When grammar-internal factors hinder an OV-to-VO shift: the case of Surgut Khanty

Lena Borise[1,2], Katalin Gugán[3], Balázs Surányi[3,4]
([1]LLF, CNRS & Université Paris Cité; [2]University of Göttingen; [3]HUN-REN Hungarian Research Centre for Linguistics; [4]Pázmány Péter Catholic University)

Preview. Language contact is often argued to lead to word order change -- e.g., the OV-to-VO shift. Though the exact mechanism of change is poorly understood, instances of lack of this change, despite the favorable conditions being present (intense contact, dominant bilingualism with a majority VO language, etc.) are worth investigating. In this paper, we present novel data from Surgut Khanty (SKh; Ugric, Uralic) and show how language-internal factors pertaining to information structure (IS) – fusion of thematic roles and IS-properties; prosodic/IS restrictions on the postverbal domain; the role of passivization – govern the distribution of postverbal constituents (PVCs) in an otherwise strictly verb-final language. We further propose that these factors have so far prevented SKh from undergoing an OV-to-VO shift despite intense contact and widespread bilingualism with a dominant VO language, Russian.

Background: SKh. SKh is a severely endangered variety of Eastern Khanty spoken in the Khanty-Mansi Autonomous Okrug in Russia by ca. 3000 people (almost exclusively bilingual with Russian, except for a small number of older speakers; Schön 2017). Since the earliest (19th c.) descriptions, the frequency of PVCs in SKh has stayed stable (found in ca. 7-9% of clauses). Typical PVCs include locatives, purpose infinitives, obliques, and definite subjects; intriguingly, very few postverbal objects have been reported (Asztalos et al. 2017).

Database. To better understanding of the properties of PVCs in SKh, we used a dedicated new database of 265 clauses containing PVCs sourced from a range of spoken narratives and dialogues and including data from 13 speakers (Schön 2024; while the total number of tokens might seem low, this is due to the overall rarity of non-verb-final clauses in SKh, as above). The database consisted of transcribed utterances accompanied by surrounding context, to assess the IS-properties of their components, as well as audio files for prosodic analysis.

Quantitative analysis. Only declarative clauses were considered for the purposes of the analysis; we excluded imperatives (n=3) and questions (n=26), as well as clauses with PVCs represented by afterthoughts (defined as having a preverbal correlate and prosodically separate; n=32) and discourse particles (n=70). This yielded the distribution of PVCs shown in (1); 'Targets' in (1) include purpose infinitives, goals of motion, and detachable directional preverbs.

(1)

Subjects, active clauses	Subjects, passive clauses	*by*-phrases, passive clauses	Direct objects, active clauses	Temporal/Locative adjuncts	Targets
14	**10**	7	1	60	26

We assessed the IS-properties of the argument PVCs by labelling them as Discourse-/Hearer-New/-Old, which yielded a four-way typology (DN, DO, HN, HO), as shown in (2).

(2)

	DN	DO
HN	6	0
HO	0	**26**

The pattern emerging from (1) is that PVCs are mostly represented by subjects, *by*-phrases (LOC-marked agents in SKh), adjuncts, and Targets. While this might seem like a heterogeneous set, we propose that all but the latter can occur as PVCs due to the IS-related properties of SKh, which also explain the prevalence of DO+HO in (2), and the nearly complete lack of object PVCs.

Proposal. In Ugric languages there is a fusion between thematic roles and the IS-status of constituents (henceforth TR-IS fusion): subjects act as topics/given information, and (direct) objects act either as foci/new information or as secondary topics; while similar tendencies are cross-linguistically common, in Ugric languages the TR-IS fusion has been grammaticalized, so

that IS notions play a role in processes that typically target particular TR roles, such as case marking, agreement, and passivization (Nikolaeva 1999; É. Kiss 2019). We propose that the TR-IS fusion (as mediated by its prosodic correlates, described below) also plays an important role in regulating word order in SKh.

First, given the TR-IS fusion, subjects act as given information/topics. The notion of topichood encompasses several distinct sub-types: primary topics, which establish/change the main topic that the sentence is about, including aboutness topics (Reinhart 1981; Givón 1983, a.o.) and contrastive topics (Kuno 1976, Büring 1999, a.o.), both of which stand in opposition with familiarity topics, which refer to the backgrounded, given material that has been part of the discourse but does not constitute the main topic of the sentence (Givón 1983, a.o). Bringing this together with the TR-IS fusion means that subjects in SKh can act as primary topics or as backgrounded material. Distribution-wise, primary topic subjects reside in the left periphery (3), while backgrounded subjects surface postverbally (4). Objects, in turn, represent either narrow foci/new information or secondary topics. Both types occupy the immediately preverbal slot but differ morphosyntactically: topical objects trigger object agreement (objective conjugation) on the verb, (3), while focused ones trigger the use of subjective conjugation, (4); the subject-object type pairs in (3-4) do not necessarily co-occur and are presented that way for the sake of conciseness.

(3) (Misha caught a pike; then…)
 Miːʃɐ sɒːrt jaqə tuːβ-təɣ.
 M. pike home bring-PST.3SG(S).SG(O)
 'Misha brought the pike home.'

(4) *βɛɬi qɒːt ɒːmtəχmən min Ljali-nat.*
 reindeer house build.PST.1DU 1DU L.-with
 'Me and Ljali built a reindeer house.'

With respect to their prosodic realization, primary topics, secondary topics and foci all bear intonational pitch accents (i.e., are accented); backgrounded material is unaccented. This means that accented material in SKh appears to the left of the verb, whereas unaccented elements are confined to the postverbal domain.

We propose that the restriction on postverbal objects stems from the fact that, acting as foci or secondary topics, objects cannot be backgrounded and unaccented. On the other hand, the prevalence of subjects in (1) follows from the fact that subjects can act as backgrounded, familiar material. Accordingly, the prediction is that if a theme can surface as a subject instead of an object, it can surface as a PVC; this prediction is borne out and explains the frequency of passivization for the sake of theme-backgrounding in SKh, (5). Conversely, if the theme becomes a primary topic, it turns into a passive subject that is placed preverbally (or dropped, in case of topic continuation), and the backgrounded *by*-phrase becomes a PVC:

(5) (You will see a cup with blood in the drawer.)
 Mɐːnə noq jiːnʲtʲɬʲəɣti βərət.
 1SG.LOC up drink.PST.PASS.3SG blood.3SG
 'I have drunk her blood.'

(6) (What happened to your eye?)
 Tʲuːtən iːɬə paβɬəmi nuŋnə.
 then out poke. PST.PASS.3SG 2SG.LOC
 'You poked [it] out then.'

Summing up. The proposal above accounts for the high frequency of active and passive subject PVCs, the availability of *by*-phrase PVCs, and the scarcity of object PVCs. With regard to temporal and locative adjuncts, the first one of the two remaining categories in (1), we hypothesize that they do not have an IS-specification, and, as such, may become backgrounded PVCs if warranted by context. Finally, Target PVCs likely require a non-IS-related explanation; we discuss them in the full version of the paper, including in the context of typological evidence (e.g., from West Asian OV languages). The small number of HN+DN tokens in (2) suggests that the process of allowing for non-backgrounded PVCs has begun in SKh, possibly as a result of contact with Russian (cf. É. Kiss 2014 on the decay of the TR-IS fusion and the resulting OV-to-VO shift in Hungarian.)

Selected references. Asztalos, E. K. Gugán & N. Mus. 2017. Uráli VX szórend: nyenyec, hanti és udmurt mondatszerkezeti változatok. *Nyelvelmélet és diakrónia* 3:30–62. **É.Kiss, K**. 2014. *The evolution of functional left peripheries in Hungarian syntax*. OUP. **Nikolaeva, I.** 1999. *Ostyak.* Lincom Europa. **Schön, Zs.** 2017. *Postpositionale Konstruktionen in chantischen Dialekten*. PhD, LMU.