

Dialectal Variation in Word Order in Selkup: A Corpus-Based Syntactic Study and the Role of Russian Influence

Josefina Budzisch (Univierstät Wien)

Selkup is a severely endangered Samoyedic (Uralic) language spoken in Western Siberia along the Ob and Yenisei rivers. It comprises four major dialect groups—Northern, Central, Southern, and Ket Selkup—that differ considerably in phonology, morphology, and lexicon. Despite this internal diversity, grammatical descriptions have generally characterized Selkup as an SOV language. At the same time, several studies report deviations from this pattern, including SVO order in the Central and Southern dialects (Kuznecova et al. 1980: 395; Kuznecova 1999; Kim-Maloney & Kovylin 2018), and OSV order with topical or definite objects (Helinski 1998: 576). These observations raise a broader typological and diachronic question: to what extent is Selkup word order variable across dialects, and how far are recent patterns attributable to sustained contact with Russian (an SVO language)?

This study addresses these questions through a corpus-based comparison of word order in all four Selkup dialect areas. The analysis is grounded in two large spoken corpora: the SLC Corpus (Budzisch et al. 2024; 59,859 tokens) and the INEL Selkup Corpus (Brykina et al. 2021; 81,498 tokens), which together span over a century of recordings. This allows for a rare combination of dialectal breadth, temporal depth, and empirical coverage in the study of Selkup syntax. The methodology integrates quantitative frequency counts with qualitative syntactic diagnostics, focusing on clause type, argument realization, verbal morphology, and negation, thus enabling a fine-grained assessment of basic word order and its information-structural conditioning.

The results support the traditional view that SOV is the unmarked order across the Selkup continuum, but they also reveal considerable dialectal asymmetries. SVO occurs in all dialects but is substantially more frequent in Central and Southern Selkup, while Northern and Ket varieties maintain more consistently SOV patterns. Illustrative corpus examples demonstrate: (1) canonical SOV order in Northern Selkup; (2) productive SVO order in Southern Selkup declaratives; and (3) OSV order in contexts involving highly topical or definite objects. These findings indicate that information structure continues to play a central role in Selkup clause organization, particularly in object fronting, but that the overall distribution of word-order types is not uniform across dialects.

- (1) *Ira qäli-m-ti apsi-ti-q-olam-ni-t.*
old.man Nenets-ACC-3SG food-TR-INF-be.going.to-CO-3SG.O
'The old man started giving food to the Nenets.' (Northern Selkup; INEL:
NEP_1965_HareParka1_flk.019)
- (2) *I:d'e nö:-mba-t pe:q-i-m.*
Itja chase-DUR-3SG.O elk-EP-ACC
Itja is chasing the elk. (Southern Selkup; INEL: TFF_1973_ItjaHuntsElk_flk.002)
- (3) *Tab-i-p qu:-la ma:di-r-gu-za:-dit [...]*
3SG-EP-ACC person-PL ask.for-FRQ-ITER-PST-3PL
'She was asked by people [...]'
(Southern Selkup; SLC: ILP_1981_StonyOldWoman_flk.3)

The observed geographical asymmetry aligns with sociolinguistic and historical factors. Central and Southern communities experienced earlier, more intensive, and more asymmetrical contact with Russian, including schooling, administration, and interethnic settlement patterns (Kuz'mina 1974; Kazakevich 2007). Against this backdrop, the increased prevalence of SVO in these dialects is most plausibly interpreted as a case of contact-induced syntactic convergence, superimposed on pre-existing

information-structural patterns. Northern and Ket dialects, by contrast, show comparatively conservative SOV usage, consistent with their more limited exposure to Russian.

Taken together, the study provides the most comprehensive empirical picture to date of Selkup word-order variation. It demonstrates how dialectal comparison and corpus-based methods can illuminate ongoing syntactic change in highly endangered Uralic languages, and it contributes to broader debates on contact-induced change, the stability of basic word order, and the interaction between information structure and syntax in mixed-contact environments.

References

- Brykina, M., Orlova, S., & Wagner-Nagy, B. (2021). *INEL Selkup Corpus* (Version 2.0). Universität Hamburg. <https://hdl.handle.net/11022/0000-0007-F45A-1>
- Budzisch, J., Behnke, A., & Wagner-Nagy, B. (2024). *Selkup Language Corpus (SLC)* (Version 2.0). Universität Hamburg. <https://hdl.handle.net/11022/0000-0007-FE08-3>
- Helimski, E. (1998). Selkup. In D. Abondolo (Ed.), *The Uralic languages* (pp. 548–579). Routledge.
- Kazakevich, O. (2007). The Northern (Taz–Turukhan) Selkups. In V. Napolskikh, A.-L. Siikala, & M. Hoppál (Eds.), *Selkup mythology (Encyclopaedia of Uralic Mythologies 4)* (pp. 20–26). Akadémiai Kiadó.
- Kim-Maloney, A., & Kovylin, S. (2018). Izmenenie bazovogo porjadka slov v central'nykh dialektach sel'kupskogo jazyka na osnove dannykh konca XX – načala XXI vv. [Change of basic word order in the central dialects of the Selkup language based on data from the late 20th and early 21st centuries]. *Tomsk Journal of Linguistics and Anthropology*, 4, 39–49.
- Kuz'mina, A. (1974). *Grammatika sel'kupskogo jazyka* [Grammar of the Selkup language]. Novosibirsk.
- Kuznecova, A., Helimski, E., & Gruškina, E. (1980). *Očerki po sel'kupskomu jazyku (Tazovskij dialekt)* [Sketches on the Selkup language (Taz dialect)] (Vol. 1). Izdatel'stvo Moskovskogo Gosudarstvennogo Universiteta.
- Kuznecova, A. (1999). Sel'kupskij jazyk s točki zrenija tipologii porjadka slov [The Selkup language from the perspective of word order typology]. In *Tipologija i teorija jazyka. Ot opisanija k ob"jasneniju. K 60-letiju Aleksandra Evgen'eviča Kibrika* [Typology and theory of language: From description to explanation. Festschrift for the 60th birthday of Aleksandr E. Kibrik] (pp. 88–98).