University of Chicago Press.

Nichols, Johanna. 1993. The Lingusitic Geography of the Slavic Expansion. In: Robert A. Maguire & Alan Timberlake (eds.), *American Contributions to the Eleventh International*

Congress of Slavists, Bratislava, August- September 1993: Literature, Linguistics, Poetics. Columbus: Slavica.

Nichols, Johanna 1995. Diachronically stable structural features. In: Andersen, Henning (ed.), *Historical Linguistics 1993. Selected Papers from the 11th International Conference on Historical Linguistics, Los Angeles 16–20 August 199*, pp. 337–355. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Nichols, Johanna & Tandy Warnow 2008: Tutorial on computational linguistic phylogeny, *Language and Linguistics Compass* 2, 760–820.

Olander, Thomas 2015: *Proto-Slavic Inflectional Morphology A Comparative Handbook.* Leiden/Boston: Brill.

Pakendorf, Brigitte. 2010. Contact and Siberian languages. In Raymond Hickey (ed.), *The Handbook of Language Contact*, 714-737. Malden: Wiley-Blackwell

Peterson, John 2017: Fitting the pieces together – Towards a linguistic prehistory of eastern-central South Asia (and beyond), *JSALL* 4(2): 211–257.

Rafiyenko, Dariya & Ilja A. Seržant 2020: Postclassical Greek. An overview. In: Rafiyenko, D. & I. Seržant (eds.), *Contemporary Approaches to Postclassical Greek.* [Trends in Linguistics series]. Berlin/New York: De Gruyter. 1-18.

Rjabinin, E. A. 1997. *Finno-ugorskie plemena v sostave drevnej Rusi*. St. Petersburg.

Róna-Tas, András 1998: The Reconstruction of Proto-Turkic and the Genetic Question. In: Johanson & Csato (eds.), *The Turkic Languages*. 67-80.

Seržant, Ilja A. 2012: The so-called possessive perfect in North Russian and the Circum-Baltic area. A diachronic and areal approach, *Lingua* 122, 356-385.

Seržant, Ilja A. 2014: The Independent Partitive Genitive in North Russian. In: Seržant, I. A. and B. Wiemer (eds.), *Contemporary Approaches to Dialectology: The area of North, Northwest Russian and Belarusian vernaculars.* [Slavica Bergensia 13]. Bergen: John Grieg AS. 270-329.

Seržant, Ilja A. 2015: Independent partitive as a Circum-Baltic isogloss, *Journal Language Contact* 8, 341-418.

Seržant, Ilja A. 2020a. Dataset for the paper "Typology of partitives", Linguistics [Data set]. *Zenodo.* http://doi.org/10.5281/zenodo.4277624

Seržant, Ilja A. 2020b. Dataset for the paper "Slavic morphosyntax is primarily determined by its geographic location and contact configuration", Scando-Slavica Journal [Data set]. *Zenodo.* http://doi.org/10.5281/zenodo.4277593

Seržant, Ilja A., forthc.-a: The Circum-Baltic Area. An Overview. In: Jan Fellerer & Neil Bermel (eds.), *Oxford Guides to the World's Languages. The Slavonic Languages*. OUP.

Seržant, Ilja A., forthc.-b: Cyclic changes in verbal indexes are not drift processes. In: Sansò, A. & L. Konnerth (eds.), *Diachrony of individual person markers.* Special issue of *Folia Linguistica Historica*.

Seržant, Ilja A., forthc.-c: Towards diachronic typology of partitives. Submitted to Giuliana Giusti & Petra Sleeman (eds.), Partitive Determiners, Partitive Pronouns and Partitive Case. [Linguistiche Arbeiten]. Berlin: De Gruyter

Seržant, Ilja A., forthc.-d: Typology of partitives. *Linguistics.* (accepted)

Siewierska, Anna. 2010. Person asymmetries in zero expression and grammatical functions. In Franc Floricic (ed.), *Essais de typologie et de linguistique générale. Mélanges offerts à Denis Creissels*, 471-485. Paris: Presses de L'Ecole Normale Supérieure.

Siewierska, Anna 2013: Verbal Person Marking. In: Dryer, Matthew S. & Haspelmath, Martin (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info/chapter/102, Accessed on 2019-01-21.)

Sinnemäki, Kaius 2020: Linguistic system and sociolinguistic environment as competing factors in linguistic variation: A typological approach, *Journal of Historical Sociolinguistics* 6(2), 1-39.

Stachowski, Stanisław 2014: Türkischer Einfluss auf den slavischen Wortschatz. In: Karl Gutschmidt, Sebastian Kempgen, Tilman Berger & Peter Kosta (eds.), *Die slavischen Sprachen: Ein internationales Handbuch zu ihrer Struktur, ihrer Geschichte und ihrer Erforschung.* 1198-1210. Berlin: De Gruyter Mouton.

Stankiewicz, Edward 1986: *The Slavic Languages. Unity and Diversity.* Berlin/ New York / Amsterdam: Mouton de Gruyter.

Thomason, Sarah Grey. 2001. *Language contact: An introduction.* Edinburgh: Edinburgh University Press.

Trudgill, Peter 2009: Sociolinguistic typology and complexification. In: Geoffrey Sampson, David Gil & Peter Trudgill (eds.), *Language Complexity as an Evolving Variable.* Oxford: OUP. 98-109.

Trudgill, Peter 2011. *Sociolinguistic typology: Social determinants of linguistic complexity.* Oxford: Oxford University Press.

Wiemer, Björn 2019: *Matrëška* and areal clusters involving varieties of Slavic. On methodology and data treatment. In: Andrii Danylenko & Motoki Nomachi (eds.), *Slavic on the Language Map of Europe*. [Trends in Linguistics, 333]. Berlin/Boston: DeGruyter. 21-62.

Wiemer, Björn & Ilja A. Seržant 2017: Diachrony and typology of Slavic aspect: What does morphology tell us? In: Bisang, Walter & Andrej Malchukov (eds.), *Unity and diversity in grammaticalization scenarios.* [Studies in diversity linguistics 16] Berlin: Language Science Press. 239-307.

Wichmann, Søren & Eric W. Holman 2009. *Temporal stability of linguistic typological features.* Munich: Lincom Europe.