

1. Intro

1.1. Debitive construction

- primarily different kinds of necessity, as well as related epistemic meanings (see, *inter alia*, Bergmane (ed.) 1959: 615, Fennel 1973, Holvoet 2007, Kalnača & Lokmane 2014);
- formed by auxiliary *būt* ‘to be’ (optional in the present affirmative, cf. Stolz 1987: 225) and the non-finite debitive *gerundive* (marked by the prefix *jā-*);
- own case frame (Standard Latvian):

all NP types	¹ DAT _{A/S} -NOMP
SAP, reflexive pro	¹ DAT _{A/S} -ACCP

(1) (Standard Latvian)

*Viņam ir jā-redz sav-s draug-s / *sav-u draug-u*
 3.SG.DAT.M AUX.PRS.3 DEB-see POSS.ADJ-NOM.SG friend-NOM.SG / *-ACC.SG *-ACC.SG
 ‘He has to see his friend.’ [Elicited]

(2) (Standard Latvian)

*Viņ-am ir jā-redz *es / mani*
 3.SG.DAT.M AUX.PRS.3 DEB-see *1SG.NOM / 1SG.ACC.SG
 ‘He has to see me.’ [Elicited]

1.2. Historical background: the origin of the gerundive

Possessive cxn with an embedded infinitival clause -> **Necessity cxn**
I have a cake to eat -> *I have to eat a cake*

Gerundive:

- from a reinforced infinitive (Seržant & Taperte 2016), cf. German: *Ich habe ein Haus zu bauen*, where *jā-* is an old subjunction (cf. Lith. *jo(g) < *jā*), “*(у меня) есть хлеб, чтобы-съесть”;
- traditionally from a generalized relativizer *jā* (*inter alia*, Endzelīns 1951: 971-3; Holvoet 1998; Wälchli 2000 since Adolphi’s *Erster Versuch Einer kurz-verfasseten Anleitung Zur Lettischen Sprache* published in 1685), “*(у меня) есть хлеб, который-съесть”;
- crucially, the P argument was originally the subject of the main clause (possessive/existential), hence, the NOM marking, cf.:

NB (Latvian is SVO):

	OV	VO
Old Latvian	↑ 92% (105)	↓ 8% (9)
Early Modern Latvian	↑ 63% (250)	↓ 37% (149)
Modern (Contemporary) Latvian	↑ 44% (1407)	↓ 56% (1783)

Table 1: Position relative to the predicate²

- The NOM marking must have been original as suggested by the origin and some few archaic examples as (3):

¹ DAT argument sometimes referred to as “(non-canonical) subject”, cf. Fennel 1973, Stolz 1987.

² This and all other statistical data stem from Seržant & Taperte 2016.

- (3) (Folklore Latvian, Šmits 1925/1937 *apud* Endzelīns 1951: 972)
tad tu man arī būsī jā-kuopj
 then **2SG.NOM** 1SG.DAT also be.FUT.2SG DEB-care
 ‘Then, I will also have to take care of **you**.’

1.3. Expansion of ACC marking onto other NP types in Colloquial Latvian

In Colloquial Latvian, the ACC marking expands onto all other NP types:

	NOM	ACC	<i>p</i> -value compared to Contemporary Latvian
Old Latvian	95% (81)	5% (4)	significant ($p < 0.05$)
Contemporary Latvian	87% (2778)	13% (412)	

Table 2: ACC vs. NOM, diachronic sketch

yielding DOM with NOM vs. ACC:

ACC is obligatory for SAP+reflexive but (Latvian, all registers)
ACC is optional with all other NP types (Colloquial Latvian only)

Further characteristics of this DOM *according to the ontology in Seržant & Witzlack-Makarevich 2016+):

- restricted (to the debitive predication only),
- NP-triggered, i.e. conditioned by NP-internal properties (as opposed to predicate/clause-triggered).

1.4. Symmetric vs. asymmetric

Morphological coding:

symmetric: X vs. Y **asymmetric: X vs. zero**

In Latvian, both ACC and NOM are morphologically marked, cf.:

	<i>o</i> -stems	<i>ā</i> -stems	<i>i</i> -stems	<i>u</i> -stems
singular	<i>nam-s</i>	<i>galv-a</i>	<i>nakt-s</i>	<i>tīrg-us</i>
plural	<i>nam-i</i>	<i>galv-as</i>	<i>nakt-is</i>	<i>tīrg-i</i>

Table 3: The nominative case endings of the main declensions in Latvian

- “traditional” DOM: asymmetric; based on prominence/salience/markedness; the morphologically marked NP is more prominent on animacy, definiteness, etc., cf., *inter alia*, Bossong (1985); but also Aissen (2003).
- Later, some have argued that symmetric and asymmetric DAM systems are conditioned by different parameters (de Hoop & Malchukov 2008; Dalrymple & Nikolaeva 2011: 19; Abraham & Leiss 2012; Iemmolo 2013).
- E.g., Iemmolo (2013) argues that:
 - symmetric DOM respond to: verbal semantics, e.g. polarity and quantification, affectedness or boundedness (aspectuality), whereas
 - asymmetric systems reflect various argument-internal properties, most prominently information-structure role, animacy, etc. (similar Abraham & Leiss 2012: 320).
- E.g. de Hoop & Malchukov (2008: 584):
 - “We have also argued that while asymmetrical differential case marking can sometimes be explained by DISTINGUISHABILITY, **symmetrical differential case marking must always be due to IDENTIFY.**” (emphasis – mine) (IDENTIFY is

primarily related to semantic roles, information structure roles of the argument, while DISTINGUISHABILITY serves to distinguish A and P)

NB: Symmetric systems are problematic for functional explanations of DOM which are crucially based on economy and markedness:

- “[...], the most natural kind of transitive construction is one where the A is high in animacy and definiteness, and the P is lower in animacy and definiteness; and any deviation from this pattern leads to a more marked construction.” Comrie (1989:128)
- ➔ *Markedness reversal* (Aissen 2003)

So, what about Latvian?

2. Synchronic variation in Latvian

2.1. Primary factors:

- Only 13% of ACC -> a significant overall preference for NOM.
- Responsible for this ratio is the combination of two factors (unfortunately no way to disentangle):

1) **strong influence of prescriptive rules** (favoring NOM)

the ratio of 87% NOM cannot be accounted for simply in terms of purification rules and standardization norms albeit, confessedly, the latter are very strong

NB: unlikely that the speakers are completely re-learning this part of grammar in the schools and universities.

2) **the very expansion of ACC is quite incipient in the language** (favoring NOM)

2.2. What about 13% of ACC? Secondary factors:

F(actor)1 Position:

	DEB-Object		Object-DEB	
ACC	150	8% (ACC)	262	19% (ACC)
NOM	1633		1145	

Table 4: The interrelation of position relative to the debitive predicative and case ($p < 0.001$)

F2 Different input types:

	ACC	NOM	<i>p</i> -value if compared to the preceding NP type	<i>p</i> -value if compared to the total average
Reciprocal pronouns	70% (7)	3		significant²
<i>wh</i> -pronouns	25% (54)	164	significant²	significant
Demonstrative pronouns	21% (64)	246	<i>not significant / chance</i>	significant
<i>viss</i> ‘all’	19% (12)	51		<i>not significant / chance</i>
3 rd person pronoun	19% (14)	59		<i>not significant / chance</i>
Indefinite pronouns	11% (16)	125		<i>not significant / chance</i>
Full NP (total)	10% (232)	1998		significant
Proper names	9% (13)	134		<i>not significant / chance</i>
Total	13% (412)	2780		

Table 5: The interrelation of the input type and case

- (4) *Effect scale of expansion of the ACC marking*
 SAP pronouns (& reflexive pronoun)^{100%} < Reciprocal pronouns^{70%} < *wh*-pronouns^{25%}
 < demonstratives^{21%} < *viss* ‘all’ & third person pronoun^{19%} < ...**13%**... indefinite pronouns^{11%} < full NPs^{10%} < proper names^{9%}

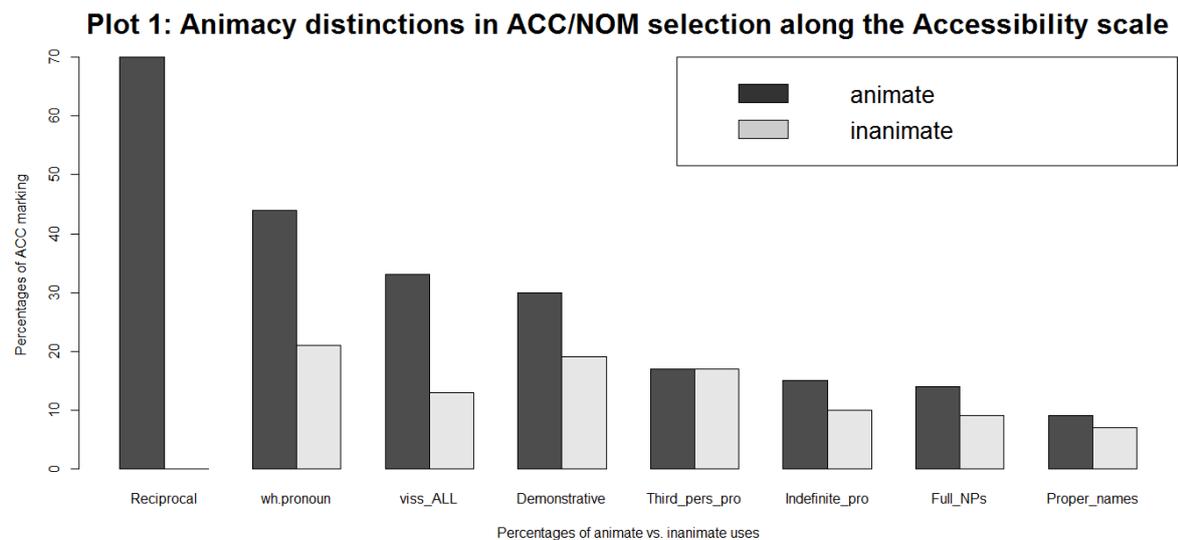
Notably, *proper names* typically score high on various scales (such as the definiteness or animacy scale) and acquire new marking earlier than many other NP types -> counter to a number of hierarchies:

- (5) *Extended Animacy Hierarchy* (Croft 2003:130)
First/second person pronouns > third person pronoun > proper names > human common noun > nonhuman animate common noun > inanimate common noun (Croft 2003: 130, for the sake of comparison the arrows are reverse here)
- (6) *Definiteness hierarchy* (Aissen 2003: 437)
Personal pronoun > proper name > definite NP > indefinite specific NP > non-specific NP

At the same time, the Effect scale in (4) largely correlates with the *Accessibility Marking Scale* in (7) as suggested in Ariel (1990: 73; 2001: 31):

- (7) *Accessibility Marking Scale* (Ariel 1990: 73)
Full name + modifier > Full name > Long definite description > Short definite description > Last name > First name > Distal demonstrative + modifier > Proximal demonstrative + modifier > Distal demonstrative (+ NP) > Proximal demonstrative (+ NP) > Stressed pronoun + gesture > Stressed pronoun > Unstressed pronoun > Cliticized pronoun > Extremely High Accessibility Markers (gaps, including pro, PRO and *wh* traces, reflexives, and Agreement)

F3 Animacy:



from this it follows that:

- the *prima facie* average of 13% of ACCs (above) is not particularly telling (the range is between 7% with inanimate proper names up to 70% with reciprocals);
- 13% results from the overall frequency of full NPs;
- the largest distance is found among the most accessible and, at the same time animate NPs
- in turn, the third person pronouns (both animate and inanimate), indefinite pronouns, all kinds of full NPs and proper names do not differ from each other to the same extent. Here, animacy is a less important factor.
- accessibility degree of a particular NP is a stronger attractor of ACC than animacy because the rows are ordered according to accessibility in the first place.

F Definiteness:

Definiteness has revealed itself as not any significant:

	Definite	Indefinite	Percentage of definite NPs
ACC	126 (264)	106 (148)	54% (64%)
NOM	1052 (1634)	945 (1144)	53% (59%)

Table 6: Definite vs. indefinite NPs; (full NPs only, all NP types of the sample in brackets)

2.3. Interim conclusions

Conditioning factors: Preverbal position + high accessibility + animacy (= saliency of an NP)

Factors with no impact: Definiteness

2.4. Verbal factor

Semantic classes	ACC	NOM	%	Odds ratio, relative to the overall average
Experiencer verbs	133	598	18%	1.5
Communication verbs	44	325	12%	0.9
Agentive verbs	167	1351	11%	0.8
Meeting verb	21	184	10%	0.8
Transfer verb	27	310	8%	0.6

Table 5: Frequencies of ACC across different verb classes

F4 Semantic class of the verb

	ACC	NOM	%	<i>p</i> -value
Experiencer verbs	74 (133)	271 (598)	21% (18%)	significant (< 0.01)
All other verbs except <i>provēt</i>	91 (259)	575 (2170)	14% (11%)	

Table 7: Frequencies of ACC across different verb classes

NB: the figures represent animate objects for both semantic classes only (whereas the respective total is given in brackets)

F5 Lexical verb

Verb	ACC	NOM	%	Verb	ACC	NOM	%
<i>ņemt</i> 'take'	15	188	7%	<i>sacīt</i> 'say'	24	158	13%
<i>ņest</i> 'carry'	8	105	7%	<i>lasīt</i> 'read'	15	96	14%
<i>rakstīt</i> 'write'	16	221	7%	<i>tēlot</i> 'feign'	13	79	14%
<i>sazvanīt</i> 'call'	7	88	7%	<i>dzirdēt</i> 'hear'	10	58	15%
<i>pirkt</i> 'buy'	12	122	9%	<i>klausīties</i> 'listen'	20	109	16%
<i>pucēt</i> 'dress, clean'	6	58	9%	<i>laist</i> 'let'	30	162	16%
<i>slēgt</i> 'close'	18	184	9%	<i>iznīcināt</i> 'destroy'	30	149	17%
<i>stūķēt</i> 'push'	10	104	9%	<i>mīlēt</i> 'love'	32	128	20%
<i>satikt</i> 'meet'	21	184	10%	<i>redzēt</i> 'see'	22	90	24%
<i>taisīt</i> 'make'	22	180	11%	<i>apvainot</i> 'offend'	18	50	26%
<i>piekaut</i> 'pummel'	5	38	12%	<i>apsaukāt</i> 'call so. offensively'	14	15	48%
<i>trenkāt</i> 'expel'	7	53	12%	<i>provēt</i> 'try'	20	12	63%
				<i>besīt</i> 'annoy'	0	7	#

Table 8: Frequencies of ACC across different lexical verbs

- different lexeme-specific factors;

- 48% with *apsaukāt* ‘to call so. offensive names’ are due to frequency-based associations (cf. Haspelmath (2008: 191)’s *complementary expected association*; von Heusinger (2008) on the basis of the Spanish DOM). The experienter verbs generally more frequently take animate objects, with 47% in our sample (animate 345 hits vs. inanimate 386 hits), than other verbs do, with 27% of all occurrences in our sample (666 animates vs. 1763 inanimates).

(8) Associations relevant for ACC vs. NOM selection:

Verbs having a higher association ratio with animate objects will generally favor ACC (even with inanimate objects), while verbs having a higher association ratio with inanimate objects will favor NOM (even with animates).

Underlying assumptions: animacy is a strong predictor which may impose implicationally coding preferences via frequency effects.

- *ap-vainot* ‘to offend’ has 91% (62/6) and *ap-saukāt* ‘to call so. offensive names’ 83% (24/5) of animate objects, in contrast to, for example, *dzirdēt* ‘to hear’ with 3% (2/66) of animate objects.
- Yet, *ap-saukāt* and *ap-vainot* have 48% and 26% of ACC, respectively, while *dzirdēt* attests only 15% of ACCs.

3. The overall account of the Latvian DOM

Motivations: *Increasing grammaticalization/ Incipient loss of the debitive’s own case frame / Canonization* (cf. Seržant 2013; Seržant & Taperte 2016)

Conditioning factors: Primary-importance factors: > *Incipientcy of the spread of ACC* > *Prescriptive rules imposing the retrograde effect* > Secondary-importance factors: > *The positional factor* > *the accessibility factor* (4) > *The animacy factor* > *The verbal factor* > *The lexical-verb factor*

⇒

the Latvian DOM is not particularly different wrt conditioning from the asymmetric DOMs such as Spanish

4. Rise of DOM / diachronic typology

DOM generally very often emerges from:

- a mixing/merger of two historically distinct constructions; as Harris & Campbell (1995: 257) put it “the pull towards consistency between subsystems [...] is stronger than that between rules” conditioning the subsystems.

- development towards consistency across subparadigms and generalized alignment is well-described for different languages (cf. Harris & Campbell 1995: 263-5 on Pamir languages (Iranian) and Kartvelian; Haig 2008).

- Paradigm outliers start gradually merging with the more frequent and, hence, more basic case frame of the same verb:

- **DOM Latvian:** debitive construction “borrows” the case-patterns from the basic transitive construction;
- **DOM in Russian:** the affirmative NOM-ACC and the predicate-negated NOM-GEN started merging already in Proto-Slavic as has been pointed out first in Klenin (1983). (Old Russian basically required GEN in the “Polish” manner)
- **DOM in Tundra Yukaghir** (isolate): cleft-construction (marked argument focus) with the marker *-len* merged with the basic transitive construction (Maslova 2008: 795-6).

- **DOM in Romance:** dislocation construction (marked topic) conflates with basic transitive (Spanish, Sicilian, Occitan) (Laca 1987; Pensado 1995; Melis 1995; Iemmolo 2010);
- **DOM in Romanian:** dislocation construction (marked topic) conflates with basic transitive (Mardale 2015);
- **DOM in Persian:** dislocation construction (marked topic) conflates with basic transitive (Iemmolo 2010);
- **DOM in Syrian Arabic:** dislocation construction (marked topic) conflates with basic transitive (Brustad 2000: 353).

(9) From dislocation to basic transitive construction:

As for X_i, Y_j VERB it/him/her_i -> *Y_j VERB as for it/him/her/X_i*

- there is no crucial difference between the Latvian DOM and the textbook DOM of Spanish in terms of the origin;
- minor difference is that in most of the DOM systems mentioned the new marker is not a structural case to begin with (but e.g. a dislocation marker ‘as for’).

5. Conclusions

⇒ Functionally, the Latvian DOM very much resembles textbook examples of asymmetric DOMs.
 ⇒ Diachronically, the Latvian DOM resembles other DOMs: the new marking is introduced into the construction in order to code direct objects.

- many DOMs have a different motivation than “markedness reversal” as per Comrie (1989), namely: **the trend to generalize across subparadigms towards a unified transitive alignment.**
- however, the introduction of the new marking – once it entered the basic transitive construction – may follow Comrie’s discrimination because it gradually acquires the syntactic function of coding direct objects in the new environment;

NB: Insignificant differences: Romance *a*, Romanian *pe*, Iranian *-rā*, Arabic *la-* are originally markers of peripheral NPs while the Latvian ACC is the old marking for direct objects.

⇒ Morphologically, asymmetric vs. symmetric is a matter of the morphological makeup of the respective case paradigm:

- Latvian simply disallows zeros in all declensional classes;
- In Finnish, the partitive/accusative DOM is symmetric in the singular but asymmetric in the plural (whereby the less prominent option is marked while more prominent (ACC) is unmarked ACC=NOM) -> **asymmetric morphologically, though functionally symmetric.**
- analogically in Russian: *я выпил коньяк-#* vs. *я выпил коньяк-а/у*: zero_{prominent} / Y_{less prominent} -> **asymmetric morphologically, though functionally symmetric.**
- moreover, in plural, *я купил конфет-ы* vs. *я купил конфет-#* -> **asymmetric morphologically, though functionally symmetric.**
- DOM in Kolyma Yukaghir (Yukaghir, isolate) is symmetric: ACC *-gele/-kele* for definite nouns and the INS *-le* for indefinite nouns with third person A’s (Maslova 2003: 93) ± functionally reminiscent of the asymmetric DOM in Biblical (and modern) Hebrew involving the DOM preposition *’et*. -> **symmetric morphologically, though functionally asymmetric.**
- even Spanish *a mí* 1SG.ACC.DOM vs. *me* 1SG.ACC is not fully X vs. zero.

⇒ symmetric vs. asymmetric seems to be rather a property of the morphological makeup of a particular case-marking class; there are three main prototypes:

1	+ECONOMY, +UNIVERSAL	the prominent meaning is coded (with more material than the more prominent)	Spanish
2	REDUNDANT, EXPLICIT	both the prominent and the non-prominent meanings are similarly coded	Latvian, Yukaghir, Finnish SG
3	+ECONOMY, -UNIVERSAL	the less prominent is coded (with more material than the more prominent)	Finnish PL

References

- Abraham, Werner & Elisabeth Leiss 2012: The case differential: syntagmatic versus paradigmatic case – its status in synchrony and diachrony, *Transactions of the Philological Society* 110(3), 316-341.
- Aissen, Judith. 2003. Differential object marking: Iconicity vs. economy. *Natural Language and Linguistic Theory* 21. 435–483.
- Ariel, Mira. 1990. *Accessing noun-phrase antecedents*. London: Routledge.
- Bossong, Georg 1985: *Empirische Universalienforschung. Differentielle Objektmarkierung in den neuiranischen Sprachen*. Tübingen: Narr.
- Brustad, Kristen E. 2000: *The Syntax of Spoken Arabic. A Comparative Study of Moroccan, Egyptian, Syrian and Kuwaiti Dialects*. Washington, D.C.: Georgetown University Press.
- Comrie, Bernard. 1989. *Language universals and linguistic typology*, 2nd ed. Oxford: Blackwell.
- Croft, William. 2003. *Typology and universals*. 2nd edition. Cambridge: Cambridge University Press.
- Dalrymple, Mary & Irina Nikolaeva. 2011. *Objects and information structure*. Cambridge: Cambridge University Press.
- Haspelmath, Martin. 2008. Creating economical patterns in language change. In *Linguistic Universals and Language Change*, Jeff Good (ed), 185–214. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199298495.003.0008
- Iemmolo, Giorgio. 2010. Topicality and differential object marking: Evidence from Romance and beyond. *Studies in Language* 34(2). 239–272.
- Iemmolo, Giorgio. 2013. Symmetric and asymmetric alternations in direct object encoding. *STUF - Language Typology and Universals* 66(4). 378–403.
- Krys'ko, Vadim B. 1994. *Razvitie kategorii oduševlennosti v istorii russkogo jazyka*. Moscow: Lyceum.
- Laca, Brenda 1987: Sobre el Uso del Acusativo Preposicional en Español. In: C. Pensado (ed.), *El Complemento Directo Preposicional*. Madrid: Visor. 61-91.
- Malchukov, Andrej L. & Helen de Hoop Tense. 2007. On fluid differential case marking: A bidirectional OT approach. *Lingua* 117. 1636–1656.
- Malchukov, Andrej L. and Helen de Hoop. 2010. Tense-, aspect- and mood-based differential case marking. *Lingua* 121(1). 35–47.
- Mardale, Alexandru 2015: Differential Object Marking in the First Original Romanian Texts. In: Hill, Virginia (ed.), *Formal Approaches to DPs in Old Romanian*. Leiden/Boston: Brill. 200-245.
- Maslova, Elena 2003a: *A Grammar of Kolyma Yukaghir*. [Mouton Grammar Library 27]. Berlin: De Gruyter.
- Maslova, Elena. 2003b. Information Focus in Relational Clause Structure. In: Tsunoda, Tasaku, & Kageyama, Taro (eds), *Voice and Grammatical Relations: Festschrift for Masayoshi Shibatani*. [Typological Studies in Language] Amsterdam: John Benjamins. 175–194.
- Melis, Chantal 1995: El Objeto Directo Personal en *El Cantar de Mio Cid*. Estudio Sintáctico-Pragmático. In: C. Pensado (ed.), *El Complemento Directo Preposicional*. Madrid: Visor. 133-163.
- Pensado, C. 1995: El Complemento Directo Preposicional: Estado de la Cuestión y Bibliografía Comentada. In: C. Pensado (ed.), *El Complemento Directo Preposicional*. Madrid: Visor. 11-59.
- Seržant, Ilja 2013: The Diachronic Typology of Non-canonical Subjects. In: Seržant, I. A. and L. Kulikov, *The Diachrony of Non-canonical Subjects*. SLCS 140. Amsterdam/Philadelphia: John Benjamins. 313-360.
- Seržant, Ilja A. & Jana Taperte 2016: Differential Argument Marking with the Latvian Debitive: a Multifactorial Analysis (together with Jana Taperte). In: *Argument Structure in Baltic*, eds. Axel Holvoet & Nicole Nau [Valency, Argument Realization and Grammatical Relations in Baltic, 3]. Amsterdam/Philadelphia: John Benjamins. 199-258.
- Seržant, Ilja & Alena Witzlack-Makarevich 2016+: Differential Argument Marking: an introduction. In: Seržant, I. A., A. Witzlack-Makarevich & K. Mann (eds.), *The Diachronic Typology of Differential Argument Marking*. [Studies in Diversity Linguistics]. Berlin: Language Science Press.
- Seržant, Ilja & Alena Witzlack-Makarevich 2016b+: Diachronic Typology of Differential Argument Marking. In: Seržant, I. A., A. Witzlack-Makarevich & K. Mann (eds.), *The Diachronic Typology of Differential Argument Marking*. [Studies in Diversity Linguistics]. Berlin: Language Science Press.
- Sinnemäki, Kaius 2014: A typological perspective on Differential Object Marking, *Linguistics* 52(2), 281-313.