

Split Noun Phrases in Finno-Ugric languages

Saara Huhmariemi, University of Helsinki
Gisbert Fanselow, University of Potsdam

CIFU XII International Congress for Finno-Ugric Studies
Symposium 6. Syntactic Structure of Uralic Languages
Oulu 2015

1 Introduction

Many or even all Finno-Ugric languages have a split NP construction as exemplified in (1) and (2). In the Split NP construction the noun phrase is divided into two parts. One part typically occupies a position at the left periphery of the finite clause, and the other part typically stays in situ; cf. Finnish (1) and Estonian (2). The construction is often also called “discontinuous noun phrase”, “split topicalization”, “split scrambling”, or “bipartite NP”.

(1) Finnish

Mieh_{iä} saapu*i* paikalle viisi
man:PL:PART arrived:3SG place:to five

‘Five men arrived.’

(2) Estonian

- a. Paju on Peeter näinud kasse.
many be:3SG Peter:NOM see:PART2 cat:PL:PART
‘Peeter has seen many cats.’
- b. Kolm on Peeter näinud kasse.
three:NOM be:3SG Peter:NOM see:PART2 cat:PL:PART
‘Peeter has seen three cats.’

In Finnish and Estonian, the construction comes in two varieties distinguished by the case: partitive (1)-(2) and elative (3). Elative split applies more freely than the partitive split. We concentrate here mainly on the partitive split.

- (3) Tenttikirjoisia minä luin kolme.
exam.book:PL:ELAT:NOM read three
‘I read three of the exam books.’

Split NPs are a frequent construction among languages. In the literature, two basic models have been proposed:

- The two parts are merged independently of each other, but related to the same argument position of the predicate. This was proposed for Warlpiri discontinuous NPs by Hale (1983). The analysis has been applied to many other languages, often coming with the idea that the two NPs are merged very closely to each other, with one of them having to undergo movement to the left periphery for various reasons (e.g. Ott, 2011).

- The two parts are merged as a single NP. Split NPs are the result of subextraction of some nominal projection inside NP, as proposed by van Riemsdijk (1989) for German.

There is evidence that both ways of deriving split NPs occur in natural languages. In German, all kinds of islands for movement can be split (as shown in (4a)), which is hardly compatible with a subextraction analysis. Splitting also occurs when a category different from NP is fronted, cf. VP fronting in (4b). There is no convincing sub-extraction analysis available for this sentence. Likewise, split NPs may come with two overt nouns, as in (4c). There is no single NP source for (4c) either.

- (4) a. Studenten haben ihm viele gesehen
students have him many seen
b. Bücher gelesen hat er keine
books read has he no
c. Vogel mag sie nur Bussarde
birds likes she merely buzzards

In Swedish (or at least for most speakers of Swedish) and Icelandic, the counterparts in (5) are ungrammatical. This suggests the construction involve independent merge of two NPs in German, but subextraction in Swedish.

- (5) a. *Studenter har många läst den här boken
students have many read this book
b. *[VP Läst böcker] har Maria många intressanta
read books has Mary many interesting
c. *Fågl-ar känner han bara till häktergal-en.
Birds knows he only to nightingale

How is the split NP derived in Finnish and Estonian?

2 A-bar movement

2.1 Landing sites

The left periphery of finite clause contains two information structural positions in both Finnish and Estonian. The lower position hosts subjects and topical/discourse-old elements. The higher position is targeted among other by wh-movement.¹ We will assume the structure of left periphery by Vilkuna (1989, 1995), Vainikka (1989) and subsequent work, where the lower subject/topic position is associated with Spec,TP and the higher position is the Spec,CP. The split noun phrases target both of these positions:

- (6) Finnish (Hakulinen et al., 2004, §902)
- a. Vastauksia tuli vajaa 3000.
answers:PAR came less than 3000
‘Less than 3000 answers arrived.’
- b. Hakijolta kutsuttiin haastattelun useita
applicants:PAR invited:PASS interview:to several:PL:PAR
‘Several applicants were invited to the interview.’

- (7) Estonian
- a. Klaase puruses viis.
glass:PL:PART broke five
‘Glasses broke five.’

¹Both positions have language-specific properties, for example, Estonian topic position can host both contrastive and topical elements and the difference is encoded prosodically (Tral, 1990, 35).

- b. Raamatuid ta ostis palju.
book:PL:PART he:NOM buy:PAST many

(8) Finnish

- Kuinka paljon Pekka kutsui vieraita?
how many Pekka invited guest:PL:PART
lit. 'How many did Pekka invite guests?'

(9) Estonian

- Kui palju Peeter sai tooli?
how many Peter:NOM get:PAST chair:PL:PART
lit. 'How many has Peter got chairs?'

Both languages permit also other landing sites for the split constituents. These landing sites are typical in other word order changes related to the information structure.

- (10) a. Pekka on kirjotaja ostanut kolme.
Pekka has book:PL:PART bought three
'As for books, Pekka has bought three.'
b. Mari sai seeni kolm.
Mari:NOM get:PAST mushroom:PL:PART three:NOM

2.2 Reflexive reconstruction

The possessive suffix of the NP reconstructs below the subject in Finnish:

- (11) a. Runojaan Pekka luki tänään kaksi.
poem:PL:PART Pekka read today two
'Pekka read two of his poems today.'
b. *?Hänen runojaan Pekka luki tänään kaksi.
his poem:PL:PART Pekka read today two

The data considered so far show that the element at the left edge has moved to the left periphery. They do not show that the pertinent movement is an instance of sub-extraction from NP. This, however, is suggested by the data in 2.3-2.6.

2.3 Islands

In Estonian and Finnish, split NPs involving a numeral are grammatical only for positions governed by a verb.

2.3.1 Contexts that permit splitting

- Direct objects: see e.g. examples (2) and (8) above.
- Subjects of unaccusatives

- (12) Finnish
a. Michiä juoksi paikalle viisi
man:PL:PART ran:3SG place to five
'Several/five men ran to the place.'

- b. Kuivia särkyi viisi.
picture:PL:PART broke:3SG five
'Five pictures got broken.'

(13) Estonian

- a. Raamatuid ilmus palju.
book:PL:PART appear:PAST many
'Books did many appear.'
b. Raamatuid ilmus kolm.
book:PL:PART appear:PAST three

– Subjects of ECM-constructions

(14) Estonian

- a. Merja nägi lapsi mängimas.
Merja:NOM saw child:PL:PART playing
'Merja saw children playing.'
b. Lapsi nägi Merja mängimas kolme.
child:PL:PART saw Merja playing three:SG:PART

(15) Finnish

- ??Lapsia näki Merja leikkimässä kolme.
child:PL:PART saw Merja playing three.

2.3.2 Contexts that exclude splitting

– Subject condition

- (16) a. Miehet juoksisivat kotiin.
men:NOM ran:3PL home to
'The men ran home.'
b. *Miehet juoksisivat kotiin viisi.
men:NOM ran:3PL home to five
c. *Miehet ovat viisi juosseet kotiin.

(17) Finnish

- a. *Opiskelijat ovat tavanneet hänet viisi.
students:NOM have met him five
b. *Opiskelijat ovat viisi tavanneet hänet.
students:NOM have five met him
c. *Opiskelijotia on viisi tavannut hänet.
students:PL:PART have five met him

(18) Estonian

- a. *Kolm ostsid üllapõlased raamatuid.
three buy:PAST:3PL student:PL:NOM book:PL:PART
'Three bought students a book.'

b. *?Iiioipiased ostsid kolm raamatuid.
 student.PL:NOM buy.PAST:3PL three book.PL:PART
 'Student bought three a book.'

(19) Opiskelijat ovat monet lukeneet kirjan.
 student.PL:NOM be:3PL many.PL:NOM read.PAST:3PL book:SG:ACC
 'Students have many read a book.'

(20) Pajlu vihastavad raamatud iga last
 many annoy:3PL book.PL:NOM every.PART child:PART
 'Many did books annoy everyone.'

–Adjunct islands

(21) a. Pekka luki kirjaa kolme tuntia.
 Pekka.NOM read book.PAR three hour.SG:PAR
 'Pekka was reading a book for three hours.'

b. *Kolme Pekka luki kirjaa tunteja.
 three Pekka.NOM read book.PAR hour.PL:PAR

c. *?Tunteja Pekka luki kirjaa kolme.
 hour.PL:PAR Pekka.NOM read book.PAR three

(22) ?Nädalavahetustel ma olen töötanud kolmel
 weekend.PL:ADE I have worked three:ADE

Again, constructions with quantifiers such as *many* show a different behavior

(23) ?Hotelleissa on Pekka työskennellyt monissa.
 hotel.PL:INE has Pekka worked many.PL:INE

(24) a. Nädalavahetustel ma olen töötanud paljudel
 weekends I have worked many'
 'Weekends I have worked many'
 b. Paljudel ma olen töötanud nädalavahetustel.
 Many I have worked weekends

–DPs in semantic cases

Semantic case DP-complements are islands for movement in Finnish.

(25) Pekka ihastui kirjaan Merjasta.
 Pekka fell.in.love book:ill.of:Merja
 *Kenestä Pekka ihastui kirjaan?

Splitting is also normally not okay

(26) *?Opiskelijoihin on Pekka tutustunut viiteen.
 student.PL:ILL has Pekka get.known five:ILL

But we observe again that the construction with *many*, at least, behaves differently

(27) ?Moihin on Pekka vienyt Merjan kokouksiin.
 many.PL:ILL has Pekka taken Merja meeting.PL:ILL
 'Many has Pekka taken Merja to meetings.'

The restrictions on partitive splitting suggest that partitive splits come about by subextraction.

2.4 Possessive prevents splitting

A genitive DP can occur above the quantifier in ordinary noun phrases, as in (28a). However the genitive DP is not allowed in split NPs (b). The constraint is reminiscent to constraint where the possessive prevents extraction (c).

(28) a. Täällä on Merjan kaksi vanhaa kirjaa.
 here is Merja:GEN two old book
 'Here are Merja's two old books.'

b. *Vanhoja kirjoja täällä on Merjan kaksi.
 old.PL:PART book.PL:PART here is Merja:GEN two

c. *Kenestä Pekka luki Merjan kirjan?
 who of Pekka read Merja's book
 *'Who did Pekka read Merja's book about?'

2.5 Noun doubling is only marginally possible

(29) *?Lemmikkejä Pekka hankki vain koiria.
 pet.PL:PART Pekka buys only dog.PL:PART
 ?Pets Pekka only buys dogs.'

(30) *Linde ta tunneb ainult väikseid linde.
 bird.PL:PART he:NOM know:3SG only small.PL:PART *bird
 ?Birds he only knows small birds.'

The absence of noun doubling is explained if partitive splits involve subextraction

2.6 VP-fronting is not generally accepted

The acceptability of VP-fronting varied extensively between informants of Finnish. But with the exception of one liberal speaker, none of our Finnish informants found splitting in VP fronting contexts really acceptable.

(31) a. ??[Ostanut autoja] hän on kolme.
 bought car.PL:PART s/he has three
 ?Bought cars she has three.'

b. ?[Autoja ostanut] hän on kolme.
 car.PL:PART bought s/he has three.

VP-fronting is reported possible in Estonian, for some but not for all speakers:

(32) raamatuid lugenud on ta pajlu
 book.PL:PART read:PART2 be:3SG he:NOM many
 'Books read he has many'

(33) a. Raamatuid lugeeda ta tahab kolm.
 books read he will three

- b. Lugeða raamatuid ta tahab kolm.
read books he will three

Preliminary conclusion: the island facts and the doubling data make it likely that partitive splits involve subextraction. The VP fronting data are less clear, but go in the same direction.

3 Variation between numerals and other quantifiers

Finnish and Estonian quantifiers permit splitting to a varying extent. As a preliminary generalization, quantifiers that take singular NP-complement, such as the numeral in (34a), are more restrictive than quantifiers that take plural complement (or both plural and singular complement). One such example is Finnish *useita* 'several' in (34b).

(34) Finnish

- a. Pekka osti viisi kirjaa.
Pekka bought five book:SG:PART
'Pekka bought five books.'
b. Pekka osti useita kirjoja.
Pekka bought several:PL:PART book:PL:PART
'Pekka bought several books.'

(35) Estonian

- a. Mari sai kolm seent.
Mari:NOM get:PAST three:NOM mushroom:PART:SING
'Mari got three mushrooms.'
b. Ta on lugenud palju raamatuid
he:NOM be:3SG read:PART2 many book:PL:PART
'He has read many books.'

This suggests that Finnish and Estonian have not only one, but two strategies for forming split noun phrases.

4 Problems of the subextraction account

Split NPs such as (3) suggest that the two phrases are generated independently because the morphological shape of the left part differs from what it would have to be in a continuous NP, cf. also Hakulinen & Karlsson (1979), Vilkuna (1996). In both Estonian and Finnish, a numeral requires a singular complement:

- (36) need kolm väikest tüdrukut
this:PL:NOM three little:SG:PART girl:SG:PART
'these three little girls'

In the split NP, the partitive NP always appears in plural.

(37)

- a. Joutsenia voi nähdä jopa viisi.
swan:PL:PART can see even five
'Swans you can see even five.'
b. *Stella voi nähdä jopa viisi joutsenia.
there can see even five swan:PL:PART
Intended: 'There you can see even five swans.'

(Vilkuna, 1996, 224)

- c. *Joutsenia voi nähdä jopa viisi.
swan:SG:PART can see even five

Such mismatches of morphological shapes occur in the split constructions of many languages. They were used as arguments for the independent merger of two NPs in the split construction (cf. Harder (1985) and subsequent work), but the argument is not compelling if

- (a) plurality is the semantic default of a noun that is quantified over, cf.

(38) I will buy 1.0 books

- (b) certain quantifiers may govern different choices for N's which they c-command but
(c) morphological shape determination follows movement, so that there are no reconstruction effects.

The split noun phrase can involve a determiner pronoun, as in (39a), or a full DP, as in (b) (see also Hakulinen & Karlsson, 1979). These elements do not reconstruct (c). It is not clear whether this constituent is sub-extracted from the DP (d).

- (39) a. Näitä Pekka osti viisi.
these:PART Pekka:NOM bought five
'Pekka bought five of these.'
b. Näitä usia kirjoja Pekka osti viisi.
these:PART new:PL:PART book:PL:PART Pekka:NOM bought five
'Of these new books, Pekka bought five.'
c. *Pekka osti näitä viisi usia kirjoja.
Pekka:NOM bought these:PART five new:PL:PART book:PL:PART
d. ?*Pekka osti viisi näitä usia kirjoja.
Pekka bought five these:PART new:PL:PART book:PL:PART

If (39d) is the base-generated form, it supports the subextraction account. The quantifier *paljon* 'lot' has similar distribution regarding NP-split, and it takes the partitive plural NP as its complement.

- (40) a. Pekka osti paljon näitä kirjoja.
Pekka bought lot these:PL:PART book:PL:PART
'Pekka bought a lot of these books.'
b. Näitä kirjoja Pekka osti paljon.
this:PL:PART book:PL:PART Pekka bought lot

5 Some speculations concerning the status of split NPs.

Languages with scrambling tend to have split NPs of the German type, in which two NPs are merged independently (German, Yiddish, Russian, Serbian, Polish, Hindi, Bengali, Japanese).

Languages with strict word order have split NPs of the Swedish type, arising by subextraction (Swedish, Spanish)

Why does Finnish not pattern with German?
Possible answer: Finnish has no A-scrambling.

In German, objects may precede a subject in TP even when they are not contrastive.
In Finnish, objects may precede a subject in TP only when contrastive.

In German, A-scrambling leads to the absence of superiority effects in multiple questions. In Finnish, we find superiority effects.

→ Finnish does not have A-scrambling → patterns with Swedish.

6 Other Finno-Ugric languages

The grammatically estimations of the languages presented here are based on data collected in the project Discontinuous Noun and Prepositional Phrases in the University of Potsdam. Finnish data was tested altogether with twelve informants, Estonian and Udmurt with three and Komi was tested with one informant.²

6.1 Other Finno-Ugric languages

Udmurt and Komi have split noun phrases, but the speaker variation is extensive. All informants of Udmurt accepted split objects and subjects of certain single-argument verbs. For two of the informants, also other types of splits were possible.

- (41) Udmurt
- a. So vit' pun'y bas'ti-i-z.
(s)he 5 spoon.ACC buy-1PST-3SG
Spoons she has bought twelve (but she has not bought any forks).
 - b. So bas'ti-i-z vit' pun'y.
(s)he buy-1PST-3SG 5 spoon.ACC
- (42) Udmurt
- a. Pun'y so tros bas'ti-i-z.
spoon.ACC (s)he many buy-1PST-3SG
 - b. Pun'y so bas'ti-i-z tros.
spoon.ACC (s)he buy-1PST-3SG many
Spoons she has bought many (but she has not bought any forks).
- (43) Udmurt
- a. Kr'iga tros pot-i-z.
book many go out-1PST-3SG
 - b. Kr'iga pot-i-z tros.
book go out-1PST-3SG many
Books have appeared many/three recently.
- (44) Komi
- a. Ul'bs-s6 Merja njöbis kuimös.
chair-P3acc Merja buy-pret3sg three.acc
'Chairs Merja bought three (but tables she bought one).'
 - b. Kuimös Merja njöbis ul'bs-s6.
three Merja bought chairs

²The data was collected by the help of Anja Arnold, Antje Saueremann, Birte Klaas-Lang, Kaja Kohler, Nale Salveste, Merja Sorsakivi, John Niemi, Jussi Niemi, Deryabin Alexander Stepanovich, Irina Reshneikova, Orosya Tanazos and Svetlana Edygarova, among others. Thank you also to Rigmā Ajanki, Lene Antonsen, Pauli Brattico, Tommi Gröndahl, Paula Kokkonen, Galina Misharina, Jukka Furma, Iona Rauhala, Taina Saikkonen and Anne Vainikka. The work by the first author was funded by Kone Foundation.

c. Kuim Merja njöbis ul'ös.

North Saami split noun phrases are similar to Estonian and Finnish in that the split constituents do not reconstruct. The structure where there is NP-internal agreement between the numeral and the NP cannot be split (a-b), whereas in the split construction, the NP has to be in the object predicative.

- (45) a. Mähite oasá golbma girji.
Mähite buys three book:SG:GEN
b. *?Golbma Mähite oasá girji.
c. Golbma Mähite oasá girjiid.
three Mähite buys book:PL:ACC

References

- Haider, Hubert. 1985. Über sein oder nicht sein: zur grammatik des pronomens sich. In Werner Abraham (ed.), *Erklärende syntax des Deutschen [studien zur Deutschen grammatik 25]*, 223–254. Tübingen: Narr.
- Hakulinen, Auli & Fred Karlsson. 1979. *Mýhysuomen lauseoppi*. Helsinki: Finnish Literature Society.
- Hakulinen, Auli, Maria Vilkuna, Riitta Korhonen, Vesa Koivisto, Tarja Riitta Heimonen & Irja Alho (eds.). 2004. *Iso suomen kielioppi*. Helsinki: Finnish Literature Society.
- Hale, Kenneth. 1983. Warlpiri and the grammar of non-configurational languages. *Natural Language and Linguistic Theory* 1. 5–48.
- Ott, Dennis. 2011. *Local instability: The syntax of split topics*. Harvard University dissertation.
- Tael, Kaja. 1990. *An approach to word order problems in estonian*.
Vainikka, Anne. 1989. *Deriving syntactic representations in Finnish*. University of Massachusetts Amherst dissertation.
- van Riemsdijk, Henk. 1989. Movement and regeneration. In Paola Benincà (ed.), *Dialect variation and the theory of grammar*, 105–136. Dordrecht: Foris.
- Vilkuna, Maria. 1989. *Free word order in Finnish: Its syntax and discourse functions*. Helsinki: Finnish Literature Society.
- Vilkuna, Maria. 1995. Discourse configurationality in Finnish. In Katalin É. Kiss (ed.), *Discourse configurational languages*, 244–268. Oxford/New York: Oxford University Press.
- Vilkuna, Maria. 1996. *Suomen lauseopin perusteet* (Kotimaisten kielten tutkimuskeskuksen julkaisu 90). Helsinki: Edita.