

# Information structure at odds with discourse factors: evidence from Finno-Ugric differential object marking

Serdobolskaya N., Toldova S.

Russian State University for the Humanities

## 1. Introduction

Studies of information structure (IS) differ as to whether the discourse properties of a given participant are relevant for establishing its IS status (cf. Lambrecht 1995, Givón 1983, Erteshik-Shir 2007, Vallduví 1992). Many languages are known where the same linguistic devices that reflect IS categories are used as discourse markers (e.g. particles), or where their use is influenced by discourse factors. This paper examines the interaction of IS with discourse factors, basing on differential object marking (cf. Bossong 1985) in four Finno-Ugric (FU) languages, Komi-Zyrjan, Udmurt, Mari and Erzya-Mordvin (data collected during fieldwork).

IS terms:

Focus: The semantic component of a pragmatically structured proposition whereby the assertion differs from the presupposition (Lambrecht 1994: 213)

(i) What have you been doing? — [I]<sub>THEME</sub> [’ve sewn [A SHIRT]]<sub>FOCUS</sub>

(ii) Where is Peter? — [Peter]<sub>THEME</sub> [went [TO SAINT-PETERSBURG]]<sub>FOCUS</sub>

Theme: the rest of the sentence. (Cf. theme vs. rheme definition in Prague linguistic school.):

(iii) What have you done with the meet? — [I]<sub>TOPIC</sub> threw out [the meet.]<sub>THEME</sub>  
How are you? — Well, [yesterday I]<sub>THEME</sub> have been to the theatre.

Topic: A referent is interpreted as the topic of a proposition if in a given situation the proposition is construed as being about this referent, i.e. as expressing information which is relevant to and which increases the addressee’s knowledge of this referent. (Lambrecht 1994: 131)

Focus domain: “a shirt” in (i); “to Saint-Petersburg” in (ii) (Lambrecht 1994), cf. focus in (Erteshik-Shir 2007)

Contrast: a choice from a set of alternatives

(iv) Who did you see, Kate or Ann? — I saw KATE.

(v) How are your sisters? — KATE is married, and ANN went to Saint-Petersburg.

## 2. Differential Object Marking in Finno-Ugric languages

### 2.1. The phenomenon

The DO can be encoded in more than one way, e.g. no marking vs. accusative marker:

UDMURT (BESERMYAN DIALECT)

(1) a. So tâl ž’ət-i-z.  
DEM fire set.fire-PRT-3

*He set the fire.*

b. So kâs-i-z tâl pu-ez.  
DEM put.out-PRT-3 fire tree-ACC

*He put out the fire (lit. fire tree).*

In languages of the world the choice of DO encoding is regulated by the following factors: aspect/negation/mood of the verb, animacy, referential properties of the DO, information structure. Cf. Moravcsik 1978, Comrie 1979, Bossong 1985, Muravyova 1992, Aissen 1998, de Hoop 2005, de Swart 2007, Dalrymple, Nikolaeva 2011; see Wickman 1955 for FU languages.

## 2.2. The choice of DO marking in FU languages

In each language under discussion, DO can be with the overt marker or unmarked (1), depending on various factors as lexical properties (some pronouns allow only one type of marking), quantification, aspect, DO animacy, referential properties, IS.

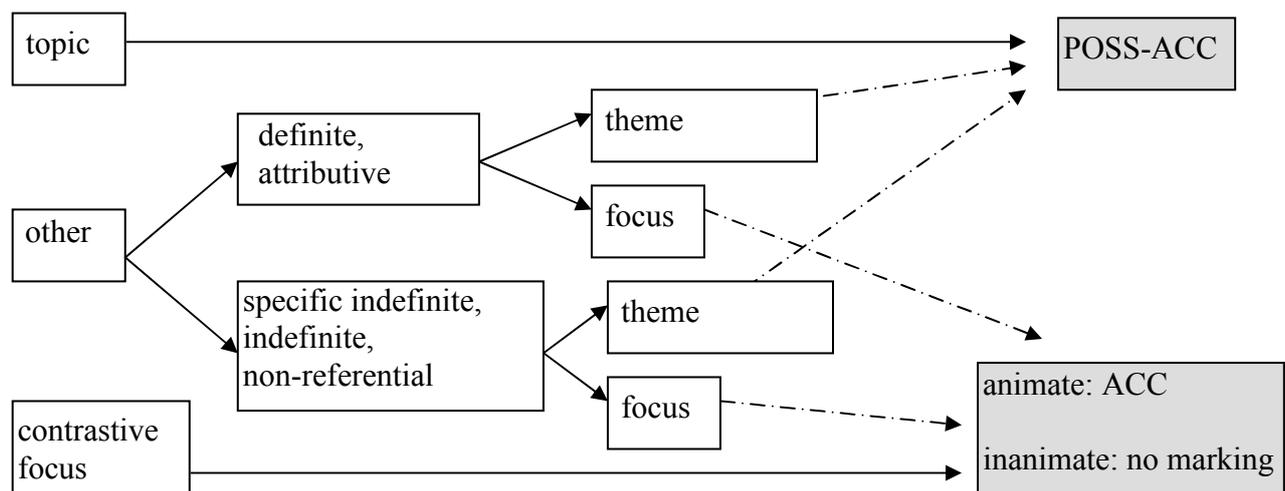
Mari (East Meadow): Ø vs. ACC

Komi-Zyrjan (Pechora) and Udmurt (Besermyan): Ø vs. ACC vs. POSS(1/2/3, Sg/Pl)

Erzya-Mordvin (Shoksha): DO: Ø vs. GEN vs. ABL

Verb: **Indefinite** (Subject) vs. **Definite** (Subject-Object) conjugation

The rules of choosing the DO marking are very complex and can be formulated in algorithms as below, for Komi-Zyrjan:



Most important factors are animacy, referential properties, and information structure. For each language under discussion, these factors pattern in a different way. See Toldova, Serdobolskaya 2012 for details.

- Animacy: animate DOs are more likely to be encoded with a DO marker

KOMI-ZYRJAN (PECHORA)

(2) Me **dərəm** / **dərəm-sə** vur-i.  
 I shirt shirt-ACC.3 sew-PRT.1SG  
*I have sewn a / the shirt.*

(3) Sij-a viž'-ə č'el'ad-əs / č'el'ad-sə / \*č'el'ad'.  
 DEM-NOM keep-PRS.3 children-ACC children-ACC.3 children  
*She nurses children / the / her children.*

- Referential properties
  - definite DOs are marked with the overt DO marker
  - indefinite DOs are unmarked
  - non-referential DOs behave differently in each language

## ERZYA-MORDVIN (SHOKSHA)

(4) a. **Son laz-y-nde ava-n' s'el'maz'e-tn'in'e.**  
 he break-PRT-3PL.O.3SG.S mother-GEN glasses-PL.DEF.GEN

б. \***Son laz-s' ava-n' s'el'maz'a-t.**  
 he break-PRT.3SG mother-GEN glasses-PL

*He broke mother's glasses.*

в. **Son laz-s' s'el'maz'a-t.**  
 he break-PRT.3SG glasses-PL

*He broke glasses.*

(5) **Mon takkam loman' ni-i-n' kardas-ne.**  
 I some man see-PRT-1SG court-TEMP

\***Mon ni-e takkam loman'-t'**  
 I see-PRT.3SG.O.1SG.S some man-DEF.GEN

*I saw a man in the court.*

- IS: for Erzya-Mordvin, Udmurt, and Komi-Zyrjan
  - topical DOs are marked with the overt DO marker
  - focused DOs remain unmarked

## UDMURT (BESERMYAN)

(6) **Č'orâg-ez so ič'i š'i-i-z.**  
 fish-ACC he little eat-PRT-3

*{Looking at the plate with the fish} As for fish, he ate a little {and as for soup, a whole plate}.*

(7) **Š'ij-em-e pot-e mar-e-ke č'ârs. – Š'i jablok!**  
 eat-NZR-1SG come-PRS.3SG what-ACC-INDEF sour eat.IMP apple

*I want to eat something sour. – {Giving an apple.} Eat an apple!*

## KOMI-ZYRJAN (PECHORA)

Basic rule: POSS-ACC marking of topical DOs and DOs in the theme. Else: ACC (for animates) and Ø (for non-animates)

Definite DOs in the topic:

(8) **Te jäv-sə soš'ed-lyš' boš't-i-n? – Me jäv-sə n'əb-i.**  
 you milk-ACC.3 neighbour-GEN2 take-PST-2 I milk-ACC.3 buy-PST.1SG

*Did you take milk from the neighbour? – I bought the milk.*

Indefinite DO in the focus:

(9) **Məj te n'əb-i-n? – Me n'əb-i jäv.**  
 what you buy-PST-2 I buy-PST.1SG milk

*What did you buy? I bought milk.*

Indefinite DO in thethetic sentence:

(10) **Kəny š'əm-ys? -- Me n'əb-i jäv.**  
 where money-P.3 I buy-PST milk

*Where is the money? I bought milk.*

Non-referential DOs: POSS-ACC marking for topical DOs, Ø for DOs in the focus:

- (11) **N'an'-sə** verəs-əj kəsjiš'-i-s vaj-ny, a e-z pi-əj.  
 bread-ACC.3 husband-P.1 promise-PST-3 bring-INF and NEG.PST-3 son-P.1SG  
*As for the bread, my husband promised to buy [it], not my son.*

- (12) Verəs kəsjiš'-i-s **n'an'** vaj-ny a e-z rynək-ti vetlädly-ny.  
 husband promise-PST.3 bread bring-INF and NEG.PST-3 market-PROL hang.about-INF  
*The husband promised to buy bread, and not to hang about the market.*

Unmarked definite DO in the focus:

- (13) Myj te asyv-na-s vəc'-i-n? – **Pač'** lomt-i, **č'el'ad'-əs** vörd-i.  
 what you morning-INESS-P.3 do-PST-2 stove stoke-PST.1SG children-ACC feed-PST.1SG  
*What have you been doing all morning? – I have stoked the stove, fed the children.*

for Mari

- if the whole VP belongs to the same information structure unit (topic domain or focus), then it can be unmarked
- else DO is obligatorily marked with the accusative.

MARI (EAST MEADOW)

- (14) — Pet'a-lan [mo-m əšt-aš] kü-l-eš?  
 Peter-DAT what-ACC do-INF must-PRS.3SG  
 — Pet'a-lan [**pareŋge** erəkt-aš] kü-l-eš.  
 Peter-DAT potato peel-INF must-PRS.3SG  
*What should Peter do? — Peter should peel the potatoes.*

- (15) — Tide materjal gəč' mo-m urg-aš lij-eš?  
 this stuff from what-ACC sew-INF become-PRS.3SG  
 — **Tuvər-əm** / \***tuvər** urg-aš lij-eš.  
 — dress-ACC / dress sew-INF become-PRS.3SG  
*What can we sew from this stuff? — We can sew a dress.*

NB. In all the four languages, word order is also used to encode IS (concrete rules differ for each language). E.g. in Komi-Zyrjan there is a tendency for the focused DOs to appear post-verbally:

KOMI-ZYRJAN (PECHORA)

- (16) a. Te myj ony vəc'-i-n? – Me pəžal-i **papu**.  
 you what morning do-PRT-2 I bake-PRT cake  
 – *What have you been doing this morning? – I've baked a cake.*

- б. Te **papu-sə** magaž'in-yš' n'əb-i-n-ny-d?  
 you cake-ACC.3 shop-EL buy-PRT-2-PL-2?  
 {Context: We are eating a cake.} *Have you bought the cake in the shop?*

### 3. Discourse factors influencing DO encoding

#### 3.1. “Counterexamples” for the rules formulated in section 2

These rules are valid for sentences pairs. If larger discourse is taken, they are overruled by discourse factors. Deviation from the rules:

(A) a new participant in (wide or narrow) focus (expected to be unmarked) is marked:

UDMURT (BESERMYAN)

(17) So žug-iš’k-on-ân zver’-jos kuč’âran-ez vand-i-z-â.

DEM beat-DETR-VN-INSTR animal-PL hawk-ACC cut-PRT-3-PL

{The characters of the tale organized a big fight. Animals fought with birds.}

During this fight an owl has been wounded (lit. the animals wounded a hawk).

{The hawk wanted to fly away, but it could not.}

(B) a participant with definite referential status and belonging to the topical domain (expected to be marked with the overt DO marker) is unmarked.

(18) Leš’t-o tâl, tâl dor-ân so-os šun-ž’ik-o. D’iš’ kwaš’t-o.

make-PRS.3PL fire fire near-INSTR he-PL warm-DETR-PRS.3PL clothes dry-PRS.3PL

{The soldier goes through the forest and sees a man in the swamp. The soldier helps him to get out, take off the clothes, wash himself.} They make a fire and warm themselves. **As for the clothes, they dry them.**

KOMI-ZYRJAN (PECHORA)

(19) [Sumka nu-iš’] Koš’t’a zev muž’-i-s.

bag carry-PTCP.ACT Kostja very get.tired-PRT-3

{The speaker and his friends were out together.} Kostja, who was carrying the bag, got tired.

(20) Kyk mužyk, Fed’a da Val’ent’in, mun-əm-a-aš’ L’agador-ə lepty-ny kerka.

two man Fedja and Valentin go-PF-ATTR-PL Laga near-ILL build-INF house

Two men, Fedja and Valentin, went to the river Laga to build a house.

**Kerka** lept-əm bər-yn kən’ešnə najas t’apn’ityšt-əm-ny.

house build-NZR back-INESS of.course they grab-PF-PL

Having built the house they, certainly, “grabbed” a little (drank alcohol).

In both sentences the DO is definite. However, they remain unmarked. The context suggests that the identification of the referent is not important for the speaker as much as the situation itself. The referent of the DO is not mentioned in the subsequent discourse.

Cf. incorporation in Chukchee:

(21) ji?emittumg-ət Ø-ŋawə-nrat-Ø-gʔat

brother-ABS:PL 3PL.SUBJ-wife-bring-PST-3PL.SUBJ

The brothers brought their wives. (Muravyova 1994: 195)

“The identification of the referent does not make the goal of the speaker” (Muravyova 1994: 195), cf. Mithun 1984.

### 3.2. Experiment / Analysis of texts

Three texts (in Russian) were given to native speakers for translation and free rendering. The texts were annotated on the basis of the following parameters:

- semantic parameters
  - Referential properties of the DO.
  - Information structure of the sentence.
  - Animacy of the DO referent.
- syntactic position of the antecedent
- discourse parameters (Givón 1984):
  - Referential distance – the distance (in clauses) between the present and the previous references of the participant.
  - Topic persistence – the number of times the referent persists as an argument in the subsequent ten clauses following the current clause. (Givón 1984: 908)
  - Number of mentions of the referent.

UDMURT (BESERMYAN)

(22) Mən-o      až'-lan'.      Až'-o      ludkeš'-ez.  
 drive-PRS.3PL    ahead-APPROX    see-PRS.3PL    hare-ACC

*{The fox, the wolf and the bear are driving together in a carriage.} They drive further. They meet a hare. {The hare says: «Please, take me with you ».}*

ERZYA-MORDVIN (SHOKSHA)

(23) Ard-y-t'      dal'se.    Aj-karšuva-sy-z'      numulu-t'.  
 drive-PRT-3PL    farther    IPF-meet-PRT-3.O.3PL.S    hare-DEF.GEN

*{The fox, the wolf and the bear are driving together in a carriage.} They drive further. They meet a hare. {The hare says: «Please, take me with you ».}*

The first “violation of rules” is attested with NPs that have high “degree” of topic continuity / discourse relevance (values of topic persistence, referential distance, and number of mentions). The second “violation of rules” is attested with NPs that have “degree” of topic continuity.

KOMI-ZYRJAN (PECHORA)

(24) A    tany mun-ə    kyš'-iš'      vər kuž'a,    kaž'al-i-s    š'už'-əs  
 and here go-PRS.3SG hunt-PTCP.ACT forest through notice-PRT-3 owl-ACC  
 i    suž'əd-i-s    ružjə-sə.    Až'-i-s    suž'    myj kyjš'-iš'  
 and reach-PRT-3 gun-ACC.3 see-PRT-3 owl that hunt-PTCP.ACT  
 ružjə    p'erj-ə                                    i    šu-ə...  
 gun    take.out-PRS.3SG                            and say-PRS.3SG

*A hunter went by, spotted the owl, and reached for the gun. The owl sees the hunter taking his gun and says...*

MARI (EAST MEADOW)

- (25) Ohotnik pünč'ö vuj-əšte šinč'-əše or'ol-əm už-eš da  
hunter pine branch-INESS sit-PTCP.ACT eagle-ACC see-PRS.3SG and  
**pəč'al-ž-əm** tud-ən ümbak vikt-a, lüj-ne-že. A or'ol-et  
gun-P.3SG-ACC he-ACC on aim-PRS.3SG shoot-DESID-3SG and eagle-EMPH  
ohotnik-ən **pəč'al** vikt-əm-əm už-eš-at,  
hunter-GEN gun aim-NZR-ACC see-PRS.3SG-EMPH  
ajdeme jük de-ne tud-lan ojl-a:  
human voic near-INESS he-DAT say-PRS.3SG

*A hunter went by, spotted the eagle, and pointed the gun at it. The eagle sees the hunter pointing his gun and says with a human voice ...*

Hence, the factor of topic continuity is relevant for the choice of DO marker: it can “overrule” the factors of referential properties and IS.

Another argument: type of anaphoric reference in the subsequent context:

MARI (EAST MEADOW)

- (26) Pet'a-n ala-mogaj **poŋg-əm** pog-əm-əž-əm da vara  
Peter-GEN INDEF-what.kind.of mushroom-ACC collect-NZR-3SG-ACC and then  
**tud-əm** lukt-ən kudalt-əm-əž-əm už-əm.  
he-ACC take.out-CONV throw.away-NZR-3SG-ACC see-NARR.1SG

*I saw Peter collect a lot of mushrooms and then throw them away.*

- (27) Pet'a-n **poŋgo** pog-əm-əž-əm už-əm no vara č'əla  
Peter-GEN mushroom collect-NZR-3SG-ACC see-NARR.1SG but then everything  
lukt-ən kudalt-əš.  
take.out-CONV throw.away-NARR

*I saw Peter collect a lot of mushrooms and then throw everything away.*

#### 4. Discussion

IS and discourse structure

de Swart: 138. Prominence (“umbrella term for... animacy, definiteness and person”) is concerned with the centrality of an entity in the discourse or with the readiness with which an entity presents itself to the speaker as a topic of conversation.

Lambrecht 1995: 114. “The syntactic structure of sentences and the assumed discourse representations of discourse referents correlate with each and this correlation is determined by ... the TOPIC and FOCUS structure of the proposition in which a referent is an argument.”

The first possibility is to analyze the IS in (A) and (B) differently, non-focus in (A) and non-topical domain in (B). This would require to redefine IS categories in such a way as to make them align with discourse properties of the NPs, hence, leading toward a non-universal approach to IS categories.

However, such a decision seems implausible:

I. In FU languages word order is also used to encode IS.

In Komi-Zyrjan, the focused constituent tends to appear post-verbally. The sentence (8), despite the non-expected choice of DO marking, yields to the word order rules: the focused NP is post-verbal. If we redefine IS categories, the word order difference is left unexplained.

II. Discourse factors do not make strong predictions, see variation in tables 1-3.

We suggest that IS is to be analyzed independently from discourse factors. The choice of DO marking is then explained by the rules that work in two steps: first, IS rules are applied, then the rules influenced by discourse factors.

### References

- Aissen J. Differential Object Marking: Iconicity vs. Economy // *Natural Language and Linguistic Theory*. 1998. 21 (3). P. 435—483.
- Bossong G. *Differentielle Objektmarkierung in den neuiranischen Sprachen*. Tübingen: Gunter Narr, 1985.
- Comrie B. Definite and animate direct objects: A natural class // *Linguistica silesiana*. 1979. 3. P. 13—21.
- Dalrymple M., Nikolaeva I. *Objects and Information Structure*. Cambridge: Cambridge University Press, 2011.
- de Swart P. *Cross-linguistic variation in object marking*. Ph.D. diss. Radboud Universiteit Nijmegen, 2007.
- Erteschik-Shir N. 2007. *Information Structure: The Syntax-Discourse Interface*. Oxford: Oxford University Press.
- Givón T. (ed). 1983. *Topic Continuity in Discourse: A Quantitative Cross-Language Study*. Amsterdam: John Benjamins.
- Lambrecht K. 1994. *Information structure and sentence form: Topic, focus, and the mental representation of discourse referents*. Cambridge, UK: Cambridge University Press.
- Moravcsik E. On the Case Marking of Objects // Greenberg J. H. (ed.). *Universals of human language*. V. 4: Syntax. Stanford: Stanford University Press, 1978. P. 249—289.
- Toldova S. Ju., Serdobolskaya N. V. 2012. *Differencirovanoe markirovanie pramogo dopolneniya v finno-ugorskix jazykax*. In *Finno-Ugorskie jazyki: Formalnyj i funkcionalnyj podxody*. Moscow: *Jazyki slavyanskix kultur*.
- Vallduví E. 1992. *The Informational Component*. New York: Garland.
- Wickman B. *The Form of the Object in the Uralic Languages*. Uppsala, 1995.