Finnish left peripheral clitics

Pauli Brattico

November 2024

1 Introduction

Finnish (a Finno-Ugric language) has a set of left peripheral enclitic particles which express yes/no questions (=*kO*), familiarity (=*hAn*), affirmative emphasis (=*pA*), de-emphasis (=*s*) and their combinations (e.g., =*kO*=*hAn*, =*pA*=*s*). We review their morphological, syntactic and sematic properties. The material is organized into three sections: this section provides an overview, Section 2 addressed a few special topics that have gained more attention in the literature, while Section 3 presents a summary analysis I think best describes the clitics as they are understood today and discusses a few alternatives. In addition to the left peripheral clitics, Finnish has a few free clitic particles (focus particles =*kin*/*kAAn* 'also/even', and =*kA*) which are not associated overtly with the left periphery and are not discussed in this article. Furthermore, other closely related Finno-Ugric languages have the same or very similar left peripheral clitics that will not be reviewed here; see (Nevis, 1990) for a review.

1.1 Syntax

Finnish is a discourse-configurational language with substantial freedom in word order that articulates information structural notions such as topic and focus. The Finnish left periphery has, however, a rigid syntactic structure. It contains a mandatory subject/topic just above the finite element (=finite main verb, auxiliary or auxiliary-like negation) and a higher C-domain which contains optional Ā-operators such as wh-pronouns and contrastive focus/topic operators (Vilkuna, 1989, 1995). In most analyses these two fields are taken to correspond to

the TP- and CP-projections of the standard generate analysis (Vilkuna, 1995)(1), a useful simplification we adopt here.[1] The unmarked word order is SVO.

(1) [$_{CP}$ Mitä$_1$ [$_{TP}$ Pekka [maala-si __$_1$]].
what.PAR Pekka.NOM paint-PST.3SG
'What did Pekka paint?'

The left peripheral clitics belong uncontroversially to the C-layer: whenever some element hosts one of the clitics, it must occur in the CP above the subject/topic (if any).[2] Moreover, we can discern two types of situations in which some element in the CP "hosts" a clitic: either it is attached to a verbal element that gets displaced to C^0 by head movement (2)a, or it is attached to a hosting element which pied-pipes a larger phrase to SpecCP (2)b–c.

(2) a. Maala-si=*ko*$_1$ Pekka __$_1$ tätä aita-a?
paint-PST.3SG=Q Pekka.NOM this.PAR fence-PAR
'Did Pekka paint this fence?'

b. [Tätä=*kö* aita-a]$_1$ Pekka maala-si __$_1$?
this.PAR=Q fence-PAR Pekka.NOM paint-PST.3SG
'Was it this (and not that) fence that Pekka painted?'

c. [Tätä aita-a=*ko*]$_1$ Pekka maala-si __$_1$?
this.PAR fence-PAR=Q Pekka.NOM paint-PST.3SG
'Was it this fence (and not the gate/that gate) that Pekka painted?'

---

[1] In some analysis the TP is dominated by a higher FinP projection (Holmberg et al., 1993; Holmberg & Nikanne, 2002), which will then become the host for the surface subject/topic. The matter is irrelevant here.

[2] The Finnish left peripheral clitics are all bound morphemes and do not have free form alternatives.

As a first approximation we can assume that these clitics may be suffixed to any word inside the pied-piping phrase, such as demonstratives (2)b, noun heads (2)c, adjectives (3)a, numerals (3)b, quantifiers (3)c, possessives (3)d, participle adjectives (3)e and semantically marked nouns (3)f (the list is not exhaustive).

(3) a. [Sitä vanha-a=*ko* aita-a]$_1$ Pekka maala-si __$_1$?
that.PAR old-PAR=Q fence-PAR Pekka.NOM paint-PST.3SG
'Was it that old (and not new) fence that Pekka painted?'

b. [Näitä kolme-a=*ko* laulu-a]$_1$ Pekka harjoitte-li __$_1$?
these.PAR three-PAR=Q song-PAR Pekka.NOM practice-PST.3SG
'Was it these three (and not two) songs that Pekka pracised?'

c. [Näitä kaikki-a=*ko* laulu-ja]$_1$ meidän pitää harjoitella __$_1$?
these.PAR all-PAR=Q song-PL.PAR we.GEN must practice?
'Was it all these songs (and not just one) that we must pracise?'

d. [Sitä naapuri-n=*ko* aita-a]$_1$ Pekka maala-si __$_1$?
that.PAR neighbor-GEN=Q fence-PAR Pekka.NOM paint-PST.3SG
'Was it that neighbor's (and not our fence) fence that Pekka painted?'

e. [Kaatu-nut-ta=*ko* aita-a]$_1$ Pekka maala-si __$_1$?
fall-PST.PRTC-PAR=Q fence-PAR Pekka.nom paint-PST.3SG
'Was it the fallen (and not the standing) fence that Pekka painted?'

f. [sitä auto-ssa=*ko* ole-va-a vikaa]$_1$ Pekka korja-si __$_1$?
that.PAR car-INE=Q be-PRTC problem Pekka.NOM fix-PST.3SG
'Was it that problem in the car (and not in the bike) that Pekka fixed?'

No other element may occur in the C-field in (2)-(3), and it is not possible to apply the head movement strategy (2)a and the phrasal movement strategy (2)b–c simultaneously in the

same sentence. If the C-field is occupied by another Ā-operator, such as a *wh*-interrogative or contrastive topic/focus phrase, separate hosts carrying the left peripheral clitics cannot be present (4)a, although it is possible to combine the various Ā-operators just in they case have the *same* host (4)b.

(4) a. *[*Mitä* aita-a]$_1$ maala-si=*han*$_2$ Pekka __$_2$ __$_1$?
what.PAR fence-PAR paint-PST.3SG=FAM Pekka.NOM
SpecCP C^0
Intended: 'Which fence did Pekka paint?'

b. [*Mitä=hän* aita-a] Pekka maala-si __?
what.PAR=FAM fence-PAR Pekka.NOM paint-PST.3SG
'Which fence did Pekka paint?'

In sum, Finnish CP has room for one element, and one group of elements that can reserve that position are those which host the left peripheral clitics.

In (2) and (3), the host together with possible pied-piped material is moved into the CP from the clause below.[3] There is a corresponding gap in the sentence. The hosts cannot remain *in situ* without producing ungrammaticality or marginality (5), which means that the clitics are normally hosted by elements located in the left periphery – hence the term "left peripheral clitics."[4]

[3] The only possible exceptions are some adverbs that might be analyzed as merged directly to the CP-domain, but even here the facts are controversial. See (Brattico & Holmberg, 2025: §3.5).

[4] Sometimes also "first position clitics" referring to the fact that the hosts are always the first elements in the clause, but more frequently "second position clitics" if we count the high complementizers that can take a still higher position (*että tämä-n aida-n=ko*…'that this-ACC fence-ACC-Q…') or if we assume that the clitic is counted as the second element after its own host. Both characterizations fits poorly with the Finnish facts, however, and the clitic may

(5) *Pekka maala-si tätä aitaa=*ko*?

Pekka.NOM paint-PST.3SG this.PAR fence-PAR=Q

Intended: ‘Did Pekka paint this fence?

It is uncontroversial that in the case of phrasal movement the operation creates Ā-chains. The same applies also to wh-interrogatives and contrastive topic/focus constituents that appear inside CP (Huhmarniemi, 2012). Thus, the left peripheral clitic hosts are a subset of a larger category I call *Ā-operators* which include also wh-interrogatives, contrastive topic/focus elements and relative pronouns (Brattico et al., 2013). The situation with head chains in sentences such as (2)a is less obvious, and the status of head movement is controversial overall (Dékány, 2018), but in this particular case there will not do great injustice to the facts if we assume that the sentence is produced by T-to-C movement, also Aux-to-C (6)a and Neg-to-C (6)b depending which word is the highest (Finnish sentential negation takes a high position in the clause just below C, see (Holmberg et al., 1993)).

(6) a. on=*ko* Pekka maalan-nut aita-a?

be.PRS.3SG=Q Pekka.NOM paint-PST.PRTC.SGfence-PAR

’Has Pekka painted the fence?’

b. ei=*kö* Pekka ole maalan-nut aita-a?

not.3SG=Q Pekka be.PRS paint-PST.PRTC.SGfence-PAR

‘Has Pekka not painted the fence?’

Several clitics can be stacked into one host (7)a, although some combinations as well as all repetitions are ungrammatical. The individual clitics contribute to the overall interpretation,

---

also get attached to a constituent inside the pied-piped host phrase (3). The term “left peripheral clitics” seems to capture the overall profile and will be used here.

in some cases in a fairly linear way. However, while stacking is possible, scattering is not (7)b.

(7) a. [Tämä-n=*kö*=*hän* aida-n] Pekka maala-si __?
this-ACC=Q=FAM fence-ACC Pekka.NOM paint-PST.3SG

b. *[Tämä-n=*kö* aida-n=*hAn*] Pekka maalasi __?
this-ACC=Q fence-ACC=FAM Pekka.NOM paint-PST.3SG

I am aware of only one category of hosts that might be plausibly analyzed as falling clearly outside of the class of Ā-operators, namely conjunctions as shown in (8).

(8) a. Jos(=*pa*) lähdettäisiin ulos?
if=POL leave.IMPSS.COND outside
'What if we went outside?'

b. Pekka mietti jos(=*ko*) lähdettäisiin ulos.
Pekka.NOM wonder.PST.3SG if=Q leave.IMPSS.COND outside
'Pekka wondered if it made sense to go out.'

c. Lähde-tään vaikka(=*pa*) Turkuun.
leave.IMPSS.PRS what.if=POL Turku
'What if we went to Turku?'

The mechanism is not general, since most conjunctions cannot be combined with the clitics (**mutta-pa* 'but-POL', ??*koska=pa* 'because-POL') and most clitics cannot be combined with conjunctions (*jos=pa* 'if-POL', **jos=han* 'if=FAM'). Also the meanings are partly idiosyncratic, so these items are perhaps best analyzed as lexicalized items.

1.2 Semantic interpretation

Four factors contribute to the semantic interpretation of sentences with the left peripheral clitics. The first factor is the clitic itself. The yes/no clitic creates regular interrogatives that must be selected as such or be used as root interrogatives. The familiarity, polarity and de-emphatic clitics create pragmatic effects, discussed below.

The second semantic factor is the choice of the hosting element. Suffixing the clitic to the demonstrative, as in (2)b, creates a different set of interpretations than if it is suffixed to the noun head (2)c, which again creates a different set of interpretations than when it is suffixed to the main verb (2)a (overlap not excluded).

The third factor is the choice of the hosting CP which determines the clause that falls inside the syntactic and semantic scope of the clitic. In example (2), the semantic scope of the question is the whole clause. If (2)a–c are embedded inside another main clause, the scope will be the embedded finite clause (9)a. If the pied-piped phrase is moved out of the embedded clause to the superordinate clause CP, the scope will contain both clauses (9)b.

(9) a. Merja kysyi (että) [tämä-n=*kö* aida-n]$_1$ Pekka maala-si__$_1$.
Merja asked (that) this-ACC=Q fence-ACCPekka.NOM paint-PST.3SG
'Merja asked if it was this (and not that) fence that Pekka painted.'

b. [Tämä-n=*kö* aida-n]$_1$ Merja väitt-i (että) Pekka maala-si __$_1$?
this-ACC=Q fence-ACCMerja claimed (that) Pekka.NOM paint-PST.3SG
'Was it this fence Merja claimed that Pekka painted?'

The fourth semantic factor is whether the same element hosts other clitic particles and/or operator features such as wh-features. There are relatively few studies addressing the semantic effects of the clitic combinations, but as a first approximation the semantic effects

of individual clitics seem to get summed together as independent contributions. For example, the fronted yes/no clitic never fails to produce an interrogative clause no matter how it is combined with other clitics and/or Ā-operators.

While the yes/no clitic creates standard interrogatives, the three other clitics (*-hAn*, *-pA*, *-s*) contribute to pragmatic interpretation and communicate background assumptions and/or conversational intentions. They are sometimes called "discourse clitics" or "tonal particles" in reference to this property. The exact semantic role may be difficult to characterize, involve ambiguity, and/or be lexicalized; what is clear is that they are all unacceptable in contexts such as textbooks and official documents.

The familiarity particle (*=hAn*) creates an impression that something in the clause, typically but not necessarily the object denoted by the host expression, is assumed to be familiar, either by local discourse context or by general knowledge (Karttunen, 1975; Hakulinen, 1976; Hakulinen & Karlsson, 1979; Nevis, 1986; Hakulinen et al., 2004: §830; Raevaara, 2004; Holmberg, 2014; Palomäki, 2016).[5] See the translations in (10) for some approximations of how native speakers might interpret the sentences.[6]

(10) Pekka=*han* maala-si aida-n.
Pekka.NOM=FAM paint-PST.3SG fence-ACC
'As you may know, Pekka painted the fence.'
'Pekka, who you might know, painted the fence.'

The familiarity that is at issue must be distinguished from information structural notions such as topic, information focus, contrastive focus and others, since they can be combined freely

[5] See Välimaa-Blum (1986) another view.
[6] Sentence (10) must be read without marked focus stress; if focus stress is positioned on the first word, an additional interpretation is generated in which the stressed element is focused contrastively.

with the familiarity clitic. Sentence (10) combines the familiarity particle with the topic, but does not feel redundant due to the possibility of the first interpretation in which the whole proposition is considered familiar. When the familiarity clitic is combined with information focus, the result is an interpretation in which the discourse-new information is assumed to be familiar on the basis of background information. Thus, sentence (11)b is more likely to be interpreted as expressing contrastive topic (contrast + familiarity) than (11)a, which is more likely to be interpreted as expressing contrastive focus.

(11) a. Aida-n$_1$ Pekka maala-si __$_1$.
fence-ACC Pekka.NOM paint-PST.3SG
'It was the fence that Pekka painted, not the wall.'

b. Aida-n=*han*$_1$ Pekka maala-si __$_1$.
fence-ACC=Q Pekka.NOM paint-PST.3SG
'But it was the fence that Pekka painted, not the wall.'
(Both fence and wall known from the discourse.)

Clitic (=*pA*) emphasizes polarity (whether negative or positive) and/or communicates surprise. It creates unnatural results when a sentence with positive polarity and the clitic is embedded inside a context of doubt. One way to translate the effect into English is to use affirmative discourse markers such as 'yes' or 'sure' (12).

(12) a. Ei ole totta että Pekka(??=*pa*) maala-si aida-n.
not be true that Pekka(=POL) paint-PST.3SG fence-ACC
'#It's not true that sure/yes, Pekka painted the fence.'

b. Väitä-n että Pekka(=*pa*) maala-si aida-n.

claim.PRS.1SG that Pekka.NOM(=POL) paint-PST.3SG fence-ACC

'I claim that sure/yes, Pekka painted the fence.'

The clitic can also be suffixed to the negation, in which case it emphasizes negative polarity.

(13) a. Ei ole totta että ei(??-pä) Pekka maalan-nut aita-a.

not be true that not(=POL) Pekka.NOM paint-PST.PRTC.SGfence-PAR

'#It is not true that sure, Pekka did not paint the fence.'

b. Väitä-n että ei-pä Pekka maalan-nut aita-a.

claim-PRS.1SG that not-POL Pekka.NOM paint-PST.PRTC.SGfence-PAR

'I claim that sure/yes, Pekka did not paint the fence.'

In some contexts, such as (14), the affirmative particle communicates surprise or unexpected results.

(14) O-n=*pA*(=*s*) Pekka maalan-nut aita-a!

be-PRS.3SG=POL=DE/EM Pekka.NOM paint-PST.PRTC.SG fence-PAR

'Wow! Pekka has really painted (a lot of) the fence.'

The affirmation particle cannot be combined with the yes/no clitic (**maala-si=ko=pa* 'paint-PST.3SG-Q-POL'), which could be explained as a clash between affirmation and question. It can be combined with the familiarity particle (*Pekka=pa=han* 'Pekka.NOM=POL=FAM').

The de-emphatic =*s* is associated with politeness, informality or overall de-emphasis, and is perhaps somewhat comparable in its discourse role to English 'by the way' or 'if I may say/ask'. Sentences such as (15)a can feel blunt if uttered in everyday context, while (15)b, which cannot be used in any official document or context, communicates a more friendly tone.

(15) a. Kuinka paljon tienas-i-t vuonna 2024?
how much earn-PST-2SG year 2024
'How much you earned in year 2024?'

b. Kuinka=*s* paljon tienas-i-t vuonna 2024?
how=DE/EM much earn-PST-2SG year 2024
'By the way/if I may ask, how much you earned in year 2024?'

The de-emphatic clitic can be combined with interrogatives and polarity particles (*kuka=s* 'who=DE/EM', *Pekka=pa=s* 'Pekka.NOM=AFF=DE/EM'). It should be noted that some authors exclude *-s* from the list of clitics and treat it as an affix forming lexicalist units (Nevis, 1986). This would mean, however, that the lexicon must contain two separate interrogative words such as *kuinka* and *kuinkas* 'how'. I take the view that *-s* must be listed in the lexicon as an independent element, but do not rule out the possibility that it is used in some contexts as part of a frozen, lexicalist item.

While in the case of the yes/no clitic the scope determines unproblematically the scope of the question, in the case of the pragmatic clitics the scope is involved in determining the thinker (e.g., speaker, main clause subject) whose background assumptions and/or politeness they convey.

1.3 Morphology

Morphologically the clitics are suffixed to the outermost position in the word, after tense and agreement (*maala-si=ko* 'paint-PRS.3SG-Q'), case suffixes (*aita-a=ko* 'fence-PAR-Q', *aida-ssa=ko* 'fence-INE-Q'), possessive suffixes (*aita-a-ni=ko* 'fence-PAR-PX1SG-Q'), number (*aida-t=ko* 'fence-PL.NOM-Q'), pronominal stems (*sinä-kö* 'you=Q'), infinitival and participial markers (*lähte-ä=kö* 'leave-A/INF-Q', *lähte-vän-kö* 'leave-VA/INF=Q', *maalaa-ma-ko* 'paint-MA.PRC=Q'), imperatives (*lähde-pä* 'leave.IMP-Q', *lähde-hän* 'leave.IMP-FAM') and some

conjunctions (*vaikka=pa* ‘though-POL’, *koska=pa* ‘because-AFF’, *jos-pa* ‘if-POL’, **mutta-pa* ‘but-POL’, **ja-pa* ‘and-POL’). Relative pronouns cannot be suffixed with any of the clitics (**jota=kO/pA/hAn/s* ‘which-Q/POL/FAM/S’). If the word hosts several suffixes, then they are stacked into a sequence after everything else (*maala-si=ko=han* ‘paint-PRS.3SG-Q-FAM), but the stacking is regulated by rules we examine further below. As these examples show, the clitics are unselective with respect to the host (one reason to regard them as something else than regular affixes). The clitic suffixation never changes the lexical category of the host stem. No further affixation is possible after the clitics. The clitics themselves are monomorphemes, and have neither alternative free forms nor can they be separated from the host.

The clitics undergo vowel harmony like regular derivational and inflectional suffices (*päät-ti=kö* ‘decide-PST.3SG-Q’ vs. *maala-si=ko* ‘paint-PST.3SG-Q’). Also stress rules support the assumption that they are affixes, creating phonological words with their hosts. On the other hand, cliticalization triggers word-initial gemination and t-assimilation rules that are absent from regular affixation. Nevis (1986), who reviews the above evidence and other facts, considers that the clitics exhibit mixed affixal and word-like behavior and concludes that they are best classified as clitics.

Although the clitics themselves can be suffixed to virtually any host independent of the latter’s lexical category, clitic combinations are restricted. The distribution of *-hAn*, *-pA* and *-kO* can be described by the following rules: no clitic can be doubled; *-kO* and *-pA* cannot be suffixed to any clitic; *-hAn* cannot be suffixed to *-s*. The distribution of *-s* is more complex. It can be suffixed to *-kO* and *-pA (16)*, but if neither is present, then the suffixation is determined nonlocally by the properties of the root stem such that interrogative root stems for example accept *-s* while many other stems do not (17).

(16) a. aita=*ko*=*s* b. aita=*pa*=*s*

fence.NOM-Q-DE/EM fence.NOM=POL=DE/EM

‘fence or…’ ‘yes, fence…’

(17) a. *aida-lta=*s* b. *Peka-lta=*s* c. mi-ltä=*s* d. kene-ltä=*s*

fence-ELA-DE/EM Pekka-ELA-DE/EM what-ELA-DE/EM who-ELA=DE/EM

‘from the fence’ ‘from Pekka’ ‘from what’ ‘from who’

2 Special topics

2.1 Syntactic status of the clitics: affixes or separate heads?

The left peripheral clitics resemble standard inflectional affixes such as number and case in that they never change the lexical category of their host category, but instead add something to the feature content of the host. The clitics are not structure-building units like passive, causative or tense. This has led to the affix hypothesis which regards them as inflectional suffixes, much like case forms (Brattico, 2021, 2022). One problem with this analysis is that the clitics still behave in many ways like independent units, which becomes especially notable when one host is suffixed with several clitics. Examination of the data (Section 3) shows that instead of creating one host category with its features “mixed” from several clitics, as predicted by the affix theory, each clitic must be treated as an independent syntactic and semantic object. The second problem is their mixed morphological behavior, already reviewed, which suggests that they should not be categorized unproblematically as affixes. An alternative is that the clitics are exponents of independent heads (Kenesei, 1994; Holmberg, 2001, 2003, 2014, 2016; Palomäki, 2016), but the fact that they do not seem to build structure or participate in selection then becomes a problem. This debate will be taken up again in Section 3.

2.2 Left peripheral clitics as Ā-operators

In the majority of recent syntactic literature the left peripheral clitics are regarded as part of a larger system of Finnish Ā-operators (Huhmarniemi & Brattico, 2013; Brattico, 2021, 2022; Brattico & Holmberg, 2025: §3), the latter which contains also wh-interrogative pronouns, relative pronouns and contrastive focus/topic operators. All these elements are unified by the fact that they (i) occupy a well-defined position inside CP, (ii) reserve that position so that nothing else occurs there, (iii) create Ā-chains, (iv) create intervention effects for other Ā-operators and (v) can all be combined together. Example (18) shows the three other Ā-operators which occupy the same position as the left peripheral clitics.

(18) a. *Mitä*$_1$ Pekka maala-si __$_1$?
what.PAR Pekka.NOM paint-PST.3SG
'What did Pekka paint?

b. Aita, *jonka*$_1$ Pekka maala-si __$_1$
fence which Pekka.NOM paint-PST.3SG
'a fence Pekka painted'

c. *AIDAN*$_1$ Pekka maala-si __$_1$, ei talo-a.
fence.ACC Pekka.NOM paint-PST.3SG not house-PAR
'It was the FENCE Pekka painted, not the house'

The wh-interrogatives are represented by wh-pronouns, relative operators by relative pronouns (a partially distinct morphological class in Finnish), and contrastive focus/topic operators by marked prosodic stress.[7] If a prosodically and morphologically unmarked

[7] A similar effect can be accomplished by in situ prosodic stress. Whether these two cases are produced by two different focus features in Finnish is difficult determine. Although the two cases can have different properties, it does not follow without further argument that two

constituent is moved to the CP-domain, native speakers will interpret it as a contrastive focus/topic operator. We could perhaps regard this interpretation as a "default" interpretation for the higher phrasal position of the Finnish finite clause, reviewed in Section 1.1.

2.3 Focus and focus sets

The left peripheral clitics can be suffixed with almost any host stem independent of its lexical category and what other suffixes may be present (Section 1.3, examples (2) and (3)). The choice of the host affects semantic interpretation, in particular it determines the *focus set* associated with the clitic (using (Rooth, 1985, 1992) as a framework). Here we demonstrate these effects by using the yes/no clitic *=kO*.

First, independent of the host the presence of *=kO* causes the underlying clause to be interpreted as an interrogative. The semantic effects that depend on the selection of the host are added to the basic interrogativization effect. If the host contains lexical content X (i.e. we exclude lexically empty items such as auxiliaries), then the default interpretation is one in which the yes/no operator selects X from a set of implicit alternatives defined by X (19).

(19) [Tämä-n=*kö* aida-n] Pekka maalasi __?
this-ACC=Q fence-ACC Pekka.NOM paint-PST.3SG
'Was it this (and not that) fence that Pekka painted?'
Focus set: 'this', 'that'

If the clitic is attached to a fronted auxiliary that lacks independent lexical content, then the interpretation is a standard yes/no question corresponding to Aux-inversion in English (20) in which the whole sentence/proposition is in focus (a subtype of "wide focus").

---

different features are a stake. The issue has not been examined for Finnish and requires further scrutiny.

(20) O-n=*ko* Pekka maalan-nut tämä-n aida-n?

be-PRS.3SG=Q Pekka.NOM paint-PST.PRTC.SG this-ACC fence-ACC

'Has Pekka painted this fence.'

Focus set: 'Pekka has/has not painted this fence.'

If the fronted verb has lexical content, both interpretations are possible (21).

(21) Vuokra-si=*ko* Pekka __ kesämöki-n?

rent-PST.3SG=Q Pekka.NOM summer.house-ACC?

A. 'Did Pekka rent the summer house? (Yes or no?)'

B. 'Did Pekka rent the summer house? (Or did he buy it?)'

A natural way to construct interpretation A is that Pekka's renting had been the topic of previous conversation, which the speaker wants to confirm, hence the focus set is 'Pekka rented/did not rent the summer house'. A natural way to construct interpretation B is to understand it as a question of how Pekka acquired the rights to occupy the house, with the set of alternatives being 'Pekka rented/bought/stole… the summer house' or, more generally, 'Pekka X the summer house' with narrow focus on X.[8] If the clitic is suffixed to the last word of a pied-piped phrase (22), then we have the two interpretation shown in (22).

(22) [Tämä-n aida-n=*ko*]$_1$ Pekka maala-si __$_1$?

this-ACC fence-ACC=Q Pekka.NOM paint-PST.3SG

A. 'Was it this fence (and not the gate) that Pekka painted?'

B. 'Was it this fence (and not that wall) that Pekka painted?'

[8] The actual focus set created by X depends on several pragmatic factors present in the transient discourse, and is possibly unlimited if considered from the point of view of what is possible in principle.

As pointed out by Sulkala & Karjalainen, p. (1992, p. 9), if we front an auxiliary but put focus stress on another element, then that element alone introduces the focus set (23):

(23) O-n=*ko* Pekka maalan-nut TÄMÄ-N aida-n?
be-PRS.3SG=Q Pekka.NOM paint-PST.PRTC.SG this-ACC fence-ACC
'Has Pekka painted THIS his fence?'

The interpretation is similar (perhaps identical) to the one created by demonstrative operator fronting and pied-piping. Wide focus interpretation is not available.

2.4 Long head movement: predicates as Ā-operators

So far we have examined constructions in which the fronted verbal host is the closest head to C^0, such as an auxiliary (20) or a finite and tensed main verb (2)a. Fronting is not restricted to the closest verbal element. Also nonfinite verbs of nonfinite complement clauses can be fronted (24):

(24) Myy-dä=*kö*$_1$ Pekka aiko-o __$_1$ koko omaisuute-nsa?
sell-A/INF=Q Pekka.NOM plan-PST.3SG whole possession-ACC.PX3
'Does Pekka intend to <u>sell</u> (and not loan) all his possessiongs?'

The fronting creates Ā-chains, as one would expect if the left peripheral clitics created Ā-operators, and moves the verbal head to C^0 (Brattico, 2022). The interpretation is one in which, in addition to the interrogativization effect, the fronted predicate determines the focus set 'Pekka planned to X all his possessions'. (24) does not have a wide focus interpretation in which the yes/no question targets the whole proposition. The fact that also nonfinite verbs (of any type) can be fronted reinforces the conclusion that the clitics are extremely unselective with respect to their host.

2.5 Interaction between =*kO* and focus stress

Holmberg (2014) reports a detailed study on the interaction between yes/no operator =*kO* and focus stress (glossed here as |FOC, assumed to be a prosodic feature), the latter which creates contrastive focus interpretations in Finnish with or without movement to CP. The first crucial observation is that if the pied-piped phrase fronted to CP contains a prosodically stressed word followed by another word with the =*kO*, the former receives the narrow focus while the yes/no morpheme only triggers interrogativization (ex. 13 in the original).

(25) OLLIN isän=kö auto-lla te ajo-i-tte __?
Olli.GEN|FOC father.GEN=Q car-ADE you.NOM drive-PST-2PL
'Was it OLLI's father's car that you took?'

If *Olli* is not stressed, an alternative interpretation is possible in which narrow focus falls on 'father' (ex. 11 in the original). Thus, placing the focus stress on the first constituent "cancels" the focus effect from the constituent with the yes/no clitic and leaves only the interrogativization effect. A similar effect was observed in (23). The second observation is that it is not possible to place contrastive stress on the constituent following the yes/no clitic (ex. 14 in the original):

(26) *Ollin isän=*kö* AUTOLLA te ajo-i-tte __?
Olli.GEN father.GEN=Q car.ALL|FOC you.NOMdrive-PST-2PL
Intended: 'Was it Olli's father's CAR that you drove?'

These and other observations led Holmberg to propose a generalization according to which the head representing =*kO* must c-command the constituents determining focus. This would make it a focus operator much like words such as *only* and *even*. Two analyses are

considered, while neither is regarded as completely unproblematic. One is that *=kO* is adjoined to maximal projections and comes to c-command everything inside ("Theory 2", discussed further in Section 3), another claims that *=kO* is a head that probes material to its specifier into which it the then cliticizes ("Theory 3"). Since no final unproblematic analysis emerges, and I concur that both analyses as well as some of the semantic judgments have issues that need to be sorted out, it is difficult to draw firm conclusions from these data. I will return to this matter briefly in Section 3.

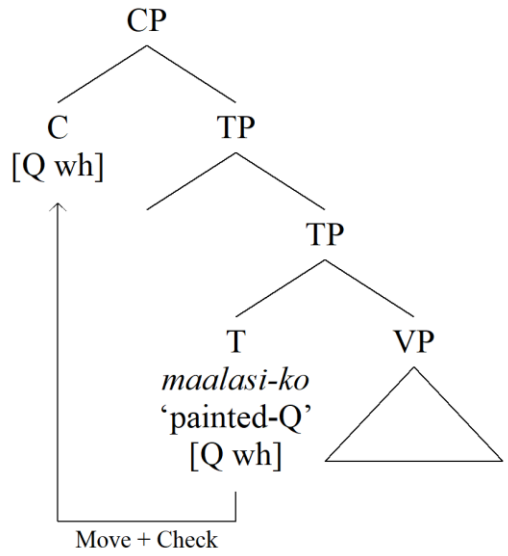
3 An analysis of the left peripheral clitics

I conclude with an analysis of the Finnish left peripheral clitics that summarizes the properties reported above. Several alternative models are considered. The proposal was formalized and tested by the computational generative grammar methodology.[9] The formalization and the computational experiment is described in the supplementary document accompanying the source code and will be ignored here.

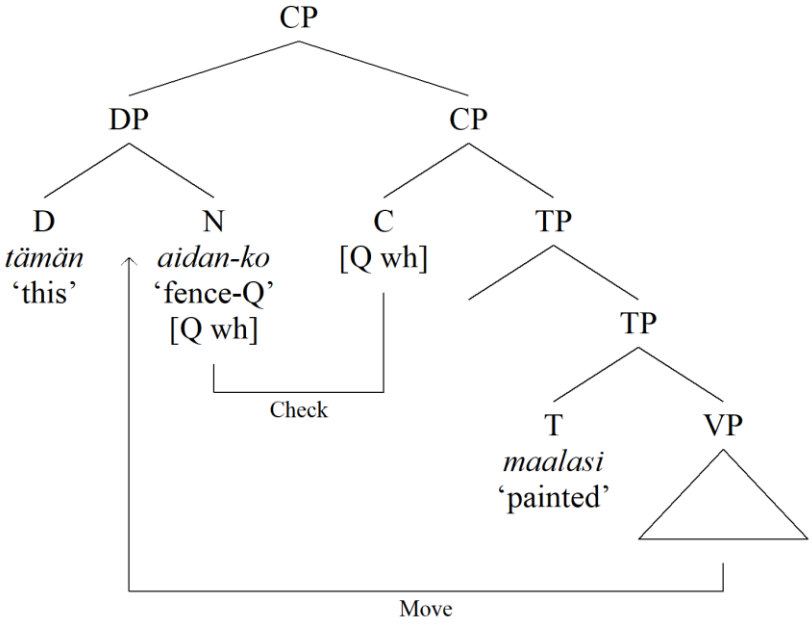
The data shows that the Finnish left peripheral clitics are part of the CP-domain, the latter which implements a system of Ā-operators that also includes wh-interrogatives, relative pronouns and contrastive focus/topic operators. Since several Ā-operators, including the yes/no clitic, affect clause typing, their presence implies properties of C^0, thus we assume that both head and phrasal fronting imply feature checking at C^0. (18) shows feature checking by head movement, (19) by phrasal movement. Q = yes/no operator creating the focus set, wh = clausal interrogative force.

[9] Links to the source code, data, input and outputs together with the supplementary document documenting the methods will be provided for the final version. The baseline model is available at https://github.com/pajubrat/parser-grammar.

(27)


CP
C
[Q wh]
TP
TP
T
maalasi-ko
'painted-Q'
[Q wh]
VP
Move + Check

(28)


CP
DP
CP
D
tämän
'this'
N
aidan-ko
'fence-Q'
[Q wh]
C
[Q wh]
TP
TP
Check
T
maalasi
'painted'
VP
Move

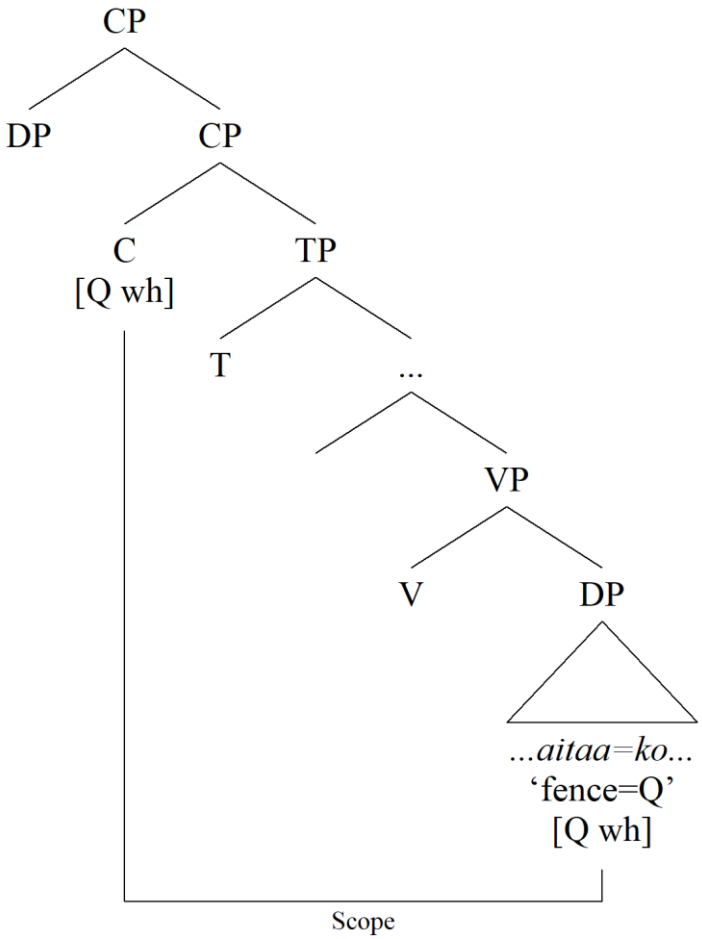
I assume that the clitic is originally placed at the moved host, while C^0 is phonologically null. This undeniably true for (26) since there is nothing between the fronted object and the subject in the surface sentence. Both Q and wh are copied to C.[10]

Kenesei (1994) and Holmberg (2001) suggest that the left peripheral clitics are exponents of feature content at C^0 (for Kenesei, Focus head F^0). Palomäki (2016) makes essentially same claim but proposes a more fine-grained left periphery in which the familiarity clitic expresses the head of ForceP. Nevis (1986) proposes that the clitics are independent sentence-initial adverbs and cliticize to their host by postsyntactic "phonological concatenation" (p. 91) due to being bound morphemes. Example (2a) supports these analyses, and they create a plausible analysis for (2c) where the C(=*kO*) is cliticized to the right edge of the phrase at SpecCP (i.e. [$_{DP}$ *tätä aitaa*]=*kO*). However, examples such as (2b)/(3) in which the host is inside the pied-piping phrase are left without explanation (Holmberg, 2014). I therefore assume, following Holmberg (2014), that the clitic is originally attached ("base-generated") to the host word and moved to CP, with or without phrasal pied-piping.

The left peripheral clitics can only occur inside the CP. This follows from an interface condition (Brattico, 2021, 2022) which was adopted here. Consider the LF-configuration (25) feeding the semantic interpretation for (24).

[10] Since the features can occur deeply inside the pied-piping phrase, we need a separate scanning process that locates them – a recursive left-to-right depth-first search was assumed.

(29)



We assume that the scope for the operator *aitaa=ko* 'fence=Q' is determined by a c-commanding C with the relevant features it acquired in virtue of the copying process described above and that Ā-operations lacking the in situ option must be paired with scope information at the LF as an interface condition. This rules out sentences in which the moved left peripheral clitics occur at any other position than the CP. The assumption that fronting copies features to C and the scope algorithm mean that the fronting is essentially a mechanism for "expressing scope."

A feature of the Finnish CP is that it cannot host phonologically overt material both at C^0 and SpecCP (Section 1.1). The existence of V2 languages and English root interrogatives suggest that this is a language-specific feature that is controlled either by a parameter or by a

lexical feature. There is virtually no research on this issue when it comes to Finnish, so the the lexical approach was adopted. When the features of C^0 are created by the phrasal movement route (24), C^0 must have a property that licenses the extra specifier position for the pied-piped phrase. If the feature (called EF, from "edge feature") is missing, the specifier positions are not projected. If we assume that the specifier behavior is a language-specific property of C^0, and that the phrasal movement route (23/24) assigns the required property to C^0, then the facts follow: head movement does not create new specifier positions, the specifier position is not licensed, and all double-filled CPs are ruled ungrammatical.

Let us consider the focus sets. We assume that when a simple or complex head is dominated by a head with an operator feature, every head of the predicate with lexical content are included in the focus set calculations. In addition, let us assume that finite T^0 (or whatever head heads the finite clause, such as Fin^0) creates a special wide focus interpretation over the whole event, possibly because it contains a polarity feature (Mitchell, 1991; Holmberg, 2003, 2014, 2016) or a feature which allows it to refer to the whole event (as assumed here). As a result, the fronted finite verb *maala-si-ko* 'paint-PST.3SG=Q' in (30) with the word-internal structure $((V\ v)\ T)^0$ creates two focus interpretations: one by T, another by V, which captures the ambiguity in its interpretation (an analysis of this type was suggested in (Brattico & Holmberg, 2025: §3) and was formalized here). This process is triggered by the yes/no operator feature at the highest head T at the LF-interface representation (31).

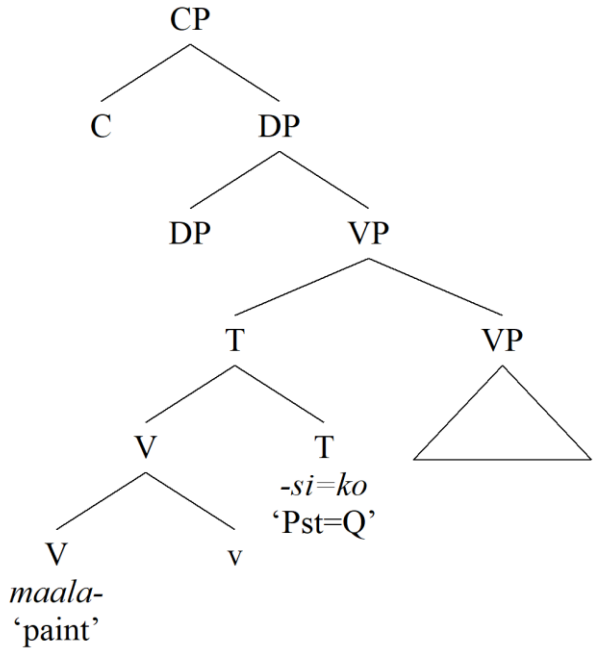
(30) Maala-si=*ko* Pekka aida-n?

paint-PST.3SG=Q Pekka.NOM fence-ACC

A. 'Did Pekka paint the fence?'

B. 'Was it painting (and not washing) that Pekka did to the fence?'

(31)



The important point about (31) is that the interpretation targets the complex head that is headed by the operator: not C^0, which is headed by null C^0, and none of the individual heads populating the base-positions. When a nonfinite verb is fronted, its focus interpretation arises in the same way at the base position (32) where the complex word is headed by the operator (here, the left peripheral yes/no clitic).

(32)	Myy-dä=*kö*	Pekka	aiko-o	~~*myy-dä-kö*~~	koko	omaisuutensa?
	sell-A/INF-Q=C	Pekka	plan-PST.3SG	sell-A/INF-Q	whole	posseions
	C (scope)			Focus set: 'sell', 'loan', …		

Since finite T is not involved, the wide focus interpretation is correctly ruled out.

The three left peripheral clitics induce pragmatic effects within their scopes (Section 1.2), the latter determined by the scope calculations just elucidated. These effects were modelled inside their own pragmatic module. Each pragmatic contribution is summed into

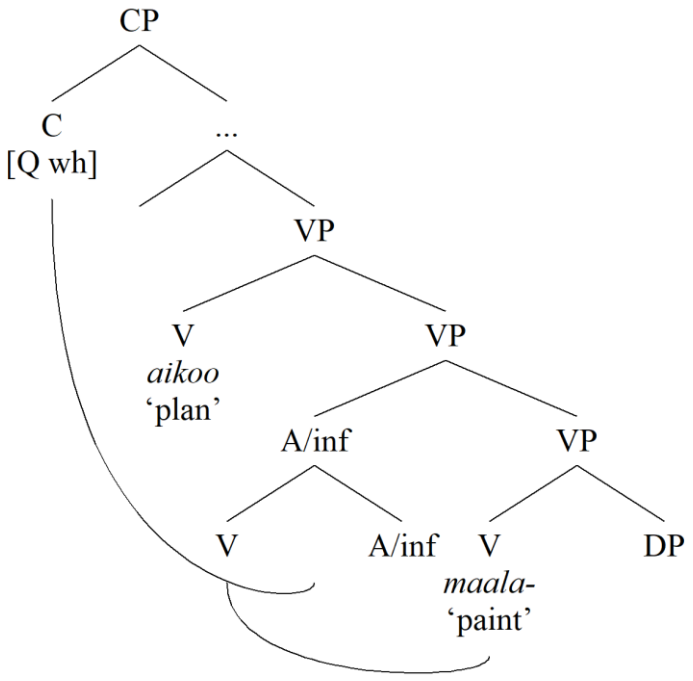
the final output produced by the model. They are triggered by dedicated pragmatic features associated with the clitics.

One property of Finnish Ā-operators is that both phrasal operators and verbal operators form Ā-chains (Section 2.4)(33).

(33) Maala-ta=*ko*$_1$ Pekka aikoo __$_1$ tämä-n aida-n?
paint-A/INF=Q Pekka.NOM plans this-ACC fence-ACC
'Does Pekka intent to paint (and not tear apart) this fence?

In an earlier study of these constructions I assumed, following (Roberts, 1993, 2010), that all Ā-operators create Ā-chains independent of whether they involve pied-piping or not. The former create phrasal Ā-chains, the latter *head Ā-chains*. I assumed, furthermore, that head Ā-chains are triggered by the left peripheral clitics which made the fronted verbal elements Ā-operators. This analysis was carried over to the present model: sentence (33) is analyzed as (34) where the relevant head chain is marked by the curved line.

(34)



The technical details of how to program Ā-chains such that they apply both to heads and phrases involve a range of interesting and challenging issues that I will put aside here; see the supplementary. By using the principles assumed so far the semantic interpretation comes out as 'Was it painting (and not, say, tearing down) that Pekka intended to do with the fence?', with narrow focus on the verb 'paint' (highlighted) under the complex head. C^0 has the operator yes/no feature which creates the scope for the operator at LF plus an interrogative force feature which gives it an interrogative force, both generated by the head Ā-chain as explained above.

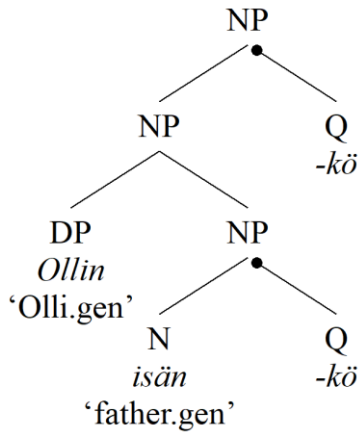
All of the above presupposes that the lexical entries of the clitics contain minimally the following features: the yes/no clitic *-kO* = interrogative wh + yes/no operator feature; familiarity clitic *-hAn* = pragmatic familiarity feature + familiarity operator feature; polarity clitic *-pA* = pragmatic polarity emphasis feature + polarity operator feature; de-emphatic clitic *-s* = pragmatic de-emphasis features + de-emphasis operator feature. The operator features are responsible for the scope mechanism and the creation of the focus sets. Some features must be marked for copying to C (they were equipped with special strength diacritic **). All left peripheral clitics must rule out the in situ strategy, hence another feature is required to separate them from the in situ operators. Pragmatic features are interpreted inside their own specialized semantic module which understands contextual notions such as 'familiarity', 'emphasis' and 'de-emphasis' (as well as topic, information focus, see (Brattico, 2021)).

I conclude by considering the syntactic status of the clitics themselves. Finnish left peripheral clitics, like clitics overall, exhibit conflicting traits: one one hand they behave like affixes and add linguistic content to the host words without altering their lexical category or selection properties. In this they differ from morphemes such as passive, causative or tense.

Yet they are also unlike inflectional suffixes in that they must be processed as if they were independent objects.

Holmberg (2014, 2016) considers a hypothesis according to which the left peripheral clitics, which are assumed to be independent heads, are right-adjoined by base-generation to their head or phrasal hosts. The fact that they are adjoined explains why their presence does not affect selection. Accordingly to this model, in *Olli-n isä-n-kö* 'Olli-GEN father-GEN-Q…' the clitic has two attachment positions (35)(three if we consider it a DP).

(35)



The higher attachment sites creates an interpretation in which the focus set is determined by 'Olli's father', the lower 'father'. This model captures also the fact that the clitics can occur inside the phrase they pied-pipe: Q could be right-adjoined to *Olli-n* to create *Olli-n-ko isän* 'Olli-GEN-Q father-GEN'.

While this analysis gets many properties right, some problems remain. As Holmberg himself acknowledges, it is not always the case that *-kO* must c-command the constituent(s) creating the focus set, an assumption Holmberg relies on to motivate the analysis. This is shown by examples such as (36), where *-kO* is attached to the adjective and the focus stress

on the demonstrative determines the focus set (see Section 2.5). The demonstrative is not c-commanded by *-kO*.

(36) TÄTÄ vanhaa=*ko* aita-a Pekka maala-si __?
this|FOC old-PAR=Q fence-PAR Pekka.NOM paint-PRS.3SG
'Was it this (and not that) old fence that Pekka painted?'

The major challenge, however, is how to restrict the distribution of the clitics. Some principle must force all left peripheral clitics to adjoin to the same host (no scattering, Section 1.1) and another to control the clitic combinations (Section 1.3). The fact that the clitics are adjoined captures neither property. The latter in particular is a challenge since the clitics could be piled on the same host without having any intrinsic relation, selection or otherwise, and even if they were adjoined to each other, adjunction does not, as this relation is usually understood, imply selection. Finally, the assumption that the clitics are adjuncts will have implications for any formal system that already includes an adjunction operation. In this particular case it is unclear, for example, what would prevent left-adjunction and what licenses the adjunction operation itself. In short, some theory of adjunction that needs to be fleshed out is presupposed in the analysis.

For the above reasons the model formalized here does not assume that the clitics build new structure; instead, they were treated as affixes. The fact that they behave like independent objects then creates a problem. To solve this issue it was assumed following Stump (1980) that the clitics correspond to independent feature bundles inside their host words. For example, when the yes/no clitic *=kO* is suffixed to a noun head such as *aita-a* 'fence-PAR', the noun head will have a set of regular noun head features followed by another set containing the features acquired and bundled together from the lexical entry of the yes/no

clitic. If the head is suffixed with second clitic, its features are represented as another bundle (37).

(37) . . . tätä aita-a=*kO=hAn*

this.PAR fence-PAR=Q=FAM

[{N, Par, Singular, 3rd Person,… }, …, {OP:Q…}, {OP:FAM…}, …]

Then we assume that the bundled features can then go into semantic interpretation as independent units. Clitic combinations, and the positioning of the clitics within the overall word formation template of Finnish, can be controlled by restricting the feature bundle sequences. Since the sequence is visible at the LF-interface, the restrictions may be syntactic or semantic or both.

The whole model was tested by using the computational generative grammar methodology against a dataset that contained the expressions discussed in this article together with systematic variations (a total of 491 expressions). The source code, dataset and all the output are available in the source code repository, which also contains instruction for replication. The model passed the test: it calculated the properties of Finnish left peripheral clitics reviewed in this article.

References

Brattico, P. (2021). A dual pathway analysis of information structure. *Lingua, 103156.*

Brattico, P. (2022). Predicate clefting and long head movement in Finnish. *Linguistic Inquiry*, *54*(4), 663–692. https://doi.org/http://dx.doi.org/10.1162/ling_a_00431

Brattico, P., & Holmberg, A. (2025). *The Syntax of Finnish*. Cambridge: Cambridge University Press.

Brattico, P., Huhmarniemi, S., Purma, J., & Vainikka, A. (2013). The structure of Finnish CP and feature inheritance. *Finno-Ugric Languages and Linguistics*, *2*(2), 66–109.

Dékány, É. (2018). Approaches to head movement: A critical assessment. *Glossa: A Journal of General Linguistics*, *3*(1). https://doi.org/10.5334/gjgl.316

Hakulinen, A. (1976). Liitepartikkelin -han/-hän syntaksia ja pragmatiikkaa. In *Reports on text linguistics: Suomen kielen generatiivista kielioppia 2* (Vol. 7, pp. 3–60).

Hakulinen, A., & Karlsson, F. (1979). *Nykysuomen lauseoppia*. Helsinki: Finnish Literature Society.

Hakulinen, A., Vilkuna, M., Korhonen, R., Koivisto, V., Heinonen, T. R., & Alho, I. (2004). *Iso suomen kielioppi*. Helsinki: Finnish Literature Society.

Holmberg, A. (2001). The syntax of yes and no in Finnish. *Studia Linguistica*, *55*, 140–174.

Holmberg, A. (2003). Yes/no questions and the relation between tense and polarity in English and Finnish. In P. Pica (Ed.), *Linguistic Variation Yearbook* (pp. 43–68). John Benjamins.

Holmberg, A. (2014). The syntax of the Finnish question particle. In P. Svenonius (Ed.), *Functional Structure from Top to Toe* (pp. 266–289). Oxford: Oxford University Press.

Holmberg, A. (2016). *The syntax of Yes and No*. Oxford: Oxford University Press.

Holmberg, A., & Nikanne, U. (2002). Expletives, subjects and topics in Finnish. In P. Svenonius (Ed.), *Subjects, Expletives, and the EPP* (pp. 71–106). Oxford: Oxford University Press.

Holmberg, A., Nikanne, U., Oraviita, I., Reime, H., & Trosterud, T. (1993). The Structure of INFL and the Finite Clause in Finnish. In A. Holmberg & U. Nikanne (Eds.), *Case and other functional categories in Finnish syntax* (pp. 177–206). Mouton de Gruyter.

Huhmarniemi, S. (2012). *Finnish A'-movement: Edges and Islands*. University of Helsinki.

Huhmarniemi, S., & Brattico, P. (2013). The Structure of Finnish Relative Clause. *Finno-Ugric Languages and Linguistics*, *2*(1), 53–88.

Karttunen, F. (1975). The Syntax and Pragmatics of the Finnish Clitic -han. *Texas Linguistic Forum*, *1*, 41–50.

Kenesei, I. (1994). Complementation in Finno-Ugric. *Manuscript*.

Mitchell, E. (1991). Evidence from Finnish for Pollock's theory of IP. *Linguistic Inquiry*, *22*, 373–379.

Nevis, J. A. (1986). *Finnish Particle Clitics and General Clitic Theory*.

Nevis, J. A. (1990). Sentential clitics in Finno-Ugric. *Folia Linguistica*, *24*(3–4). https://doi.org/10.1515/flin.1990.24.3-4.349

Palomäki, J. K. (2016). *The pragmatics and syntax of the Finnish –han particle clitic* [Ph.D. dissertation]. University of Georgia.

Raevaara, L. (2004). »Mitäs me sovittais». S-partikkelin sisältävien hakukysymysten tehtävistä. *Virittäjä*, *108*(4), 531–558.

Roberts, I. (1993). *Verbs and diachronic syntax. A Comparative History of English and French*. Amsterdam: Kluwer Academic Press.

Roberts, I. (2010). *Agreement and Head Movement: Clitics, Incorporation, and Defective Goals*. Cambridge, MA.: MIT Press.

Rooth, M. (1985). *Associations with Focus*.

Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics*, *1*(1), 75–116. https://doi.org/10.1007/BF02342617

Stump, G. T. (1980). An Inflectional Approach to French Clitics. *OSU WPL*, *24*, 1–54.

Sulkala, H., & Karjalainen, M. (1992). *Finnish*. London: Routledge.

Välimaa-Blum, R. (1986). The Discourse Function of the Finnish Clitic–HAN: Another Look. In P. Lilius & M. Saari (Eds.), *Proceedings of the Sixth International Conference of Nordic and General Linguistics in Helsinki, August 18-22* (pp. 471–481).

Vilkuna, M. (1989). *Free word order in Finnish: Its syntax and discourse functions*. Helsinki: Finnish Literature Society.

Vilkuna, M. (1995). Discourse Configurationality in Finnish. In K. É. Kiss (Ed.), *Discourse Configurational Languages* (pp. 244–268). Oxford: Oxford University Press.