Contrasting indefinite pronouns in Northern Tungusic languages: data from spoken corpora

From typological (Haspelmath 1997, 2013; Beekhuizen et al. 2017; van der Auwera & Van Alsenoy 2018) and contrastive studies (Van Alsenoy & van der Auwera 2015) it is known that languages differ considerably in how they express indefinite pronouns, not only in what formal elements are involved, but also in the way in which the semantic space of indefinites is divided between different forms. However, most studies have focused on negative indefinites in particular rather than on the domain of indefinite pronouns as a whole, nor is there much information about the extent of variation between dialects of one language or between closely related languages. In this study, we investigate how indefinite pronouns are formed in three Northern Tungusic languages spoken in Russia – Evenki (evn), Even (eve), and Negidal (neg) – in order to answer the following questions: Do these languages differ in the markers they use to derive indefinite pronouns? Do they differ in their division of the semantic space? Are there differences at the dialectal level? Can we discern patterns in this variation? We address these questions using oral corpora of 11 different Evenki dialects (Däbritz & Gusev 2021, Kazakevich et al. 2023, Mamontova et al. 2023) belonging to two distinct geographical groups, Western Evenki and Eastern Evenki; two geographically distinct Even dialects, namely Lamunkhin Even from Central Yakutia and Bystraja Even from Kamchatka (Pakendorf et al. 2010); and Negidal (Pakendorf & Aralova 2017), in which we searched for various Russian indefinite pronouns (e.g. кто-нибудь 'somebody, non-referential', кто-то 'somebody, referential'). We classified our results following Haspelmath's (1997) functions, even though they have been criticized for conflating syntactic and semantic functions and for not distinguishing between the meanings of the items themselves and the meanings obtained in context (Denić et al. 2022, van der Auwera & Van Alsenoy 2011). However, we did not distinguish between "specific known to the speaker" and "specific unknown to the speaker" since it is hard to identify the knowledge base of the speaker in narratives (cf. Beekhuizen et al. 2017).

The results show considerable variability in the three languages contrasted here (1-5), both with respect to the markers of indefinite pronouns and to the semantic distinctions they make. Evenki and Even dialects spoken in the western range of these languages' distribution make a two-way distinction between negative and free choice indefinites vs. specific and non-specific indefinites (Table 1). Eastern Evenki dialects and Negidal, in contrast, make a three-way distinction between negative and free choice indefinites vs. specific indefinites vs. non-specific indefinites. The Bystraja dialect of Even spoken on Kamchatka stands out by using the same marker (namely the clitic =da) to express non-specific and negative indefinites while using two different markers to express specific indefinites and free choice indefinites.

In a later step, narrative data from more Tungusic languages and dialects will be included in order to obtain insights into the historical development of this domain. In particular, it will be important to assess the relative contribution of independent innovations vs. contact-induced changes to the variation we observe.

Table 1. Main markers of indefinite pronouns (languages/dialects arranged from West to East)

	W. Evenki	Lamunkhin Even	E. Evenki	Negidal	Bystraja Even
specific	=βal	=gol	=ka / =la	=ka	=βut / =βutta
non-specific	=βal	=gol	=βal	=βal	=da
free-choice	=da	=da	=da	=da	?=βul
negative	=da / =mat	=da	=da	=da	=da

Examples:

(1)   *ŋinaki-r-βun*   **e:-βa=βal**   *amta-ra*
     dog-PL-POSS.1PL.EXCL   what-ACC=INDF   feel-NFUT[3PL]
     'Our dogs felt **something**' (Western Evenki;
     Arch_Lebedeva_1952_Uchami_Bukharev_BearCub_new_a182)

(2)   *tar*   *etiken*   *dọlda-tɕa*   **iak=gol**   *korgo-d-di-βe-n…*
     that   old.man   hear-PST[3SG]   what=INDF   growl-PROG-PRS.PTCP-ACC-3SG
  'That old man heard that **something** was growling...' (Lamunkhin Even; RDA_shatun: 027)

(3)   **e:kun-ma=ka**   *tolkit-t͡ʃa*
     what-ACC=SPEC.INDF   dream-PST[3SG]
     '(He) dreamt about **something**.' (Eastern Evenki; Mengrundya-mata_0)

(4)   *i*   **e:kun-ma=ka**   *ulgutɕa:n-da-j*   *ɲeko-tɕa*   *bi-si-m*
     and.R   what-ACC=SPEC.INDF   tell-PURP-REFL.SG   do-PST.PTCP   AUX-NFUT-1SG
     'And I wanted to tell **something** (more).' (Negidal; AET_grandmother: 123)

(5)   *nan*   *etike:n*   **ia=βutta=si**   *tor-duk*   *ga-βa:t-ta-n…*
     and   old.man   what=SPEC.INDF=EMPH   earth-ABL   take-HAB-NFUT-3SG
     'And the old man took **something** from the ground ....' (Bystraja Even; NAT_vojna_013)

References:

Beekhuizen, Barend, Julia Watson & Suzanne Stevenson. 2017. Semantic Typology and Parallel Corpora: Something about Indefinite Pronouns. *Proceedings of the 39th Annual Conference of the Cognitive Science Society*, 112-117.

Däbritz, Chris Lasse & Gusev, Valentin. 2021. *INEL Evenki Corpus. Version 1.0.* https://hdl.handle.net/11022/0000-0007-F43C-3.

Denić, Milica, Shane Steinert-Threlkeld & Jakub Szymanik. 2022. Indefinite pronouns optimize the simplicity/informativeness trade-off. *Cognitive Science* 46. e13142.

Haspelmath, Martin. 1997. *Indefinite pronouns*. Oxford: Clarendon Press.

Haspelmath, Martin. 2013. Indefinite pronouns. In Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*, (http://wals.info/chapter/46).

Kazakevich, Olga et al. 2023. The Evenki language corpus. https://gisly.net.

Mamontova, Nadezhda et al. 2023. The Evenki language corpus. https://corpora.iea.ras.ru/corpora.

Pakendorf, Brigitte, Dejan Matić, Natalia Aralova & Aleksandra Lavrillier. 2010. https://Hdl.Handle.Net/1839/00-0000-0000-000F-2DFE-0.

Pakendorf, Brigitte & Natalia Aralova. 2017. http://hdl.handle.net/2196/00-0000-0000-000F-E387-1.

Van Alsenoy, Lauren & Johan van der Auwera. 2015. Indefinite pronouns in Uralic languages. In Matti Miestamo, Anne Tamm & Beáta Wagner-Nagy (eds.), *Negation in Uralic languages*, 517–546. Amsterdam, Philadelphia: John Benjamins.

van der Auwera, Johan & Lauren Van Alsenoy. 2011. Mapping indefinites: towards a Neo-Aristotelian map. *Selected papers on theoretical and applied linguistics* 19.

van der Auwera, Johan & Lauren Van Alsenoy. 2018. More ado about nothing: on the typology of negative indefinites. In Ken Turner & Laurence Horn (eds.), *Pragmatics, Truth and Underspecification: Towards an Atlas of Meaning*, 107–146. Leiden: Brill.