

Participial free adjuncts in early Slavic treebanks: the discourse properties of the dative absolute

In this paper I exploit data from the TOROT Treebank (Eckhoff & Berdicevskis 2015) to substantiate the intuition that the functions of the early Slavic dative absolute (DA) are better understood at the discourse-structural level (Worth 1994; Corin 1995; Collins 2004, 2011). On the basis of New Testament Greek, Bary & Haug (2011: 5) have shown that participial clauses behave to a large extent like grammaticalised rhetorical relations (GRRs): their discourse effects on sub-sentential units are comparable to those between elementary discourse units within formal frameworks of discourse interpretation, particularly Segmented Discourse Representation Theory (SDRT) (Asher & Lascarides 2003).

In this paper the notion of GRR and the corpus evidence thereof (Haug 2012) are used to maximally exploit the existing annotation on the Old Church Slavonic (OCS) treebank and narrow down the number of constructions potentially competing with the DA at a finer-grained discourse level. Under a GRR view, DAS are found to be identifiable as *framing* participles (cf. Bary & Haug 2011: 11-13; Haug 2012: 299-301), on the basis of indirect evidence emerged from lexical, morphosyntactic and information-structural annotation on OCS texts, as well as their aligned Greek parallels. SDRT's diagnostics for rhetorical relations (Asher & Lascarides 2003; Asher et al. 2007) are then used to confirm the patterns thereby discovered and to produce a more granular classification of discourse properties not captured by GRRs. In particular, among *frames_{GRR}*, DAs most prototypically introduce *Background_{SDRT}*, in contrast with the much more diversified relations conveyed by conjunct participles. This approach especially helps motivate the persistence of the construction in later Slavic text, in which syntactically 'non-canonical' or 'unruly' DAs (Collins 2011) appear to become more frequent.

The methodological issues which arise when investigating the same phenomenon in texts with shallow annotation is self-evident. A comparison between translated and original Slavic texts can be crucial for a proper understanding of early Slavic constructions which do not stand out as either slavish Greek calques (e.g. *eže*-nominalised infinitive, cf. MacRobert 1986) or genuinely Slavic phenomena (e.g. *possessive constructions*, cf. Eckhoff 2018). In TOROT, however, only OCS texts have deep annotation, as well as richly annotated Greek parallels that provide vital insights when carrying out quantitative analyses. Later Church Slavonic and early Slavic original texts are instead only partially represented in the corpus, and have overall much shallower annotation. This is at best limited to morphology (automatically performed, but in need of post-correction), and (generally patchy) manually-added syntactic annotation. Since speedy deep annotation is unfeasible through manual tagging, an alternative strategy has been employed to tackle our specific case study in treebanks with abysmally different annotation depths. This is shown through two case studies on datasets with shallower annotation than the one available for OCS. The first exploits mixed (O)CS and Old Russian data, from a treebank which also contains several levels of annotation but has no information-structural tagging nor available Greek parallels, so that Slavic independence of Greek has to be evaluated indirectly on the basis of internal evidence (i.e. synchronic consistency and distributional differences with competing constructions). The second case study uses data from a new, unannotated text (from the 14th-century Middle Bulgarian *Bdinski Sbornik*) which has been specially uploaded onto TOROT and automatically annotated for PoS and morphology (using the model presented in Scherrer et al.

2018), and only the sentences containing potential DAs have then been strategically annotated with syntactic information. All potentially prototypical DA's configurations are isolated by making predictions on the basis of the frequencies calculated on OCS across several variables (e.g. position in the sentence; number and types of clauses intervening between the DA and the matrix clause; lemma-tense correlations; aspectual shifts). This is mainly achieved by extracting highly predictable configurations first, allowing for closer inspection only of less prototypical uses. This approach (a synchronic one at the outset) is also meant to have the effect of aiding a diachronic analysis by noting how the DA's prototypical uses have expanded with respect to those emerged from the OCS dataset. The paper is structured so that a primary focus is given to the methodological challenge of testing existing theories by means of indirect evidence emerged from treebanks with different levels of annotation, including newly added (therefore yet unannotated) ones. Collaterally, the case study on DA is unravelled by stressing the role that combined corpus data and formal frameworks can have in shedding new light on long-standing language-specific debates.

References

- Asher, Nicholas & Lascarides, Alex. 2003. *Logics of conversation*. Cambridge: Cambridge University Press.
- Asher, Nicholas, Laurent Prévot & Laure Vieu. 2007. Setting the background in discourse. *Discours* 1. DOI: <https://doi.org/10.4000/discours.301>
- Bary, Corien & Dag Haug. 2011. Temporal anaphora across and inside sentences: The function of participles. *Semantics and Pragmatics* 4. 1-56.
- Collins, Daniel E. 2011. The pragmatics of “unruly” dative absolutes in Early Slavic. In Eirik Welø (ed.), *Indo-European syntax and pragmatics: contrastive approaches*, Oslo Studies in Language 3(3). 103–130.
- Corin, Andrew R. 1995. The dative absolute in Old Church Slavonic and Old East Slavic. *Die Welt der Slaven* 42. 251–284.
- Eckhoff, Hanne M. & Aleksandrs Berdicevskis. 2015. Linguistics vs. digital editions: The Tromsø Old Russian and OCS Treebank. *Scripta & e-Scripta* 14–15. 9-25.
- Eckhoff, Hanne M. 2018. Quantifying syntactic influence: Word order, possession and definiteness in Old Church Slavonic and Greek. *Diachronic Slavonic Syntax: The Interplay between Internal Development, Language Contact and Metalinguistic Factors*. Berlin/Boston: Mouton De Gruyter. 29-62.
- Haug, Dag. 2012. Open verb-based adjuncts in New Testament Greek and the Latin of the Vulgate. In Cathrine Fabricius-Hansen & Dag Haug (eds.), *Big events and small clauses*, 287-321. Berlin: Mouton de Gruyter.
- MacRobert, Mary. 1986. Foreign, naturalized and native syntax in Old Church Slavonic. *Transactions of the Philological Society* 84(1). 142-66
- Scherrer, Yves, Susanne Mocken & Achim Rabus. 2018. New Developments in Tagging Pre-modern Orthodox Slavic Texts. *Scripta & e-Scripta* 18. 9-34.
- Worth, Dean S. 1994. The dative absolute in the Primary Chronicle: Some observations. *Harvard Ukrainian Studies* 18. 29–46.