FOCUS POSITION IN SOV ~ SVO-VARYING LANGUAGES – EVIDENCE FROM ENETS, NGANASAN, AND DOLGAN

Chris Lasse Däbritz
University of Hamburg

Abstract: It is well known that the basic word order pattern of a language is closely intertwined with the syntactic realization of argument focus constituents. SVO languages exhibit a focus position at the sentence's right periphery, SOV languages exhibit an immediately preverbal focus position. The study at hand examines both the basic word order patterns and the syntactic realization of focus in Enets, Nganasan and Dolgan. The major outcome is that Nganasan and Dolgan are much more flexible with respect to their basic word order pattern and, in consequence, exhibit both an immediately preverbal focus position and a right-peripheral focus position, whilst Enets realizes argument focus constituents almost exclusively immediately preverbally.

1. Introduction

Two maxims are often named for both Samoyedic and Turkic languages: 1) Samoyedic and Turkic languages are SOV[1] languages (Tereščenko 1973: 283ff., Johanson 1998: 57), and 2) according to the first maxim, Samoyedic and Turkic languages exhibit an immediately preverbal focus position (Tereščenko 1973: 284; Johanson 1998: 59). It

---

[1] The abbreviations SOV and SVO refer to the linear order of the main constituents of a sentence (S = subject, O = object, V = verb). Two short comments are in order here: First, *subject* and *object* on the one hand and *verb* on the other hand clearly refer to different levels of the language system, this labeling is, thus, partly inconsistent. Second, not only subjects and objects are examined here, but also adverbials etc., thus, a better label would be SXV and SVX, X standing for a constituent fulfilling any syntactic function. For the sake of better comprehension, the traditional labels SOV and SVO are nevertheless kept.

is widely acknowledged that information structure in general and focus-background-structure in particular interact with word order (Krifka and Musan 2012: 35, Surányi 2016: 422, 429). Moreover, it has been observed since a long time that there is a connection between the basic word order of a language and its information structure in that SVO languages tend to place focused constituents postverbally at the sentence's right periphery, whilst SOV languages tend to place focused constituents immediately preverbally (Dezső 1978: 7–8, Kim 1988: 150).

The languages under discussion here, i.e., Enets[2] and Nganasan (< Northern Samoyedic < Samoyedic < Uralic) as well as Dolgan (< North-Siberian Turkic < Siberian Turkic < Turkic), are traditionally classified as SOV languages, thus showing an immediately preverbal focus position (Tereščenko 1973: 284–285, Artem'ev et al. 2013: 75–76). However, recently it has been shown that Nganasan and Dolgan do not exhibit a rigid SOV word order (Wagner-Nagy 2019: 363, Stapert 2013: 247), whereas Enets is reported to exhibit indeed a rigid SOV word order (Siegl 2013: 361). This immediately leads to the question whether word order variation also leads to variation of the placement of focused constituents, the hypothesis being that Enets shows exclusively an immediately preverbal focus position, whilst in Nganasan and Dolgan postverbally realized focused constituents can be observed.

The aim of this study is to examine the basic word order patterns of the named languages and its impact on coding information structural relations. The main attention will be drawn to the syntactic expression of focus, i.e. in other words, to the syntactic focus position in these languages.

The material used for this study comes from the following corpora: *Nganasan Spoken Language Corpus (NSLC)* (Brykina et al. 2018), *INEL Dolgan Corpus* (Däbritz et al. 2019), *The Digital Corpus of Enets* (Khanina and Shluinsky in prep.). From each corpus, approx. 50 texts with 2 500 to 5 000 sentences/information units and 20 000 tokens were taken and annotated for both topic-comment-structure and focus-background-structure.

---

2   Enets is used here as a cover term for both varieties Forest Enets and Tundra Enets. If differences in the described phenomena are not explicitly stated, both dialects/languages behave similarly.

The study at hand is structured as follows: In Section 2, the theoretical background of this study is provided. Section 3 deals with the basic word order patterns in the three languages under discussion, Section 4 pertains to the focus position. In Section 5, conclusions are drawn and some outlook on upcoming problems and questions is given.

2. Theoretical background

As research on information structure is manifold and diverse, often lacking a clear terminological base, some comments on the theoretical background of this study are in order here. *Information structure* is understood as "[…] the formal expression of the pragmatic structuring of a proposition in a discourse" (Lambrecht 1994: 5). Hence, information structure is an integral part of the language system – the *langue* in the Saussurian sense – and not an idiolectal, optional or even stylistic phenomenon beyond this system. This can be easily proven by an example:

(1)   a. *Who is this woman wearing a blue top?*
      b. *#It's a blue TOP[3], which my mother is wearing.*

Although (1b) is a perfectly grammatical sentence in English, native speakers of English most probably will not accept it as appropriate in the given context. The reason for this is that (1a) asks for the identity of a woman unknown to the producer of this utterance, and uses the clothes of this woman for clarifying which woman in the room – there are assumingly more than one – is meant. The producer of (1b), however, does not answer this question, but assumes the woman's identity – his/her mother – to be given in the context, whereas s/he points out this woman's clothes. An appropriate answer to (1a) would instead be the following:

(2)   *It's my MOther, who is wearing the blue top.*

---

[3]   The use of majuscules indicates the intonatorily "highlighted" syllable of a clause (see below).

If information structure is part of the language system itself, then there have to be rules how to encode it. These rules may be and often are language-specific, in either case the encoding of information structure is, thus, not up to idiolectal and individual variation. Therefore, a solid theory of information structure as well as a coherent terminology is needed.

Within the theoretical framework used here, information structure is assumed to be realized on three levels: 1) topic-comment-structure, 2) focus-background-structure, 3) information status.[4] *Topic* is understood as the part of the sentence which something is predicated about; the *comment* is the predication about the topic. *Focus* is understood as the piece of the information in the sentence that the speaker conceives and marks as important and informative for the hearer; *background* is the piece of information in the sentence that the speaker conceives and marks as less important and informative (Molnár 1991: 58, Junghanns 2002: 13). Surely, it is often difficult to decide which piece of information is most important and informative from the speaker's perspective. But exactly because of this difficulty – which not only the researcher, but also the hearer of an utterance is confronted with – there is a need to mark this piece of information linguistically. Though differently formulated, this understanding of the notion *focus* is in line with other approaches such as Krifka and Musan (2012) and Lambrecht (1994). According to Krifka and Musan (2012: 7) "[f]ocus indicates the presence of alternatives that are relevant for the interpretation of linguistic expressions". That in turn means that one of those alternatives is chosen when being marked as focus, e.g. the color of the speaker's new car in *I bought a RED car*. Here, the speaker probably had already told that s/he had bought a car and now s/he specifies its color. According to Lambrecht (1994: 207) "the focus is what makes an utterance into an assertion". By specifying the color of the car, i.e. linguistically focusing the relevant constituent, the speaker adds new information[5] to the

---

[4]   The idea of making up an information structural trichotomy was presented e.g. by Molnár (1991: 58–70). She distinguishes a topic-comment-structure, a focus-background-structure and a theme-rheme-structure and connects these structures to the levels of communicative structuring in the sense of Bühler (1934).

[5]   Note, that "new information" here relates to the level of semantics, i.e. whether a certain fact/circumstance is known by the speaker, and does not relate to the information status of the relevant referent (see below for the latter).

hearer's knowledge. And in accordance to Grice's (1975: 46) maxims of communication – especially: "Be relevant!" – the speaker will add such information that s/he considers to be relevant and informative.

On the level of information status discourse referents are marked as *new* (new for both discourse and hearer), *accessible*/*inferable* (new in discourse, but old for hearer) and *given* (old for both discourse and hearer) (Prince 1992: 299, Götze et al. 2007: 151). Often, *topic* is associated with givenness in research on information structure, and *focus* is associated with new, not aforementioned information. This is often true, but not necessarily, as the following example shows:

(3)   a. *What are your plans for the evening?*
      b. *I don't know yet.*
      c. *Well, Diana and Mary wanted to go to the cinema. Nick will have a party at home. And Peter hasn't answered my messages yet.*
      d. *And where is Paul going?*
      e. *Paul will probably go to the CInema.*

In (3e) the speaker tells something about Paul, thus, *Paul* is the topic of the sentence. S/he points out that Paul will go the cinema, so *to the cinema* is the focus of the sentence. Both *Paul* and *the cinema* are given in this context, as they are explicitly mentioned before. Hence, a given referent, the cinema, belongs to the comment here, and given information is focused. This shows that the assumption of a three-partite system of information structure is justified.

This study deals with the level of focus-background structure. There are several types of focus (cf. Krifka and Musan 2012: 6–21), only non-contrastive focus (in other terms: denotation focus, information focus) is discussed here. Focus is marked on different levels of the language, the most important being prosody and syntax (Krifka and Musan 2012: 34). On a prosodic level, focus is associated with pitch accent (ibid., Chen 2012: 257–259). On a syntactic level, focus is associated with a focus domain that contains the focal information within the clause (Lambrecht 1994: 214). This implies that the scope of the focus domain can vary, Lambrecht (1994: 223) distinguishes *sentence focus*, *predicate focus* and *argument focus* (in other terms: *wide focus*, *intermediate focus* and *narrow focus* resp.). Roughly spoken, sentence focus scopes over the entire clause, predicate focus scopes over the VP (and possibly

dislocated constituents from it), whereas argument focus scopes over a single constituent smaller than VP, e.g. the subject or object NP.

The distinctions of the scope of focus can heuristically be shown with question-answer-pairs:

(4)   Sentence focus:
    a. *What's up? Why are you so happy?*
    b. [FOC *My brother will marry his girlfriend.*]

(5)   Predicate focus:
    a. *What's with your brother? How is he doing?*
    b. *My brother* [FOC *will marry his girlfriend*].

(6)   Argument focus:
    a. *Whom will your brother marry?*
    b. *My brother will marry* [FOC *his girlfriend*].

Speaking about the *focus position* of a language, mostly the position of argument focus is meant.[6] Hence, the article at hand will examine the position of argument focus in Enets, Nganasan and Dolgan. As this is obviously intertwined with patterns of word order, the basic patterns of word order of these languages will be discussed in the next section.

3. Word order in Enets, Nganasan and Dolgan

As it has been stated already in the introduction, Enets, Nganasan and Dolgan are traditionally analyzed as SOV languages. Sentences with this word order can be found easily for all of them:

(7)   Forest Enets (SiNI_20090719_GoldenerFisch_nar.091[7])
    Modʼ   tɛχɛ   tɔrse   kare-ðo-d   tɔðu-ta-ðʔ.
    1SG   there   such   fish-DST-OBL.2SG   bring-FUT-1SG
    'I will bring there such fish for you.'

---

[6]   Eventually, this can be transferred to cases of predicate and sentence focus as well (cf. Junghanns 2002: 30), but this goes beyond the scope of this paper.

[7]   The references are made as follows: Speaker_DateOfRecording_Title_Genre.NumberOfSentence; the transcription is slightly adapted, e.g. the dental fricative is written as <ð>, not as <z>.

(8)   Tundra Enets (TuSU_20090816_Leben_nar.031)
       Šeɔ-χɔa   sɔjða   eneče?   udʼa-ku-ðɔ-d   teða-da.
       who-EMPH   good   person   meat-DIM-DST-OBL.2SG   bring-FUT.3SG
       'Some good person will bring you some meat.'

(9)   Nganasan (KBD_71_Tent_nar.003–005)
       Təgətə   mənə   ni-mə   kubu-j   nəðə-?ki-?ə.
       then   1SG   woman-1SG   skin-ACC.PL   scrape-RES-AOR.3SG
       '[Then I work during three days. There must be 25 wooden poles.] Then
       my wife scrapes skins.'

(10)   Dolgan (SuAA_20XX_Birth_nar.059)
       Onton   bi͡ar-i-n   ilbij-er.
       then   belly-3SG-ACC   stroke-PRS.3SG
       'Then she [= a midwife] strokes over the belly [of the woman giving
       birth].'

   Whereas Enets is assumed to be a quite rigid SOV language (Siegl
2013: 361[8]), Nganasan (Wagner-Nagy 2019: 363) and Dolgan (Stapert
2013: 247) are reported to show considerable variation of SOV and
SVO. Whilst Wagner-Nagy (2019) does not quantify this observation
for Nganasan, Stapert (2013: 247) states that almost a quarter of transi-
tive sentences in her material show the pattern (S)VO[9]. In the analyzed
material here, instances of (S)VO can easily be found for Nganasan and
Dolgan:

(11)   Nganasan (JSM_080212_Hibula_flkd.091)
       Təti   tahari͡aa   nʼakəlʼi-?ə   ŋəmsu-ðə-mtu,
       that   now   take-AOR.3SG   meat-DST-ACC.3SG
       hu͡antu-ə-bta-?a   təndə-j   ŋana?san-ə-j.
       trickery-VBZ-CAUS-AOR.3SG   that-ACC.PL   human-EP-ACC.PL
       'Now he took his meat, he tricked those people.'

---

[8]   Siegl (2013) only describes Forest Enets in this respect; analyses of the syntax of Tundra
      Enets are hardly available.
[9]   Dolgan, as all the languages under examination, exhibits extensive subject pro-drop,
      therefore S does not have to be overtly realized.

(12)   Dolgan (PoXN_19701118_Chopochuka_flk.005)

kïrsa-lar,   *tûok-tar*   *hîe-bit-ter*   *araj*   *baliɡ-i-n.*
polar.fox-PL   what-PL   eat-PST2-3PL   PTCL   fish-3SG-ACC
'[So he went once,] polar foxes or so have eaten his fish.'

There are also some instances of (S)VO word order in both Enets varieties, however, by far not to the same extent as in Nganasan and Dolgan. Nevertheless, two examples are given here:

(13)   Forest Enets (RoSA_20080824_ImSumpfVersunken_nar.081)

Sɔbarta-š   *šiʔ*   *koli*   *miʔ-χoð.*
drag.out-3SG.PST   1SG.ACC   peat.bog   inside-ABL
'He dragged me out of the peat bog.'

(14)   Tundra Enets (KoIP_20080716_HochzeitTroick_nar.061)

Či,   *mod'ina*   *Irka-baʔ*   *čua*   *Majka*   *nen.*
well   1PL   Irka-1PL   marry.AOR.3SG   Majka   with
'Well, our Irka married Majka.'[10]

In order to quantify these observations, a sample set of approx. 500 transitive sentences was analyzed counting the word order pattern. Only those sentences were counted where the pattern was unambiguous (i.e., sentences with e.g. a direct object before the verb and an indirect object after the verb were excluded). Moreover, sentences with a topical object appearing in the beginning of the sentence were excluded for obvious reasons, too. The following table shows the relative occurrences of (S)OV and (S)VO word order in the analyzed material of the languages under discussion:

Table 1. SOV and SVO in the analyzed languages

Language/Variety	(S)OV	(S)VO
Forest Enets	94.5%	5.5%
Tundra Enets	95.4%	4.6%
Nganasan	83.4%	16.6%
Dolgan	85.7%	14.3%

---

[10]   The wedding of two women is unexpected in the given context. Indeed, the text is not about a real wedding, but about some kind of joke and funny situation in which it seemed to be as if these two women had married.

The table clearly shows that Enets on the one hand and Nganasan and Dolgan on the other hand behave differently with respect to their basic word order. In the following, some thoughts on possible reasons for this shall be presented.

On the first glance one could come up with the following syllogism:

proposition 1:	Nganasan and Dolgan show deviations from their original(?) SOV word order.
proposition 2:	Russian exhibits SVO word order.
proposition 3:	Nganasan and Dolgan are under Russian influence.
conclusion:	The instances of SVO word order in Nganasan and Dolgan can be explained by Russian influence.[11]

This is the most common pattern for explaining phenomena in minority languages spoken in the Russian Federation, which are seen as deviations from a standard pattern. This may be true in several cases, but here the role of Enets as a tertium comparationis cannot be neglected. If the explanation of the Nganasan and Dolgan word order pattern via Russian influence held, then it would have to be explained why Enets behaves differently. The logical conclusion would be either that the analyzed Enets material is "purer" in terms of lacking Russian influence or that there is significantly less Russian influence on Enets than on Nganasan and Dolgan in general.

The former explanation can be rejected easily by comparing the metadata of the three corpora: Whereas all Enets material stems from the 1990s and 2000s, the Nganasan and Dolgan material stems from a longer period of time, reaching back to the 1970s in the Nganasan case and even to the 1930s in the Dolgan case. As the Russian influence can be assumed to have become stronger during the last century, the Enets data should be more influenced by Russian than the Nganasan and Dolgan data. This is indeed the case (the Enets material exhibits e.g. much more code-switching, code-copying etc. than the Nganasan and Dolgan material), wherefore the more rigid SOV pattern in Enets can hardly be explained by lesser Russian influence.

---

[11]   This explanation has been given e.g. by Tereščenko (1973: 295–296) and Artem'ev et al. (2013: 76).

As for the intensity of Russian influence on the three languages, some comments are in order, too. Khanina et al. (2018) showed very convincingly that the original Enets territory was located much more to the south than today's territory. Still in the 17th century some Forest Enets were reported to dwell around the fortress of Mangazeya (Khanina et al. 2018: 112). Forced by migrations of both neighboring people and Russian settlers, the Enetses moved northwards to the territories where they live now (ibid.). Always having been located on the shores of the river Yenisey, at least some contacts with Russians are more than probable. The Nganasans, however, led their (semi-)nomadic life inside the Taimyr Peninsula, being geographically separated from the Russians i.a. by the Enets (Wagner-Nagy 2019: 4–5). Only in the 1960s ~ 1970s the Nganasans moved into the settlements Volochanka and Ust'-Avam, where they were encouraged to live, and started to have more intensive contact with Russians (Wagner-Nagy 2019: 8–9). The Dolgans, finally, take an intermediate position to that extent that their ethnicity arose from the mixture of Tungus (Evenki), Turks (Yakuts), Russian settlers (*zatundrinskie krest'jane)* and few Samoyeds (Enets) in the 18th and 19th centuries (cf. Dolgix 1963 for details). Settling along the Khatanga Trading Way, the Dolgans always had at least some contacts with Russians (ibid.).

Putting these ethno-geographical facts together, one would expect Nganasan being least influenced by Russian, Enets being most influenced by Russian and Dolgan somewhere in between. Be it correct or not in general, in the case under discussion here, the explanation of Russian influence obviously fails. Therefore, one has to state that language-external factors hardly can explain the different word order patterns in Enets on the one hand and Nganasan and Dolgan on the other hand.

Describing the variation of SOV and SVO word order in Nganasan and Dolgan, one has basically two options: Either one of the two patterns is regarded as basic and the other one as derived, or both patterns are regarded as basic. Considering the first option, one would intuitively assume the SOV pattern as basic and the SVO pattern as derived, as the former outnumbers the latter significantly[12]. However, analyzing

---

12   Surely, it has not necessarily to be the case that the more frequent phenomenon is the underived one. But at least from the point of view of language economy this seems to plausible, if no other linguistic facts point to the opposite scenario.

the occurrence of an SVO pattern as a derivation of the original SOV pattern, the question comes up what motivates this derivation. Possible motivations of changes in word order are manifold – definiteness, aforementionedness or focusing/backgrounding of referents only being a couple of them. Comparing the Nganasan examples (9) and (11), e.g., one can see from the wider context that the direct object in (9) is not aforementioned, whereas in (11) it is. Comparing the Dolgan examples (10) and (12), one could assume that the direct object in (10) is part of the (predicate) focus domain, whereas in (12) it is backgrounded (the sentence implicitly answers the question what has happened to fish). However, as for now there seem to be no clear patterns in the analyzed material with respect to these or similar functional motivations of a change of the word order, which surely does not mean that they do not exist.

The other possibility is to assume a flexible SOV ~ SVO word order pattern for these two languages. This would be in line with the observations of Haider (2018, 2020) who analyses verbal phrases not necessarily being strictly head-initial or head-final in one language. Instead, he proposes a flexible position of the head of VP e.g. for Slavic languages (Haider 2018: 26). According to Haider (2018: 10), "flexible position of the head" does not mean arbitrary variation, but the possibility of both head-initial and head-final structures in certain morphosyntactic constructions. However, this approach has a severe shortcoming either: It remains unclear when and how the direction of branching is decided upon the derivation of a clause.

Having discussed the basic word order patterns of the languages under discussion, one can conclude that Enets exhibits a quite rigid SOV word order – not without exception, though – whereas Nganasan and Dolgan exhibit a flexible SOV ~ SVO word order. By now, it cannot be finally stated whether the SOV ~ SVO-varying languages Nganasan and Dolgan are in fact SOV languages with derived SVO structures, or whether they are SOV ~ SVO-flexible languages in the sense of Haider (2018). Further research on this topic is highly desired. In the following section, the above observations will be related to the syntactic realization of argument focus in these languages.

4. Focus position in Enets, Nganasan and Dolgan

As it has been said in the introduction, the basic word order pattern of a language seems to predict the position of argument focus within a clause: SVO word order predicts a postverbal and right-peripheral focus position whereas SOV word order predicts an immediately preverbal focus position (cf. Dezső 1978, Kim 1988). Following the assumption of an SOV ~ SVO flexible type, one would conclude that also the focus position varies in these languages.

Given the fact that all languages under investigation exhibit mostly SOV structures, it is not surprising that instances of immediately pre-verbally realized focused constituents can easily be found:

(15)   Forest Enets (BoAS_20090715_WieManEinLassoBenutzt_nar.059
  Nɛ-ʔ   *ɔbu*   *šeda-ʔ,*   *nɛʔ*   [FOC *pɛɛ*]
  woman-PL   what   make-AOR.3PL   woman-PL   shoe
  šeda-gɔɔ-ʔ.
  make-DUR-AOR.3PL
  'What do women make, women make SHOES.'

(16)   Tundra Enets (TuSU_20090816_UmzugAusVoroncovo_nar.085–087)
  [FOC *Sal'ɔnij-χɔn*]   *a-abi-aʔ.*   [FOC *D'eru-χɔn*]   *a-abi-aʔ.*
       Sol'onyj-LOC   be-HAB-1PL   forest-LOC   be-HAB-1PL
  '[Sometimes we go up the river.] We are [then] in SoL'Onyj. We are in the FOrest.'

(17)   Nganasan (ASS_161024_Life_nar.011–014)
  Təni   [FOC *biiʔ klasə*]   *s'ati-d'iə-m*   *hoðə-tə-sa.*
  there   ten class   finish-PST-1SG   write-IPFV-INF
  '[I studied in school, in Volochanka. Until the eigth class. Then I went to Dudinka.]
  There I finished TEN classes [i.e. the tenth grade], writing [i.e., learning].'

(18)   Dolgan (AnMS_1972_GoodSovietTimes_nar.055)
  Onuga   *ühüs*   *emȋe*   *ủol-um*   [FOC *avijpor-ka*]
  then   third   again   son-PO1SG   airport-DAT/LOC
  üleli:-r.
  work-PRS.3SG
  'Then my third son works at the AIRport.'

In all languages, this pattern is the most common one (see below for numbers). Nevertheless, there are once more severe differences between both Enets varieties on the one hand and Nganasan and Dolgan on the other hand. The latter languages exhibit regular instances of post-verbally realized argument foci, cf. the following examples:

(19)   Nganasan (ChND_041212_Girl_flkd.042)
       ŋanuə   *maðə-Ɂə-Ɂ*   *malə-baða-təɁ*
       real   tent-EP-AUG-PL   build.tent-INFER-3PL.REFL
       [FOC *turku*   *bərə*].
              lake.GEN   edge.LAT
       '[Big tents are standing there.] The tents are apparently standing on the SHORE of the lake.'

(20)   Dolgan (AkEE_1990_PearlBeard_flk.014)
       Context: a man has found an ice hole and wonders what may be inside.
       Ojbon-ton   *bik-pit*   [FOC *uː*   *ičči-te*].
       ice.hole-ABL   lean.out-PST2.3SG   water   master-3SG
       'Out of the ice hole the master of WAter leaned out.'

In both Enets varieties this pattern can be found, but rather exceptionally. Nevertheless, two examples are given here:

(21)   Forest Enets (SiNI_20080823_MärchenLeben_nar.248)
       TɔɁ,   *mu-go-š*   *mu-go-j-ðuɁ*
       well   take-DUR-CVB   take-DUR-PL.OBJ-AOR.3PL.OBJ
       [FOC *Dud'inka-χan*].
              Dudinka-LOC
       'Well, as to buying, they buy it [= fox furs] in DuDINka.'

(22)   Tundra Enets (TuZA_20080723_Life_nar.021)
       kan'e-ðɔɁ   [FOC *ribak-Ɂa*].
       leave-AOR.1SG   fisherman-TRL
       '[Then, when I worked at the collective farm 'Northern',] I became a FIsherman.'

Comparing the position of argument focus in the analyzed material on a quantitative base, the claims made above can be proven.

The following table shows both absolute and relative numbers of the occurrences of preverbal and postverbal focus positions in the analyzed material, together with the figures from Table 1:

Table 2. Word order and focus position in the analyzed languages

Language/ Variety	(S)OV	(S)VO	Preverbal Foci		Postverbal Foci	
Forest Enets	94.5%	5.5%	223	94.1%	14	5.9%
Tundra Enets	95.4%	4.6%	194	96.5%	7	3.5%
both FE and TE	95.1%	4.9%	417	95.2%	21	4.8%
Nganasan	83.4%	16.6%	430	86.7%	66	13.3%
Dolgan	85.7%	14.3%	584	86.2%	93	13.8%

If one compares the numbers of SOV ~ SVO variation to the numbers of preverbally and postverbally realized argument foci, the resemblances are striking: In both Enets varieties the amount of sentences exhibiting SVO word order as well as postverbally realized argument foci mount to ca. 5%; in Nganasan and Dolgan, however, in either case it is ca. 13–16 %. Hence, there is certainly a connection of word order patterns and the focus position. SOV ~ SVO variation of the basic word order pattern obviously leads to variation of the focus position, too.

Finally, these outcomes shall be evaluated against a broader typological and theoretical background. First, Nganasan and Dolgan as SOV ~ SVO flexible languages are typologically by no means uncommon, cf. e.g. Skopeteas (2012: 129) for the Kartvelian language of Georgian or Dum-Tragut (2009: 555) for the Indo-European language of Eastern Armenian. As for Georgian, Skopeteas (2012: 140) also reports a connection of basic word order pattern and focus position which is very similar to the Nganasan and Dolgan case. This leads to the question why these two phenomena (basic word order and focus position) are so closely intertwined.

According to the theory of *Functional Sentence Perspective* a sentence within a discourse aims at contributing to the discourse and is, therefore, oriented towards the part of it that carries the highest degree of *communicative dynamism* (Firbas 1992: 5–6). The communicative dynamism of a constituent of the clause "measures" to what extent this constituent brings the communication forward. According to Firbas

(1992) the communicative dynamism of the constituents of the clause is increasing from the left to the right. This observation leads to such principles as *given before new*, *topic before comment* or *non-important before important*. Within the framework used here, the latter principle is, of course, the relevant one. In other words, focused constituents in the sense of Molnár (1991), Lambrecht (1994) and Junghanns (2002) carry the highest degree of communicative dynamism in a clause and, thus, tend to be realized at the sentence's right periphery. Junghanns (2002: 30) – working on Slavic languages, which are mostly SVO languages – calls this the *Default Principle of Focus Realization* (DPFR). As for SOV languages, Kim (1988: 149) argues that this principle basically works in SOV languages, too, but conflicting with the well-known *verb-final constraint* (VFC). In optimality theory terminology one could say now that VFC outranks DPFR in SOV languages; as the latter is nevertheless active, focused constituents now take the "second-best" right peripheral position, which is then immediately preverbal. That means that the basic principle of focus realization in SVO and SOV languages as well as in SVO ~ SOV varying languages is exactly the same. That in turn explains easily why the correlation of word order patterns and the surface focus position is so close in these languages and, thus, why SVO ~ SOV variation in languages like Nganasan or Dolgan leads to a variation of the focus position, too.

However, this does not tell anything about basic word order in Nganasan and Dolgan so far. Once more, the two scenarios *SVO as derivation from SOV* and *SOV ~ SVO flexibility* shall be discussed here. Given that word order and focus position structurally belongs together, the first scenario is plausible: SVO structures are – for reasons still to be found – derived from underlying SOV structures. In these derived SVO structures, argument focus is then realized postverbally, because there is no more VFC preventing it. The second scenario, however, copes with a problem: If we assume a flexible head of the VP and if we assume that the postverbal, sentence-final position is the best one for argument foci (DPFR), then it is hard to explain why only some 15% of argument foci are realized in that position. This holds especially true, as the VFC cannot be combined with a word order type that is not necessarily verb-final. Therefore, one would need to find an explanation why still the majority of argument foci is realized preverbally, although the VFC cannot apply. That in turn means that DPFR and VFC together with

the realization of argument foci in Nganasan and Dolgan make it more likely that the observed SVO structures are indeed derivations from an underlying SOV structure.

5. Conclusion

The starting point of this study was the observation that focus realizations in Nganasan and Dolgan seemed to be not as straightforward to analyze as e.g. in Enets. Both immediately preverbally and postverbally realized argument foci were observed which surely was unexpected from the traditional point of view that SOV languages do exhibit only a preverbal focus position. However, it could be shown that on the one hand Nganasan and Dolgan exhibit SOV ~ SVO variation (which supports the observations made in Wagner-Nagy (2019) and Stapert (2013)) and that on the other hand both Enets varieties behave clearly – and even in spite of heavy Russian influence – like rigid SOV languages. In what followed, it could be shown also by quantitative means that variation of the basic word order pattern and a flexible focus position are apparently closely intertwined. Finally, it was argued that this is not unexpected regarding the fact that both SVO and SOV languages realize minimally focused constituents according to a principle of right-peripheral focus realization. Whereas in SVO languages this also leads to a surface right-peripheral focus position, in SOV languages this principle is outranked by the verb-final constraint whence minimally focused constituents are realized in the second-best right-peripheral position, namely immediately preverbally. An immediately upcoming question is whether or how verb-initial and object-initial languages fit into this pattern. But as Samoyedic and Turkic languages surely do not show this pattern, this question goes beyond the article at hand and remains a desideratum for further research.

Address

Chris Lasse Däbritz
Universität Hamburg, Institut für Finnougristik/Uralistik
Max-Brauer-Allee 60
D-22765 Hamburg
E-mail: chris.lasse.daebritz@uni-hamburg.de

Abbreviations

1, 2, 3 – 1st, 2nd, 3rd person, ABL – ablative, ACC – accusative, AOR – aorist, AUG – augmentative, CVB – converb, DAT – dative, DIM – diminutive, DST – destinative, DUR – durative, EMPH – emphatic, EP – epenthetic, FUT – future, GEN – genitive, HAB – habitual, INF – infinitive, INFER – inferential, IPFV – imperfective, LAT – lative, LOC – locative, OBJ – objective conjugation, OBL – oblique, PL – plural, PRS – present tense, PST – past tense, PTCL – particle, REFL – reflexive, RES – resultative, SG – singular, TRL – translative, VBZ – verbalizer

References

Artem'ev, Nikolaj M., Tat'jana S. Nazmutdinova, and Žanna P. Spiridonova (2013) *Dolganskij jazyk. 3. Sintaksis.* Sankt-Peterburg: Almaz-Graf.

Brykina, Maria, Valentin Gusev, Sándor Szeverényi, and Beáta Wagner-Nagy (2018) *Nganasan Spoken Language Corpus (NSLC).* Archived in Hamburger Zentrum für Sprachkorpora. Version 0.2. Publication date 2018-06-12. Available online at <http://hdl.handle.net/11022/0000-0007-C6F2-8>. Accessed on 30.03.2019.

Bühler, Karl (1934) *Die Darstellungsfunktion der Sprache.* Jena: Fischer.

Chen, Aoju (2012) "The prosodic investigation of information structure". In Manfred Krifka, and Renate Musan, eds. *The expression of information structure*, 249–286. (The Expression of Cognitive Categories, 5.) Berlin, Boston: De Gruyter.

Däbritz, Chris Lasse, Eugénie Stapert, and Nina Kudryakova (2019) "INEL Dolgan Corpus". Version 1.0. Publication date 2019-08-31. Archived in Hamburger Zentrum für Sprachkorpora. In Beáta Wagner-Nagy, Alexandre Arkhipov, Anne Ferger, Daniel Jettka, and Timm Lehmberg, eds. *The INEL corpora of indigenous Northern Eurasian languages*. Available online at <http://hdl.handle.net/11022/0000-0007-CAE7-1>. Accessed on 30.03.2019.

Dezső, László (1978) "Towards a typology of theme and rheme: SOV languages". In Maria-Elisabeth Conte, Anna Giacalone Ramat, and Paolo Ramat, eds. *Wortstellung und Bedeutung*. Akten des 12. Linguistischen Kolloquiums Pavia 1977. Vol. 1, 3–11. Tübingen: Max Niemeyer.

Dolgix, Boris O. (1963) "Proisxoždenie Dolgan". *Sibirskij Ėtnografičeskij Sbornik* 5, 92–141. Moskva: Izdatel'stvo Akademii Nauk SSSR.

Dum-Tragut, Jasmine (2009) *Armenian: modern Eastern Armenian.* (London Oriental and African language library, 14.) Amsterdam and Philadelphia: John Benjamins.

Firbas, Jan (2002) *Functional sentence perspective in written and spoken communication.* Cambridge: Cambridge University Press.

Götze, Michael, Thomas Weskott, Cornelia Endriss, Ines Fiedler, Stefan Hinterwimmer, Svetlana Petrova, Anne Schwarz, Stavros Skopeteas, and Ruben Stoel (2007) "Information structure". In Stefanie Dipper, Michael Götze, and Stavros Skopeteas, eds. *Information structure in cross-linguistic corpora,* 147–187. (Interdisciplinary Studies on Information Structure, 7.) Available online at <https://publishup.uni-potsdam.de/opus4-ubp/frontdoor/deliver/index/docId/1294/file/ISIS07.pdf>. Accessed on 07.12.2020.

Grice, H. Paul (1975) "Logic and conversation". In Peter Cole and Jerry L. Morgan, eds. *Syntax and semantics.* Vol 3: *Speech acts*, 41–58. New York: Academic Press.

Haider, Hubert (2018) *Phrase-structure based typology*. Preprint. DOI: https://doi.org/10.13140/rg.2.2.34341.76008

Haider, Hubert (2020) "VO/OV-base ordering". In Richard Page and Michael Putnam, eds. *Handbook of Germanic linguistics.* Cambridge: Cambridge University Press. Available online at <https://www.uni-salzburg.at/fileadmin/multimedia/Linguistik/Chapter-16-HH-VO-OV-base-ordering-AE_01.pdf>. Accessed on: 15.03.2018.

Johanson, Lars (1998) "The structure of Turkic". In Éva Á. Csató and Lars Johanson, eds. *The Turkic languages*, 30–66. London: Routledge.

Junghanns, Uwe (2002) *Informationsstrukturierung in slavischen Sprachen: Zur Rekonstruktion in einem syntax-zentrierten Modell der Grammatik.* Habilitation dissertation. University of Leipzig.

Kim, Alan (1988) "Preverbal focusing and type XXIII languages". In Michael Hammond, Edith Moravcsik, and Jessica Wirth, eds. *Studies in syntactic typology*, 147–169. Typological Studies in Languages 17. Amsterdam and Philadelphia: John Benjamins.

Khanina, Olesya, Yuri Koryakov, and Andrey Shluinsky (2018) "Enets in space and time: a case study in linguistic geography". *Finnisch-Ugrische Mitteilungen* 42, 109–135.

Khanina, Olesya and Andrey Shluinsky (in prep.) *The Digital Corpus of Enets.* Unpublished.

Krifka, Manfred and Renate Musan (2012) "Information structure: Overview and linguistic issues". In Manfred Krifka and Renate Musan, eds. *The expression of information structure,* 1–43. (The Expression of Cognitive Categories, 5.) Berlin, Boston: De Gruyter.

Lambrecht, Knud (1994) *Information structure and sentence form: topic, focus, and the mental representation of discourse referents.* (Cambridge Studies in Linguistics, 71.) Cambridge: Cambridge University Press.

Molnár, Valéria (1991) *Das Topik im Deutschen und Ungarischen.* (Lunder germanistische Forschungen, 58.) Stockholm: Almqvist & Wiksell.

Prince, Ellen (1992) "The ZPG letter: subjects, definiteness, and information-status". In Sandra Thompson and William Mann, eds. *Discourse description: diverse analyses of a fund raising text,* 295–325. Amsterdam: Benjamins.

Siegl, Florian (2013) *Materials on Forest Enets, an indigenous language of Northern Siberia.* (Suomalais-Ugrilaisen Seuran Toimituksia, 267.) Helsinki: Suomalais-Ugrilainen Seura.

Skopeteas, Stavros (2012) "The information structure of Georgian". In Manfred Krifka and Renate Musan, eds. *The expression of information structure*, 127–158. (The Expression of Cognitive Categories, 5.) Berlin, Boston: De Gruyter.

Stapert, Eugénie (2013) *Contact-induced change in Dolgan: an investigation into the role of linguistic data for the reconstruction of a people's (pre)history.* Utrecht: LOT.

Surányi, Balázs (2016) "Discourse-configurationality". In Caroline Féry and Shinichiro Ishihara, eds. *The Oxford handbook of information structure*, 422–440. Oxford: Oxford University Press.

Tereščenko, Natal'ja M. (1973) *Sintaksis samodijskix jazykov.* Leningrad: Nauka.

Wagner-Nagy, Beáta (2019) *A grammar of Nganasan.* Leiden: Brill.

Kokkuvõte. Chris Lasse Däbritz: Fookuse asend SOV ~ SVO variatsiooniga keeltes – tõendus eenetsi, nganassaani ja dolgaani keeltest. On üldiselt teada, et ühe keele põhiline sõnajärjestus on tihedas seoses (kitsalt) fookustatud konstituentide süntaksiga. SVO keeltes on (kitsas) fookus realiseeritud lause paremas perifeerias, SOV keeltes on (kitsas) fookus realiseeritud vahetult verbi ees. Selles artiklis uuritakse nii põhilist sõnajärjestust kui ka fookuse süntaksit eenetsi, nganassaani ja dolgaani keeltes. Uurimuse kõige olulisem tulemus on see, et nganassaani ja dolgaani keeltes on põhiline sõnajärjestus tunduvalt paindlikum kui eenetsi keeles. Sellepärast realiseeritakse nganassaani ja dolgaani keeltes (kitsalt) fookustatud konstituendid nii vahetult verbi ees kui

ka lause paremas perifeerias, samas kui eenetsi keeles realiseeritakse nad ainult vahetult verbi ees.

Märksõnad: infostruktuur, fookuse asend, SOV ~ SVO variatsiooniga keeled, eenetsi keel, nganassaani keel, dolgaani keel

Аннотация. Крис Лассе Дэбриц: Позиция фокуса в языках с вариативным порядком слов SOV ~ SVO – данные энецкого, нганасанского и долганского языков. Как известно, базовый порядок слов в языке определяет синтаксическую позицию фокусных аргументов. В языках SVO фокусные составляющие ставятся в конце предложения, в языках SOV – непосредственно перед глаголом. В данной статье рассматриваются базовый порядок слов и позиция фокуса в энецком, нганасанском и долганском языках и показывается, что в нганасанском и долганском языках порядок слов намного более свободен и, соответственно, фокусные аргументы могут располагаться как перед глаголом, так и в конце предложения; а в энецком языке позиция фокуса – почти всегда непосредственно перед глаголом.

Ключевые слова: информационная структура, позиция фокуса, языки с вариативным порядком слов SOV ~ SVO, энецкий язык, нганасанский язык, долганский язык