Information structure in Tundra Yukaghir and typology of focus structures

Elena Maslova, University of Bielefeld and Stanford University

1. Introduction

The morphosyntactic encoding of information structure in Tundra Yukaghir is known for intrinsic interaction between the semantic/syntactic role of NP, on the one hand, and grammatical properties of constructions that signal its information-structure role, on the other. Thus, Tundra Yukaghir provides an interesting test case for Lambrecht's (1994) theory of *focus-structure types*, which is intended to capture universal principles of such interactions. The analysis of Tundra Yukaghir paradigm of focus structures presented here demonstrates that Lambrecht's *subject-oriented* classification of focus-structure types must be enhanced in such a way as to accommodate the special role of *object* in structuring transitive propositions, or, more generally, to take into account cross-linguistic variation in the degree of grammaticalization of subject-topic correlation.

2. The encoding of semantic and pragmatic macro-roles in Yukaghir

The morphosyntactic means of encoding information structure in Tundra Yukaghir are confined to the core of finite clause, that is, to the finite verb (V) and its core arguments, S (sole core participant of intransitive clause), A and O ("active" and "inactive" participants of transitive clause), and distinguishes two grammaticalized pragmatic macro-roles, *Topic* and *Focus*. These terms (with capital initial letters) are intended to refer to *language-specific* grammatical phenomena, which can nonetheless be assumed to instantiate *cross-linguistic types* of information-structuring strategies captured by the theoretical notions of *topic* and *focus* (with small initial letters)[1]. The basic facts underlying this assumption have been described in the earlier works on Yukaghir Focus system (Krejnovič 1958 :36-49, 131-155 ; 1982 :184-216, 232-263 ; Comrie 1992 ; Maslova 1997 ; Mel'čuk

---

1. An appropriate bibliography of literature behind these notions would certainly exceed the space reserved for this paper as a whole ; in view of this, I have limited references to relevant theoretical literature to the absolute minimum.

2001 :306-311) and are only briefly outlined in the following para-graphs ; the next section is devoted to a more detailed discussion of the relation between Yukaghir Topic and Focus and their cross-linguistic counterparts.

The term «Focus» is justified primarily by grammaticalized asso-ciation of Focus role with canonical «focus-presupposition» contexts, where the information about the situation being described is (directly or indirectly) «activated» by the time of utterance or can be viewed as a part of its extra-linguistic context, the referential identity of one par-ticipant being the only unknown piece of information about the situa-tion (Comrie 1992 ; Maslova 1997). More precisely, if *X* is a semantic macro-role involved in the Focus-marking system, then the *X*-Focus construction is obligatory for questions with question words in this role and for (expected) answers to such questions, as well as for utter-ances where *X* is in the focus of contrast (Chafe 1976 :33-38). These contexts correspond to so called «narrow focus», or *argument-focus structure* in Lambrecht's classification (1994 :228-232) ; this usage of Focus constructions is exemplified in (1) and (5c) below.

The use of term «Topic» for the other member of the grammatical opposition under discussion is justified by its association with the op-posite class of discourse environments : if the *X* participant of the situation referred to by a clause *C* is encoded as Topic in the previous finite clause *C'*, it must also be encoded as Topic in *C*, i.e., *continuous topics* obligatorily receive Topic encoding. The requirement that (the referent of) *X* must be a *Topic* of *C'* is essential, since two consecutive finite clauses can have coreferential Foci, as in the following exam-ple[2] :

---

2. The following notations are used in idiomatic translations : parentheses () enclose (the summary) of preceding discourse context, relevant text-external information is given in brackets []. Curly brackets with subscript {}$_{FOC}$ enclose the focus part of the sentence. Examples are taken from text corpus published in (Maslova 2000) ; texts are referred to by roman numbers, sentences within each text, by Arabic numbers. Abbreviations :   0 – submorph ; 1 - first per-son ; 1|2 – Locutor ; 2 - second person ; 3 - third person ; ABL – Ablative ; ACC – Accusative ; AFF – Affirmative ; ANR - Action Nominalizer ; AT - At-tributive ; AUG - Augmentative ; DAT - Dative ; DS - Different-Subject ; EXST - existential verb ; FOC - Focus ; HCR - Hypochoristic ; INFR - Inferential ; INGR - Ingressive ; INTJ - Interjection ; IPFV - Imperfective ; LOC - Locative ; NOM - Nominative ; OF - O-Focus ; ORNT - orientative (postposition) ; OT -

(1) *tung   gode   ta :t   l'e-reng   wal'har-uol*
     this   person   so   EXST-SS :IPFV   half-TRNSF
     wangt'i :-t'-a :-m.   purege+la :her ma :rqa-n
     search-VEN-INGR-3.OT   west+ORNT   one-AT
     qa :li-t'e   gode-k   l'ie-l'el-u-l.
     horrible-AT   person-FOC   EXST-INFR-0-SF
     tang   gol-le   wangt'i :-t'e-mle
     that   person-FOC   search-VEN-OF.3
     «Now this man$_i$ {began to look for a partner}$_{FOC}$. {There was a formidable man$_j$}$_{FOC}$ in the west. He$_i$ went to look for {that man$_j$}$_{•FOC}$» (II : 6-8)

The *syntactic role* of NP is determined by the combination of its semantic and pragmatic macro-roles ; there are no other (voice or voice-like) oppositions which would allow for distinct morphosyntactic encoding of identical propositional structures. Formally, the Focus syntactic roles are distinguished from Topics by two properties : (i) they must be filled by overt NPs for the construction to function as a finite clause[3], and (ii) the Focus NP must precede the finite verb. To put it the other way round, the term Topic subsumes syntactic roles that can be linked to previously established «discourse topics» without overt nominal reference. Thus, the semantic clause core consists of an obligatory «communicative core» (indicated by curly brackets in schemes (2) and (4) below) and, possibly, one or two Topic NPs.

  Morphological means employed to encode the pragmatic role of a core participant differ depending on its semantic macro-role. The pragmatic role of S is signaled by its own case form and by the verb inflection. In the S-Topic (ST) construction, S takes the Nominative case and controls Number and Person suffixes on the verb. In the S-Focus (SF) construction, S takes one of the Focus case forms[4] (see Table 1) and controls only Number agreement. The pragmatic role of

O-Topic ; PL - Plural ; PROG - Progressive ; SF - S-Focus ; SG - Singular ; SS - Same-Subject ; ST - S-Topic ; STAT - Stative ; TRNSF - Transformative ; VEN - Venitive (go in order to).

3. The verb forms used in Focus constructions function as non-finite forms if not preceded by Focus NPs.

4. See (Krejnovič 1982 :232-263) on the distribution of Focus case forms.

S is marked on the verb by suffixes *-j(e)-* (ST) and *–l* (SF). This formal opposition is schematically represented in (2) and illustrated by examples in (3).

(2) <u>Intransitive clause core :</u>
<u>ST :</u> ST-NOM **{V-NUMBER(S)-*j(e)*-PERSON(ST)}**
<u>SF :</u> **{SF-FOC V-NUMBER(S)-*l*}**

(3) a. *uo-gi   me-kelu-Ø-j-Ø*
         child-3SG   AFF-come-SG-ST-3
         «(She was sitting and waiting for her son). Her son
         {came}_{FOC}.» (VI :24)

    b. *ta :t+l'ie-nu-da-ha   ki-n   köde-k   kelu-ngu-l.*
        SO+EXST-PROG-3-DS   two-AT person-FOC come-PL-SF
        «While he was living like that, {two men came}_{FOC}. » (V :6)

The paradigm of transitive information-packaging variants encodes two binary oppositions, A-Topic (AT) vs. A-Focus (AF) and O-Topic (OT) vs. O-Focus (OF) ; there can be no more than one Focus NP per clause. The OF role is encoded by the Focus case, i.e., in the same way as SF ; the AF role is incompatible with the nominal Focus markers and takes *zero* case marking, which consistently differs from the Nominative form only for third-person pronouns (see Table 1). The AT role is encoded by the Nominative case. The case form of OT is governed by Person hierarchy (Locutor >> Non-Locutor) : if A has a higher rank than O, then the latter takes the Nominative form, otherwise, one of the Accusative forms (see Table 1).

Table 1. Case marking of A, S, O

	"Zero" form **A**	Nominative **A, S, O**	Focus **S, O**	Accusative **O**
Pronouns				
1-2 person	*met* "I"	*met*	*met-**ek***	*met-**ul*** *met-**qane***
3 person	*tude* "(s)he"	*tude-l*	*tude-l*	*tude-**hane***
Nouns	*ile* "deer"	*ile*(***-ng***)	*Ile* *ile-**le(-ng)*** *ile-**k***	*ile-**le(-ng)*** *ile-**hane***

The AT participant controls Person suffixes on the verb ; with the exception of 2PL form, the form of Person suffix depends on the pragmatic role of O. The AF verb form contains no Person markers ; as in the case of intransitive clauses, the Number agreement with A is retained independently of its pragmatic role. The paradigm of transitive packaging variants is summarized in (4) and illustrated by examples in (5).

(4) Transitive clause core (Tundra Yukaghir)
<u>AT/OT :</u> AT-NOM OT-NOM/ACC {V-NUMBER(A)-PERSON(AT) :OT}
<u>AT/OF :</u> AT-NOM {OF-FOC V-NUMBER(A)-PERSON(AT) :OF}
<u>AF/OT :</u> OT-NOM/ACC {AF-Ø V-NUMBER(A)-Ø}

(5) a.   *me-pun'i-**m***
     AFF-kill-OT.3
     «(He saw a stag and shot.) He {killed}_{FOC} it. » (X :7)

  b.   *t'a :rt'eqa :n   kelu-l-pe-de   kiejie   tand'e-hat*
     Ch.   come-ANR-PL-3SG before   merchant-ABL
     *ma :rqa-n   kerewe-d-uo-**k***   *pun'-**mele***
     one-AT   cow-AT-child-FOC   kill-OF.3
     «Before they arrived, Charchehan {had killed one calf}_{FOC} of the merchant's. » (V :89)

  c.   *tand'e :-pul   pun'-ngu*
     merchant-PL   kill-PL(AF)
     [A cow was killed.] «{The merchants}_{FOC} killed it.» (V :39)

A noteworthy property of this system (referred to as *quasi-ergativity* in (Krejnovič 1982 :232)) is the identity of Focus forms of S and O, as opposed to A. Moreover, one of the Focus markers, *-le*(*ng*), can also serve as an Accusative marker if both A and O are Non-Locutors (in such sentences, the pragmatic role of O is signaled only on the verb). As will be shown in the next section, this formal alignment is matched by functional affinity between the SF and OF constructions, as opposed to the AF construction.

3. Focus structures

The goal of this section is to clarify the relation between the Topic vs. Focus distinction in Tundra Yukaghir and cross-linguistic categories of *topic* and *focus* that have emerged from analysis of other languages. In accordance with a long-standing grammatical tradition, the notion of *topic* is taken to capture the presumably universal strategy of sentence construal which allows the speaker to present a proposition as a piece of information *about* a previously established discourse referent. The concept of *focus* is used here in its relatively broad sense, as reflected in Lambrecht's definition of focus as "the semantic component of a pragmatically structured proposition whereby the assertion differs from the presupposition" (Lambrecht 1994 :213). A constituent is said to be *in-focus* if it contains a focal component ; the *focus articulation* of a contextualized sentence is the set of its in-focus constituents. A *focus structure* is "a conventional association of a focus meaning with a sentence form" (Lambrecht 1994 :222), i.e., a *construction* with a certain information-structure semantics. The focus structure of a sentence delimits the range of possible focus articulations, but need not determine its focus articulation uniquely, i.e., different focus articulations can be subsumed under the same focus structure.

In this section, Tundra Yukaghir information-packaging options sketched in Section 2 are described as focus structures, i.e., in terms of sets of focus articulations compatible with opposed constructions. The *focus meaning* of a construction is the semantic invariant common for all compatible articulations. It should be noted that each of the relevant constructions can be, in principle, instantiated by distinct linear structures (within the limits outlined in section 2), which can further delimit the range of possible focus readings. However, since all linear variants of a construction *inherit* its focus meaning (in the sense of Construction Grammar), the latter can be described in its own right, and this is what is attempted here.

The following grid classifies nine possible *focus articulations* into three focus types defined by Lambrecht (1994 :221-225), on the one hand, and according to transitivity and the pragmatic role of O, on the other. The borders delimit sets of focus articulations subsumed under the same focus structure in Tundra Yukaghir (see below for some important qualifications).

(6)   Intransitive   Transitive

		O in-focus	O in-presupposition
predicate-focus	I_0 = S $\{V\}_{FOC}$	F_0 = A$\{OV\}_{FOC}$	T_0 = AO $\{V\}_{FOC}$
sentence-focus	I_1 = $\{SV\}_{FOC}$	F_1 = $\{AOV\}_{FOC}$	T_1 = O$\{AV\}_{FOC}$
argument-focus	I_2 = $\{S\}_{FOC}$V	F_2 = A$\{O\}_{FOC}$V	T_2 = O$\{A\}_{FOC}$V

The *intransitive* focus structures fit a well-known cross-linguistic pattern : the SF construction subsumes both *S-in-focus* meanings, S-focus (I_2) and sentence-focus (I_1), whereas the predicate-focus (I_0) structure is represented by the ST construction, i.e., the semantic impact of SF encoding is essentially similar to that of subject accentuation in English, subject inversion in Russian, etc. The *sentence-focus* (or *thetic*) sentences are used to introduce a new referent (see (3b) and the second sentence of (1)) and/or to describe an unexpected event (with S distinct from the current discourse topic) :

(7) *anme   it'uo-da-hane   en'ie-gi   jaqla :q u :-nu-l*
     DP   look.at-3-DS   mother-3   further   go-PROG-SF
     lat'il   wel'i :-reng
     fire   carry-SS :IPFV
     «Suddenly he saw {his mother going ahead of him}$_{FOC}$, carrying firewood on her back. » (I :47)

The unmarked status of predicate-focus structure is manifested by its compatibility with discourse environments where the referent of S does not qualify for the status of «pragmatically available» topic. For example, the ST of (8) is the first and quite unexpected reference to (a specific set of) trees, which are unidentifiable for the addresses :

(8) *aka :-pe-ng,   alhad'a : sa :-pe-ng   me-jewlu :-ngi*
     brother-PL-NOM   extremely   tree-PL-NOM   AFF-nice-PL :ST(3)
     [A : Have you brought our mother's clothes ? B :] «Elder brothers, {(these) trees were so charming...}$_{FOC}$ (I put clothes on every tree.») (I :143)

In such cases, the predicate-focus articulation imposed by the ST construction is in conflict with the discourse-determined activation status of S referent. As in many other languages, such conflicts are determined mainly by constraints on the *propositional* contents of a thetic sentence, which impede the use of SF construction for a broad class of propositions with in-focus predicates. For such propositions, the in-focus status of predicate triggers the predicate-focus structure independently of «topic-worthiness» of S in the given context.

Turning now to transitive clauses, the AF construction unambiguously signals the argument-focus articulation (T_2), that is, it marks the remainder of the clause («open proposition» $q(x, O)$) as its pragmatic presupposition and is available only if $q(x, O)$ can be assumed to be known by the listener(s), as in (5c) above[5]. The OF construction signals *in-focus* status of O, but does not impose any other constraints on the focus interpretation of a sentence and so subsumes *all* focus types distinguished by Lambrecht, argument-focus (F_2 ; see the last clause in (1)), predicate-focus (F_0 ; example (5b)) and sentence-focus (event-reporting) sentences (F_1). The latter articulation is exemplified by sentence (9), which is intended to inform a stranger about the cause of speaker's grief and involves no pragmatic presuppositions :[6]

(9) *ma :rq-uo-d'e mit   uo   korel   bun'i-l-ngin'   l'e-**mle***
    one-STAT-AT 1PL child   ogre   kill-ANR-DAT   EXST-OF.3
    «{An ogre is going to kill our only son}_FOC. » (VI :86).

The AT/OT construction is obligatory if O is construed as topic. In (5a) above, both A and O are continuous topics (T_0 articulation). Example (10) is a topic-comment sentence with O in the topic role and in-focus A (T_1 articulation). The topic of this sentence is the speaker (the S participant of the initial medial clause) ; the A participant (the speaker's uncle) is never mentioned elsewhere in the narrative :

---

5. Sentences (9) and (10) exemplify Topic encoding of in-focus A in the context of in-focus predicate.
6. This sentence can be said to involve a pragmatic presupposition like "the speaker has a son", but this is irrelevant in the present context.

(10) *tang peldudie-ha l'e-l-ha   met   t'umuot'ie wie-n*
  that   old.man-LOC EXST-1|2SG-DS 1SG   uncle   another-AT
  lukunburebe-ha l'e-reng   kode-ngin' tadi-l'el-u-m.
  land-LOC   EXST-SS :IPFV person-DAT give-INFR-0-OT.3
  «While I lived at that old man's, I$_{TOP}$ {was given to somebody in
  marriage by my uncle, who lived in another place}$_{FOC}$.' (IX :11)

On the other hand, the AT/OT construction is grammatically *un-marked*, as witnessed by its compatibility with F_0 articulation, where the O participant is *in-focus*. The following sentence expresses the speaker's hypothesis about the reason why her daughters are absent ; the only possible focus reading is F_0, yet the AT/OT construction is used :

(11) *e,   ma :rqall'eha t'a :j-le   lawi :-t'e-kodi-l'el-nga*
  Intj   together   tea-ACC   drink-VEN-HCR-INFR-PL :OT(3)
` «They {probably went to drink tea together.}$_{FOC}$» (I :346)

However, Topic encoding of focal O is possible only if A is construed as topic and the finite verb is in-focus, i.e., only for F_0 articulation. Thus, the AT/OT construction signals that V is in-focus and (at least) one of the core participants is a topic ; in other words, this is a topic-comment (predicate-focus) structure that is *underspecified* for the semantic role of topic.
 To sum up, Tundra Yukaghir Focus system distinguishes three focus structures with the following grammatical meanings :

Topic-comment (or *predicate-focus*) structure is represented by the ST
 and AT/OT constructions and encodes V as (a part of) comment
 about one of the core participants of the situation.
S/O-in-focus structure (SF and OF constructions) encodes S/O as a
 part of focus domain and is unmarked for the scope of this domain.
A-focus structure (AF construction) encodes only narrow focus on A.

The semantic domains of predicate-focus and S/O-in-focus structures overlap, hence the variation in encoding of the most common focus articulation of transitive clauses (F_0). However, the predicate-focus structure clearly serves as the unmarked (default) option and is at least

twice more frequent than the O-in-focus structure, even though the O-in-focus *meaning* is quite common.

4. Conclusion

In conclusion, it seems interesting to contrast the Yukaghir paradigm of focus structures with the classification of *focus-structure types* proposed by Lambrecht, which are claimed to be manifested by distinct formal categories cross-linguistically. The focus-structure types correspond to three types of focus articulations introduced in scheme (6) above. Apart from the argument-focus structure, which corresponds to the canonical notion of «narrow» focus, these types are defined by Lambrecht in terms of *subject vs. predicate* opposition, where «predicate» corresponds to VP (rather than V).

As I hope to have shown, the Yukaghir paradigm of focus structures shows both remarkable affinities with Lambrecht's system of focus-structure types and significant discrepancies. For example, the Yukaghir topic-comment structure fits Lambrecht's notion of *predicate-focus* structure, but with two important qualifications : (i) the topic of a sentence need not correspond to the subject, but can also be represented by the object ; (ii) this structure is incompatible with the argument-focus reading with O in the focus role (cf. Lambrecht 1994 :297-306). As shown by this example, such discrepancies «disappear» if the subject vs. predicate opposition in the definitions of focus-structure types is replaced with the opposition of *any* core participant and the main verb (plus, possibly, peripheral components). Intuitively, the fact that the unmarked predicate-focus structure in Yukaghir subsumes O-topic sentences along with A-topic sentences is directly related to its lack of passive, and the possibility of OF encoding of transitive thetic sentences is related to the non-existence of VP in Yukaghir. This observation suggests that the «subject-prominence» of Lambrecht's theory, in its present form, effectively confines its applicability to languages where the «default» subject-topic correspondence is *grammaticalized* both syntagmatically (as a well-defined opposition between subject and VP) and paradigmatically (as a paradigm of voice-like constructions).

It seems, then, that the typological validity of Lambrecht's theory of focus-structure types will greatly increase if it is acknowledged that the set of *potential topics* of unmarked predicate-focus structure is language-specific and the focus-structure types are redefined with this

cross-linguistic variable in mind. In this way, the theory will retain its explanatory power with regard to information structure in «subject-prominent» languages, but will also be easily adjustable to other language types.

References

– CHAFE (Wallace), «Givenness, contrasitiveness, definiteness, subjects, topics and point of view.» In : LI (Charles) (ed.) *Subject and Topic*, p. 25-56. New York et al. Academic Press. 1976.

– COMRIE (Bernard), «Focus in Yukagir (Tundra dialect).» In : ARONSON (H.I.) (ed.) *The Non-Slavic Languages of the USSR*, p. 55-70. Chicago Linguistic Society, University of Chicago, 1992.

– KREJNOVIČ (Yeruhim), *Jukagirskij yazyk* (*The Yukaghir Language*). Moscow – Leningrad, Nauka, 1958.

– KREJNOVIČ (Yeruhim), *Issledovanija i materialy po jukagirskomu jazyku* (*Investigations and Materials on the Yukaghir Language*). Leningrad, Nauka, 1982.

– LAMBRECHT (Knud), *Information Structure and Sentence Form*. Cambridge University Press, 1994.

– MASLOVA (Elena), Yukaghir Focus System in a Typological Perspective. *Journal of Pragmatics 27* (1997), p. 457-475.

– MEL'ČUK (Igor), *Communicative Organization in Natural Language. The Semantic-Communicative Structure of Sentences*. Studies in Language Companion Series. Amsterdam – Philadelphia : John Benjamins, 2001.