STRUCTURAL CASE ASSIGNMENT, THEMATIC ROLES AND INFORMATION STRUCTURE*

Pauli Brattico

Abstract. A formal model is presented which deduces the structural case assignment profile of Finnish by relying on graph-theoretical paths and intervention. More complex and controversial features such as nonlocal dependencies, adverbial case marking, case competition, DP-internal case patterns and interaction of case with agreement, aspect and polarity are also calculated from the model. Abstract Case and case concord were eliminated. The model links case assignment with word order, thematic roles and information structure. The results suggest that case might play a role in language comprehension.

*Acknowledgments. The research reported in this article is based on a rather long series of attempts by the present author at formulating a structural case theory for Finnish, most of it in (formal and information) collaboration with Anne Vainikka. This background and my intellectual debt to her is reflected throughout the article. Writing of this article benefitted from the thorough comments of two anonymous *SL* reviewers.

1. Introduction

Finnish structural case assignment has evoked a considerable amount of controversy in the linguistic literature. The most interesting features of the system are its nonlocal dependencies (Ikola, 1950, 1986, 1989; Linden, 1956; Wiik, 1972; Toivonen, 1995; Brattico, 2009, 2012, 2014; Vainikka & Brattico, 2014; Anttila & Kim, 2017), sensitivity to agreement, aspect and polarity in addition to syntactic position (Timberlake, 1975; Itkonen, 1976, 1977; Carlson, 1981; Vainikka, 1988, 1989; Reime, 1993; Heinämäki, 1984, 1994; Kiparsky, 1998, 2001; Nelson, 1998; Megerdoomian, 2000, 2008; Csirmaz, 2005, 2012; Anttila & Kim, 2011; Huumo, 2013; Brattico, 2020b), adverbial case marking (Maling, 1993; Vainikka & Maling, 1996) and layered case assignment and case competition (Nelson, 1998; Brattico, 2010, 2011). Moreover, Finnish exhibits at least fifteen different case forms, with four separate

structural case forms alone. Thus, the empirical footprint of the Finnish structural case assignment is considerable.

I present a formula which captures the structural case assignment signature of this language. The formula is based on the hypothesis that overt morphological case features in the linguistic input are checked against lexical elements ("case assigners") by means of a graph-theoretical upward path dependency, developed on the basis of Kayne's (1983, 1984). connectedness hypothesis. Path dependencies are limited by intervention, however, which confines case checking into well-defined domains. Abstract Case plays no role in the model. The case checking mechanism is then embedded inside a Python-based recognition grammar that maps input sentences into syntactic and semantic representations. The resulting algorithm detects oddball arguments in noncanonical positions and attempts reconstruction on the basis of overt case information, filtering out failed solutions before they are forwarded to the semantic systems for interpretation. These mechanisms, which operate during language comprehension, correlate morphological case forms with thematic roles and information structure.

Section 2 introduces the basic properties of Finnish structural case assignment, key principles of the analysis and provides a few examples towards justification. Section 3 reports a computational experiment where the model is applied to a dataset containing the structural case assignment signature of Finnish. Most of the concrete empirical data concerning Finnish case assignment can be found from this section. Section 4 contains the conclusions.

2. Case assignment, upward paths and intervention

2.1 Background

In many languages the morphological form of nominal words varies as a function of their morphological, syntactic and semantic context. For example, the two forms *he* ~ *him* are traditionally classified as the two case forms, nominative and accusative, respectively, of the one and the same underlying masculine singular pronoun. The term *case assignment* is used in this connection to refer to the process of selecting some case form, given a context. For example, English prepositions can be said to assign the accusative to their complement since they control its distribution (e.g., *to him*, **to he*). A linguistic theory of case assignment is concerned with specifying the necessary and sufficient conditions for the distribution of some

or all case forms, in one or several languages. Here we are interested in developing a theory of Finnish case assignment in this sense.

Finnish has fifteen nominal case forms, of which four are structural (1)(these numbers depend slightly on the theoretical prism used in the analysis). Further comments concerning this system are provided below. All examples are in Finnish unless otherwise stated.[1]

(1) a. *Nominative*

Merja/	lapse-t/	hän	hävis-i.[2]
Merja.NOM	child-PL.NOM	he.NOM	disappear-PST.3SG

'Merja/children/he disappeared.'

b. *Partitive*

Pekka	ihaile-e	Merja-a/	lapsi-a/	hän-tä
Pekka.NOM	admire-PRS.3SG	Merja-PAR	child.PL-PAR	he-PAR

'Pekka admires Merja/children/him.'

c. *Accusative (n-accusative ACC(N), t-accusative ACC(T)*)

Pekka	näk-i	Merja-n/	lapse-t/	hän-et.
Pekka.NOM	see-PST.3SG	Merja-ACC(N)	child-PL.ACC(T)	he-ACC(T)

'Pekka saw Merja/children/him.'

d. *Accusative (zero-accusative ACC(0), t-accusative ACC(T))*

Me	näh-tiin	Merja/	lapse-t/	hän-et.
we.NOM	see-PST.IMPASS	Merja.ACC(0)	child-PL.ACC(T)	he-ACC(T)

---

[1] Abbreviations: 0 = no agreement or the default agreement; 1, 2, 3 = first, second and third person; ACC = accusative case, any form; ACC(0) = zero accusative, homophonous with the nominative in singular; ACC(N) = n-accusative, homophonous with the genitive in singular; ACC(T) = the t-accusative (assigned for pronoun and plural direct objects); A/INF = A-infinitival (corresponding loosely to the English *to*-infinitival); ARG = case assigner; CAU = causative morpheme/head; GEN = genitive; MALLA = MALLA-infinitival (i.e. one particular MA-infinitival, such as *juokse-malla* 'by running'); NEG = negation; NOM = nominative; IMPASS = impersonal passive; O/o = object of the main clause/embedded infinitival clause; PAR = partitive; PL = plural; PRS = present tense; PRTCPL = participle; PST = past tense; S/s = subject of the main clause/embedded infinitival clause; SG = singular; V/v = verb of the main clause/embedded infinitival clause; VA/INF = VA-infinitival (a 'propositional' complement clause); ±VAL = licenses (+) or blocks (–) overt agreement. Both stems and affixes undergo frequent morphophonological changes in Finnish. The segmentations provided in the glosses should be considered as approximations. See Table 2 for what I take to be the unmarked lexical case forms.

[2] Combination of the plural subject *lapset* 'children' with the third person singular finite verb *hävisi* 'disappeared.3SG' is grammatical in colloquial Finnish. Formal, written Finnish requires plural agreement on the verb.

‘We saw Merja/children/him.’

e. *Genitive*

Pekka	käsk-i	Merja-n/	laste-n/	hän-en	tulla.
Pekka.NOM	order-PST.3SG	Merja-GEN	child.PL-GEN	he-GEN	came.

‘Pekka ordered/asked Merja/children/he to come.’

The nominative (1a) can be regarded as the canonical subject case. It is assigned to the grammatical subjects of both canonical intransitive and transitive clauses. The partitive and accusative (1c-d) are the canonical direct object cases and tend to represent objects or patients. The genitive is assigned to the subject of the infinitival clause in the example (1e) but has many more uses. The system is complicated by the existence of the three accusative forms: the t-accusative ACC(T) (for plurals and pronouns[3]), n-accusative ACC(N) and the zero-accusative ACC(0) (for singular full DPs). I will have much to say about these three forms later. The case forms are summarized in Table 1. See note 1 for a comment concerning the glossing conventions.

Table 1. Case marking of full argument DPs and pronouns.

	FULL DPS		**PRONOUNS**	
	SINGULAR	PLURAL	SINGULAR	PLURAL
NOM	*talo* ‘house’	*talo-t* ‘houses’	*minä, se* ‘I, it’	*me, ne* ‘we, it.PL’
GEN	*talo-n*	*talo-jen*	*minun, sen*	*meidän, niiden*
ACC	*talo* (ACC(0)) *talo-n* (ACC(N))	*talo-t* (ACC(T))	*minu-t* (ACC(T)) *se-n* (ACC(N))	*meidät* (ACC(T)) *ne* (ACC(0))
PAR	*talo-a*	*talo-ja*	*minu-a, si-tä*	*meitä, niitä*

The two accusative forms, the zero-accusative (ACC(0)) and the n-accusative (ACC(N)), are homophonous with the nominative and genitive, respectively. This pattern is restricted to full

[3] An anonymous *SL* reviewer points out, correctly, that the t-suffix of a nonpronominal argument could be analysed as the bare plural marker, since the same form is used for nominative plurals (see Table 1). Accordingly, the argument would bear an “unmarked case” (presumably the nominative and/or the zero-accusative). This alternative is not a notational variant of the present analysis and must be tested rigorously against the same dataset. For pronouns, the t-form represents the unambiguous accusative case (*hän* ‘he.NOM’ ~ *hän-et* ‘he-ACC(T)).

singular DPs: pronouns and plurals have their own unique t-accusative forms (ACC(T)) in all contexts where the zero-accusative and n-accusative are attested.[4] Most descriptive grammars assume that the zero-accusative is the nominative, the n-accusative the genitive. So far I have been unable to calculate the data from a system that makes this simplification, so the issue was left for future research. Consequently, the n-accusative will be glossed as ACC(N), the zero-accusative as ACC(0).

2.2 The hypothesis

We propose two principles regulating the distribution of the Finnish case forms listed in (1) and Table 1. First, overt morphological case forms are checked against sets of lexical features instead of a single feature or a single case assigner head. This will capture situations in Finnish where several syntactic and/or semantic factors (e.g., aspect, polarity, agreement) affect one case form. The assignment dependency between the assigner and assignee is then defined by means of a graph-theoretical *path*. Suppose a case assignee α requires checking by lexical features $F = \{f_1 \ldots f_n\}$; then

(2) *Case checking and feature intervention*

F checks α if and only if *F* occurs inside an *upward path* from α such that there is no closer nonempty set *G*, $G \subset F$, inside the same path;

(3) *Upward path*

the upward path from α contains all constituents that dominate α and their immediate daughters.

We imagine the case assignee searching for a suitable case checker by 'scanning through the path'. The search continues until the case assignee encounters either a full match of features *F*, leading into checking, or partial match *G*, $G \subset F$, leading into failure; or reaches the end of the structure, which will also lead into failure. Intuitively case forms are licensed inside the "syntactic scopes" of lexical elements, where the notion of syntactic scope is defined by (2-3) and the relevant lexical elements by *F*. I will use the term "government" when referring to the syntactic scope from the point of view of the assigner. Case checking establishes that the case

[4] There are a few exceptions to this rule. For example, the accusative forms of the nonhuman pronouns *ne* 'it.PL' and *se* 'it.SG' are the zero form (*ne* 'it.ACC(0)') and the n-form (*se-n* 'it.ACC(N)'), respectively.

form and its grammatical context match. To illustrate both the terminology and analysis, consider (4a-b).

(4) a. Pekka [$_{\text{NegP}}$ **e-i** voitta-nut *kilpailu-n/ kilpailu-a.]
Pekka.NOM not-3SG win-PST.PRTCPL competition-ACC(N) competition-PAR
'Pekka did not win the competition.'

b. Pekka [$_{\text{AuxP}}$ **o-n** voitta-nut kilpailu-n/ *kilpailu-a.]
Pekka.NOM be-PST.3SG win-PST.PRTCPL competition-ACC(N) competition-PAR
'Pekka did win the competition.'

These data show that the Finnish partitive-accusative alteration is in some way sensitive to polarity. The accusative cannot be governed by the negation, while the partitive has the opposite profile. The principles (2-3) proposed above define the relevant checking configuration, shown in (5).

(5) Pekka [$_{\text{NegP}}$ e-i [voitta-nut *kilpailu-n.]]
Pekka.NOM not-3SG win-PST.PRTCPL competition-ACC(N)
└──────────────────────────────┘

The dependency is formed by creating a path from the case assignee to the assigner through the phrase structure. The assigner, in turn, is defined by a set of features F, in this case features that have to do with polarity (the details are examined later). As a consequence, the direct object appears inside the syntactic scope of the negation, and the latter is said to govern the former. Standard local case assignment, such as that holding between a preposition and its complement, is modelled as a special case of the same dependency. If the case form and the grammatical context match, we say that the case features are checked.

2.3 Case forms and their features

2.3.1 Introduction

Next, we specify the feature sets F involved in the mechanism. These features define the elements that will govern case assignees in our model. Suppose we want to say that arguments with the direct object cases must be governed by transitive verbs. We would then use F to define what we mean by "transitive verb," most likely by means of at least two features 'being a verb' and 'being transitive'. Similarly, if we want to capture (4-5) by linking the

Finnish direct object case forms to polarity, F will define what we mean by "polarity."[5] The definitions must be provided in an explicit form so that the system can be tested rigorously.

2.3.2 Partitive

Vainikka (1988, 1989, 1993, 2003) has argued that the Finnish partitive behaves like a "default complement case." It occurs in the complement position of prepositions (*kohti talo-a* 'towards house-PAR'), numerals (*kolme talo-a* 'three house-PAR'), participle adjectives (*talo-a ostava* 'house-PAR buying'), noun heads (*kasa sukki-a* 'stack sock-PAR'), quantificational elements (*paljon sukki-a* 'many socks-PAR'), and further encodes aspectual properties when occurring as a direct object (4). For an explicit argument that the Finnish partitive constitutes a structural complement case, not semantic or inherent case, see (Vainikka & Maling, 1996). While the notion of "complement" does not occur in (2-3), the preposition is inside the upward path generated from the case assignee. Furthermore, the cases just mentioned are unified by the fact that none of the lexical items assigning the partitive agree in phi-features with the case assignee. I will show in this article that the facts follow if we assume that the partitive is checked against non-agreeing case assigners.

This requires that we define the class of case assigners. Case assigners will be distinguished from all other lexical items by the lexical feature +ARG. This will prevent determiners, conjunctions, complementizers, numerals and many other case-neutral lexical items, lacking this feature, from governing case forms. Whether a lexical item can exhibit agreement is marked by the lexical feature +VAL: +VAL allows the head to exhibit overt agreement, −VAL prohibits it. Hence, we will assume PAR ~ +ARG, −VAL which says that the partitive DP must occur inside the syntactic scope of a non-agreeing (−VAL) case assigner (+ARG)(i.e. F in (2-3) will be {−VAL, +ARG}). We show that this calculates the correct results and subsumes Vainikka's default complement rule. To illustrate, consider the Finnish adposition data (6).

(6)	a.	lähellä	minu-a/	b.	minu-n	lähellä(-ni)/	c.	*lähellä-ni	minu-a
		near	I-PAR		I.GEN	near(-PX/1SG)		near-PX/1SG	I-PAR
		'near me'			'near me'			'near me'	

[5] The matter is nontrivial, since several other grammatical devices, such as those expressing nonveridicality or epistemic hesitation, trigger the same effect. See (Kaiser, 2002, 2003; Thomas, 2003) for discussion. What seems to me to be at stake is grammaticalization of "sensitivity to (non)veridicality" in the sense of Giannakidou (1998).

Some Finnish adpositions have two forms, one that assigns the partitive to the argument at the complement position (6a), another assigning the genitive to the specifier position (6b). When the genitive is assigned, the adposition exhibits optional phi-agreement with its argument. Agreement is not possible if the argument is marked for the partitive case (6c). Rule PAR ~ +ARG, −VAL is designed to capture generalizations of this type in our dataset.

2.3.3 Accusative (three forms)

In addition to the partitive, direct objects of verbs and deverbal predicates can be assigned the accusative. Its presence correlates with telic properties of the event denoted by the verb phrase (e.g., Carlson, 1981; Vainikka, 1989; Heinämäki, 1984, 1994; Kiparsky, 1998; Thomas, 2003; Csirmaz, 2012)(7).

(7) a. Pekka pes-i hevos-en.
Pekka.NOM wash-PST.3SG horse-ACC(N)
'Pekka washed the (whole) horse.'

b. Pekka pes-i hevos-ta.
Pekka.NOM wash-PST.3SG horse-PAR
'Pekka washed the horse (but the horse did not necessarily become clean).'

Let us assume, following Kiparsky (1998), that (7a-b) are distinguished from each other by whether the event denoted by the verb phrase includes an end point ('complete action')(7a) or not ('incomplete action')(7b), and that the feature representing the relevant distinction in the lexicon is ASP:BOUNDED. The feature is part of a verbal head (V, v), possibly a separate Asp head. To capture (7), we assume that the accusative rule refers to this aspectual feature. The hypothesis is illustrated in (8).

(8) Pekka pesi hevos-en.
Pekka washed$_{[+ASP]}$ horse-ACC(N)
└──────┘

There is a complication, however. The accusative, when licensed by aspect in the manner illustrated in (8), can take several forms depending on whether the upward path contains an agreeing predicate (9).[6]

[6] Sentence (9a) is part of written Finnish, (9b) colloquial Finnish. The Institute for the Languages of Finland writes that (b) is "frequently used" but not "recommended" due to its

(9)	a.	Me	pes-i-**mme**	hevos-en/	*hevonen/	hevos-et.
		We.NOM	wash-PST-1PL	horse-ACC(N)	horse.ACC(0)	horse-PL.ACC(T)
		‘We washed the horse.’				
	b.	Me	pest-**iin**	*hevos-en/	hevonen/	hevos-et
		we.NOM	wash-PST.IMPASS.0	horse.ACC(N)	horse.ACC(0)	horse-PL.ACC(T)
		‘We washed the horse.’				

To capture (9), I assume that overt agreement enters into the feature sets checked by the accusative. Since plural full DPs are not affected, the rule is restricted to singular full DPs. Polarity is also relevant, as already shown by (4). Consequently, we add polarity to the accusative rule. Finally, both the agreement and polarity effects are nonlocal. This is an uncontroversial feature of the Finnish case system (for recent work, see Vainikka & Brattico, 2014; Anttila & Kim, 2017). Example (10) shows how the main clause agreement affects direct object case forms inside an infinitival complement clause α.

(10)	a.	Me	halus-i-mme	[α rakenta-a	*talo/	talo-n.]
		we.NOM	want-PST-**1PL**	build-A/INF	home.ACC(0)	home-ACC(N)
		‘We wanted to build a house.’				
	b.	Me	halut-tiin	[α rakenta-a	talo/	*talo-n.]
		we.NOM	want-PST.IMPASS.**0**	build-A/INF	home.ACC(0)	home-ACC(N)
		‘We wanted to build a house.’				

The first person plural agreement in the main clause affects the direct object case forms inside the infinitival complement clause. The case assigner and assignee are separated from each other by at least four grammatical heads and the infinitival clause boundary α. This type of nonlocality is an intrinsic feature of the path mechanism (2-3): the upward path may continue until either there is intervention or no more structure.

In sum, to calculate the distribution of all accusative forms in Finnish we will refer to four features: case activity, aspect, agreement and polarity, all which must be checked by (2-3).

---

colloquial character (http://www.kielitoimistonohjepankki.fi/ohje/345, retrieved 4. 8. 2020). Use of the standard from in colloquial speech feels hypercorrect, whereas the use of the impersonal form in standard written Finnish is (still) “not recommended” but not grammatically offensive either, not to me at least.

2.3.4 Genitive and nominative

Vainikka suggested that the Finnish genitive case is a "default specifier case." The genitive is assigned to what looks to be the specifier positions of adpositions (*minun lähelläni* 'I.GEN near'), infinitival complement clauses (*Pekka käski minun lähteä* 'Pekka ordered I.GEN to.leave'), nouns (*minun auto* 'I.GEN car'), participle adjectives (*minun löytämä* 'I.GEN found', i.e. something found by me), and certain finite constructions, such as the modal construction (*minun täytyy lähteä* 'I.GEN must leave'). The present approach is incompatible with her proposal because there is no upward path from the specifier to its head. On the other hand, it is not uncommon that at least one element in a chain headed by the genitive argument occurs in a potential checking position. Brattico (2020a) proposed that the genitive is checked at the base position of the chain headed by the genitive argument. To illustrate, consider the modal construction (11).

(11) Minun täyty-y lähte-ä
I.GEN must-PRS.0 leave-A/INF
'I must leave.'

There is no overt licencing structure inside the upward path from the genitive argument at the preverbal subject position. On the other hand, the subject receives its thematic role from the A-infinitival verb *lähte-ä* 'leave-A/INF', which suggests that it reconstructs into the infinitival phrase. If we assume GEN ~ +ARG, −FIN and allow the genitive argument to check its case against the A-infinitival head inside the reconstructed position (thus at $\_\__1$ in [DP_1 [*must*…[A/inf [$\_\__1$ *leave*]]]]), the dataset can be calculated correctly. Nominative case can then be handled by the rule NOM ~ +ARG, +VAL, +FIN which checks it against agreeing finite verbs from the base position of the subject A-chain (12).

(12) Minä$_1$ T_{fin} $\_\__1$ v löysin avaim-en.
I.NOM T v found key-ACC(N)
[+ARG] 'is a case checker'
[+VAL] 'can (and often does) exhibit phi-agreement'
[+FIN] 'is finite'

This generalization will capture nominative case checking in connection with grammatical subjects.

2.3.5 Summary

The complete feature system is summarized in Table 2.

Table 2. Structural case symbols, their morphological forms and checking features

Case	Suffix	Feature set	Example
NOM	-0	{+ARG, +VAL, +FIN}	*Pekka nukkuu* 'Pekka.NOM sleeps'
PAR	-(t)A	{+ARG,−VAL}	*Pekka söi omenaa* 'Pekka.NOM ate apple.PAR'
GEN	-n	{+ARG,−FIN}	*Merja näki Pekan lähtevän* 'Merja saw Pekka.GEN to.leave'
ACC(T)	-t	{+ARG, +ASP:BOUNDED}, {−NEG}	*Merja näki hän-et* 'Merja saw he-ACC(T)' *Merja osti kuka-t* 'Merja bought flower-PL.ACC(T)' *Me nähtiin hän-et* 'We saw.IMPASS he-ACC(T)' *Me näimme hän-et* 'We saw.1PL he-ACC(T)' *Me ei nähty *hän-et/hän-tä* 'We did not see he-ACC(T)/he-PAR'
ACC(N)	-n	{+ARG, +ASP:BOUNDED}, {−NEG}, {+PHI}	*Me näimme talo-n* 'We saw.1PL house-N/ACC' **Me nähtiin talo-n* 'We saw.IMPASS house-N/ACC'
ACC(0)	0	{+ARG, +ASP:BOUNDED}, {−NEG}, {−PHI}	*Me nähtiin talo* 'We saw.IMPASS house-0/ACC' **Me näimme talo* 'We saw.1PL house-0/ACC'

FIN = finiteness; ARG = case assigner; VAL = phi-agreement; NEG = negative polarity; PHI = overt phi-agreement, ASP:BOUNDED = aspectual boundedness.

A few details concerning the Table 2 require a further comment. First, pronouns and plural DPs are marked by the t-accusative form (ACC(T)) that does not check +PHI. It is an empirical fact that this form is not sensitive to agreement, while it is sensitive to polarity and aspect. See the examples in Table 2, rightmost column.

Second, −NEG and ±PHI occur inside separate sets in Table 2. This is because according to the more or less standard theory of Finnish finite clause structure (e.g., Mitchell, 1991; Holmberg et al., 1993; Manninen, 2003; Huhmarniemi, 2012), aspect, polarity and agreement occur inside different heads: aspect at a verbal head (v, V, Asp), agreement at finite T or Fin, and polarity at Neg. Each must therefore be checked by a separate dependency established between the case assignee and the corresponding head.

Third, there is no binary distinction between structural and semantic cases. The accusative is sensitive to both syntax and semantics: while aspect and polarity can be said to be semantic, verbal phi-agreement is a formal property. The Finnish accusative represents a "mixed" case. The nominative, partitive and genitive can be said to be structural: only formal features, agreement (±VAL), finiteness (±FIN) and case activity (±ARG), are relevant.

3. Simulation experiment

3.1 Introduction

Next we verify that the logical consequences of the hypothesis converge with empirical observations. There are two ways to do this. One is to formulate the theory as an enumerative grammar that generates sentences and their meanings by using the linguistic mechanisms, principles and lexical resources posited in the theory. We then verify that the theory generates only grammatical and/or acceptable sentences and further provides them with correct or at the very least plausible syntactic and semantic analyses. This corresponds to a literal generative grammar. An alternative is recognition grammar, which analyses input sentences instead of generating them. Thus, instead of deriving sentences from a given set of lexical items, recognition grammars derive syntactic and semantic analyses from surface sentences. A recognition grammar was used in this study.

To this end, the analysis was embedded inside a minimalist-oriented Python-based language processing algorithm that maps linguistic inputs into syntactic and semantic representations (Brattico, 2019a). The resulting model was tested against a dataset containing Finnish sentences exhibiting possible and impossible case configurations. The background model is explained in Section 3.3. For a general discussion of the computational methodology, see (Brattico, 2021b).

3.2 Dataset (test corpus)

Recognition grammars are tested by feeding them with sentences, both grammatical and ungrammatical. A *test corpus* was created for this purpose. The test corpus used in the present study contains almost the whole structural case assignment signature of Finnish. Contents of the test corpus are listed in Appendix A. The test sentences were linear lists of bare phonological words without morphosyntactic or syntactic tagging or analyses. All words were normalized (e.g., capitals and punctuation removed), while some words were disambiguated when testing specific lexical items for an otherwise ambiguous word. Disambiguation blocks irrelevant parsing derivations but has no impact on the evaluation of the analysis. Virtually the whole case assignment signature was included. Special complex constructions exhibiting labile case alternations where both the zero-accusative and n-accusative are possible were excluded from the dataset. See (Anttila & Kim, 2017). Some predicative copular sentences

were tested but the examination was not systematic due to the controversial and to me still unclear nature of this class.

3.3 Procedure

The test sentences were fed into the Python based recognition grammar that was assumed as a syntactic background theory in this study. The algorithm creates an idealized brain model for the speaker of any language that it uses to model processing in that language. It maps input sentences into phrase structure representations and interprets them semantically. Figure 1 illustrates the information flow. The underlying grammatical theory is minimalist in orientation.

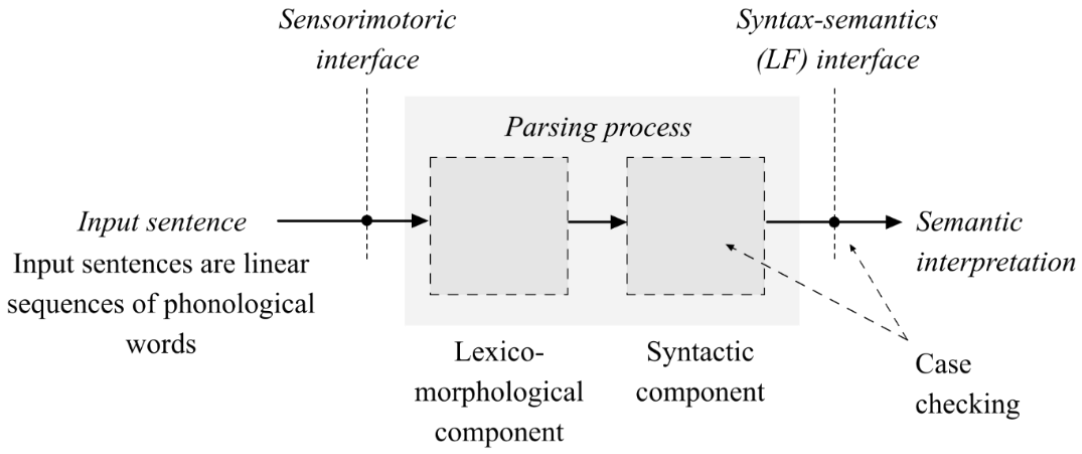


Figure 1. Syntactic background theory and the position of the proposed case checking mechanisms within the architecture. See the main text for explanation.

Each input sentence is processed through the *lexico-morphological component*, retrieving lexical items on the basis of the phonological words present in the input, and the *syntactic component* which generates parsing solutions on the basis of the lexical items it receives from the former. For example, a sentence such as *the horse ran past the barn* is mapped into a parsing solution [[$_{DP}$ *the horse*] [$_{VP}$ *ran* [$_{PP}$ *past the barn*]]] where the terminal elements are lexical items. Lexical elements are sets of features. The case checking principles (2-3) function as a filter before the parsing solution is forwarded to semantic interpretation. In addition, the algorithm detects oddball arguments in the input that cannot check their case features and attempts to reconstruct them into canonical thematic positions where the features can be checked (Brattico, 2020a). We can think of overt morphological case forms as guiding the parser towards plausible solutions, increasing its 'error tolerance'. This mechanism then

feeds an independent pragmatic pathway that links noncanonical word orders with information structural interpretations (such as topic, focus)(Brattico, 2021a). We are interested in whether the resulting model is able to separate grammatical case configurations from the ungrammatical ones and provide the former with plausible syntactic and semantic interpretations. Since the model uses case forms to guide reconstruction, we are also interested in how the proposed case checking mechanism handles noncanonical word orders.

3.4 Results: Observational adequacy

First we compare the grammaticality judgments provided by the model with grammaticality intuitions provided by a native speaker (here, the author). The model and native speaker judgments were compared by an automatic file comparison tool. The model judged 290 out of 293 constructions correctly. Correct judgment means that the grammaticality judgment of the model matched with that of a native speaker. The three errors were: a spurious reconstruction inside a complex noun phrase leading the model to accept an ungrammatical expression (13a); an illegitimate reconstruction of a rightward genitive argument (13b); a partitive-marked adverbial (13c). These are all judged ungrammatical by native speakers, but wrongly accepted by the model.[7]

(13) a. *Ne sukka-a$_1$ kaksi __$_1$ hävi-si. (#162)
those sock.SG-PAR two.0 disappear-PST.3SG
Intended: ‘Those/the two socks disappeared.’

b. *Pekka sano-i __$_1$ lähte-vän Merja-n$_1$. (#234)
Pekka.NOM say-PST.3SG leave-VA/INF Merja-GEN
Intended: ’Pekka said that Merja will leave.’

c. *Pekka nukku-i koko päivä-ä. (#293)
Pekka.NOM sleep-PST.3SG all day-PAR
Intended: ‘Pekka slept all day.’

[7] An anonymous *SL* reviewer pointed out a fourth possible problem: the sentence *Pekka on Peka-n* ‘Pekka.NOM is Pekka-GEN’ (#274) was wrongly marked as ungrammatical in the test corpus and has a reading ‘Pekka owns/possesses Pekka’. The reason it was marked ungrammatical is because the intended reading is the identity statement ‘Pekka is Pekka’. If we assume that ‘Pekka owns/possesses Pekka’ involves the same (feature-wise identical) copular verb used in identity statements, then this specimen must be added to the list of problems. Predicative copular sentences were not tested systematically in this study, however.

Further exploration of these errors after this article was finalized suggests that (13a-b) are most likely irrelevant to case checking and involve problems that have to do with the reconstruction algorithm, while (13c) detected a residuum problem in the adverbial case checking mechanism.

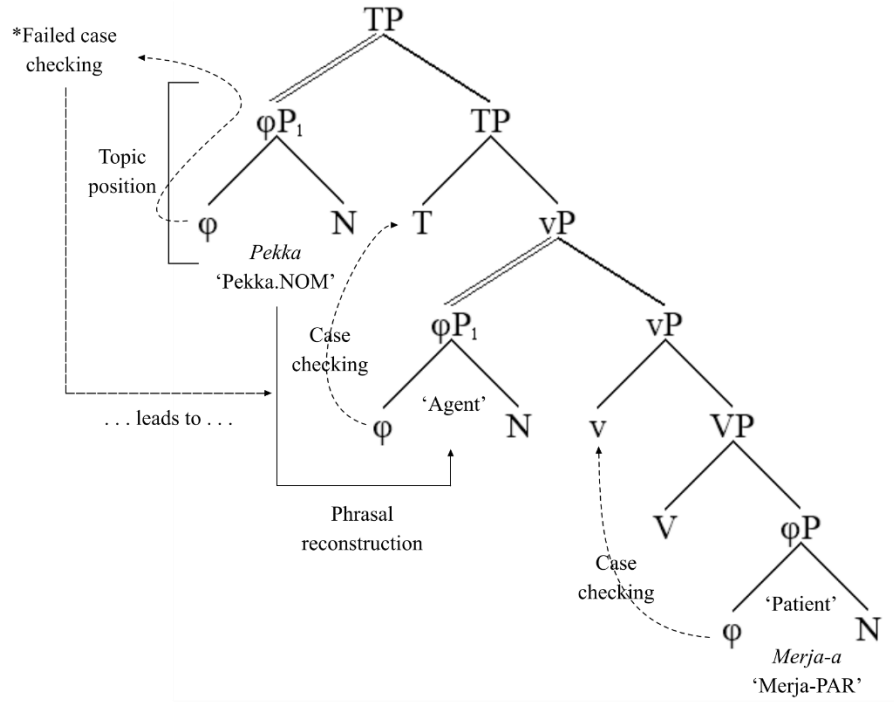
3.5 Results: Descriptive adequacy

3.5.1 Canonical and noncanonical finite clauses

Next we examine if the model calculates analyses and interpretations that are linguistically plausible and match with the syntactic and semantic interpretations elicited from native speakers.

We begin by considering the processing of a canonical transitive finite clause *Pekka ihaile-e Merja-a* 'Pekka.NOM admire-PRS.3SG Merja-PAR' (sentence #5 in the test corpus) with a canonical nominative subject and canonical partitive object. The model judges the input as grammatical and calculates (14).

(14)



The underlying phrase structure images were generated by the algorithm while some text and other symbology were added by the author to facilitate readability. The original figures generated by the model are available online (see Appendix B). Case checking dependencies established by (2-3) are notated by dashed arrows, reconstruction by solid arrows. The model

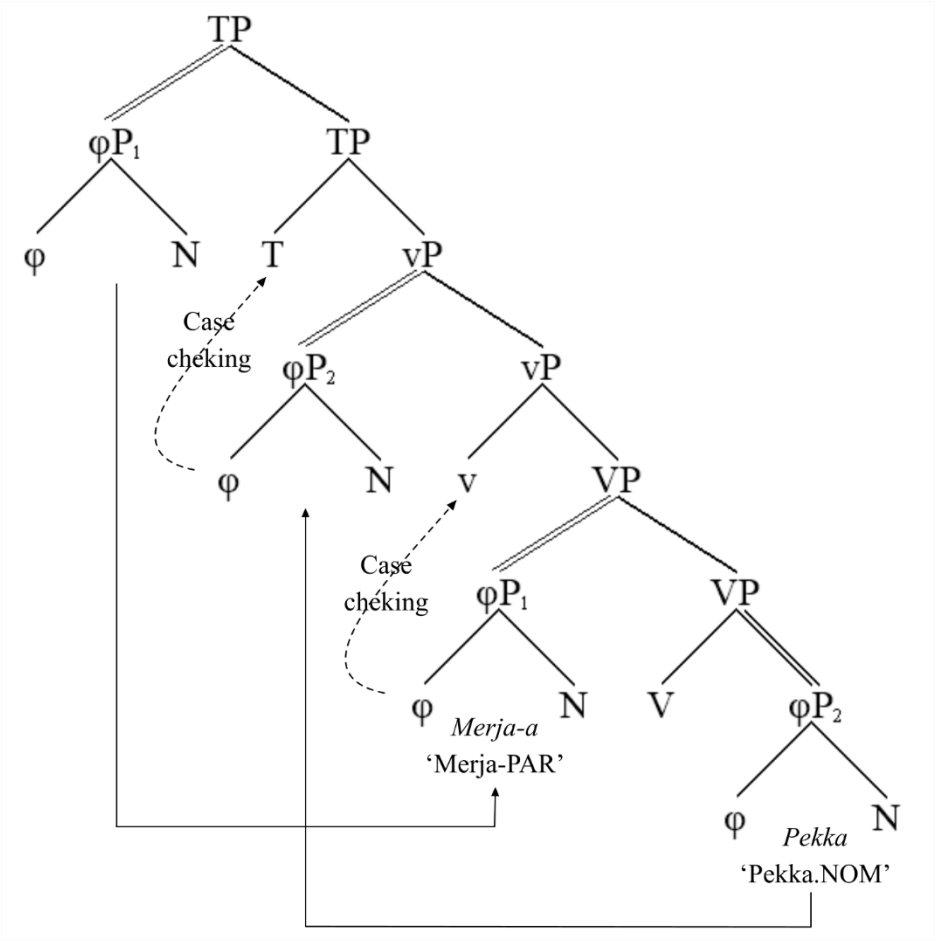
generates binary-branching asymmetric bare phrase structure representations of the form α = [A B] where A and B are the immediate left and right constituents of α, respectively. They are created by operation *Merge*, which joins the two constituents A and B to form a new constituent [A B]. We can think of [A B] as a cognitive chunk of two previously assembled elements. Constituents A and B can be primitive or complex. A primitive constituent has no daughters; complex constituent has two. The system, apart from the asymmetry, is based on the bare phrase structure model proposed by Chomsky (2001, 2008). Since the phrase structure system supports a recognition grammar, these representations are constructed from the input feed of lexical items, in this case the feed is a linear string of words /Pekka/ ∗ /ihailee/ ∗ /Merjaa/ 'Pekka.NOM admire.PRS.3SG Merja.PAR'.

Let us consider case checking. The subject *Pekka* is marked for the nominative case in the input. The features F = {+ARG, +FIN, +VAL} associated with the nominative cannot be checked at the surface position SpecTP. This alerts the system that the constituent might be in a wrong position, and consequently triggers reconstruction which brings the argument from SpecTP into SpecvP where F can be checked against T. Finite verbs are both finite (+FIN) and show agreement (+VAL), hence we capture the connection between nominative case, finiteness and agreement. The reconstructed SpecvP position is associated with an *agent* interpretation during semantic interpretation. The preverbal SpecTP, on the other hand, can be occupied by almost anything in Finnish as long as it constitutes the topic of the sentence or is otherwise topical (Vilkuna, 1995; Holmberg & Nikanne, 2002; Brattico, 2019b; Huhmarniemi, 2019a). The whole reconstruction operation therefore pairs the subject with two semantic attributes: topic and agent. The partitive object *Merja-a* 'Merja-PAR', on the other hand, was merged directly into the postverbal position where it checks +ARG and −VAL against v. It will be interpreted as the *patient*. There is no object agreement in Finnish, hence v has −VAL. No reconstruction occurred, and therefore the patient argument was not linked with a special information structural interpretation.[8]

[8] The steps corresponding to this description, in fact almost all descriptions provided in the main article, can be found from the raw output files produced by the algorithm. The thematic ('agent', 'patient') and information structural interpretations ('topic', 'focus') cited in the main text can be found from the lines 162 and 170, respectively, of the results file. The moment the grammatical subject was recognized to be in a wrong position during processing was recorded into the derivational log file, line 1401, which reads `[φ Pekka] failed [ARG][FIN][VAL]. Reconstructing [φ Pekka]...Topicalization...Done`. The result of the reconstruction is in the next line.

Consider next how the model reacts to a noncanonical OVS structure *Merja-a ihaile-e Pekka* ‘Merja-PAR admire-PRS.3SG Pekka.NOM’ (#11) where both the thematic agent and patient occur in unexpected noncanonical positions. The thematic patient is in the preverbal topic position, while the grammatical subject occurs postverbally. OVS sentences are grammatical in Finnish. Neither argument can check their case features at the surface positions. The calculated result is (15).

(15)



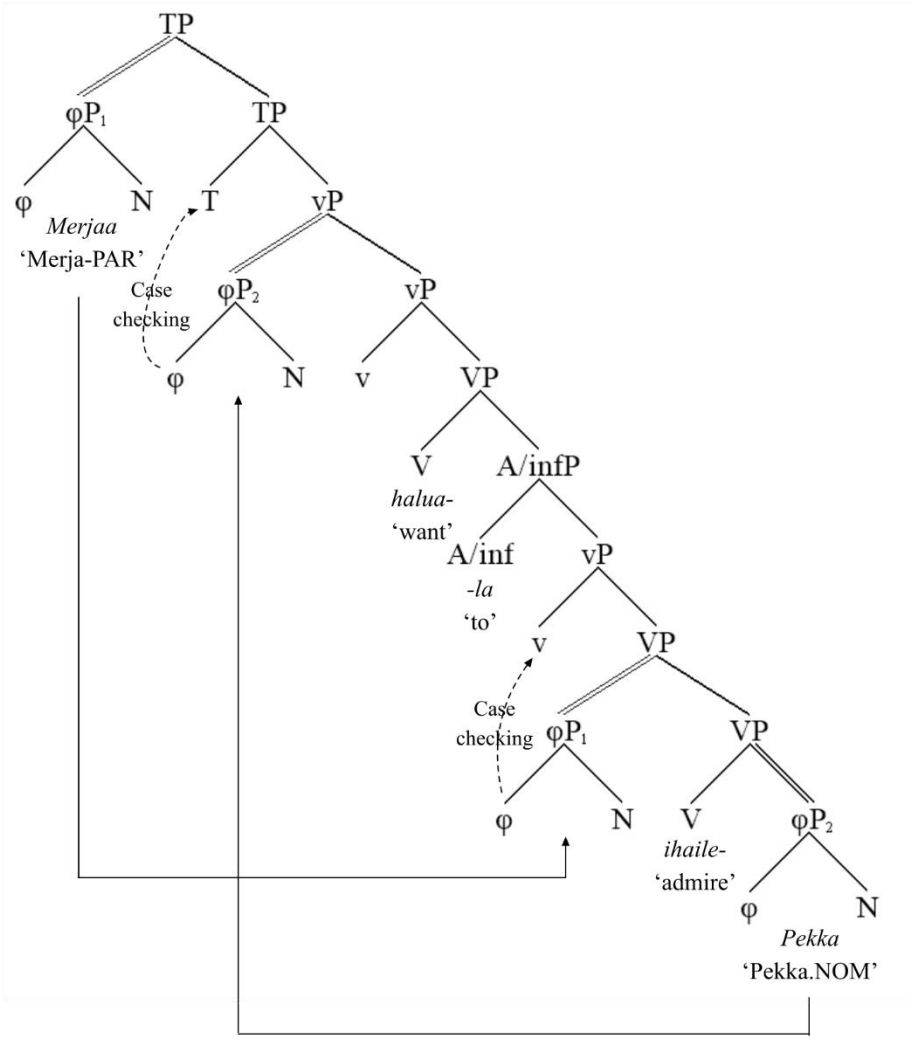
The postverbal argument reconstructs to SpecvP, the preverbal partitive argument to VP. This assigns them the correct thematic roles ‘agent’ and ‘patient’, respectively. The direct object is further interpreted as the topic, while the postverbal subject is interpreted as representing the information focus (that is, new information). In sum, then, the postverbal grammatical subject is interpreted as the focus/agent, the preverbal direct object as topic/patient. Case information allows for an argument to retrieve its thematic role even if it encodes topic/focus information by appearing in an unexpected position in the input.

The same reasoning applies if the grammatical subject occurs further to the right, as in the sentence (16).

(16) Merja-a halu-si [ihail-la Pekka.] (#14)
Merja-PAR want-PST.3SG admire-A/INF Pekka.NOM
'When it comes to Merja, it was Pekka who wanted to admire her.'

The model calculates (17).

(17)



The nominative subject is reconstructed from the rightmost/lowest position in the clause to SpecvP, while the partitive argument goes to CompVP. Reconstruction is case-based. The infinitival clause itself is derived by combining the v/VP-shell with an infinitival head corresponding to the overt infinitival suffix *-(t)A*. This provides an analysis of Finnish infinitivals, according to which their syntactic and semantic structure mirrors closely the overt

morphological composition in the input, an idea that goes back to Koskinen (1998). Because the grammatical subject occurs further to the right, it is interpreted as the marked focus. For the details concerning how the algorithm generates topic and focus interpretations, see (Brattico, 2021a).

The model was tested with all logically possible word order variations for the transitive clause (sentences #8-13 in the test corpus, see examples (18)) and for the clause containing a complement infinitival (#14-37, 87-93, some examples (19)), which it calculated correctly (S=main clause subject, O=main clause object, s = embedded clause subject, o=embedded clause object, V=main clause verb, v=embedded infinitival verb).

(18) a.	Pekka	Merja-a	ihaile-e. (#8)		
	Pekka.NOM	Merja-PAR	admire-PRS.3SG		
	S	O	V		
b.	*Ihaile-e	Pekka	Merja-a. (#12)		
	admire-PRS.3SG	Pekka.NOM	Merja-PAR		
	V	S	O		

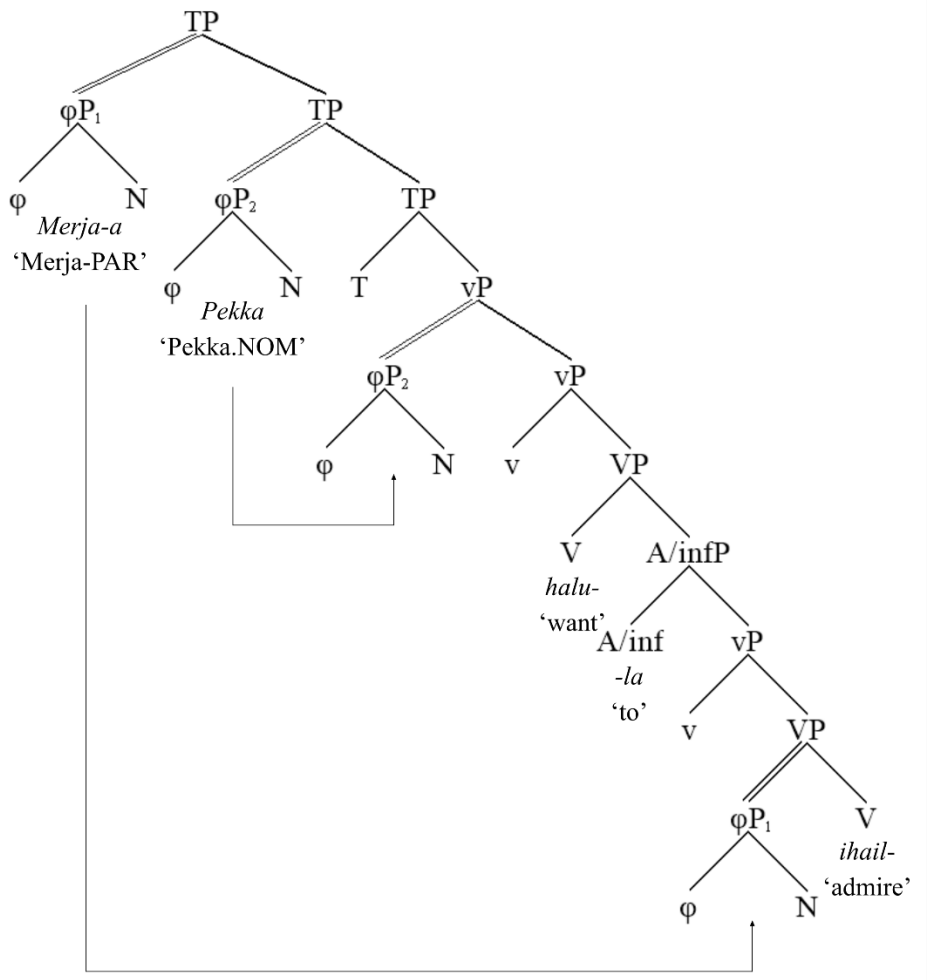
(19) a.	Merja-a	halus-i	Pekka	ihail-la. (#15)	
	Merja-PAR	want-PST.3SG	Pekka.NOM	admire-A/INF	
	o	V	S	v	
b.	Merja-a	Pekka	halu-si	ihail-la (oSVv, #16)	
	Merja-PAR	Pekka.NOM	want-PST.3SG	admire-A/INF	
	o	S	V	v	
c.	Pekka	halu-si	Merja-a	ihail-la. (SVov, #18)	
	Pekka.NOM	want-PST.3SG	Merja-PAR	admire-A/INF	
	S	V	o	v	

Canonical verb-initial clauses are ungrammatical in Finnish (18b), as correctly judged by the model.[9] A preverbal object (19a,b) is always interpreted as the marked topic, and is correctly interpreted as such by the model. If both the object and subject are fronted, they are interpreted as topics (19b). Sentence internal fronting is registered as creating secondary topics (19c), but whether this is semantically correct is difficult to judge. Example (20)

[9] (18b) is grammatical if the finite verb is focused contrastively and stressed prosodically, in which case it is in the CP-domain (Vilkuna, 1989, 1995).

illustrates how the model calculates multitopic constructions (19b). Both topics are reconstructed, based on their overt case forms, to the correct thematic positions.

(20)



Ungrammatical case combinations, correctly ruled out, are sentences #45-79, 94-125 in the test corpus. Some of these sentences are illustrated in (21). All possible case combinations were tested.

(21)	a.	*Merja	ihaile-e	Merja. (#45, 48)
		Merja.NOM	admire-PRS.3SG	Merja.NOM/ACC(0)
	b.	*Merja	ihaile-e	Merja-n. (#46, 47)
		Merja.NOM	admire-PRS.3SG	Merja-GEN/ACC(N)
	c.	*Merja-a	ihaile-e	Merja-a. (#49)
		Merja-PAR	admire-PRS.3SG	Merja-PAR

When a wrong case form is encountered, reconstruction is always attempted, hence we must make sure that there are no spurious reconstruction sites. In the case of (21a), for example, the model reconstructs both arguments to SpecvP where their nominative case features are checked by finite T, but correctly rejects this configuration because only one argument can be theta-marked at this position and *admires* lacks an obligatory patient argument. The result is rejected at the syntax-semantics interface. I tested also sentences that had too many or too few case-marked arguments (#38-44, 81-86), which the model correctly classified as ungrammatical. Some examples are provided in (22).

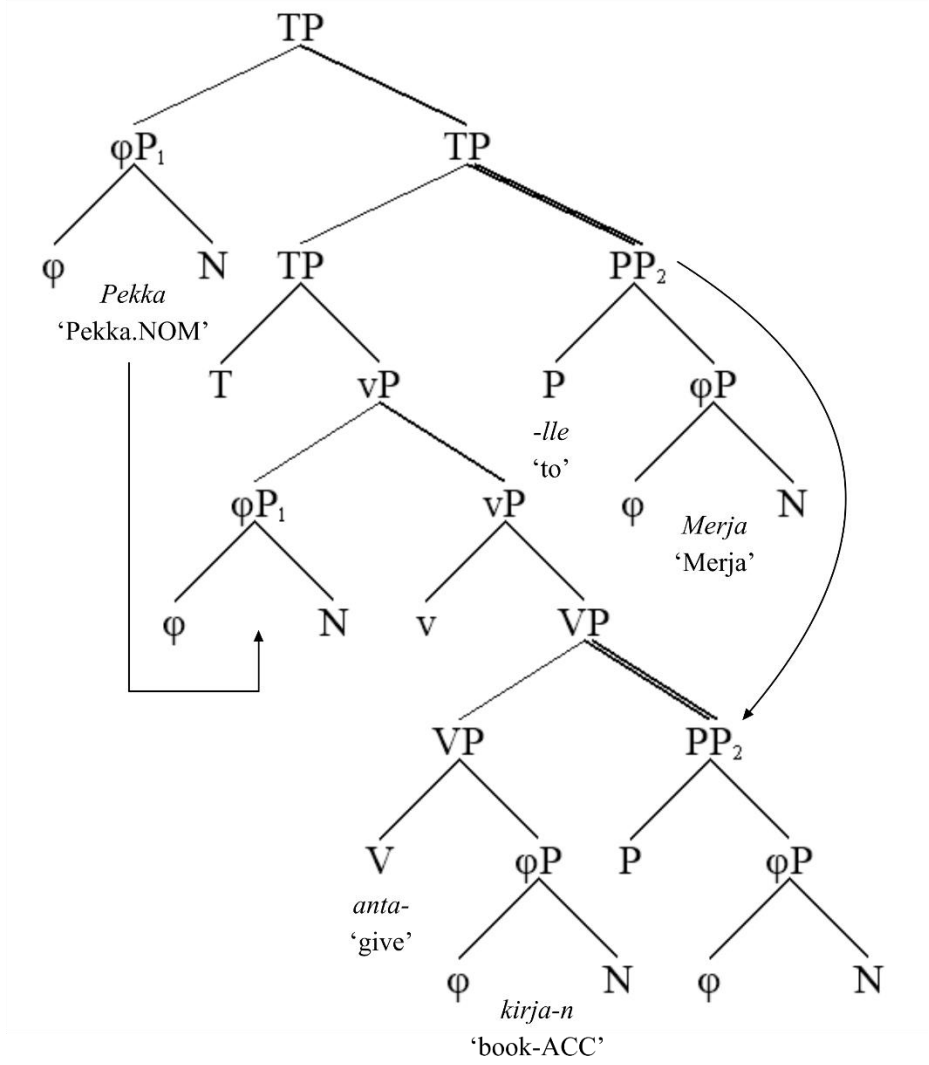
(22) a. *Merja-a/ *Merja-n nukku-u. (#38, 39)
Merja-PAR Merja-GEN/ACC(N) sleep-PRS.3SG

b. *Merja/ *Merja ihaile-e. (#40, 41)
Merja.NOM Merja.ACC(0) admire-PRS.3SG

c. *Pekka Pekka nukkuu. (#80)
Pekka.NOM Pekka.NOM sleep-PRS.3SG

Finnish semantic cases, although not in the focus in this study, deserve a comment. Example (23) illustrates how the model analyses ditransitive clauses such as *Pekka antoi kirjan Merjalle* ‘Pekka.NOM gave book-ACC(N) Merja-ALL’ (#6) that contains an allative argument *Merja-lle* ‘to Merja’. Allative is one of the Finnish semantic cases.

(23)



This analysis follows Nikanne (1993), who proposed that Finnish semantic cases such as the allative are checked by a phonologically covert preposition. The preposition selects for a φP complement. The parser attached the preposition phrase to a high right position and then reconstructed it into a lower position inside the VP.[10] The reconstruction mechanism is almost

[10] Labeling shows that the model treats the rightward PP argument as an adjunct (adjunct attachment is marked by the double line in the phrase structure images generated by the algorithm). Adjuncts are analyzed as geometrical constituents of the hosting phrase structures that are invisible for labeling and several other grammatical dependencies taking place in the hosting structure, being "pulled out" into a secondary processing pipeline (Brattico, 2020a). They increase the dimensionality of the phrase structure geometry by adding independent 'syntactic planes' to the representation. For technical implementation, see Brattico (2019a: 4.6, 7.1.2, 7.2.3).

identical to the one that reconstructs thematic arguments: P must be linked with a lexical feature (or several) that controls its syntactic distribution and semantic interpretation.

Finnish licenses partitive subjects in an experiencer construction (24).

(24) Pekka-a pelo-tta-a. (#140)
Pekka-PAR fear-CAU-PRS.3SG
'Pekka feels frightened.'

Partitive arguments cannot be licensed at the subject position according to the analysis proposed in this article, so the sentence looks problematic. The algorithm nevertheless accepts these sentences and reconstructs the partitive subject inside the VP (25).

(25) [$_{TP}$ Pekka-a$_1$ [$_{TP}$ T [$_{VP}$ Cau^0 [$_{VP}$ __$_1$ pelkää-]]]] (#139)
Pekka-PAR -tta- fear-
'(Something, not mentioned) causes Pekka to fear.'

The causative morpheme has +ARG and −VAL checking the partitive. The calculated output agrees with the style of analysis proposed by Huhmarniemi (2019b, 2019c) and Pylkkänen (2002). This construction was also tested with ungrammatical case configurations (#139-143). In general, partitive preverbal subjects cannot be reconstructed to SpecvP if finite T is encountered locally in the path; they must have a lower reconstruction site.[11]

3.5.2 Accusative

All accusative case forms (t-accusative, n-accusative and the zero-accusative) presuppose that the verb or deverbal phrase containing the case marked direct object has a specific aspectual interpretation. The data is repeated in (26).

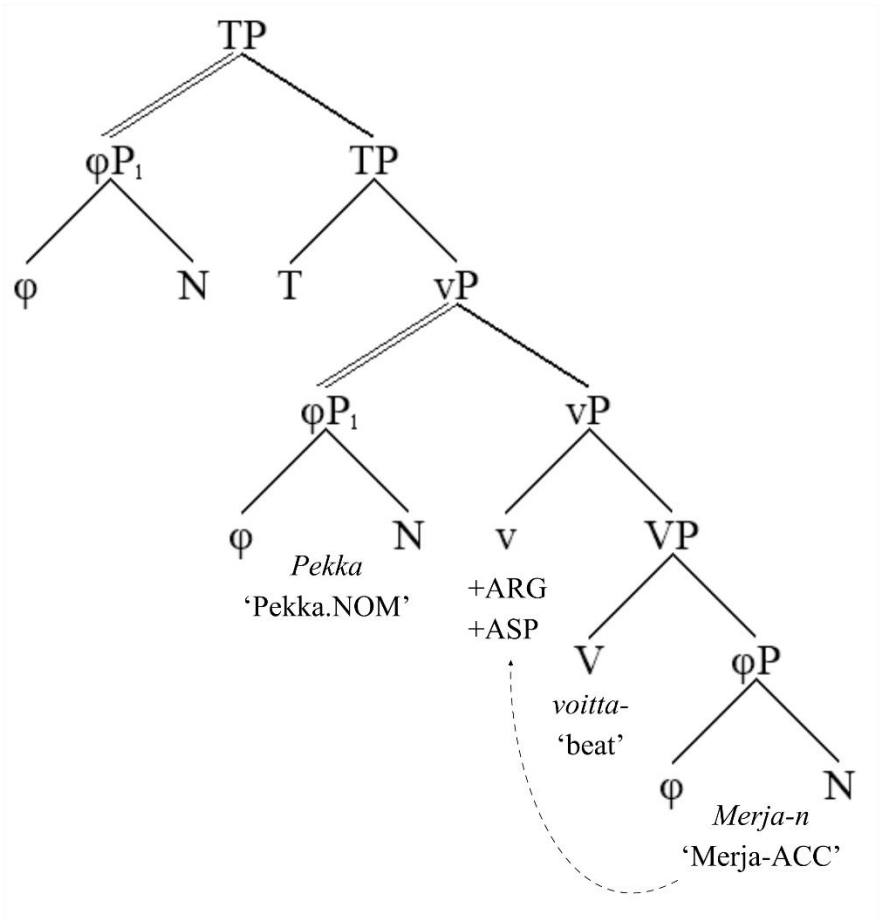
(26) a. Pekka pes-i hevos-en. (#166)
Pekka.NOM wash-PST.3SG horse-ACC(N)
'Pekka washed the (whole) horse.'

b. Pekka pes-i hevos-ta. (#167)
Pekka.NOM wash-PST.3SG horse-PAR
'Pekka washed the horse (but the horse did not necessarily become clear).'

[11] The model interpreted the sentence as a reflexive 'Pekka caused himself to fear' (see line 1600 in the results file). This problem, irrelevant here, was left for a future study.

Presence of the accusative direct object correlates with an interpretation where the whole horse was washed and washing reached an endpoint. It was assumed that the accusative is checked by ASP:BOUNDED that is part of the verb or some verbal head (or perhaps all verbal heads). Aspectually non-telic verbs such as *tönäistä* 'to nudge', which accept only the partitive when occurring without further modifiers, do not have ASP:BOUNDED, correctly rendering the accusative ungrammatical in the dataset.[12] If the verb is ambiguous, the feature may be present or absent, which was handled by lexical ambiguity. These assumptions are illustrated in (27), which shows how the algorithm calculates *Pekka voitti Merja-n* 'Pekka.NOM beat Merja-ACC(N)' (#163). I abbreviate ASP:BOUNDED as +ASP. +ASP is part of the lexical entry of the verb 'win/beat'.

(27)



[12] A telic reading is possible if a further modifier is added to the VP, as in *Pekka tönäisi miehe-n päin seinää* 'Pekka pushed/nudged man-ACC(N) against wall'. Aspectual properties of the sentence depend on the VP as a whole. Due to the lack of a more general theory of aspect, "readings" were implemented in this study by relying on lexical ambiguity.

Since the accusative rule requires checking by +ARG and +ASP, the dependency cannot look past any head with +ARG. This causes partial match and intervention. If the accusative occurred together with a non-telic verb, lacking +ASP, the presence of +ARG would terminate checking and lead into rejection (see #171). This captures locality. I have assumed above that the relevant features are part of the small verb v. Case marking of adverbials (Section 3.5.5) suggests that both features can be part of V and T; perhaps all verbal constituents can and do host aspectual features. When the accusative case form is matched with an aspectually unbounded verbal element, the results file contains an aspect field which reads "Aspectually bounded."[13] This records the fact that the model interpreted the construction as denoting an aspectually bounded event. This information is calculated in the semantic system.

All verb types were successfully tested, both with grammatical and ungrammatical case forms (#163-180), with some examples shown in (28).

(28) a. Pekka pes-i Merja-n/ Merja-a. (#166, 167)
Pekka.NOM wash-PST.3SG Merja-ACC(N) Merja-PAR
(Ambiguous, both interpretations possible.)
b. Pekka tönäi-si *Merja-n/ Merja-a (#171, 165)
Pekka.NOM push-PST.3SG Merja-ACC(N) Merja-PAR
(Telic reading not possible.)

Achievement verbs such as (29) constitutes a problem, however.

(29) Pekka voitt-i Merja-n/ *Merja-a. (#163, 164)
Pekka.NOM beat-PST.3SG Merja-ACC(N) Merja-PAR
'Pekka beat Merja.'

Rule PAR ~ +ARG, −VAL does not refer to aspect and ignores the fact that 'win/beat' must contain a culmination point. The model is unable to rule out partitive direct objects in connection with aspectually bounded events. Sentence (29) with the partitive direct object has a coerced or anomalous reading in which 'an event that occurred in an instant is ongoing'. This should be judged ungrammatical, in my view. There is, however, a second reading analogous to 'Pekka won money', which translates into something like 'Pekka won part of Merja, such as a piece of her hair'. The sentence is grammatical under this interpretation. The

[13] For the sentence #163, example (27), this information can be found the line 2665 in the results file.

problem is how the algorithm could predict what the intended reading is while it is still parsing. It seems, moreover, that the difference between *Pekka voitti raha-a/Merja-a* 'Pekka won money-PAR/Merja-PAR' can only be established once the whole sentence has been parsed and further evaluated in relation to a larger context by accessing knowledge of the world (e.g., what is money, who is Merja, and so on). I therefore propose that the partitive case checking mechanism is not sensitive to aspect, but the semantic system is. Specifically, when the partitive is connected by an upward path to a lexical item marked for ASP:BOUNDED, the semantic component alerts the language-external systems that (29) is possibly aspectually anomalous. This information occurs in the aspect field in the results file which reads "aspectually anomalous."[14] This means that we account for the contrast (29) in the semantic component.

There are two empirical arguments supporting this hypothesis. First, the partitive can occur with an achievement verb if the object is in the plural (30).

(30) Pekka voitt-i kilpailu-i-ta.
Pekka.NOM win-PST.3SG competition-PL-PAR
'Pekka won competitions.'

This is interpreted to mean that Pekka won several competitions. The sentence is grammatical. In addition, it is well-known that the aspectual properties of the sentence depend on the properties of the whole VP, not just on the verb, which suggests that at least some aspect computations target the parsed output structure. This motivates further the hypothesis that some aspecutal computations take place in the semantic component that has access to the larger context. See Kiparsky (1998) for discussion.[15]

Let us consider polarity and agreement. The accusative constitutes a positive polarity case and is ungrammatical inside negative polarity contexts (31).

(31)

a. Pekka e-i voitta-nut *Merja-n/ Merja-a. (#184, 181)

[14] For sentence #164, example (29b), this information can be found from the line 2699 in the results file. The matter is slightly more complex due to the fact that the presence of the negation higher up in the clause neutralizes the semantic anomaly effect. The algorithm takes this complication into account.

[15] An alternative is to assume that ASP represents a general verbal-aspectual feature that licenses the accusative and only biases the interpretations towards telic readings. Whether the event is interpreted as bounded or not is created in the semantic component by synthesizing information from multiple sources.

Pekka.NOM not-3SG beat-PST.PRTCPL Merja-ACC(N) Merja-PAR
'Pekka did not beat Merja.'

b. Pekka e-i halun-nut voitta-a *Merja-n/ Merja-a.
Pekka.NOM not-3SG win-PST.PRTCPL beat-A/INF Merja-ACC(N) Merja-PAR
'Pekka did not want to beat Merja.'

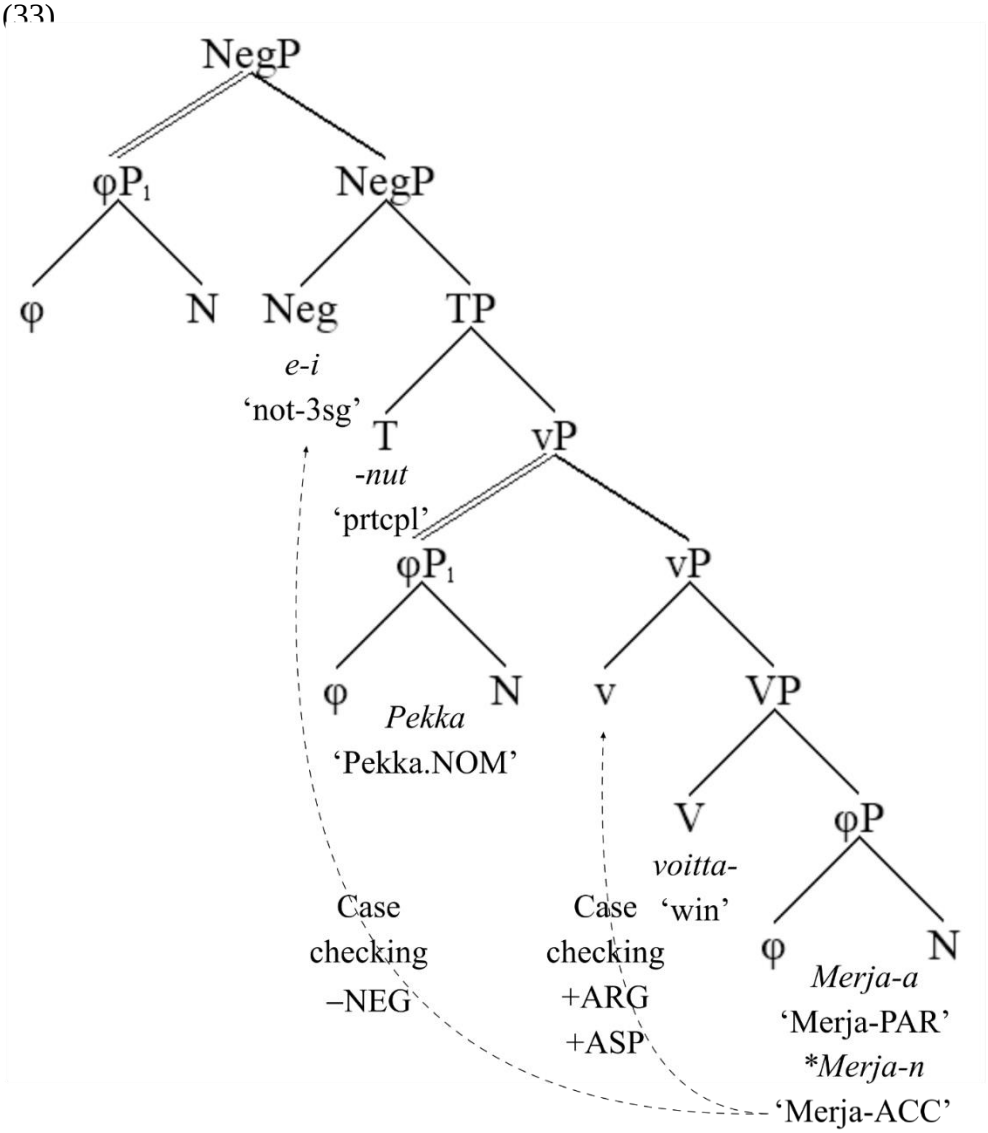
The mechanism checks the accusative case against −NEG (i.e. that the path does not contain the negative polarity feature). The test sentences are #182-194 in the test corpus, probing grammatical and ungrammatical case combinations with and without noncanonical word orders. A few examples are provided in (32).

(32) a. *Pekka e-i voitta-nut Merja. (#182, 185)
Pekka.NOM not-3SG win-PST.PRTCPL Merja.NOM/ACC(0)

b. *Pekka e-i voitta-nut Merja-n. (#183, 184)
Pekka.NOM not-3SG win-PST.PRTCPL Merja-GEN/ACC(N)

Crucially, since only −NEG is checked, partial feature match is impossible, and the dependency becomes nonlocal. This is illustrated by (33), which shows the phrase structure analysis and case checking dependencies generated for (32a).

(33)



Aspect (+ASP, +ARG) and polarity (–NEG) are checked separately by v and Neg. Some word order variations, affecting information structure, were also tested (34).

(34) a.	Merja-a	e-i	voitta-nut	Pekka. (#186)	
	Merja-PAR	not-3SG	win-PST.PRTCPL	Pekka.NOM	
	O	Neg	V	S	
b.	Merja-a	e-i	Pekka	voitta-nut. (#187)	
	Merja-PAR	not-3SG	Pekka.NOM	win-PST.PRTCPL	
	O	Neg	S	V	
c.	Pekka	e-i	Merja-a	voitta-nut. (#188)	
	Pekka.NOM	not-3SG	Merja-PAR	win-PST.PRTCPL	
	S	Neg	O	V	

d.	Merja-a	Pekka	e-i	voitta-nut. (#189)
	Merja-PAR	Pekka.NOM	not-3SG	win-PST.PRTCPL
	O	S	Neg	V
e.	Pekka	Merja-a	e-i	voitta-nut. (#190)
	Pekka.NOM	Merja-PAR	not-3SG	win-PST.PRTCPL
	S	O	Neg	V

I also tested combinations of noncanonical word orders and ungrammatical case forms (#191-194).[16]

Let us consider agreement. The nominative-looking zero-accusative (ACC(0)) is grammatical if the direct object is not c-commanded by an agreeing predicate, whereas the genitive-looking n-accusative (ACC(N)) occurs when there is overt agreement, an analysis that goes back to Timberlake (1975) and Reime (1993). These assumptions cover the following data (#195-211):

(35)

a.	Me	löys-i-mme	avaim-en/	*avain.	
	we.NOM	found-PST-1PL	key-ACC(N)	key.ACC(0)	
	‘We found the key.’				
b.	Me	löydet-tiin	*avaim-en/	avain.	
	we.NOM	found-PST.IMPASS	key-ACC(N)	key.ACC(0)	
	‘We found the key.’				
c.	Pekka	tönäis-i	*Merja-n/	Merja-a/	*Merja.
	Pekka.NOM	push-PST.3SG	Merja-ACC(N)	Merja-PAR	Merja.ACC(0)
	‘Pekka pushed Merja.’				
d.	Me	e-i	löydet-ty	avain-ta/	*avaim-en/	*avain.
	we.NOM	not-3SG	find-PST.PRTCPL	key-PAR	key-ACC(N)	key-ACC(0)
	‘We did not find the key.’					

[16] Crossing all possible word orders with all possible case forms creates 600 test sentences that must be tested in a separate study focusing on this specific issue. Thus, the present model was not verified exhaustively when it comes to more complex sentences, such as the infinitival constructions discussed in the main text.

Presence of the specific accusative forms ACC(0) and ACC(N) require that ±PHI is checked. Because the checking relations are based on one feature, long-distance effects are also captured (#212-229)(36).

(36)

a. Me halut-tiin [voitta-a Merja / *Merja-n.] (#216, 226)
we.NOM want-PST.IMPASS win-A/INF Merja.ACC(0) Merja-ACC(N)
'We wanted to win Merja.'

b. Me halus-i-mme [voitta-a *Merja / Merja-n.] (#227, 218)
we.NOM want-PST-1PL win-A/INF Merja.ACC(0) Merja-ACC(N)
'We wanted to win Merja.'

(37) Me e-i haluttu [voitta-a *Merja / *Merja-n / Merja-a.]
We.NOM not-3SG want win-A/INF Merja.ACC(0) Merja-ACC(N) Merja-PAR
'We did not want to win Merja.' (#228, 229, 217)

The checking mechanism explores the structure until either one of the relevant features ±NEG/±PHI is encountered or there is no more structure.[17] Uncanonical words orders were also tested (#219-225).

Pronouns, which take an unambiguous *t*-suffix in these contexts, were correctly judged and analyzed. Plural direct objects are assigned the t-accusative forms. These facts are captured in the lexicon: accusative pronouns and plural DPs (*häne-t* 'he-ACC(T)', *auto-t* 'car-PL.ACC(T)') map into ACC(T) that does not require checking against ±PHI. However, polarity and aspect are still relevant (38).

(38) a. Pekka pes-i hän-et. (#168)
Pekka.NOM wash-PST.3SG he-ACC(T)
'Pekka washed him.'

[17] The current rules allow polarity and agreement-based accusative rules to penetrate finite clause boundaries. This could be prevented by stipulating a finite clause restriction to the upward path mechanism, positing polarity and agreement features to the finite clause boundary where they would cause an intervention, or by relying on the phase theory (Chomsky, 2000, 2001) which stipulates locality domains, CP among them. I excluded this issue from the present study, however, because as a matter of fact the negation effect does penetrate the finite clause boundary. One example is *Pekka ei uskonut että Merja voittaa kilpailu-a/kilpailu-n* 'Pekka not believe that Merja wins competition-PAR/competition-ACC', where the negation in the main clause licenses the partitive inside the embedded finite clause.

b. *Pekka tönäis-i hän-et. (#172)
Pekka.NOM push-PST.3SG he-ACC(T)

c. *Pekka e-i voitta-nut hän-et.
Pekka.NOM not-3SG win-PST.PRTCPL he-ACC(T)

d. Me löydet-tiin hän-et.
we.NOM found-PST.IMPASS he-ACC(T)
'We found him.'

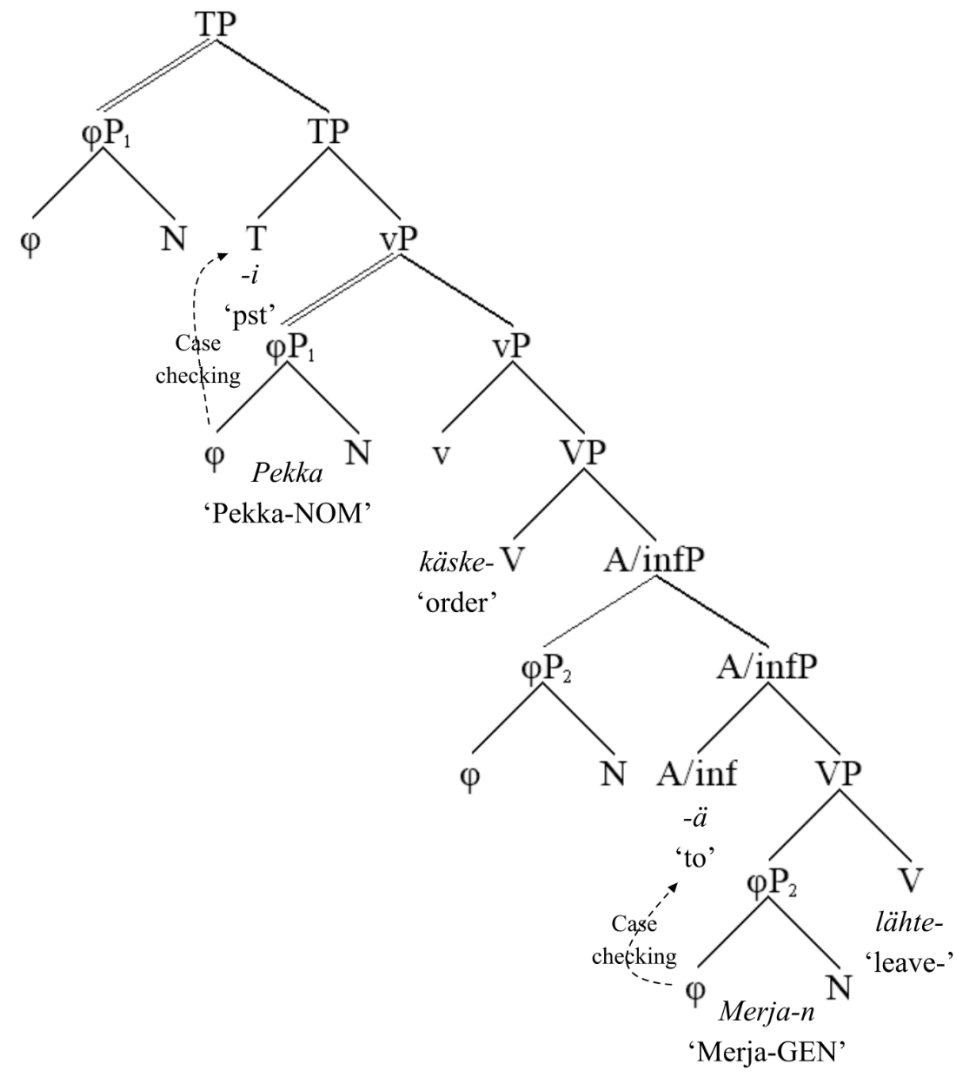
3.5.3 Genitive

Vainikka (1989, 1993, 2011) showed that the Finnish genitive behaves like a default specifier case. This generalization is quite powerful, but incompatible with the analysis proposed here. We derive it by assuming that final case checking is applied after reconstruction. Brattico (2020a) applied this analysis to the A-infinitival (39).

(39) Pekka käsk-i [Merja-n lähte-ä.] (#230)
Pekka.NOM order-PST.3SG Merja-GEN leave-A/INF
'Pekka ordered Merja to leave.'

In that study, the genitive was associated with −FIN that was checked against the infinitival head A/inf *-(t)A-* after the argument reconstructed to SpecVP. Rule GEN ~ +ARG, −FIN calculates essentially the same output.

(40)



The preverbal subject is reconstructed to SpecvP, where it checks the nominative, while the infinitival subject reconstructs to SpecVP inside the infinitival where it receives a thematic role. The genitive is checked against the A/inf head. The relevant test sentences are #230-242, which cover grammatical and ungrammatical case forms together with noncanonical word orders. Some of these sentences are illustrated in (41).

(41) a. Pekka sano-i Merja-n lähte-vän. (#231)
Pekka.NOM say-PST.3SG Merja-GEN leave-VA/INF
'Pekka said that Merja would leave.'

b. *Pekka käsk-i Merja/Merja-a lähte-ä. (#235, 236)
Pekka.NOM order-PST.3SG Merja.NOM/Merja-PAR leave-A/INF

c. *Pekka käsk-i vieraa-t lähte-ä.
Pekka.NOM order-PST.3SG visitor-PL.ACC(N) leave-A/INF

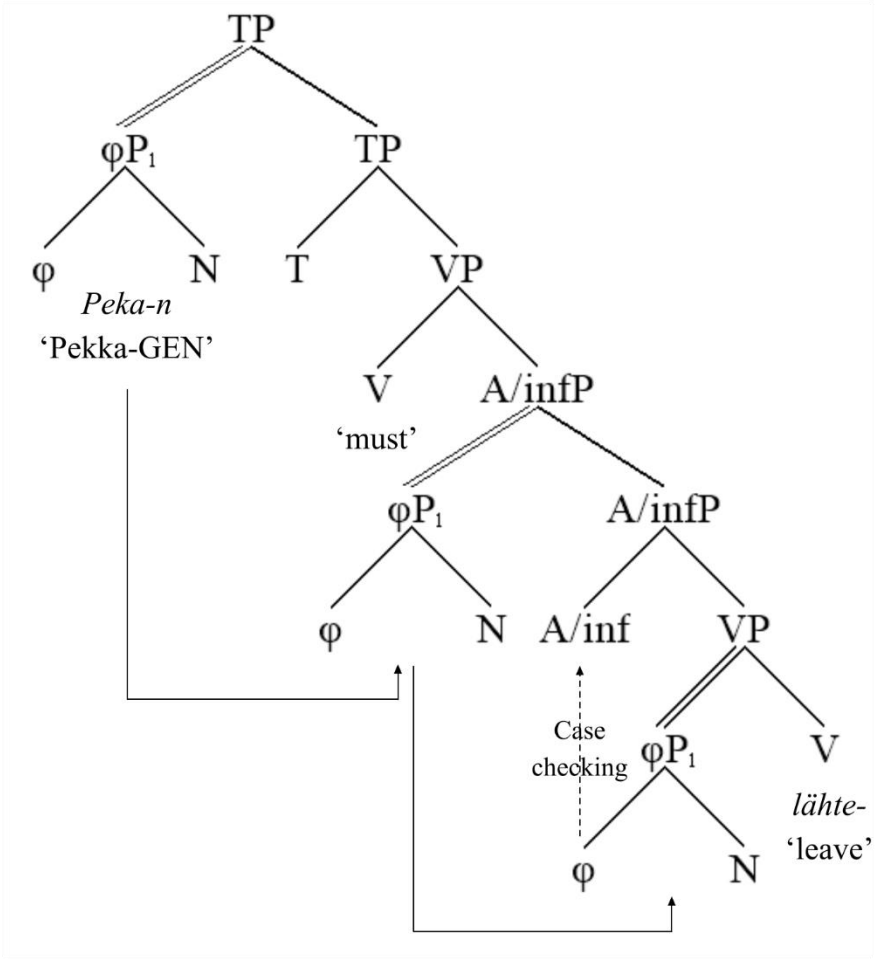
The assumption that the genitive subject reconstructs into VP is not ad hoc: it links the argument with a thematic role inside the VP.

The Finnish modal construction (42) seems at first to violate the genitive rule GEN ~ −FIN, +ARG. The genitive is in the subject position of the finite modal verb, which cannot, by the analysis proposed here, check it. Finite verbs can only check the nominative case.

(42) Peka-n täyty-y lähte-ä. (#232)
Pekka-GEN must-PRS.0 leave-A/INF
'Pekka must leave.'

The model judges these sentences correctly as grammatical and returns (43). The genitive argument is reconstructed successively-cyclically to the SpecVP inside the A-infinitival, where it checks the genitive at the lowest position of the three-member chain.

(43)



This structure cannot be judged outright implausible, since the genitive argument is the thematic subject of the infinitival ('who is leaving') and the construction is monoclausal, sustaining thematic positions for one standard set of arguments and adverbials. The meaning is approximately 'must: Merja to leave'.

3.5.4 Noun phrase

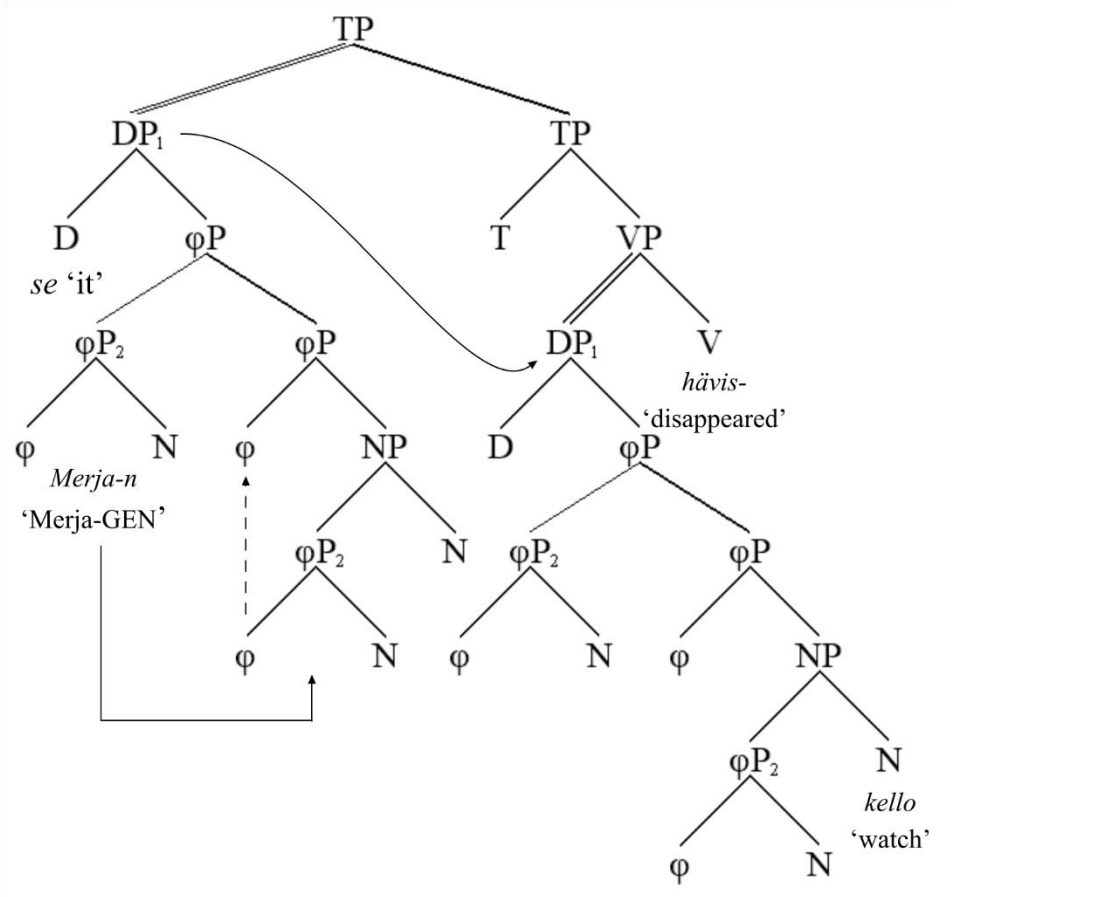
The genitive has also a possessive role (44) inside noun phrases.

(44) Peka-n avain
Pekka-GEN key
'Pekka's key'

This construction violates the GEN ~ −FIN, +ARG rule since the genitive is not governed locally by an infinitival head. In fact, there appears to be no governing head. One solution is to assume that the Finnish noun head decomposes into '*n* + root' structure (Pylkkänen, 2002;

Brattico, 2005; Brattico & Leinonen, 2009) and that the genitive is checked against *n* (*n* = nominalizer, as in *juokse-minen* ‘run-ing’, possibly a zero morpheme). Another possibility is that the genitive case is checked against φ that is part of all noun phrases. I assumed the latter in this study. The algorithm calculates (45) for the input sentence *se Merja-n kello hävisi* ‘the/that Merja-GEN watch disappeared’ (#243, variations 244-250).

(45)



The possessor *Merja-n* ‘Merja-GEN’ reconstructs to NP, where it checks the genitive case against φ.

Adpositions were tested both with grammatical and ungrammatical case configurations. The basic cases are (46a-b), where the argument occurs either in the complement position of the adposition (46a) or as its specifier (46b).

(46) a. lähellä Pekka-a. (#125)
near Pekka-PAR
‘near Pekka’

b. Peka-n lähellä. (#127)

Pekka.GEN near
'near Pekka'

The model reconstructs the genitive argument to CompPP, where the preposition checks its case. It looks contradictory that the preposition checks both the partitive and genitive, but on closer inspection we find that two lexical items are at stake. First, the adposition assigning the genitive exhibits overt phi-agreement, while the adposition assigning the partitive does not:

(47) a. *lähellä-ni minu-a/ b. minun lähellä-ni.
near-PX/1SG I-PAR I.GEN near-PX/1SG
'near me' 'near me'

Second, the agreeing form requires that the genitive argument occurs at its specifier position (48a); this is not true of the non-agreeing form (48b).

(48) a. minu-n lähellä/ *lähellä minu-n. (#138)
I-GEN near near I-GEN
'near me'
b. minua lähellä/ lähellä minua (#126, 125)
I.PAR near near I.PAR
'near me' 'near me'

This means that the agreeing adposition must have an additional EPP property forcing the genitive argument to SpecPP.[18] In sum, adpositions checking the genitive case have the agreement/EPP profile, while adpositions checking the partitive do not. Ungrammatical case configurations were tested by #128-138.

One feature that complicates the analysis of the Finnish noun phrase is the behavior of numerals. Finnish cardinal numerals fall into two paradigms. The first contains bare singular numerals that assign the partitive inside the hosting noun phrase. These facts follow if these numerals have +ARG and −VAL. The model calculates (49)(#144, variations 145-149).

[18] EPP stands for "extended projection principle" and was originally used to refer to the requirement that the English finite clause must have a grammatical subject (Chomsky, 1981, 1982). The notion was later generalized so that it can denote any head projecting an obligatory extended nonthematic specifier. The syntactic background theory used in this study models it as an unspecific and nonthematic specifier lexical selection feature [SPEC:*].

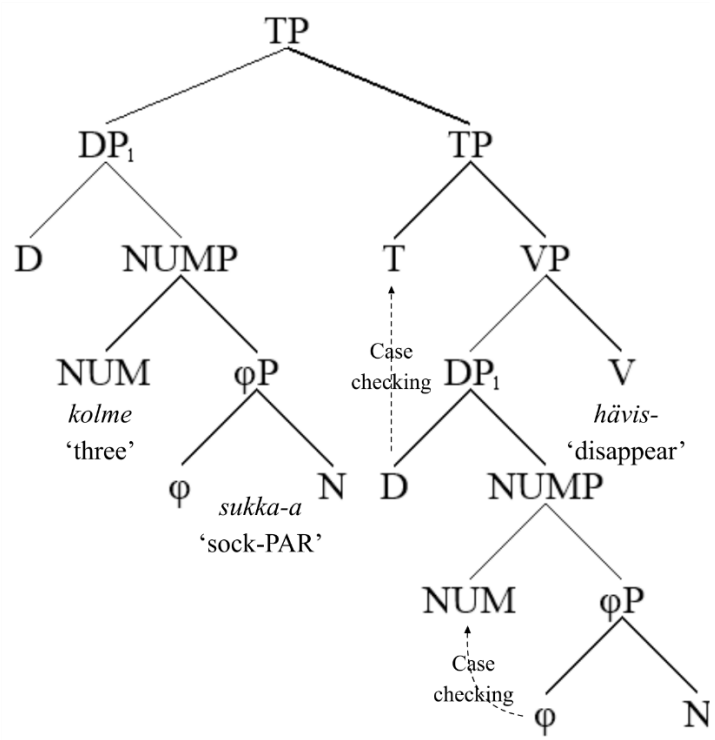
(49) kaksi sukka-a
two.SG.0 sock.SG-PAR
'two socks'

Numerals in the second group inflect like adjectives (or are adjectives) and do not assign cases to the elements inside their syntactic scope (50) (#145).

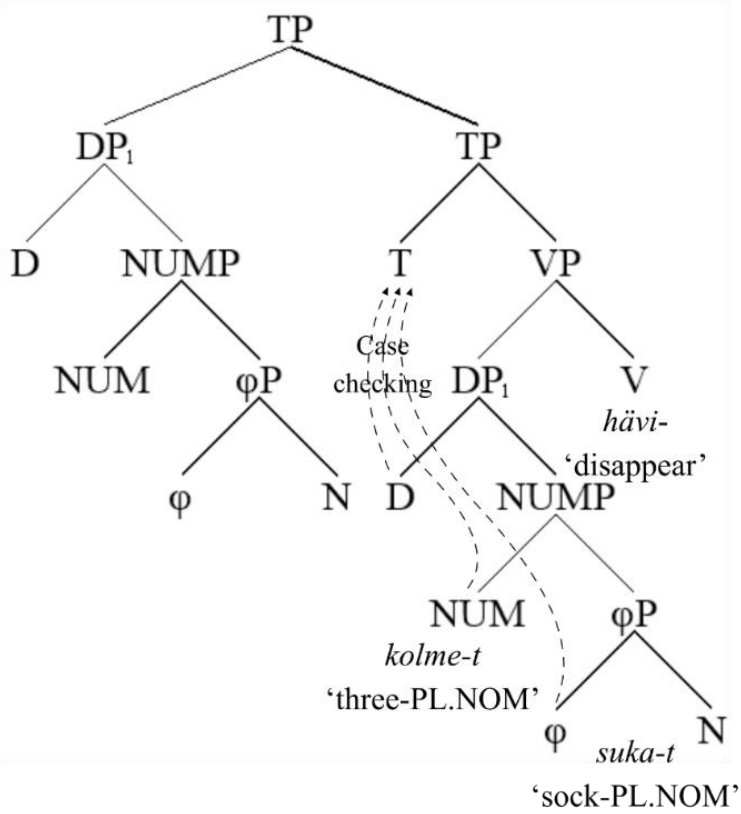
(50) [Ne kahde-t suka-t/ *sukka-a] hävis-i-vät.
those.NOM two-PL.NOM sock-PL.NOM sock-SG.PAR disappear-PST-3PL
'Those two pairs of socks (=4 socks) disappeared.'

If these numerals do not have +ARG, then case checking ignores them and finds the relevant assigners from the main clause (e.g., T, v)(#144-147). These properties were correctly derived. Example (51) illustrates the first group, (52) the second.

(51)



(52)



In the example (51), the nominative at D is checked by the head in the hosting (main) clause, here finite T, while the case at the noun head is checked by the numeral. This results in the discontinuous case profile. In (52), all nominal words check their cases against a local (main) clause element, resulting in the homogeneous case pattern. The distinction between case concord and case assignment disappears: all nominal words check their case features independently. What has made the behavior of the numerals subject to some debate in the previous literature is the fact that the numerals in the first group only occur in contexts where the hosting DP is assigned either the nominative or the accusative case. If the DP is assigned either the genitive or any of the lexico-semantic cases, the numeral-partitive pattern disappears (53).[19]

[19] Recall that for singular full DPs the genitive and the n-accusative are homophonous, which has led some researchers to analyze them as the same genitive case. Accordingly, the genitive becomes a direct object case for a restricted class of object DPs. As pointed out in the main text, however, the accusative and the genitive behave differently in the numeral context: the latter, but not the former, neutralizes the numeral-partitive pattern. This motivates in part the assumption that two different cases are at stake.

(53)

a.	Pekka	usko-i	[kahde-n	suka-n	hävin-neen.] (#148)
	Pekka.NOM	believe-PST.3SG	two.SG-GEN	sock.SG-GEN	disappear-PST.VA/INF
	’Pekka said that the two socks had disappared.’				
b.	*Pekka	usko-i	[kaksi	sukka-a	hävin-neen.] (#158)
	Pekka.NOM	believe-PST.3SG	two.0	sock-PAR	disappear-PST.VA/INF

Brattico (2010, 2011), following a tradition in the Slavic linguistics that exhibits a similar phenomenon (e.g., Babby, 1987), analyzed this pattern by relying on case competition. The “weak cases” accusative and nominative are outperformed by the “strong cases” genitive and the lexico-semantic cases. A case competition analysis for Finnish structural case assignment was also presented by Nelson (1998). However, these data follow if we assume that the bare numerals exhibit an ambiguous NOM/ACC case form. This prevents them from appearing in any other context and derives (53) and correctly rules out constructions with the wrong case forms (#148-153, 257-260). Ungrammatical word orders (54) are also correctly ruled out, with the exception of (54d) that the model wrongly accepts. I was not able to solve the root problem with (d). The analysis predicts a phantom reconstruction.

(54) a.	*Sukka-a$_1$	ne	kaksi	_$_1$	hävisi. (#160)	
	sock-PAR	those	two		disappeared	
b.	*Kaksi$_1$	ne	_$_1$	sukka-a	hävisi. (#161)	
	two	those		sock-PAR	disappeared	
c.	*[Kaksi	sukka-a]$_1$	ne	_$_1$	hävisi. (#162)	
	two	sock-PAR	those		disappeared	
d.	*Ne	sukka-a$_1$	kaksi	_$_1$	hävisi. (#162, wrongly accepted)	
	those	sock-PAR	two		disappeared	

3.5.5 Special constructions

Some special constructions that have played a major role in Finnish case theories were added to the dataset. The impersonal passive construction, shown again in (55a), is one.

(55)

a.	Me	löydet-tiin	avain/	*avaim-en.
	we.NOM	found-PST.IMPASS	key.ACC(0)	key-ACC(N)
	‘We found the key.’			

b. Me löys-i-mme *avain/ avaim-en.
we.NOM found-PST-1PL key.ACC(0) key-ACC(N)
'We found the key.'

This effect is captured by the rule which associates the two accusative forms with phi-agreement (±PHI). The impersonal passive form (*löydet-tiin* 'found-PST.IMPASS') is created by a special impersonal functional head replacing standard v, following the analysis proposed by Manninen & Nelson (2004).

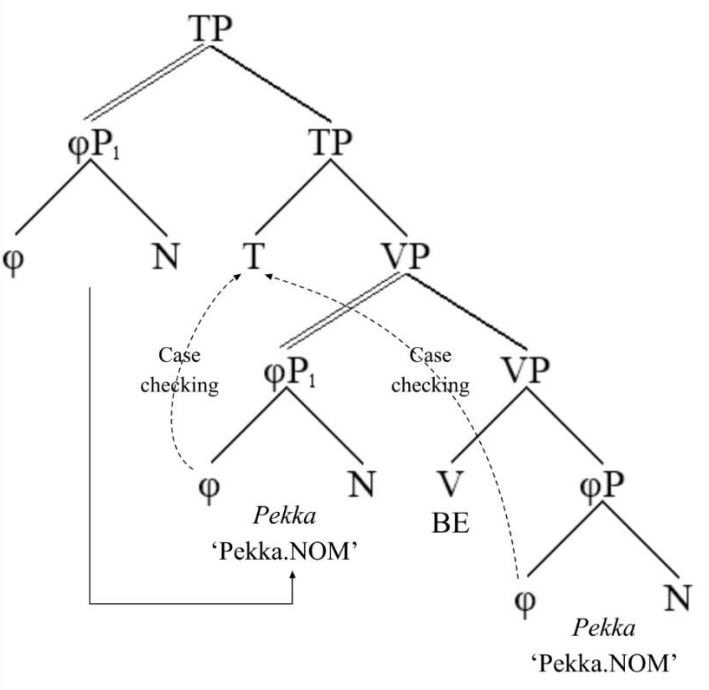
(56)
a. Löydet-tiin avain.
find-PST.IMPASS key.ACC(0)
'A key was found (by a collection of people).'
b. Avain löydet-tiin.
key.ACC(0) found-PST.IMPASS
'The key was found (by a collection of people).'

The test sentences are #260-265 in the test materials.

Copular constructions (57) were correctly judged as grammatical and calculated as (58).

(57) Pekka o-n Pekka. (#271)
Pekka.NOM be-PRS.3SG Pekka.NOM
'Pekka is Pekka.'

(58)



The grammatical subject reconstructs to SpecVP, the complement of the copular verb remains in situ. Both check their nominative cases by T. The test sentences are #271-280. Copular and predicative constructions have complex and controversial properties that have never been studied systematically by using the underlying recognition grammar and were not therefore examined here.

Finnish has one possible raising construction (59a), which should be compared to the non-raising variant (59b).

(59) a. Merja$_1$ näyttä-ä __$_1$ lähte-vän. (#266)
Merja.NOM seem-PRS.3SG leave-VA/INF
'Merja seems to be leaving.'

b. Pekka näk-i Merja-n lähte-vän.
Pekka.NOM see-PST.3SG Merja-GEN leave-VA/INF
'Pekka saw Merja leaving.'

The model cannot reconstruct the preverbal subject to the infinitival clause SpecVP because it cannot check the nominative at that position. The position is associated with the genitive (59b). If we assume that 'seem' projects a thematic role, then the subject stops at its specifier

position and the thematic agent for ‘leave’ will be determined by control,[20] effectively making (59) an obligatory control (OC) construction and not a raising structure. Although the existence of this derivation makes the model observationally adequate (#266-270), I judge that the analysis is not obviously correct and an anonymous *SL* reviewer concurs. The model does not allow an argument to change its case during reconstruction, which prevents it from reconstructing a nominative marked argument (59a) into a genitive position (59b).

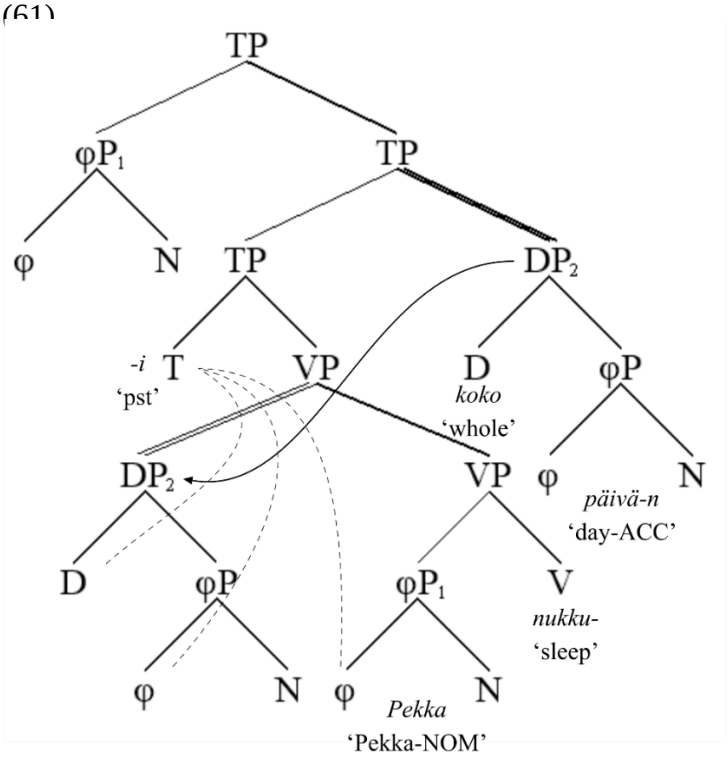
Finnish DP-adverbials can be marked with the direct object cases (60)(Maling, 1993).

(60) Pekka nukku-i [koko päivä-n.] (#281)
Pekka.NOM sleep-PST.3SG all day-ACC(N)
‘Pekka slept all day.’

Adverbial case marking follows from the proposed analytic principles because the relevant lexical features occur inside the upward paths generated from adverbs. Example (60) shows that the accusative case occurs in connection with intransitive verbs, communicating that the event has a fixed duration. The aspectual feature could be at T or V. If it is at T, the model calculates (61).

[20] See line 4916 of the result file which reports that the agent of the infinitival verb ‘to leave’ is the main clause subject. The control process responsible for this interpretation is recorded also into the derivational log file, lines 210175-210177.

(61)



The adverbial is reconstructed to SpecVP where it checks the accusative case against T. The aspectual feature is at T, creating the telic interpretation. If neither T nor V has the aspectual feature, the sentence is judged ungrammatical. Similarly, the accusative is ungrammatical if the clause is negated, grammatical if the adverbial is in the partitive (62a). Also the zero-accusative is correctly licensed (62b). These forms are correctly checked against properties of T and Neg, both which appear inside the upward path from the reconstructed adverbial.

(62)

a.	Pekka	e-i	nukku-nut	*koko	päivä-n /	koko	päivä-ä.
	Pekka	not-3SG	sleep-PST.PRTCPL	all	day-ACC(N)	all	day-PAR
	'Pekka did not sleep all day.'						
b.	Me	nukut-tiin		koko	päivä /	??koko	päivän.
	we	slept-PST.IMPASS		all	day.ACC(0)	all	day-ACC(N)
	'We slept all day.'						

The analysis succeeds in deriving the relevant pattern (#281-292) with the exception of **Pekka nukkui koko päivä-ä* 'Pekka.NOM slept all day-PAR' (#293) which the model judges wrongly as grammatical. I do not know at present how to solve this issue. Furthermore, the accusative object in (62b) is regarded as ungrammatical in the underlying test corpus, based

on my own grammaticality judgment, although this variant is sometimes used (see Anttila & Kim, 2011).

Finnish VP-fronting raises nontrivial questions. This matter came up in the present study because some of the word order permutations generated VP-fronting analyses. Since verb-initial clauses are ungrammatical in Finnish (e.g., #24-29), this phenomenon is limited to sentences that contain fronted infinitivals. Example (63) shows one.

(63) [$_{A/infP}$ ihail-la Merja-a]$_1$ halus-i __$_1$ Pekka. (#21)
admire-A/INF Merja-PAR want-PST.3SG Pekka.NOM
'To admire Merja, Pekka wanted.'

To me (63) is grammatical, and was marked as such in the test corpus, but does not seem to have any natural context of use. The fronted infinitival does not have a clear topic reading. The model analyses sentences like this by reconstructing the A-infinitival from the preverbal subject position into the complement position of 'want', and furthermore by reconstructing the postverbal grammatical subject to SpecvP. Several variations were tested. Example (64) contains a few examples (I ignore reconstruction of the grammatical subject).

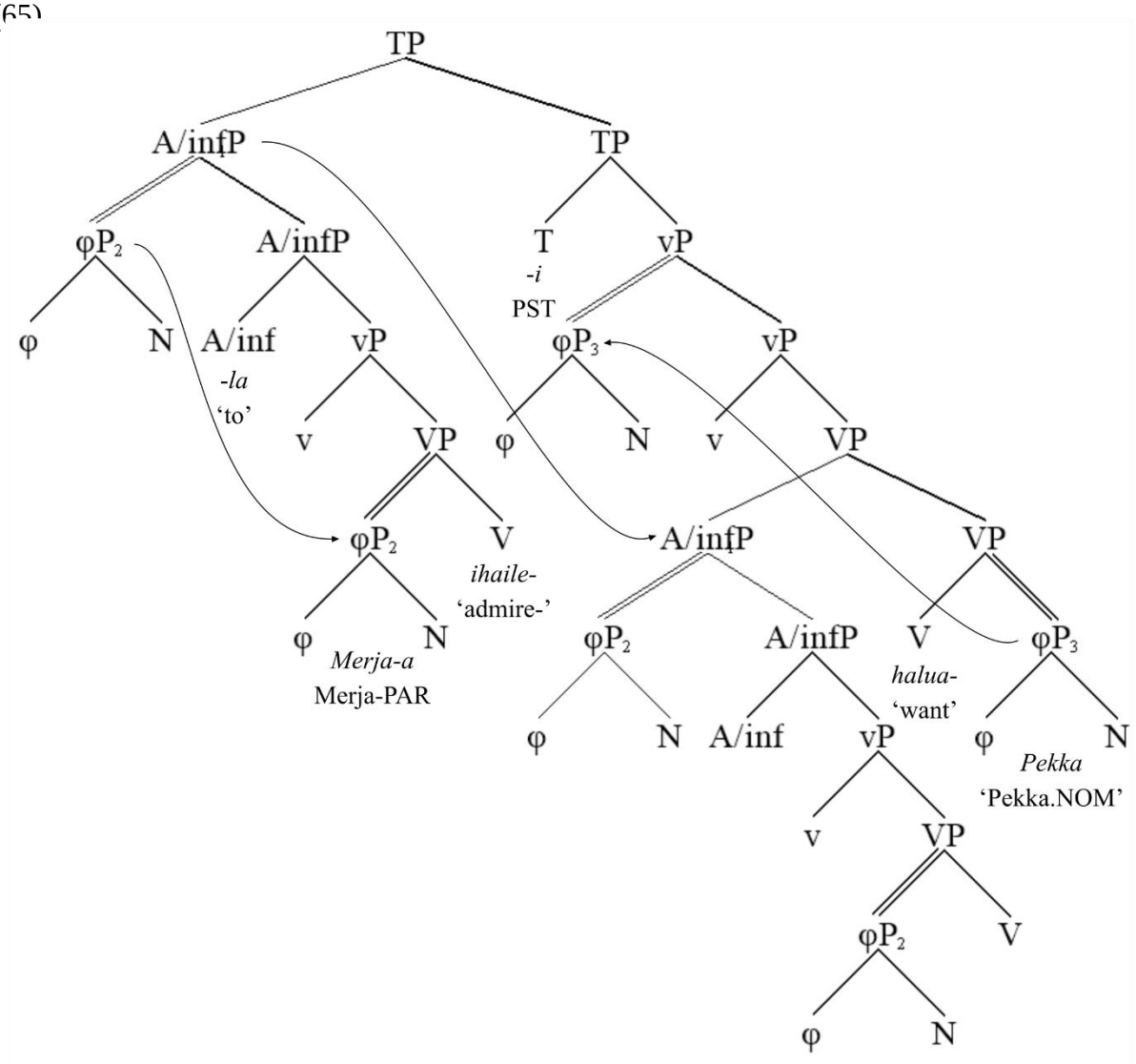
(64) a. [Merja-a$_1$ ihail-la __$_1$]$_2$ halus-i __$_2$ Pekka. (#20)
Merja-PAR admire-A/INF want-PST.3SG Pekka.NOM
o v V S

b. Pekka [Merja-a$_1$ ihail-la __$_1$]$_2$ halus-i __$_2$. (#22)
Pekka.NOM Merja-PAR admire-A/INF want-PST.3SG
S o v V

c. [Merja-a$_1$ ihail-la __$_1$]$_2$ Pekka halus-i __$_2$. (#30)
Merja-PAR admire-A/INF Pekka.NOM want-PST.3SG
o v S V

A calculated full analysis of (64a) is shown in (65).

(65)



Since fronted infinitivals do not elicit clear topic interpretations, the model was designed so that they were excluded from calculations involving information structure. This is the reason they do not appear in the marked topic and marked focus fields in the output.

4. Conclusions

Finnish structural case assignment was explored by developing a formal model that judges and analyses sentences involving the nominative, partitive, accusative (three forms) and the genitive case. The model was observationally and descriptively adequate over a representative test corpus. Some remaining problems were noted. The proposal in a nutshell is that morphological case forms are linked with a notion of syntactic scope defined by an upward path mechanism. If the surface position of the element does not satisfy its case requirements,

reconstruction is attempted. Both abstract Case and case concord were eliminated. No binary distinction was made between syntactic and semantic cases; in Finnish some case forms such as the accusative exhibit a mixed profile. Locality properties were captured by relying on feature intervention.

References

ANTTILA, A., & KIM, J.-B. 2011. On structural case in Finnish and Korean. *Lingua* 121:100–127.

ANTTILA, A., & KIM, J.-B. 2017. Locality and variation in Finnish structural Case. *Natural Language & Linguistic Theory* 35:581–634.

BABBY, L. 1987. Case, prequantifiers, and discontinuous agreement in Russian. *Natural Language and Linguistic Theory* 5:91–138.

BRATTICO, P. 2005. A category-free model of Finnish derivational morphology. *SKY Journal of Linguistics* 18:7–46.

BRATTICO, P. 2009. Long-distance Case assignment and the theory of phases. *Biolinguistica Fennica Working Papers* 1:79–104.

BRATTICO, P. 2010. One-part and two-part models of nominal case: Evidence from case distribution. *Journal of Linguistics* 46.1:47–81.

BRATTICO, P. 2011. Case assignment, case concord, and the quantificational case constructions. *Lingua* 121:1042–1066.

BRATTICO, P. 2012. Long distance case assignment in Finnish. *Studia Linguistica 66*: 245-284.

BRATTICO, P. 2014. Long distance Case assignment and intervention. *Lingua* 148:309–336.

BRATTICO, P. 2019a. *A computational implementation of a linear phase parser. Framework and technical documentation (version 9)*. Technical software documentation.

BRATTICO, P. 2019b. Subjects, topics and definiteness in Finnish. *Studia Linguistica* 73:1–38.

BRATTICO, P. 2020a. Finnish word order: does comprehension matter? *Nordic Journal of Linguistics* 44.1:38–70.

BRATTICO, P. 2020b. On some problems of rule ordering in Finnish. *Acta Linguistica Academica* 67.2:193–218.

BRATTICO, P. 2021a. A dual pathway analysis of information structure. *Lingua*: 103156.

BRATTICO, P. 2021b. Computation and the justification of grammars. *Finno-Ugric Languages and Linguistics* 10.1–2:51–74.

BRATTICO, P., & LEINONEN, A. 2009. Case distribution and nominalization: Evidence from Finnish. *Syntax* 12.1:1–31.

CARLSON, L. 1981. Aspect and quantification. *Tense and aspect.* Syntax and semantics 14. eds. Philip Tedeschi & A. Zaenen, 31–64. New York: Academic Press.

CHOMSKY, N. 1981. *Lectures in government and binding: The Pisa lectures*. Dordrecht: Foris.

CHOMSKY, N. 1982. *Some concepts and consequences of the theory of government and binding*. Cambridge, MA.: MIT Press.

CHOMSKY, N. 2000. Minimalist Inquiries: The Framework. *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*. eds. R. Martin, D. Michaels, & J. Uriagereka, 89–156. Cambridge, MA.: MIT Press.

CHOMSKY, N. 2001. Derivation by Phase. *Ken Hale: A life in language*. ed. M. Kenstowicz, 1–37. Cambridge, MA.: MIT Press.

CHOMSKY, N. 2008. On Phases. *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud.* eds. C. Otero, R. Freidin, & M.-L. Zubizarreta, 133–166. Cambridge, MA.: MIT Press.

CSIRMAZ, A. 2005. *Semantics and phonology in syntax*. MIT. PhD diss.

CSIRMAZ, A. 2012. The case of the divisible phase: Licensing partitive case in Finnish. *Syntax* 15.3:215–252.

GIANNAKIDOU, A. 1998. *Polarity sensitivity as (non)veridical dependency*. Amsterdam: John Benjamins.

HEINÄMÄKI, O. 1984. Aspect in Finnish. *Aspect bound: A voyage into the realm of Germanic, Slavonic and Finno-Ugrian Aspectology*. eds. C. de Groot & H. Tommola, 153–178. Berlin: Mouton De Gruyter.

HEINÄMÄKI, O. 1994. Aspect and boundedness in Finnish. *Tense, aspect, and action: Empirical and theoretical contributions to language typology*. eds. C. Bache, H. Basbøll & C.-E. Lindberg, 207–234. Berlin: Mouton De Gruyter.

HOLMBERG, A., & NIKANNE, U. 2002. Expletives, subjects and topics in Finnish. *Subjects, Expletives, and the EPP.* ed. P. Svenonius, 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. *Case and other functional categories in Finnish syntax*. eds. A. Holmberg & U. Nikanne, 177–206. Berlin: Mouton de Gruyter.

HUHMARNIEMI, S. 2012. *Finnish A´-movement: Edges and islands*. University of Helsinki. PhD diss.

HUHMARNIEMI, S. 2019a. The movement to SpecFinP in Finnish. *Acta Linguistica Academica* 66.1:85–113.

HUHMARNIEMI, S. 2019b. Tunnekausatiivilauseen argumenttirakenne I. NP aiheuttajana. *Virittäjä* 123.3:373–427.

HUHMARNIEMI, S. 2019c. Tunnekausatiivilauseen argumenttirakenne II. Lausemainen aiheuttaja. *Virittäjä* 123.4:516–547.

HUUMO, T. 2013. Incompleteness: On the many faces of object, hide-and-seek with the Finnish partitive. *Folia Linguistica* 47.1:89–111.

IKOLA, O. 1950. Infinitiivin objektista. *Virittäjä* 54, 468–474.

IKOLA, O. 1986. Infinitiivin objektin muodosta. *Kielikello* 3:217–223.

IKOLA, O. 1989. Mistä infinitiivin objektin horjuvuus johtuu? *Virittäjä* 95:372–408.

ITKONEN, T. 1976. Syntaktisten vaikutusyhteyksien luonteesta. *Virittäjä* 80:52–81.

ITKONEN, T. 1977. Tutkimuksen vaihtoehdoista ja vähän paradigmoistakin. *Virittäjä* 81:78–89.

KAISER, E. 2002. Case Alternation and NPIs in questions in Finnish. *WCCFL 21 Proceedings of the 21st West Coast conference on formal linguistics*. eds. L. Mikkelsen & C. Potts, 194–207. Somerville, Mass.: Cascadilla Press.

KAISER, E. 2003. A question of case. *Nordlyd* 4:694–707.

KAYNE, R. 1983. Connectedness. *Linguistic Inquiry* 14.2:223–249.

KAYNE, R. 1984. *Connectedness and Binary Branching*. Dordrecht: Foris Publications.

KIPARSKY, P. 1998. Partitive case and aspect. *The Projection of arguments: Lexical and compositional factors*. eds. M. Butt & W. Geuder, 265–307. Stanford: CSLI Publications.

KIPARSKY, P. 2001. Structural case in Finnish. *Lingua* 111:315–376.

LINDEN, E. 1956. Infinitiivin objektista: Uuden objektisäännön soveltamisesta. *Virittäjä* 60: 414–420.

MALING, J. 1993. Of nominative and accusative: the hierarchical assignment of grammatical Cases in Finnish. *Case and other functional categories in Finnish syntax.* eds. A. Holmberg & U. Nikanne, 49–74. Berlin: Mouton.

MANNINEN, S. 2003. *Small phrase layers: A study of Finnish manner adverbials*. Amsterdam: John Benjamins.

MANNINEN, S., & NELSON, D. 2004. What is a passive? The case of Finnish. *Studia Linguistica* 58.3:212–251.

MEGERDOOMIAN, K. 2000. Aspect and Partitive objects in Finnish. *WCCFL 19:*

Proceedings of the 19th West Coast conference on formal linguistics. eds. Billerey & Lillehaugen, 316–328. Somerville, Mass.: Cascadilla Press.

MEGERDOOMIAN, K. 2008. Partitive Case and Aspect. *Foundational issues in linguistic theory*. eds. R. Freidin, C. P. Otero & M. L. Zubizarreta, 73–104. Cambridge, MA.: MIT Press.

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

NELSON, D. 1998. *Grammatical Case assignment in Finnish*. New York: Routledge.

NIKANNE, U. 1993. On assigning semantic cases in Finnish. *Case and other functional categories in Finnish syntax*. eds. A. Holmberg & U. Nikanne, 75–87. Berlin: Mouton de Gruyter.

PYLKKÄNEN, L. 2002. *Introducing arguments*. Cambridge, MA.: MIT Press.

REIME, H. 1993. Accusative marking in Finnish. *Case and other functional categories in Finnish syntax.* eds. A. Holmberg & U. Nikanne, 89–109. Berlin: Mouton de Gruyter.

THOMAS, R. 2003. *The partitive in Finnish and its relation to the weak quantifiers*. University of Westminster. PhD diss.

TIMBERLAKE, A. 1975. The nominative object in Finnish. *Lingua* 35:201–230.

TOIVONEN, I. 1995. *A study of Finnish infinitives*. Brandeis University. BA thesis.

VAINIKKA, A. 1988. *Case and verbal agreement as argument numbering*. Manuscript.

VAINIKKA, A. 1989. *Deriving syntactic representations in Finnish*. University of Massachusetts Amherst. PhD diss.

VAINIKKA, A. 1993. The three structural cases in Finnish. *Case and other functional categories in Finnish syntax.* eds. A. Holmberg & U. Nikanne, 129–159. Berlin: Mouton de Gruyter.

VAINIKKA, A. 2003. Postverbal case realization in Finnish. *Generative approaches to Finnic and Saami linguistics*. eds. D. Nelson & S. Manninen, 235–266. Stanford: CSLI Publications.

VAINIKKA, A. 2011. Genitive case in Finnish reconsidered. *Biolinguistica Fennica Working Papers* 2:1–32.

VAINIKKA, A., & BRATTICO, P. 2014. The Finnish accusative: long distance case by agreement. *Linguistics* 52.1:73–124.

VAINIKKA, A., & MALING, J. 1996. Is partitive case inherent or structural? *Studies on the syntax and semantics of partitive and related constructions.* ed. J. Hoeksma, 179–208. Berlin: Mouton de Gruyter.

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 University Press.

WIIK, K. 1972. Suomen akkusatiiviobjektin muoto. *Turun Fonetiikan Laitoksen Julkaisuja* 12.

Pauli Brattico
Triggiano, Italy
Pauli.j.brattico@gmail.com

Conflicts of interest

No conflicts of interest

Appendix A.

Contents of the test corpus file

MAJOR GROUPS	SUBGROUP (# OF CONSTRUCTION TYPES)	EXAMPLES
1. Nominative and partitive	1.1.1 Grammatical, canonical (4)	*Pekka ihailee Merja-a* (SVO) Pekka.NOM admires Merja-PAR 'Pekka admires Merja.' *Pekka antoi kirja-n Merjalle* (SVO-IO) Pekka.NOM gave book-ACC to.Merja 'Pekka gave a/the book to Merja.'
1.1. Finite clause	1.1.2 Noncanonical (32)	*Merja-a ihailee Pekka* (OVS) Merja-PAR admire Pekka.NOM 'It is Pekka who admires Merja.' *Merja-a halusi ihailla Pekka* (OVvS) Merja-PAR wanted to.admire Pekka.NOM 'It was Pekka who wanted to admire Merja.'
	1.1.3 Wrong case forms (43)	**Merja-a nukkuu* Merja-PAR sleeps **Merja ihailee Merja-n* Merja-NOM admires Merja-GEN
	1.1.4 Superfluous arguments (6)	**Pekka ihailee Merjaa Merjaa* Pekka.NOM admires Merja-PAR Merja-PAR **Pekka Pekka ihailee Merja-a* Pekka.NOM Pekka.NOM admires Merja-PAR
1.2. Infinitivals	1.2.1 Grammatical, canonical (2)	*Pekka halusi ihailla Merja-a* Pekka.NOM wanted to.admire Merja-PAR 'Pekka wanted to admire Merja.' *Pekka nukahti lukemalla kirja-a* Pekka.NOM sleep by.reading book-PAR 'Pekka fell asleep by reading a/the book.'
	1.2.2 Noncanonical (5)	*Merja-a halusi ihailla Pekka* Merja-PAR wanted to.admire Pekka.NOM 'It was Pekka who wanted to admire Merja.' *Kirja-a lukemalla nukahti Pekka* book-PAR by.reading sleep Pekka.NOM 'It was by reading a/the book that Pekka fell asleep.'
	1.2.3 Noncanonical + wrong case (32)	**Pekka-a halusi ihailla Merja-a* Pekka-PAR wanted to.admire Merja-PAR **Peka-n halusi ihailla Merja-a* Pekka-GEN wanted to.admire Merja-PAR
1.3 Adpositions	1.3.1 Grammatical (3)	*lähellä Pekka-a* near Pekka-PAR 'near Pekka' *Peka-n lähellä* Pekka-GEN near 'near Pekka'
	1.3.2 Wrong case forms (6)	**lähellä Pekka* near Pekka.NOM **Pekka lähellä* Pekka.NOM near
	1.3.3 Superfluous arguments (4)	* *lähellä Pekka-a Pekka-a* near Pekka-PAR Pekka-PAR **Pekka-a Pekka-a lähellä* Pekka-PAR Pekka-PAR near
	1.3.4 Ungrammatical, wrong order (1)	**lähellä Peka-n* near Pekka-GEN
1.4 Partitive subjects	1.4.1 Grammatical, canonical (1)	*Pekka-a pelottaa* Pekka-PAR feels.frightened 'Pekka feels frightened.'
	1.4.2 Ungrammatical (4)	**Pekka pelottaa* Pekka.NOM feels.frightened **Peka-n pelottaa* Pekka.GEN feels.frightened
1.5 Numerals	1.5.1 Grammatical, canonical (6)	*ne kaksi sukka-a hävisi* those.0 two.0 sock.sg-PAR disappeared 'those two socks disappeared.' *Pekka löysi ne kaksi sukka-a*' Pekka found those.0 two.0 sock.SG-PAR

		‘Pekka found those two socks.’
	1.5.2 Wrong case forms (13)	**ne kaksi sukka hävisi*
		those.0 two.0 sock.NOM disappeared
		**ne kaksi suka-n hävisi*
		those.0 two.0 sock-GEN disappeared.
2. Accusative Case	2.1.1 Grammatical, aspect-related (6)	*Pekka voitti Merja-n*
		Pekka.NOM won Merja-ACC
		‘Pekka won/beat Merja.’
		Pekka pesi Merja-a
		Pekka.NOM washed Merja-PAR
		‘Pekka washed Merja.’
2.1 Accusative and aspect	2.1.2 Wrong case form, aspect-related (5)	**Pekka tönäisi Merja*
		Pekka.nom pushed Merja.nom
		**Pekka tönäisi Merja-n*
		Pekka.NOM pushed Merja-GEN
	2.1.3 Noncanonical (3)	*Merja-n voitti Pekka*
		Merja-ACC won Pekka.NOM
		‘It was Merja who Pekka won/beat.’
		Merja-a pesi Pekka
		Merja-PAR washed Pekka.NOM
		‘It was Merja who Pekka washed.’
	2.1.4 Wrong case form, aspect-related (4)	**Merja tönäisi Pekka*
		‘Merja.NOM pushed Pekka.NOM’
		**Merja-n tönäisi Pekka*
		‘Merja-GEN pushed Pekka.NOM’
2.2 Accusative and negation	2.2.1 Grammatical, canonical (1)	*Pekka ei voittanut Merjaa*
		Pekka.NOM not.3sg won Merja-PAR
		‘Pekka did not win/beat Merja.’
	2.2.2 Wrong case forms (4)	**Pekka ei voittanut Merja*
		Pekka.NOM not won Merja.NOM
		**Pekka ei voittanut Merja-n*’
		Pekka.NOM not won Merja-GEN
	2.2.3 Noncanonical (5)	*Merjaa ei voittanut Pekka*
		‘Merja-PAR not won Pekka.NOM’
		‘It was Pekka who didn’t win/beat Merja.’
		Merja-a Pekka ei voittanut
		Merja-PAR Pekka.NOM not won
		‘When it comes to Merja and Pekka, he didn’t win her.’
	2.2.4 Wrong case forms, noncanonical (4)	**Merja-n ei voittanut Pekka*
		Merja-ACC not won Pekka.NOM
		*Merja ei voittanut Pekka’
		Merja.0ACC not won Pekka.NOM
2.3 Accusative and agreement	2.3.1 Grammatical, canonical (2)	*Me löysi-mme avaime-n*
		we.NOM found-1pl key-ACC
		‘We found a/the key.’
		Me löydet-tiin avain
		we.NOM found.IMPASS key.0ACC’
		‘We found a/the key.’
	2.3.2 Noncanonical (2)	*Avaime-n löysi-mme me*
		key-ACC found-1SG we.NOM
		‘It was us who found the key.’
		Avain löydet-tiin me
		key.0ACC found.IMPASS we.NOM
		‘It was us who found the key.’
		(possibly marginal or ungrammatical?)
	2.3.3 Wrong case forms (5)	**Me löydettiin avaime-n*
		we.NOM found.IMPASS key-ACC
		**Me löysimme avain*’
		‘we.NOM found-1PL key.0ACC’
2.4 Negation and agreement	2.4.1 Grammatical, canonical (1)	*Me ei löydetty avain-ta*
		we.NOM not found key-PAR
		‘We did not find a/the key.’
	2.4.2 Noncanonical (3)	*Avain-ta me ei löydetty*
		key-PAR we.NOM not found
		‘As for the key, we did not find it.’
		Me ei avain-ta löydetty
		we.nOM not key-PAR found
		‘We didn’t find the key.’
	2.4.3 Wrong case forms (4)	**Me ei löydetty avain*
		we.NOM not found key.NOM
		**Me ei löydetty avaime-n*
		we.NOM not found key-GEN
2.5 Long-distance effects	2.5.1 Grammatical, V + infinitival (7)	*Pekka halusi voittaa Merja-n*
		Pekka.NOM wanted to.win Merja-ACC’
		‘Pekka wanted to win/beat Merja,ä

		Me ei haluttu voittaa Merja-a we.NOM not want to.win Merja-PAR 'We didn't want to win/beat Merja.'
	2.5.2 Noncanonical, V + infinitival (7)	*Merja-n halusi voittaa Pekka* Merja-ACC wanted to.win Pekka.NOM 'It was Pekka who wanted to beat Merja'
		Merja haluttiin voittaa me Merja.0ACC want.IMPASS to.win we.NOM 'It was us who wanted to beat Merja,ä
	2.5.3 Wrong case forms (2)	**Me haluttiin voittaa Merja-n* we.NOM want.IMPASS to.win Merja-ACC
		**Me ei haluttu voittaa Merja-n* we.NOM not wanted to.win Merja-ACC
3. Genitive case	3.1.1 Grammatical, canonical (3)	*Pekka käski Merja-n lähteä* Pekka.NOM ordered Merja-GEN to.leave' 'Pekka ordered Merja to leave.'
		Peka-n täytyy lähteä Pekka.GEN must to.leave 'Pekka must leave.'
3.1 Infinitival subject	3.1.2 Noncanonical (2)	**Pekka käski lähteä Merja-n* Pekka.NOM ordered to.leave Merja-GEN
		**Pekka sanoi lähtevän Merja-n* Pekka.NOM said to.leave Merja-GEN
	3.1.3 Wrong subject case (4)	**Pekka käski Merja lähteä* Pekka.NOM ordered Merja.NOM to.leave
		**Pekka käski Merja-a lähteä* Pekka ordered Merja-PAR to.leave
	3.1.4 Wrong S case, noncanonical (4)	**Pekka käski lähteä Merja* Pekka.NOM ordered to.leave Merja.NOM
		**Pekka käski lähteä Merja-a* Pekka.NOM ordered to.leave Merja-PAR
3.2 Possessive use	3.2.1 Grammatical, canonical (1)	*Se Merja-n kello hävisi* that Merja-GEN watch disappeared 'That Merja's watch disappeared.'
	3.2.2 Noncanonical (3)	**se kello Merja-n hävisi* that watch Merja-GEN disappeared
		**se kello hävisi Merja-n* that watch disappeared Merja-GEN
	3.2.3 Wrong case forms (4)	**se Merja kello hävisi* that Merja.NOM watch disappeared
		**se Merja-a kello hävisi* that Merja-PAR watch disappeared
3.3 Adpositions	3.3.1 Grammatical, canonical (2)	*Merja nukkui Pekan lähellä* Merja.NOM slept Pekka-GEN near 'Merja slept near Pekka.'
		Merja nukkui lähellä Pekkaa Merja.NOM slept near Pekka-PAR 'Merja slept near Pekka.'
	3.3.2 Noncanonical (1)	**Merja nukkui lähellä Peka-n* Merja.NOM slept near Pekka-GEN
3.4 Possessor + numeral	3.4.1 Grammatical, canonical (2)	*Ne kaksi Merja-n puhelinta hävisi* those two Merja-GEN phones disappeared 'Those two Merja's phones disappeared.'
		Ne Merja-n kaksi puhelinta hävisi those Merja-GEN two phones disappeared 'Those two Merja's phones disappeared.'
	3.4.2 Wrong case (4)	**Ne kaksi Merja puhelinta hävisi* those two Merja.NOM phones disappeared
		**Ne kaksi Merja-a puhelinta hävisi* those two Merja-PAR phones disappeared
4. Special constructions	4.1.1 Grammatical, canonical (4)	*Me löysimme avaimen* we.NOM found-1PL key-ACC 'We found a/the key.'
		Me löydettiin avain we.NOM found.IMPASS key.0ACC 'We found a/the key.'
4.1 Impersonal passive	4.1.2 Wrong object case (2)	**Me löysi-mme avain* we.NOM found-1PL key.0ACC
		**Me löydet-tiin avaime-n* we.NOM found-IMPASS key-ACC
4.2 Raising constructions	4.2.1 Grammatical, canonical (1)	*Merja näyttää lähtevän* Merja.NOM seems to.leave 'Merja seems to leave.'
	4.2.2. Wrong case (4)	**Merjaa näyttää lähtevän* Merja-PAR seems to.leave
		**Merjan näyttää lähtevän*

		Merja-GEN seems to.leave
4.3 Copular constructions	4.3.1 Grammatical, canonical (2)	*Pekka on Pekka* Pekka.NOM is Pekka.NOM 'Pekka is Pekka.' *Merjalla on Pekka* Merja.ALL is Pekka.NOM 'Merja has Pekka.'
	4.3.2 Wrong case (8)	**Pekka on Pekkaa* Pekka.NOM is Pekka-PAR **Pekka on Pekan* Pekka.NOM is Pekka-ACC(N)
5. Adverbials and case marking	5.1 Grammatical, canonical (5)	*Pekka nukkui koko päivä-n* Pekka.NOM slept all day-ACC 'Pekka slept all day.' *Pekka ei nukkunut koko päivä-ä* Pekka.NOM not slept all day-PAR 'Pekka did not sleep all day.'
	5.2 Wrong case forms (7+1)	**Pekka ei nukkunut koko päivän* Pekka not slept all day-ACC **Pekka ihailee Merja-a koko päivää* Pekka.NOM admires Merja-PAR all day-PAR

Appendix B

Source code and the raw input/output data

The Python source code for the underlying algorithm is currently (September 2022) maintained at www.github.com/pajubrat/parser-grammar. To replicate the study exactly as it was reported in this article, branch "Structural-Case-Study-7c" should be used. The master branch contains the most recent version. The raw input/output data files for this project can be found from the directory /language data working directory/study-7-c-structural-case/ and are as follows:

File	**Contents**
`case_corpus.txt`	Original dataset
`case_corpus_grammaticality_judgments_FINAL.txt`	Grammaticality judgments by the model
`case_corpus_grammaticality_judgments_NATIVE.txt`	Grammaticality judgments by native speaker (author)
`case_corpus_log_FINAL.txt`	Derivational log file with detailed derivations of all sentences in the dataset
`case_corpus_results_FINAL.txt`	Calculated outputs for all grammatical sentences, containing a syntactic analysis and semantic (also pragmatic) interpretation
`case_corpus_resources_FINAL.txt`	Psycholinguistic and complexity metrics
`case_corpus_simple_log_FINAL.txt`	A simplified representation of the derivations
`phrase_structure_images.zip`	All phrase structure images created by the algorithm for all grammatical sentences
`derivations.zip`	A compressed file containing the derivational log file
`Additional documentation.pdf`	A technical supplementary document containing a more detailed examination of the results