French, Polish and Czech converbs: A contrastive corpus-based study

The present study aims at investigating similarities and differences between three non-finite verb forms specialized in converbal function, i.e. in adverbial subordination, as defined in Haspelmath (1995) and Ross (2021): the French *gérondif*, the Polish *imiesłów przysłówkowy* and the Czech *transgressive*. Previous studies have shown that these three forms display important syntactic and semantic similarities (Čermák et al. 2020, Nedjalkov 1995, König & Auwera 1990): they are canonical (strict) converbs, i.e. forms limited to converbal (adverbial) function, they are same-subject, and their semantic interpretation is based on contextual factors and the three forms share almost identical proportions of different meanings, with the basic meaning of accompanying circumstance prevailing. Yet despite such similarities, the three forms differ significantly in frequency: 4,000 ipm in Polish, 1,700 ipm in French and only 200 ipm in Czech (in contemporary fiction, see Nádvorníková, forthcoming). We argue that these differences are due to diachronic distinctions in the evolution of the three forms and we investigate the impact of these differences on their (non)-equivalence in contemporary language.

The Czech converb differs from its counterpart in Polish (and in the other Slavic languages) because of its archaistic morphology, requiring agreement with the subject of the main clause. This particularity is due to a sociolinguistically motivated normative intervention in the 19th century, because of which the Czech transgressive acquired a strong stylistic mark (bookish or even archaistic) and its frequency dropped dramatically during the 20th century (from 4,000 ipm to 200 ipm in contemporary language). By contrast, the Polish converb, whose adverbialization was accepted by the norm, is well attested in contemporary language. However, despite these differences in frequency, the two Slavic forms share a clear morphological delimitation vis-à-vis the present participle forms, displaying full adjectival morphology and limited to adnominal use. This situation contrasts with French, where the *gérondif* and the *participe présent* had undergone categorial blending (see Vangaever 2022), and only during the 18th century the *gérondif* came to be clearly distinguished formally by the adposition *en* (Bourciez 1946). As a result, today the functions of the two forms partially overlap, especially when the present participle is used ad-clausally. Thus, the three forms display important contrasts in *valeur* within the paradigm of adverbial subordination (both finite and non-finite).

We explore the (non-)equivalence of the three forms and the competing constructions in large data from the InterCorp parallel corpus (https://intercorp.korpus.cz/), limited to contemporary fiction. The analysis is conducted on all texts in the language pairs and on manually analyzed samples of 500 occurrences of converbs in the six directions of translation. The analysis of parallel concordances shows that the Czech converb is strongly contested by finite verb forms in coordinate and subordinate clauses (5% converbs vs. 70% finite verbs, in translations both from French and Polish). By contrast, the Polish converb represents 60% of equivalents of both Czech and French converbs, which indicates that converbs are a more important means of adverbial subordination in Polish than in Czech. Finally, the analysis of translations from Slavic into French reveals two types of systemic constraints: (i) The French converb does not convey negative accompanying circumstance and explicit anteriority, but this difference does not have an important impact on the data, since the two meanings represent less than 5% of occurrences of Slavic converbs. (ii) More crucially, the converb (*gérondif*) and the *participe présent* represent each 30% of translation equivalents of Slavic converbs and the data indicates that the distribution of the two forms displays important semantic and syntactic differences: the present participle favours ad-clausal use, semantically only loosely related with the main clause and tending towards a participant-oriented content (see Schultze-Berndt & Himmelmann 2004), whereas the *gérondif* is semantically more integrated into the main clause, with a tendency to an 'event-oriented' content. These results suggest that Slavic converbs have a broader range of uses than the French converb.

References

Bourciez, Édouard. 1946. *Éléments de linguistique romane*. Paris: Librairie C. Klincksieck.

Čermák, František & Alexandr Rosen. 2012. The case of InterCorp, a multilingual parallel corpus. *International Journal of Corpus Linguistics*. 17(3), p. 411–427.

Čermák, Petr et al. 2020. *Complex words, causatives, verbal periphrases and the gerund: Romance languages versus Czech (a parallel corpus-based study)*. Praha: Karolinum. Available at: http://hdl.handle.net/20.500.11956/117388.

Długosz-Kurczabowa, Krystyna & Dubisz, Stanislaw. 2006. *Gramatyka historyczna języka polskieho*. Warszawa: Wydawnictwa Uniwersytetu Warszawskieho.

Dobrovský, Josef. 1819. *Lehrgebäude der Böhmischen Sprache*. Prag: Gottlieb Haase. Available at: https://vokabular.ujc.cas.cz/moduly/mluvnice/digitalni-kopie-detail/DobLehr1819.

Halmøy, Odile. 2003. *Le gérondif en français*. Paris: Ophrys.

Halmøy, Odile. 1997. *Ça va sans dire, mais ça va mieux en le disant – Regard sur la concurrence sans + infinitif/gérondif négatif*. In: Välikangas, O. – Härmä, J. (eds). *Où va le français ?* Amsterdam: De Werelt, p. 27–40.

Haspelmath, Martin. 1995. The converb as a cross-linguistically valid category. In: Haspelmath & König, eds., *Converbs in Cross-Linguistic Perspective. Structure and Meaning of Adverbial Verb Forms – Adverbial Participles, Gerunds*. Berlin: Mouton de Gruyter, p. 1–57.

Havu, Eva & Pierrand, Michel. 2006. Participe présent et co-verbialité. In: Nölke, H. (eds). *Grammatica. Festschrift in honour of Michael Herslund*. Bern: Peter Lang, p. 137–151.

König, Ekkehard & Johan van der Auwera. 1990. Adverbial participles, gerunds and absolute constructions in the languages of Europe. In: Bechert, J. – Bernini, G. – Buridant, C. (eds). *Toward a Typology of European Languages*. Berlin – NY: Mouton de Gruyter, p. 57–95.

Nádvorníková, Olga. 2021. Stylistic normalisation, convergence and cross-linguistic interference in translation: The case of the Czech transgressive. In Bisiada, Mario (ed.), *Empirical Studies in Translation and Discourse*. Berlin: Language Science Press, p. 53-93. Available at: https://langsci-press.org/catalog/book/296

Nádvorníková, Olga. Forthcoming. Český přechodník jako konverbum: Korpusová analýza překladových a nepřekladových textů [The Czech Transgressive as Converb: A corpus-based analysis of translated and non-translated texts]. Praha: Vydavatelství FF UK.

Nedjalkov, Igor V. 1995. Some typological parameters of converbs. In: Martin Haspelmath & Ekkehard König (eds), *Converbs in cross-linguistic perspective. Structure and meaning of adverbial verb forms – Adverbial participles, gerunds*, Berlin: Mouton de Gruyter, p. 97–137.

Ross, Daniel. 2021. *Pseudocoordination, serial verb constructions and multi-verb predicates: The relationship between form and structure*, PhD. Thesis. Urbana-Champaign: University of Illinois at Urbana-Champaign. Available at: http://hdl.handle.net/2142/113888

Schultze-Berndt, Eva & Nikolaus P. Himmelmann. 2004. Depictive secondary predicates in crosslinguistic perspective, *Linguistic Typology* 8(1), p. 59–131.

Vangaever, Jasper. 2021. *Categories under pressure: the gerund and the present participle from Late Latin to Old French*, PhD Thesis. Gent: Universiteit Gent. Available at: https://biblio.ugent.be/publication/8720139