Additional documentation

for the article “Structural case assignment, thematic roles and information structure in Finnish” published in *Studia Linguistica*

October 2022

1 Introduction

This document is a technical supplementary to the article "Structural case assignment, thematic roles and information structure in Finnish," published in *Studia Linguistica*.[1] It addresses technical details of the syntactic background theory, hypothesis, methods and results. The document was originally composed as an official supplementary to the main article, but was removed during the review process.

2 Syntactic background theory

The syntactic background theory is a Python-based information processing model that allows the researcher to test linguistic hypotheses and theories rigorously (Brattico, 2019).[2] The program generates an idealized brain model of the speaker of any language and uses it to process test sentences in that language and provides them with grammaticality judgments, derivations, syntactic analyses, semantic interpretations and various psycholinguistics metrics, as determined by the linguistic hypotheses developed by the researcher. The model is a combination of abstract cognitive principles of language comprehension ("performance") and grammatical principles ("competence"), the latter comprising

---

[1] Brattico (2022c). How to cite this article: Brattico, P. (2022). Additional documentation for the article "Structural case assignment, thematic roles and information structure" published in *Studia Linguistica*.

[2] To date the model has been applied to the analysis of pied-piping and Ā-chains (Brattico et al., 2022; Brattico & Chesi, 2020), free word order (Brattico, 2020), case assignment (Brattico, 2020, 2022c), information structure (Brattico, 2021a), control and agreement (Brattico, 2021c), head movement (Brattico, 2022b), binding (Brattico, 2022d) and nonfinite clause structure (Brattico, 2022a). Most of this literature focuses on Finnish.

syntax and several aspects of semantic interpretation.[3] Morphology has a limited role. For the notions of linguistic competence and performance, see (Chomsky, 1965). Because the model works from phonological input strings towards abstract semantic interpretations, it can be said to implement a recognition grammar, a system that analyses sentences rather than generates them.[4]

A recognition grammar presupposes a parser. For example, an input string such as *the horse raced past the barn* will be mapped into a syntactic representation [[$_{DP}$ *the horse*][$_{VP}$ *raced* [$_{PP}$ *past the barn*]]] by the parser component, where the latter contains lexical items, not phonological words. The generative engine therefore consumes input sentences, not arbitrary collections of lexical elements chosen by the speaker by free will.[5] Since the parser component must explore all possible ambiguities in order to guess what the speaker intended to communicate, the model tries all prima facie plausible ways of combining the lexical items in the input to find out which if any of them form syntactically and semantically well-formed representations. The competence module performs filtering by considering which, if any, of the resulting parses agree with the principles of language, again as specified by the researcher and/or the linguistic theory. Well-formed solutions appear in the output of the model, corresponding to a situation where the speaker "understood" the input sentence in certain ways. The case checking hypothesis proposed in the main article constitutes one of the cognitive principles applied to the parsing output.

This architecture presupposes that there is a lexicon, a mapping from phonological words into lexical items. The former are phonological strings such as /horse/, the latter sets of features such as category labels, case information, semantic labels, and others. The *lexico-morphological module*, which performs this conversion, is applied to each phonological word in the input before the element is integrated with

[3] Use of formal computer programs for purposes of verifying scientific linguistic hypotheses should not be confused with attempts at writing engineering parsers. While the analysis includes a parser, it must replicate responses and semantic intuitions elicited from native speakers, including grammaticality judgments. The general framework was inspired by the information processing framework of (Marr, 1982). For a more detailed description of the computational methodology used in this study, see (Brattico, 2021b).

[4] Recognition grammars and enumerative grammars are equivalent under very weak assumptions. It is possible to create one from the other, and vice versa. Both are generative grammars: they define a set of expressions, "expressions" understood in a broad sense to include semantic and pragmatic properties.

[5] The model follows (Phillips, 1996), who proposed that the generative engine bleeds lexical items in a left-to-right order, mirroring language comprehension and the parsing process. Another closely related proposal is the so-called top-down grammar, where grammatical principles regulate top-down derivations. See (Chesi, 2012). The present model uses top-down derivations, although not exclusively.

the syntactic representation currently being assembled in the syntactic system. Thus, the information flows from the input into the lexico-morphological component, and from there into syntax and finally semantics. Differences between languages (and hence differences between idealized brain models of speakers of languages) are located in the lexicon. All lexical information is provided in external lexicon files, which the model reads before processing begins.[6] The whole computational system, excluding the lexicon, is universal.[7]

The model provides each grammatical input sentence with a syntactic analysis and semantic interpretation (or several, if ambiguous). Semantic interpretation is only attempted if a syntactically well-formed representation is received. This means that we distinguish syntactic computations that are internal to the syntactic (parsing) module, and semantic computations that bleed the former. The simplest possible architecture is one where there is only a single interface between syntax and semantics. We could begin for example from the assumption that each syntactically well-formed and fully processed syntactic representation is forwarded to a separate semantic module for interpretation. The socket that connects the two is usually called the *syntax-semantics interface* ("LF interface" in the generative theory). Although the actual model we are working with here is slightly more complex, it does posit a syntax-semantics interface in this sense. The following figure, repeated from the main article, illustrates the architecture:

[6] If the lexicon contains lexical items from Finnish, English and Italian, the script generates models for speakers of these languages. These models are applied to the processing of the input sentences in those languages. The process is largely automatic, but the researcher must be aware that the Finnish input sentences are processed by slightly different model than sentences in English, depending on what differences has been coded into the respective lexicons.

[7] The term Universal Grammar (UG) is often used in this connection. It refers to language-specific linguistic universals grounded in human biology. While there certainly exist language-specific language universals, I am not sure if we need to designate them specifically. In fact, the introduction of the minimalist program (Chomsky, 1995, 2008) made part of the term obsolete by positing supramodal economy principles. Although the whole computational system proposed here is universal, and part of human biology, it would be possible but perhaps pointless to speculate at this point which aspects of the algorithm are language-specific.

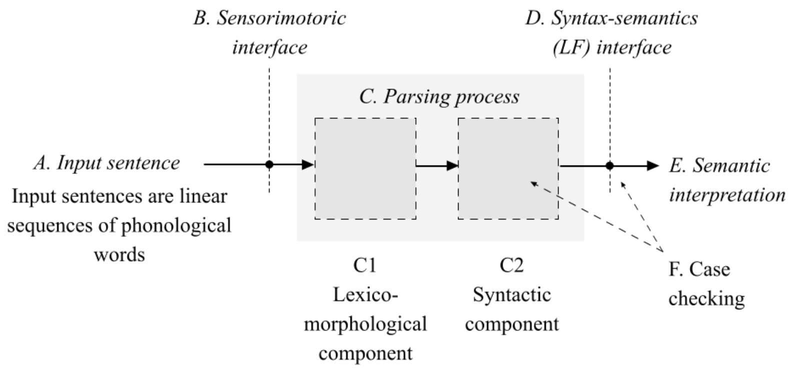


Figure 14. Architecture of the information processing model. See the main text for explanation.

This diagram shows the input, lexico-morphological component, syntactic parser and the syntax-semantic interface. Case checking operates inside the syntactic component and at the syntax-semantics interface. Case features are generated from the phonological words in the input by the lexico-morphological component.

Mapping between phonological words and lexical items is complicated by the existence of derivational and inflectional morphology. Not all phonological words map into primitive lexical items. The lexico-morphological component handles this issue by mapping polymorphemic words into linear sequences of primitive lexical items, and then by feeding these sequences into the syntactic component. It therefore performs morphological decomposition. A finite tensed verb, for example, could be decomposed into a linear sequence V#T_{fin} where both V (standing for the verb stem) and T_{fin} (finite tense) refer to primitive lexical items. Thus, the lexical entry for the finite verb contains V#T_{fin}, while both V and T_{fin} have their own primitive lexical entries retrieved by the lexico-morphological module. We can think of the lexical entry for the finite verb as a morphological chunk. Derivational morphemes map into independent lexical items, while inflectional morphemes (e.g., case forms) map into features inside those lexical items. Which are which are determined in the lexicon.[8] All these operations take place inside the morphological component and do not involve higher-level syntactic operations, such as the generation of hierarchical representations.

Reconstruction plays a role in this study. It refers to a process where a noncanonically positioned element is adjusted or realigned so that it can be interpreted at the syntax-semantics interface. A

[8] It is assumed that morphological sequences could be produced by a separate morphological parser, but this component remains unimplemented and is simulated currently by providing morphological decompositions directly in the input.

noncanonically positioned element is recognized as such by relying on an operation that detects violations of grammatical well-formedness. Case checking is one of these operations. The element is then reconstructed into a position where this error (and possible other errors) is eliminated. For example, a grammatical OVS order in Finnish triggers reconstruction for both arguments since they are unable to check their case features at the first pass parse positions. The arguments are reconstructed into canonical positions at the syntax-semantics interface, where checking (hopefully) succeeds. The operation that forwards parsing solutions generated directly from the input into the syntax-semantic interface via a sequence of reconstructions is called *transfer*.[9] The operation implements a function from phrase structure representations into phrase structure representations and operates like a cognitive reflex, adjusting or normalizing linguistic information presented in the input.

Reconstruction is in turn linked with information structure (IS), notions such as 'topic' (what is old information) and 'focus' (what is new information). The model uses a dual pathway architecture, where the information structure is created by synthesizing information from two processing pathways: syntactic and pragmatic (Brattico, 2021a). The syntactic pathway processes information structure that has been grammaticalized, say by means of discourse clitics, while the independent pragmatic pathway, mostly based on a psychological attention allocation mechanism, makes guesses concerning topics and focus on the basis of canonical and noncanonical word orders. Since rich case forms make room for reconstruction and hence noncanonical word orders, these noncanonical word orders can in turn be used to communicate information structure. The ultimate IS interpretation is generated inside the semantic system and appears in the field "Semantic interpretation" in the output of the model.

3 The hypothesis

3.1 Introduction

The present work builds on (Brattico, 2020), where it was hypothesized that the rich case assignment signature of Finnish is connected to its free word-order profile. Case forms are a mechanism for recovering thematic roles. The hypothesis was implemented by assuming that arguments that cannot check their cases in the first pass parse position undergo reconstruction, a process which locates a syntactic position that is compatible with the case form and then remerges (copies, realigns) the element

[9] Reconstruction can be viewed as a reverse-engineered transformation, in the parlance of the generative theory. Instead of triggering movement operations from canonical positions into surface (or s-structure) positions, we detect elements that appear at oddball positions in the output provided by the parser and find the gaps.

into that position. Reconstruction is a syntax-internal operation applied before the syntax-semantics interface, which means that the component responsible for semantic interpretation will find referential arguments at their canonical thematic positions. In a language that does not exhibit rich case forms, such as English, thematic roles are determined by surface order.

Implementing case checking in a recognition grammar requires that we begin from overt case forms present in the input. Case forms appearing the input were linked with lexical case features. The resulting syntactic case features are then mapped from the overt case forms of phonological words. When a nominal expression contains several phonological words, each with some case form, the lexical items corresponding to those phonological words will come to have these same case features. A Finnish nominal expression such as *piene-n auto-n* ‘small-acc car-acc’ will be represented by [$_{NP}$ *small car*] where each lexical item *small* and *car* has the same accusative case feature. Case concord does not exist as a separate phenomenon. An additional consequence is that surface forms that lack overt case features, such as the subject and object of an English sentence *John admires Mary*, will either lack case features, as is the case in the present model, or have the same feature.

Once we get case features into lexical items, we can design an algorithm that checks them inside the syntactic or semantic component (or both, as is the case here).[10] There exists a number of approaches to case checking and nominal case in general that could serve as a starting point. Here a system was adopted which checks lexical case features against other lexical elements’ features such that the former appear inside the syntactic scope of the latter. Case forms express, ultimately, syntactically represented relations between words.

[10] In the main article, features requiring checking were listed in connection with the case assignees, but they must have an additional diacritic specifying that they require checking. The required diacritic is a symbol TAIL, followed by feature(s) requiring checking. Thus, [TAIL:ARG] in the lexicon means that the lexical item must be checked against ARG. In most generative theories available today features that require checking are specified as “uninterpretable versions” of the target features. How this is implemented is not relevant to any empirical matter discussed in this work. The term “tail” is essentially a synonym for “upward path” that was originally conceptualized as an inverse probe-goal dependency and later generalized following (Kayne, 1983, 1984).

3.2 Syntactic scope and the upward path

The analysis requires that we define "syntactic scope." Suppose F is a set of lexical features, as discussed in the previous section, and α is a lexical element requiring checking by F. We define case checking by (1) and (2).

(1) 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;

(2) Upward path

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

When α can form a path to X, we say that α appears under the *syntactic scope* of X, and that X *governs* α. The notion of upward path (2) requires that we have "domination" and "immediate daughter" available in the theory. They are defined by relying on an asymmetric binary branching bare phrase structure, following the model proposed in (Chomsky, 2008).[11] Thus, if β = [A, B], then β dominates both A and B, which are both the immediate daughters of β. Now suppose the case assignee α in (2) is inside B, so that we have β = [A, [$_B$…α…]]. Then, according to (2), the upward path from α includes both B and β, which dominates it, but also A, which is an immediate daughter of β.[12] In a typical case checking configuration, A will check case features of α. Example (3) shows how this scheme applies to Finnish prepositions.

(3)	[$_{PP}$	kohti	[$_{NP}$…uusi-a	seikkailui-ta …]]
		towards	new-par	adventures-par
		X	Y	Z

---

[11] The representation differs from Chomsky's formulation in that the present model asymmetric. In [A B], A is the *left constituent*, B is the *right constituent*. Labels are determined at runtime by a labelling algorithm, intuitively as follows: Suppose we target α; if α is primitive, it will be the label; otherwise if the left constituent is primitive, it will be the label; otherwise if the right constituent is primitive, it will be label; otherwise apply the rule recursively to the right constituent.

[12] Notice how this dependency mirrors c-command: under all reasonable definitions for c-command, A c-commands α inside β.

The adjective word *uusi-a* ‘new-par’ inside the NP checks its case against the preposition, which is an immediate constituent of the PP that dominates it. This corresponds to $\beta = [A, [_{B}\ldots\alpha\ldots]]$, with $\alpha = Y$. The same scheme applies to the noun head. Furthermore, both Y and Z occur inside the syntactic scope of X, and X governs both Y and Z, in our terms.

Condition (1) requires that the case checker contains a set of features F determined by the case assignee. We can think of F as specifying the lexical element(s) that provide(s) the relevant syntactic scope for some particular case form. The condition says that all features must be checked. If there is partial match, or no match, checking fails. In addition, there is an intervention condition which determines that partial match that occurs closer to the case assignee leads into rejection. This requires that we define “closer.” The algorithm implements (1) and (2) by scanning from α upwards via the dominance relations, going one node at a time. The relevant distance metric therefore counts dominating nodes. Thus, if $\beta = [A\ [B\ \alpha]]$, B is closer to α than A and will be examined before A is examined. It follows from (1)-(2) that if B has a subset of features from F, but not all, while A has them all, the checking configuration fails.

3.3 Lexical features

The mechanism elucidated in the previous section requires that we specify the features that require checking. The first relevant observation about Finnish concerns Vainikka’s partitive and genitive generalizations (Vainikka, 1989, 1993, 2003) which say that these cases are assigned on the basis of their structural positions. The partitive constitutes a “default complement case,” while the genitive is the “default specifier case,” to borrow the terms used by Vainikka. This generalization, which is descriptively very powerful, all by itself suggests that we should not use individual functional heads (T, v, P) as case assigners.

The fact that not all functional-lexical heads participate in case checking while they still have complements and specifiers has received less interest in the literature, but must be addressed nonetheless. The distinction between “case active” and “case inactive” heads is represented by feature +ARG (from “pair this element with an *arg*ument”[13]) in the model, such that those lexical items which

---

[13] Feature ±ARG has functional motivation that does not play a role in this study, but is relevant when examining the formalization. It connotes “argument” and means that the lexical item must establish a connection with a (referential) argument; these elements are, in other words, what one would intuitively call predicates. The feature triggers control dependencies, for example (Brattico, 2021c).

participate in case checking are marked for +ARG.[14] I return to the functional motivation of this feature further below.

Another factor that regulates case assignment is agreement.[15] It turns out that the lexico-functional heads which assign the partitive to their complements never exhibit overt agreement with that complement. Furthermore, it is well-known that nominative checking is related of agreement; and this is true of Finnish well, where a necessary condition for finite verbal agreement is nominative checking. Genitive arguments can cause agreement, but do not do it necessarily. Whether a head can exhibit agreement is represented in the model by feature ±VAL, from "valuation." This controls agreement reconstruction in the algorithm. Features ARG and VAL give us the following classification of lexical items.

Table 1. Classification of Finnish lexicon on the basis of two features ARG and VAL.

	–ARG (non-predicates)	+ARG (predicates)
–VAL (no agreement)	Conjunctions, particles *mutta* 'but', *ja* 'and', *ehkä* 'maybe'	Control predicates *lähteä* 'to leave', *kohti* 'towards'
+VAL (Agreement)	Elements exhibiting concord *pieni* 'small, *kolme-t* 'three-pl'	Agreeing predicates *lähde-n* 'leave-prs.1sg'

Thus, most lexical elements (perhaps all) have some combination of these features.[16] Since the nominative and genitive assignment is sensitive to finiteness, ±FIN is also involved. This will give us the rules for nominative (+FIN, +VAL, +ARG), partitive (–VAL, +ARG) and genitive (–FIN, +ARG), as explained in the main article. The accusative represents a mixed case, responding to syntactic and

[14] There is no grammatical law that forces case checking to coincide with ARG. The generalization is implemented by the case checking features. Several examples of feature checking were posited in this study that do not involve ARG, such as NEG and PHI.

[15] I use the term agreement to refer to systematic φ-feature covariation between a predicate and a local argument. Frozen forms (no covariation) and φ-concord (variation is not with a whole DP argument but with a local head, typically inside an argument) are excluded.

[16] Currently the features are listed in connection with each independent lexical entry. If the whole lexicon is classified in terms of these features, then some more general issue must be at stake.

semantic features (ASP:BOUNDED, +ARG, –NEG, ±PHI); since this was discussed in detail in the main article, I will not repeat the justification.

There are several technical details about this system that are not essential for the linguistic analysis but useful to know when examining full formalization. One question is what happens when a lexical item has no specification for a given feature. It is assumed that both unmarked ARG and +ARG correspond to positive specification (form "ARG" is used), and the absence of this feature corresponds to the negative specification. There is no –ARG as a feature. Because there is no case rule referring to –ARG, the issue never arises. Perhaps the feature exists for some other reason. The agreement feature VAL is handled differently: there is both +VAL and –VAL. This is because the partitive checks –VAL. This means that we have three values: positive (+VAL), negative (–VAL) and "unspecified." The current rules interpret an "unspecified" as a head that does not show agreement but does not check the partitive either. Another nontrivial issue is how to calculate the accusative, which constitutes a positive polarity case and reacts to the "absence of negative polarity." This was notated as –NEG, but what are we checking exactly? The implication is that we must check the absence of something. The current checking rule specifies that "the accusative case is ungrammatical if +NEG is checked." The actual feature is [*POL:NEG] 'negative polarity' which says that checking [POL:NEG] leads into rejection (the asterisk represents the fact that checking implies rejection). There is an alternative: require checking against a positive polarity feature. This presupposes that all non-negative clauses have the required positive polarity feature.

4 Methods

4.1 Procedure and replication

The study can be replicated by downloading the source code from the code repository and running it on a local computer by writing one command to the command prompt.[17] The downloaded source code is a

[17] The online source code repository contains several branches of the code, one which is the master branch and is the latest version of the algorithm, not the one that was used in this study. The version that was used in this study is stored in the branch "study-7-c-structural-case." The researcher must download files from this branch in order to replicate the study. Running the code from the master branch will execute the latest version that by all likelihood differs from the one used in this study. That being said, the latest version includes the mechanisms proposed here and should handle the present dataset as well.

Python script that runs the study without any intervention on researcher's part.[18] The local computer must have Python (3.x) installed. The study is replicated in Windows by opening the command prompt, navigating to the local directory containing the downloaded and extracted files, and writing "python lpparse" into the command prompt. This command asks the Python interpreter ("python") to read the source code ("lpparse") and execute it.[19] The simulation will pick up sentences from the test corpus, which is part of the downloaded package, process each test sentence, and writes the output into several output files. It should produce an output that is identical to the one produced in this study.

4.2 Output files

The algorithm takes sentences from the test corpus and produces several output files, which are then used to evaluate the correctness of the theory or hypothesis.

One file contains all the sentences from the test corpus with grammaticality judgments calculated by the model, where ungrammatical sentences are marked with the asterisk (*). This file can be used to evaluate observational adequacy by comparing it (preferably by using automatic file comparison tools) with a similar file where grammaticality judgments are provided by a native speaker. Calculated grammaticality judgments are in the file *case_corpus_grammaticality_judgements_FINAL.txt* and native speaker output in *case_corpus_grammaticality_judgments_NATIVE.txt*.

The most important output file, however, is *case_corpus_results_FINAL.txt*, which contains each input sentence linked with a syntactic, semantic and psycholinguistic analysis calculated by the model. Figure 3 shows a screenshot from this file when the input sentence is *John admires Mary*.

[18] The software package that gets downloaded from the repository is not a general-purpose tool that can be used out of the box for creating grammatical theories or performance models without experience in programming. It is a linguistic theory formalized in Python. Most of the code is designed to give correct results in concrete calculations against large datasets, and most of it Finnish. Given the controversial nature of any hypothesis in modern linguistics, researchers who want to use this methodology in their own work must write their own model from scratch.

[19] Notice that in order for Windows to recognize the command in Figure 1, Windows PATH variable needs to be specified so that it knows where to find the Python installation. Instruction on how to do this can be found from the Internet.

```
90  3.  John admires Mary
91
92      [[D John]:1 [T [__:1 [v [admire [φ Mary]]]]]]
93
94      Semantics:
95      Recovery: [{'Agent of v([D John])'}, {'Patient of admire([φ Mary])'}]
96      Aspect: []
97      DIS-features: []
98      Operator bindings: []
99      Semantic space:
100     Speaker attitude: ['Declarative']
101     Assignments:
102     [D John] ~ 2, [φ Mary] ~ 8, Weight 1
103     Information structure: {'Marked topics': [], 'Neutral gradient': ['[D John]', '[φ Mary]'], 'Marked focus': []}
104     spellout structure: [[D John] [T(v,V) φ(N)]]
105     surface structure: [[D John] [T [v [admire [φ Mary]]]]]
106     s-structure: [[D John] [T [v [admire [φ Mary]]]]]
107     LF structure: [[D John]:2 [T [__:2 [v [admire [φ Mary]]]]]]
108
109     Discourse inventory:
110     Object 2 in GLOBAL: [D John]
111     Object 8 in GLOBAL: [φ Mary]
112
113     Resources:
114     Total Time:1205, Garden Paths:0, Memory Reactivation:0, Steps:13, Merge:6, Move Head:2, Move Phrase:2,
115     A-Move Phrase:2, A-bar Move Phrase:0, Move Adjunct:0, Agree:3, Phi:5, Transfer:3, Item streamed into syntax:7,
116     Feature Processing:0, Extraposition:0, Inflection:12, Failed Transfer:0, LF recovery:2,
117     LF test:5, Filter solution:5, Rank solution:2, Lexical retrieval:19, Morphological decomposition:3,
118     Mean time per word:401, Asymmetric Merge:28, Sink:9, Tail test:0, Case checking:0,
```

Figure 3. Screenshot from the results file generated by the algorithm when it processed sentence *John admires Mary*.

The first item (line 92) is the final syntactic analysis generated at the syntax-semantics interface, followed by semantics attributes. The first semantic item (line 95) represents thematic roles, thus *John* was interpreted as the agent, *Mary* the patient. Aspect field contains information about the aspect, as explained in the main article. Currently this contains information concerning the accusative-partitive alteration in Finnish and has no other use. DIS-features lists interpretations of various discourse features, not relevant here since the test corpus does not contain sentences that have them. Operator bindings lists operator-variable bindings, for example, scope information in the case of interrogative clauses. Speaker attitude represents the propositional attitude of the speaker towards the proposition.

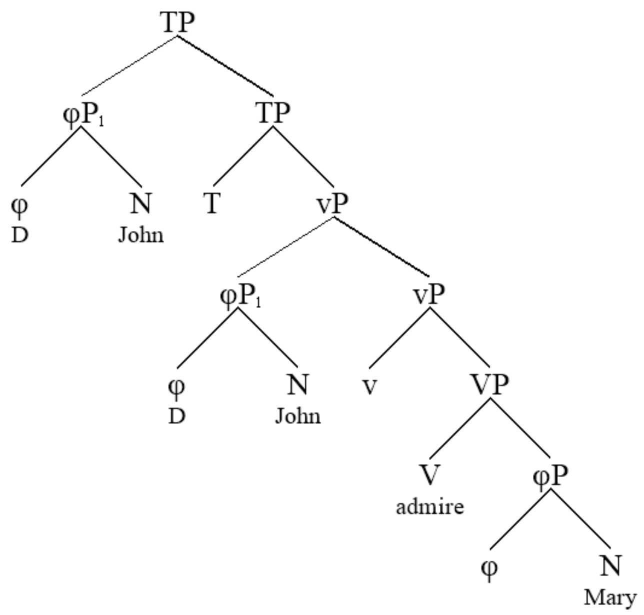
When the parser encounters a referential argument in the input sentence, such as *John* or *Mary*, it projects a corresponding semantic entity into the language-external global discourse inventory which stores the entities the sentence "is about." In this case, two entities were projected into existence: John and Mary, the persons (or representations thereof). These projected entities are listed under Discourse inventory below (lines 109-111). They are processed by global cognition, not the linguistic processing pathway. The two objects are numbers as Object 2 and Object 8 and these numbers are referred to in the assignment field (line 102), where it is specified that this sentence has exactly one possible interpretation, one where *John* refers to John and *Mary* to Mary. This becomes nontrivial when we take the binding theory and pronouns into account, but the issue is irrelevant here.

Next there is an information structure field which contains information about topics and information focus. This field contains three lists: marked topics, neutral gradient, and marked focus. Marked topics are objects that were specifically interpreted as representing old information; marked focus elements were interpreted as representing new information; and the neutral gradient is a gradient where the first elements are interpreted as more “topic-like” than the rest. This field is generated by a pragmatic module that listens the information flow inside the main language pipeline and tries to guess how the information structure was constructed. This is relevant when we examine noncanonical word orders in Finnish, which are linked with information structure. The system is based on a dual pathway architecture, where the final information structure interpretation is synthesized from parallel syntactic and pragmatic processing pathways. See (Brattico, 2021a).

Information structure is followed by three syntactic representations giving a snapshot of how the derivation was generated inside the syntactic module. Finally, the last item is resource consumption, which lists the number of computational operations (garden-paths, Merge, Move, Agree etc.) that were required to find and compute the first acceptable syntactic solution. If the sentence was ambiguous, then the remaining solutions are listed after these fields.

In some cases, the syntactic structures generated by the model are hard to read from a symbolic representation alone. The algorithm generates phrase structure trees for each solution in PNG format, which can be found from the folder */phrase structure images*. The phrase structure image generated by the algorithm for *John admires Mary* is shown in (4).

(4)



Notice that since these figures were generated by the model, they can be considered raw data and not artistic illustrations. Subscripts represent chains: the same DP John occurs in two places SpecTP and SpecvP, both they are marked by the subscript 1, meaning that this is the same proper name *John*. If the input sentence, say sentence #3, had several syntactic/semantic analyses, they are listed as (3a), (3b) and so on. To find the syntactic analyses for sentence #20 locate files *(20a).png*, *(20b).png* etc. from the image folder. If there is no image, then the sentence was judged ungrammatical and no analysis was found during the simulation. It is possible to run the simulation with or without drawing images; image generation decreases performance.

The results file provides very little information on how the sentence was derived. Moreover, if the input sentence is ungrammatical, no information is provided since no analyses were found. The whole derivation is recorded into the derivational log file, in this case *case_corpus_log_FINAL.txt*. The algorithm writes this file at runtime as it processed the sentence. Figure 4 shows a screenshot from this file, when the input sentence is again *John admires Mary*.

```
----------------------------------------------------------------------------------------------------------
#3. John admires Mary
['John', 'admires', 'Mary']

    1. ['John', 'admires', 'Mary']

        Next item: John. Lexical retrieval for "John": Morphological decomposition for John-#D#sg#3p#def#hum#m: <John-,D$,sg$,3p$,def$,hum$,m$>.
        Next item: "m$": Inflectional feature m$...Added ['LANG:EN', 'PHI:GEN:M', 'PHON'] into transient memory buffer...Done.
        Next item: hum$. Lexical retrieval for "hum$": Inflectional feature hum$...Added ['LANG:EN', 'PHI:HUM:HUM', 'PHON'] into transient memory buffer...Done.
        Next item: def$. Lexical retrieval for "def$": Inflectional feature def$...Added ['LANG:EN', 'PHI:DET:DEF', 'PHON'] into transient memory buffer...Done.
        Next item: 3p$. Lexical retrieval for "3p$": Inflectional feature 3p$...Added ['LANG:EN', 'PHI:PER:3', 'PHON'] into transient memory buffer...Done.
        Next item: sg$. Lexical retrieval for "sg$": Inflectional feature sg$...Added ['LANG:EN', 'PHI:NUM:SG', 'PHON'] into transient memory buffer...Done.
        Next item: D$. Lexical retrieval for "D$": Adding inflectional features ['LANG:EN', 'PHI:DET:DEF', 'PHI:GEN:M', 'PHI:HUM:HUM', 'PHI:NUM:SG', 'PHI:PER:3',

        Next item: John-. Lexical retrieval for "John-": Done.

    2.  Consume /John/
        One solution due to sinking.Done.
        Ranking results:
        (1) [D↓+ John]
        Sinking John into D = D(N)

        Next item: admires. Lexical retrieval for "admires": Morphological decomposition for admire-#v#T/fin#[-s]: <admire-,v$,T/fin$,[-s]$>.
        Next item: "[-s]$": Inflectional feature [-s]$...Added ['LANG:EN', 'PHI', 'PHI:GEN:F', 'PHI:GEN:M', 'PHI:NUM:SG', 'PHI:PER:3', 'PHON'] into transient mem
        Next item: T/fin$. Lexical retrieval for "T/fin$": Adding inflectional features ['LANG:EN', 'PHI', 'PHI:GEN:F', 'PHI:GEN:M', 'PHI:NUM:SG', 'PHI:PER:3', '

    3.  Consume /T/
        Filtering and ranking merge sites...Filtering...Done. Ranking...Bottom-up baseline ranking...('T/fin')-Comp selection for D(N)...(-100) Done.
        Ranking results:
        (1) [D(N)↓+ T]
        (2) [D(N) + T]
        Considering [D(N) + T]...Transferring left branch D(N)...Result: [[D John] T]...Done.
```

Figure 4. A screenshot from the derivational log file.

Figure 4 shows the beginning of the derivation, where phonological words are mapped into lexical items (/John/ ~ D+N, /admires/ ~ T, v, V) and merged into the phrase structure. Once all items have been consumed, the result is passed to the syntax-semantics interface after several reconstruction operations. These operations are shown in Figure 5.

```
808   Trying spellout structure   [[D John] [T(v,V) φ(N)]]
809       Checking surface conditions...Done.
810       Transferring to LF...
811               1. Head movement reconstruction...Reconstruct Mary from φ(N)...=[[D John] [T(v,V) [φ Mary]]]...Re
812                   = [[D John] [T [v [admire [φ Mary]]]]]
813               2. Feature processing...Done.
814                   = [[D John] [T [v [admire [φ Mary]]]]]
815               3. Extraposition...Done.
816                   = [[D John] [T [v [admire [φ Mary]]]]]
817               4. Floater movement reconstruction...Done.
818                   = [[D John] [T [v [admire [φ Mary]]]]]
819               5. Phrasal movement reconstruction...Done.
820                   = [[D John]:2 [T [__:2 [v [admire [φ Mary]]]]]]
821               6. Agreement reconstruction...T acquired PHI:GEN:M from __:2 inside its sister...T acquired PHI:N
822                   = [[D John]:2 [T [__:2 [v [admire [φ Mary]]]]]]
823               7. Last resort extraposition...Done.
824                   = [[D John]:2 [T [__:2 [v [admire [φ Mary]]]]]]
825       Done.
826       LF-legibility check:
```

Figure 5. Transfer to the syntax-semantics interface.

The structure generated from the input words waiting for transfer is represented on line 808, which then undergoes several reconstruction steps (1-7). Case-based reconstruction, discussed earlier in this document, is one line 817. Nothing happens because English arguments are not marked for case. The result is represented at line 824, which is subjected to several LF-legibility checks (826). Case checking, among other tests, is performed here. If these tests pass, as is the case here, the result is passed to semantic interpretation, which is illustrated in Figure 6.

```
827         Interpreting TP globally:
828             Project [D John]: (1, QND)(2, GLOBAL)
829             "v" with ['PHI:DET:_', 'PHI:NUM:_', 'PHI:PER:_'] was associated with 1. [D John]    (alternatives:
830             Project predicate 'v': (3, PRE)(4, GLOBAL)
831             "admire" with ['PHI:DET:_', 'PHI:NUM:_', 'PHI:PER:_'] was associated with 1. [φ Mary]    (alternat
832             Project predicate 'admire': (5, PRE)(6, GLOBAL)
833             Project [φ Mary]: (7, QND)(8, GLOBAL)
834         Denotations:
835             [D John]~['2']
836             [φ Mary]~['8']
837         Assignments:
838             Assignment {'1': '2', '7': '8'} accepted.
839         Calculating information structure...Done.
840         Solution accepted at 1205ms stimulus onset.++
841
842     Semantic interpretation:
843     Recovery: [{'Agent of v([D John])'}, {'Patient of admire([φ Mary])'}]
844     Aspect: []
845     DIS-features: []
846     Operator bindings: []
847     Semantic space:
848     Speaker attitude: ['Declarative']
849     Assignments:
850     [D John] ~ 2, [φ Mary] ~ 8, Weight 1
851     Information structure: {'Marked topics': [], 'Neutral gradient': ['[D John]', '[φ Mary]'], 'Marked focus':
852     spellout structure: [[D John] [T(v,V) φ(N)]]
853     surface structure: [[D John] [T [v [admire [φ Mary]]]]]
854     s-structure: [[D John] [T [v [admire [φ Mary]]]]]
855     LF structure: [[D John]:2 [T [__:2 [v [admire [φ Mary]]]]]]
```

Figure 6. Screenshot from the derivational log file recording semantic interpretation.

The interpretation is recorded on lines 827-839, and the result on 840 which shows that the solution was accepted. It is followed by the summary (lines 842-855), which is also visible in the results file. This is followed by a list of the contents of all lexical items used in the derivation and a detailed list of all of semantic objects that were projected into existence.

Once the full solution is calculated, the model backtracks and tries to find other solutions. From the derivational log file we find that the model tried also (5) but rejected them.

(5) a. $[_{\varphi P}$ [D John] [[T [v admire]] [φ Mary]]]

b. $[_{DP}$ D [John [T [v [admire [φ Mary]]]]]]

c. $[_{DP}$ D [John [[T [v admire]] [φ Mary]]]]

None of these solutions pass the LF-legibility test. For example, the last was rejected because “John has wrong complement [[T [v admire]] [φ Mary]]” (line 1100, derivational log file). The fact that only these solutions were tried means that several other logically possible combinations were rejected before anything was send to the syntax-semantics interface. Indeed some combinations were considered so ill-formed that the parser rejects them without submitting anything to the semantic system.

Sometimes the whole derivation involves several branches that are hard to pick up from the linear log file. A separate simple log file (*case_corpus_simple_log_FINAL.txt*) is created to make the derivation more transparent. The whole derivation for the English sentence *John admires Mary*, as it was recorded into the simple log file, is as follows.

```
1    D
     D + John
2    D(N)
     D(N) + T
3    [[D John] T]
     [[D John] T] + v
4    [[D John] T(v)]
     [[D John] T(v)] + admire
5    [[D John] T(v,V)]
     [[D John] T(v,V)] + ϕ
6    [[D John] [T(v,V) ϕ]]
     [[D John] [T(v,V) ϕ]] + Mary
7    [[D John]:2 [T [__:2 [v [admire [ϕ Mary]]]]]] <= accepted
8    [[D John] [[T [v admire]] ϕ]]
     [[D John] [[T [v admire]] ϕ]] + Mary
9    [D(N) T]
     [D(N) T] + v
10   [D(N) T(v)]
     [D(N) T(v)] + admire
11   [D(N) T(v,V)]
     [D(N) T(v,V)] + ϕ
12   [D(N) [T(v,V) ϕ]]
     [D(N) [T(v,V) ϕ]] + Mary
13   [D(N) [[T [v admire]] ϕ]]
     [D(N) [[T [v admire]] ϕ]] + Mary
```

This shows how the lexical elements were merged, what solutions were generated and considered, and which solutions were rejected. Notation X(Y) refers to a syntactically complex head, such as *admires* = T(v, V) = a finite tensed (T) transitive (v) verb (V).

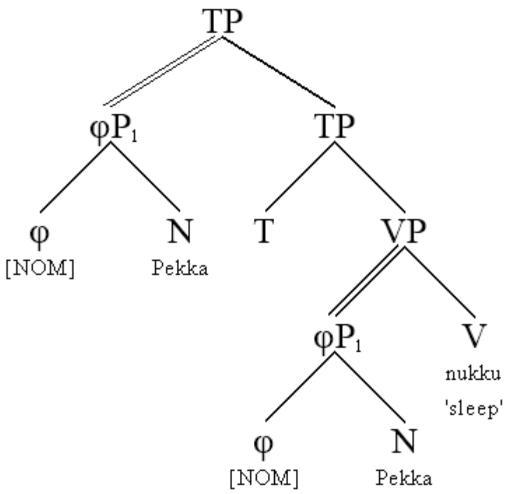
5 Results

5.1 Intransitive, transitive and ditransitive clauses

5.1.1 Intransitive clause

We begin with a group of standard finite clauses modified by valency and word order, most of which are grammatical (sentences #4-13). The model analyses an intransitive clause *Pekka nukku-u* 'Pekka.nom sleep-prs.3sg' (#4) as shown in (6).

(6)



The tensed finite verb *nukku-u* 'sleep-prs.3sg' is decomposed into two lexical items T (tense, finiteness, mood) and V (verb stem) by the lexico-morphological component, which provides that *nukkuu* 'sleeps' contains two constituents V#T (= verb, finite = finite verb) and feeds these two elements into the syntactic parser component, in a well-defined order T ∗ V, where they first form a complex head $T(V)^0$ (lines 1166-1170, log file) reconstructed into two separate heads T and V (line 1176, log file).[20] The

[20] Thus, the mapping is: phonologically complex word W ~ morphological decomposition A#B ~ syntactically complex head $A(B)^0$ ~ separate heads in syntactic structure […A…[…B…]]. Why not to map A#B directly into […A…[…B…]]? The first version of the model never produced any syntactically complex heads (Brattico & Chesi, 2020). Complex heads were posited in order to capture Finnish long head movement (Brattico, 2022b).

same mechanism generates full φPs from proper name: /Pekka/ ~ φ#N ~ φ(N) ~ $[_{\varphi P}$ φ(__$_1$) N$_1$] (phonology ~ lexicon ~ syntax~ LF interface) by a trivial head chain. These operations, taken together, take compressed sensorimotoric objects as input, two phonological words in this case, and expands them into representations that the semantic systems can interpret. The preverbal subject topic reconstructs from SpecTP into SpecVP. I return to phrasal reconstruction below.

Let us consider how case information enters into the analysis, keeping the above remarks in mind. The referential argument *Pekka* is marked by the nominative case in the lexicon. The feature appears on φ. This information originates in the lexicon, where /Pekka/ is decomposed as N#φ…#NOM and where NOM refers to Finnish nominative case and ends up inside φ as a lexical feature. All inflectional features that appear in the linear morphological decompositions are inserted inside the previous lexical item (here φ) as features and do not project independent morphemes.[21] The sequence is N#φ#...#NOM ~ N#$\varphi_{[NOM]}$, where "…" contains other inflectional features. The same operation handles agreement on the finite verb, thus T#3sg ~ $T_{[3sg]}$, which means that the third person agreement features reside inside T in (6) where they are later checked against the grammatical subject.[22] All structural case features were inflectional; semantic case features mapped into adpositions, following (Nikanne, 1993) and discussed below.

Final case checking takes place at the syntax-semantics interface. The operation targets case features at the lower φ and tries to check them. As specified in the main article, the NOM that appears in the lexical decomposition is mapped to TAIL:FIN,ARG,VAL which occurs inside φ as a lexical feature and requires checking against {FIN, ARG, VAL} during syntactic processing. Checking is done against finite T. Checking is also attempted at SpecTP, but it fails: there are no heads in the upward path. Failure triggers chain creation: the argument is reconstructed from the topic position into SpecVP where case checking succeeds (line 1182, log file). We can think of the case form as guiding referential arguments into their designated thematic scope positions, adding a layer of error tolerance into the algorithm.

This raises the question of why referential arguments appear at "wrong" positions in the first place. In Finnish, almost any argument can occur in the preverbal position where it receives the topic reading.

[21] If this were not the case, the model would regard NOM as an independent morpheme and create a separate "nominative head" into the syntactic representation. This is not completely implausible and could be simulated by making the required changes into the lexicon.

[22] If we specified that '3sg' refers to a morpheme, not an inflectional feature, separate Agr head would be generated. This is not be completely implausible; see (Pollock, 1989).

The operation is not semantically inert. It is used by the language system to carry information structural information. This insight goes back to (Vilkuna, 1989, 1995) who showed that Finnish finite clauses are composed of three distinct syntactic "fields" she called K-field (contrast), T-field (topic) and V-field (verb, event). We assume that the T-field, representing topics, is represented by SpecTP. This follows (Vainikka, 1989; Vilkuna, 1995). In addition to representing topics, the preverbal subject position *must* be filled with some constituent in a canonical finite clause. Canonical verb initial clauses are ungrammatical and are correctly judged as such by the model (#12-13, 24-29). Finnish exhibits a topic-based EPP, as argued by (Holmberg & Nikanne, 2002).

Many generative theories assume that intransitive declarative clauses project a further C-projection, so the analysis would be CP + TP + VP. Analysis (6) does generate empty C-projections. There is no phonological material in the input to prompt it. Phonologically empty categories, properties or dependencies are almost never hallucinated into the structure; the model projects information that is present in the input.

Intransitive clauses were tested with wrong case forms, listed in (7). They were correctly analysed as ungrammatical.

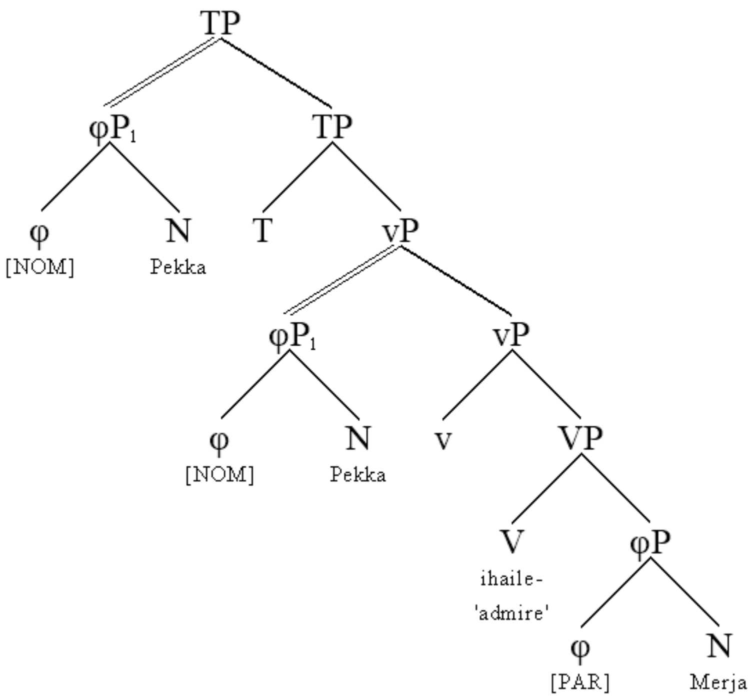
(7)	a.	*Merja-a	nukkuu. (#38)
		Merja-par	sleeps
	b.	*Merja-n	nukkuu. (#39)
		Merja-gen	sleeps
	c.	*Merja-n	nukkuu. (#40)
		Merja-acc(n)	sleeps
	d.	*Merja	nukkuu. (#41)
		Merja.acc(0)	sleeps

These arguments could not be reconstructed into positions where the respective case features could be checked.

5.1.2 Transitive clause

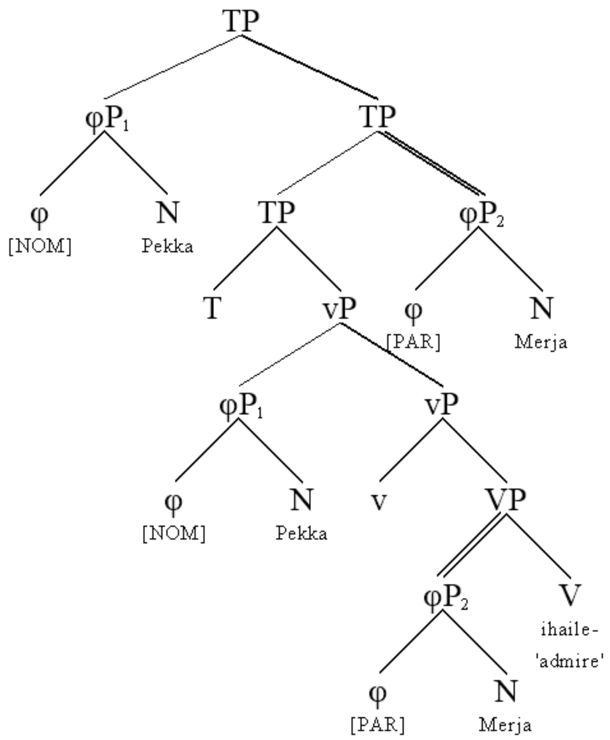
Transitive clauses have two referential arguments, while the transitive verb is decomposed into V#v#T in the lexicon where v represents the small verb that is part of all transitive verbs. A canonical transitive clause in Finnish is analysed as (8)(#5). Notice the appearance of v as a separate head in the structure. The direct object is assigned the partitive; accusative constructions, which involve more complex properties, are examined in Section 5.2.

(8)



Both arguments are extracted by trivial chains N#φ ~ $φ(N)^0$ ~ [φP φ($\_{_1}$) N_1], and the transitive finite verb as V#v#T ~ $T(v, V)^0$ ~ [$_{TP}$ …T [$_{vP}$ …v V…]]. Inflectional case features will be inside φ, as shown above, and verbal agreement inside T. Nominative case is checked as below, after the topic has been reconstructed to SpecvP where it receives a thematic role. The partitive, which constitutes a direct object case in Finnish (in fact, a "default" complement case), is checked against {ARG,–VAL} at V or v, here against V (in fact both V and v have these features, see lines 1448-1449 in log file). There is a second analysis, however, shown in (9).

(9)

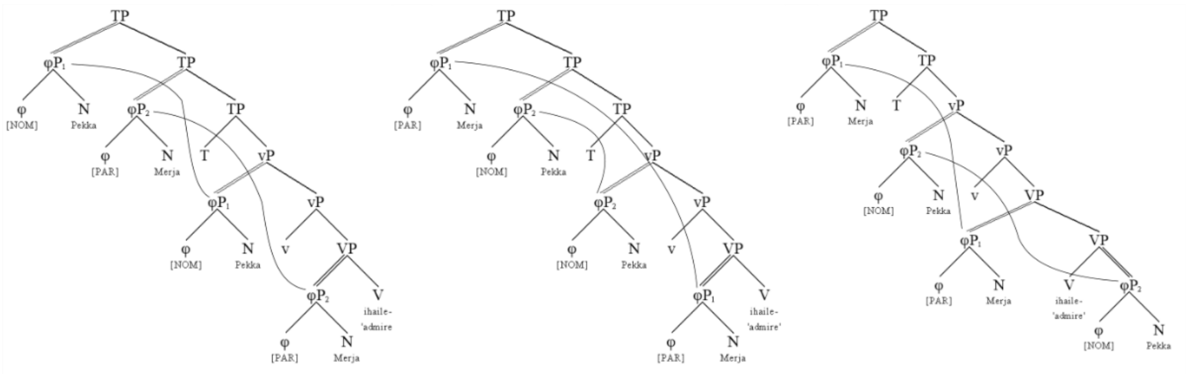


This derivation, lines 1525-1574, is generated when the parser attempts to merge the direct object *Merja-a* ‘Merja-par’ into the high right adjoined position. The grammatical principles currently implemented in the syntactic background theory are able to reconstruct or “rescue” the direct object from this position by reconstructing it inside the VP (see chain 2, line 1546 in log file) where it receives a thematic role and checks the partitive case, and the resulting representation passes at the syntax-semantics interface. It is difficult (at least, to me) to find empirical reasons to support or reject this analysis, so it was left into the model. This analysis does not exist in English due to the lack of overt case forms.[23]

Most permutations of the arguments of finite clauses are grammatical in Finnish, and are derived by reconstructing arguments into thematic positions on the basis of their case forms. Example (10) summarizes the analyses for SOV, OSV and OVS sentences #9-11.

[23] With the exception of pronouns. It is an interesting empirical and theoretical question why pronominal case forms do not license pronominal word order permutations in English.

(10)

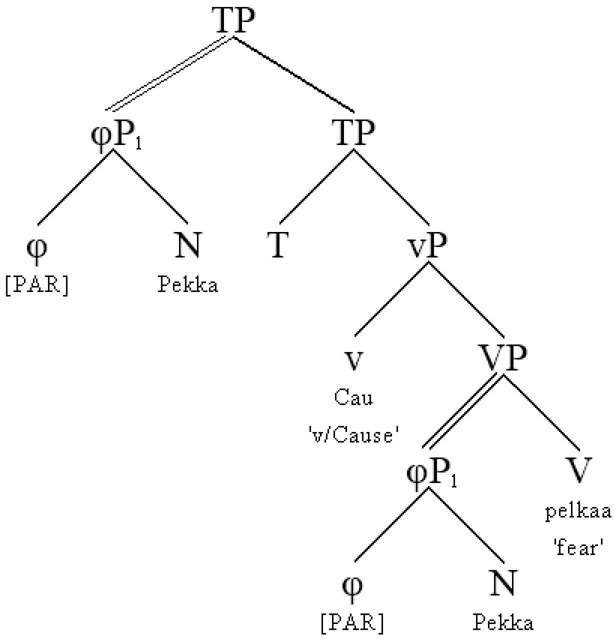


SOV (#9) OSV (#10) OVS (#11)

The derivations are self-explanatory. Information structure calculations require a comment. When both the subject and object are fronted into the T-field (#9, 10), they are interpreted as topics (lines 315, 347 in results file). Therefore, these sentences are multitopic constructions. They could be translated as 'As for both Pekka and Merja, they …' In the case of OVS, the direct object is interpreted as the topic, grammatical subject as the information focus. To me these interpretations appear linguistically plausible, and can be supported by empirical argument. Their properties were first discussed by Vilkuna (1989).

The exceptional experiencer construction with a partitive subject is sentence #139, with impossible case forms #140-143. The experiencer construction is analysed as (11).

(11)



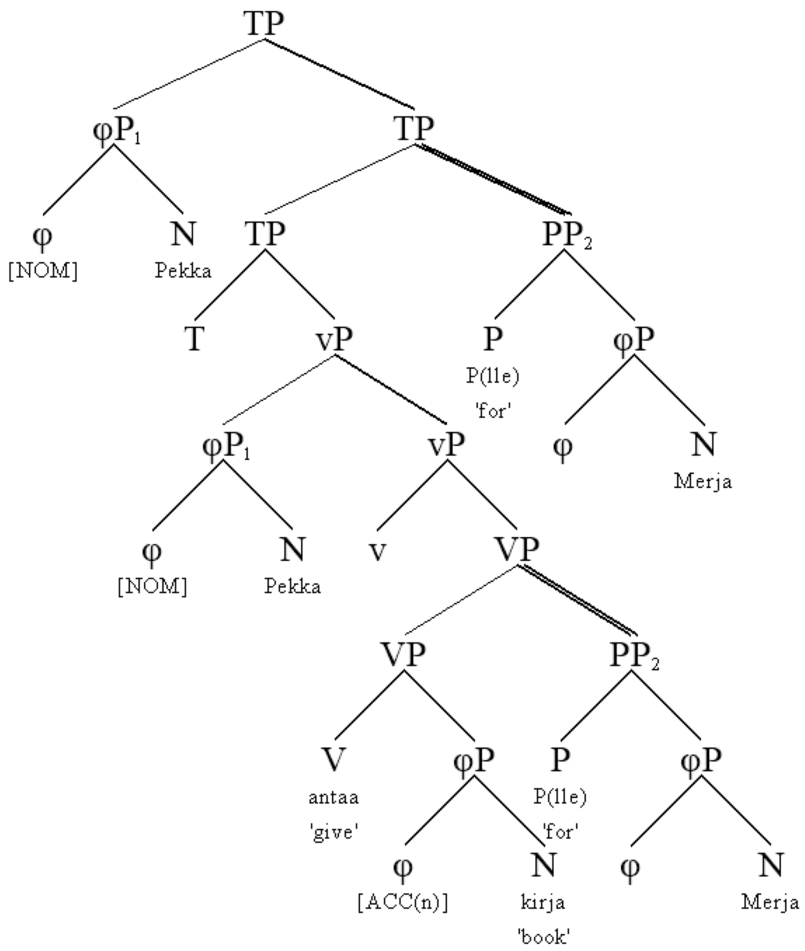
Notice how the partitive experiencer subject is reconstructed into SpecVP where it checks the partitive case against v (here, corresponding to the causative morpheme *-tta-* in the finite verb).

In English, most referential arguments are not marked for case and do not have lexical case features. Hence thematic roles depend on surface word order alone. Sentence #3 is an English transitive clause *John admires Mary*. If we reserve the positions of the referential arguments, their thematic roles will also flip. Furthermore, both arguments are collected under “Neutral gradient” in the results file when it comes to the information structure (line 103, results file). This means that neither the subject nor the object were interpreted as marked topics or marked focus elements, instead what happened was that *John* was interpreted “more topic-like” than *Mary*. The neutral gradient orders the argument in terms of their topicality without triggering marked topic or focus readings.

5.1.3 Ditransitive clause

A ditransitive clause such as *Pekka antoi kirjan Merjalle* ‘Pekka.nom give-pst.3sg book-acc Merja-all’ (#6) is analysed as in (12).

(12)



The PP *Merja-lle* ‘Merja-all’ (all = allative case, roughly ‘to’ or ‘for’) is right-adjoined into TP and then reconstructed inside the VP (the analysis where semantic cases decompose into PPs comes from (Nikanne, 1993)). Both semantic and structural cases are targeted by case-based reconstruction, and because semantic cases are decomposed into PPs, this means that DPs and PPs are both treated in the same way. Since both the direct object and the indirect object can be right-adjoined, the Finnish ditransitive clause has several surface ambiguities collected into the Figure 7 below. At the syntax-semantics interface, however, they all map into the same canonical form due to reconstruction. We can perhaps think of these ambiguities as existing only “briefly” during processing.

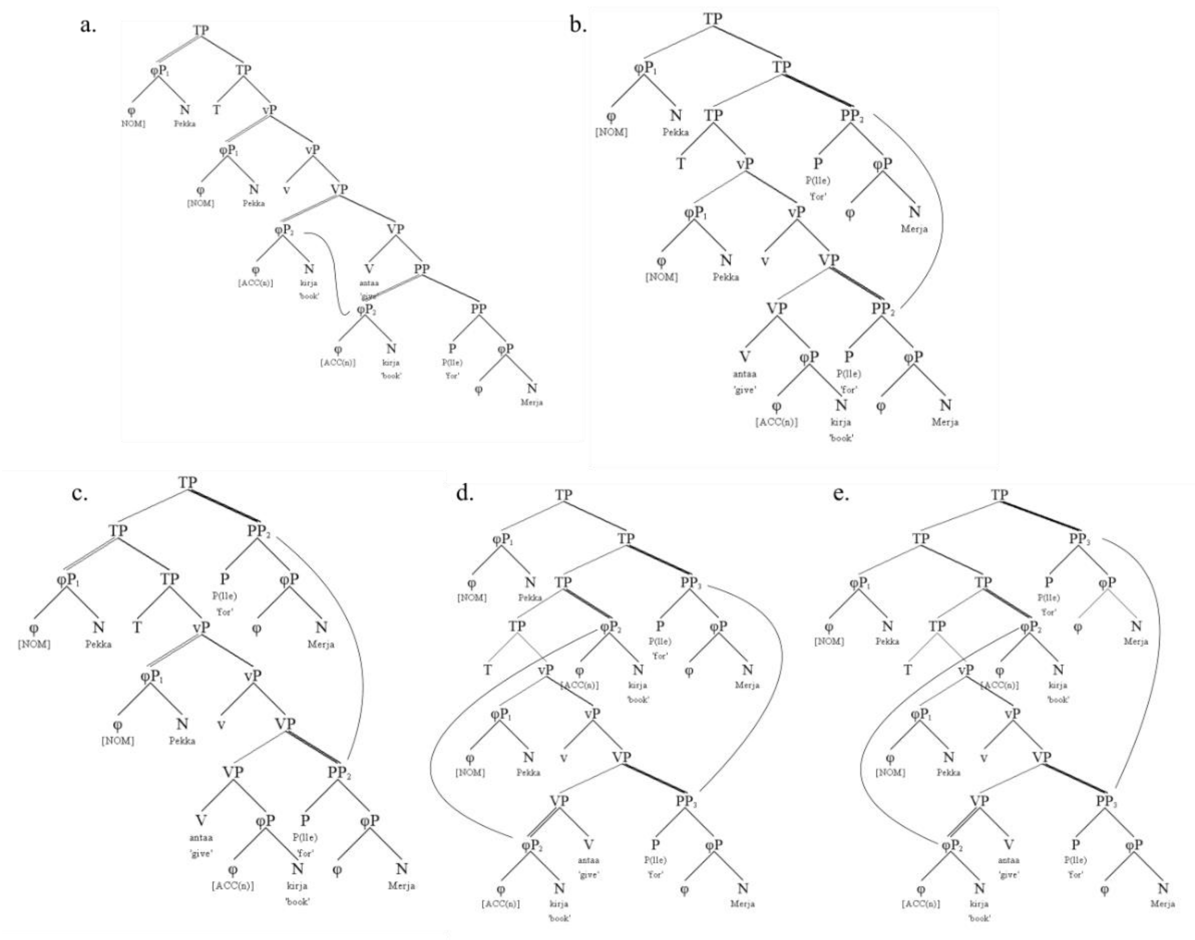


Figure 7. Five surface ambiguities of Finnish ditransitive clauses.

Solution (a) is the first one found and thus listed in the results file (lines 191-230). Noncanonical words orders are tested by sentences #8-13, with the ungrammatical verb-initial sentences being #12-13 that have already been discussed.

Transitive and ditransitive sentences with missing arguments were #42-44 and extra arguments #80-85, and were correctly judged as ungrammatical. Sentences with too many arguments are relevant here, since we must make sure that there are no spurious reconstruction possibilities. Tests that involved wrong case forms are more relevant, but due to the large amount of possible case form combinations there were rather many (#45-79). I will not list them here, suffice it to say that in the case of transitive clauses I considered all possible combinations of case forms (*nom/nom, *nom/gen, … *acc(0)/acc(0)).

5.2 Accusative

5.2.1 Aspect

Direct objects of transitive finite clauses as well as deverbal constructions can be assigned also the accusative case in addition to the general complement case partitive. The distribution of the accusative is narrower than the partitive. The fact that the accusative is a verbal case is captured by assuming that all accusative forms (n-accusative, zero-accusative, t-accusative, see below) check the aspectual feature that only appears in connection with verbs and verbal stems. The accusative is for this reason ruled out by the model in all other constructions. Adpositions, for example, do not have aspectual features (*kohti talo-a* 'towards house-par', **kohti talo-n* 'towards house-acc').

The relevant aspectual feature denotes a bounded event, following (Kiparsky, 1998, 2001). The feature will be abbreviated as +ASP (ASP:BOUNDED in the raw input/output). The feature can be part of the lexical entry of any verb or, to make testing easier, it can also be provided in the input string as an inflectional morpheme (no such morpheme exists in Finnish, however, so this option exists only for testing purposes). A verb that has lexical +ASP licenses but does not require the accusative case on current assumptions (i.e. presence of +ASP does not violate the partitive rule). Verbs that have both bounded and unbounded readings are handled in this study by assuming that the lexicon contains two variants of the same verb.[24] A verb that does not have +ASP cannot take an accusative direct object. The range of options is summarized in the Table below.

Table 1. Three verb classes as a function of presence/absence of ASP.

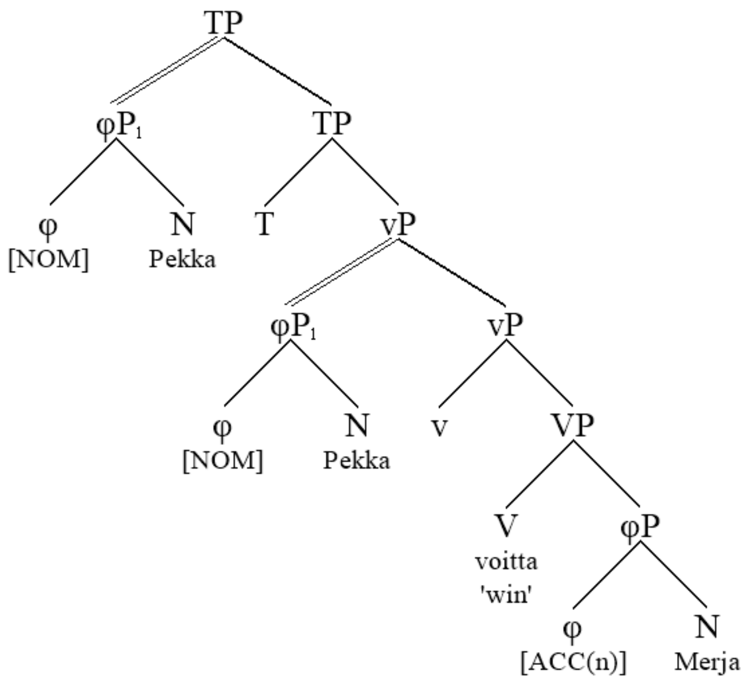
feature	Verb class	Examples
−ASP (NO ASP)	Non-telic verbs, only partitive direct objects	Pekka rakastaa Merja-a/*?Merja-n 'Pekka loves Merja-par/Merja-acc.'
+ASP	Telic verbs, accusative and partitive, the latter generating special reading[a]	Pekka voitti Merja-n/??Merja-a 'Pekka beat Merja-acc/Merja-par'
±ASP	Ambiguous verbs, both partitive and accusative, both non-telic and telic	*Pekka pesi Merja-n/Merja-a* 'Pekka washed Merja-acc/Merja-par'

[a] Captured in the semantic component, as explained in the main article.

[24] This assumption is incorrect since it doubles the size of the verbal lexicon.

A simple accusative transitive clause *Pekka voitti Merja-n* ‘Pekka.nom won Merja-acc(n)’ (#163) is analysed as any regular transitive clause, but with the aspectual feature at the verb. The n-accusative appears here for reasons discussed below; all accusative forms are checked against {+ASP, +ARG} however and could occur in the direct object position in this sentence

(13)



The field “Aspect” in the results file reads “Aspectually bounded” (line 2659) to register the fact that the event was interpreted as telic. The partitive direct object is also possible, since v-V has {ARG, –VAL}, but the aspect field now reads “Aspectually anomalous” (line 2699). This registers the fact that the sentence *Pekka voitti Merja-a* ‘Pekka won Merja-par’ (#164) has a nonstandard reading, for example, one where Pekka won a piece of Merja (comparable to *Pekka voitti kulta-a* ‘Pekka won gold-par, i.e. the first prize’). An ambiguous verb *pesi* ‘wash’ (#166, 167) is interpreted as aspectually bounded only under a reading where the verb is marked for +ASP. Notice that since ambiguity was handled by lexical ambiguity, several syntactic analyses are generated that differ only in terms of one lexical item (see lines 2831-2843 in the results file). Verbs that do not have +ASP cannot take accusative direct objects (#171, 172, 173). Since the accusative direct object cannot be checked, reconstruction is attempted. Reconstruction was tested by sentences #174-180, with both grammatical and ungrammatical case configurations. They were correctly calculated.

5.2.2 Semantics of aspect

It is clear to a native speaker that when the verb phrase denotes an event, the presence of the accusative direct object elicits the bounded reading. A bounded reading means that the event endpoint is included into the denotation of the verb phrase. If the object is not case-marked by the accusative, the endpoint is excluded.[25] If the verb phrase does not denote an event, however, the accusative-partitive alteration is usually said to represent "incompleteness" in the sense shown in (14)-(16).[26]

(14)

a.	Peitto	peittä-ä	vauva-a.
	blanked	cover-prs.3sg	baby-par
	'The blanked covers some of the baby (not necessarily the whole baby).'		
b.	Peitto	peittä-ä	vauva-n.
	blanked	cover-prs.3sg	baby-acc(n)
	'The blanked covert the whole baby.'		

(15)

a.	Näe-n	hän-et./	Näe-n	hän-tä.
	see-prs.1sg	he-acc(t)	see-prs.1sg	he-par
	'I see him.'		'I see a part of him/I see him occasionally.'	
b.	Tiedä-n	se-n./	*Tiedä-n	si-tä.
	know-prs.1sg	it-acc(n)	know-prs.1sg	it-par
	'I know it.'		?	

---

[25] It is neither presupposed nor implied that the endpoint was not reached; rather, it is excluded from the denotation of the verb phrase. This could be because the event is still ongoing, the agent decided to terminate it without reaching the endpoint, or because the speaker does not know if the endpoint was reached. The endpoint refers to a discontinuous event boundary after which it is not logically possible to continue the action denoted by the verb. If we consider the denotation as an event that evolves over time, it marks the outer edge of the event interval.

[26] The semantic interpretation of the whole clause relies on the properties of the whole verb phase and not just on the lexico-semantic features of the verb (Csirmaz, 2005, 2012; Kiparsky, 1998; Megerdoomian, 2000, 2008; Verkuyl, 1972). The aspectual feature alone does not suffice to determine an interpretation for the whole verb phrase and, conversely, semantic properties arising from the whole verb phrase cannot be used in any trivial way to infer aspectual properties of the verb.

(16) Funktio	saavutta-a	nolla-a./	nolla-n.
function.nom	reach-prs.3sg	zero-par	zero-acc(n)

‘The function approaches zero’ (partitive)
‘The function reaches zero’ (accusative)

This proposal leads us back to another traditional view, which regards accusative objects as “total objects,” an object/participant that “is totally affected by the action of the verb” whereas the partitive means it is “partially affected” (Thomas, 2003, p. 25). Yet notions such as object affectness, total objects, maximal affectness or totally affected are imprecise, and their relation to aspect boundedness is not obvious. Consider (17), discussed by Huumo (2013).

(17) Pekka	tönäis-i	Merja-a./	*Merja-n.
Pekka.nom	nudged-pst.3sg	Merja-par	Merja-acc

‘Pekka nudged/pushed Merja.’

Why is Merja not affected enough to justify the accusative? There are also accusative-assigning relations that do not affect the object at all, such as ‘to see’, ‘to solve’, and many two-place relations such as ‘hit’ that seemingly affect the object yet always take the partitive. Furthermore, abstract mathematical relationships exhibit the partitive-accusative alteration yet nothing is literally affected (18).

(18) a. Funktio saavutta-a nolla-a.
function reaches-prs.3sg zero-par
‘The function closes on zero (e.g., at infinity).’

b. Funktio saavutta-a nolla-n.
function reache-prs.3sg zero-acc
‘The function reaches/crosses zero.’

These relations do not affect the zero (as an object) in any way, they do not evolve necessarily in time, and the function can have a well-defined endpoint ($y = 0$) at ∞. Partial affectness and incompleteness remain obscure. A third possibility discussed in the literature homogeneity (Csirmaz, 2012; Kiparsky, 1998; Krifka, 1989; Link, 1998; Rothstein, 2004; Tenny, 1994). A homogeneous property is one which applies to a part of an object if it applies to the object itself (for various formulations, see Rothstein, 2004: 10-12), hence one could use divisibility as the underlying property, such that the partitive signals that the object or the event is divisible or homogeneous.

Many of the notions discussed above, especially boundedness and homogeneity, seem to be grounded in topology or topological intuitions. Topology constitutes a set-theoretical generalization of metric spaces that is concerned with boundaries and continuity. Thus, a fully homogeneous/divisible object is one that has a discrete topology, meaning that it consists of stuff that has no intrinsic shape and can be divided or organized in whichever way without losing essence. Boundedness, whether it occurs in connection of events or spatial relations, is one of the most fundamental topological distinctions, one between internal parts of an object and its boundary. Topological properties have an additional merit that they apply to both spatiotemporal events and spatial configurations. In the case of the blanket's covering the baby, configurations where the blanket covers the whole baby constitutes a logical boundary in this sense. Consider the problem with aspectually bounded punctuated events that takes partitive objects, such as nudging. If we zoom in around the centerpoint of nudging, we find further instances of the same relation. The relation is punctual in the sense of being arbitrarily small yet still continuous in topological sense.[27] This is not the case of punctual achievements such as 'to win' or 'to find' which require accusative objects. Winning occurs in a single event that is not true in all of its spatiotemporal neighborhoods. It is false at the discontinuous moment before the single point at which winning occurs. Thus, perhaps what matters is not whether there is a single point ('nudging' vs. 'winning'), but whether the event denotes a discontinuous point ('winning') or an arbitrary small but continuous area surrounding the point ('nudging'). Nonspatial, abstract relations can be evaluated against one-dimensional time. 'To think about a problem' is continuous in relation to time, hence it takes the partitive, while 'to solve a problem' is not and is compatible with the accusative.

(19)	Pohd-i-n	ongelma-a./	Ratkais-i-n	ongelma-n.
	think-pst-1sg	problem-par	solve-pst-1sg	problem-acc
	'I thought about the problem'		'I solved the problem.'	

In conclusion, it seems possible that boundedness is ultimately a topological notion that applies to events, spatial configurations and even mathematical relationships, and is grammaticalized into the Finnish case system.

5.2.3 Polarity

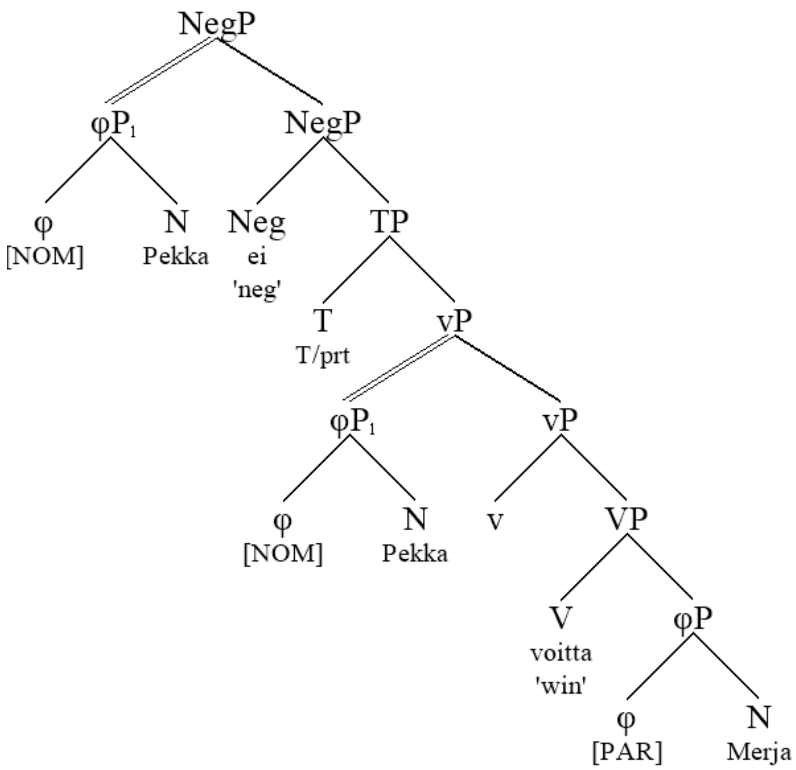
In addition of representing aspect, the accusative is also connected with polarity. It behaves like a positive polarity item and becomes ungrammatical under the scope of any negative polarity element such as the sentential negation particle *e-* 'not' used in the present study (20).

[27] A continuous function maps neighbourhoods of every f(x) into a neighbourhood of x. This means that arbitrary small changes can be induced to the output by performing sufficiently small changes in the input.

(20)	Pekka	e-i	voitta-nut	Merja-a/	*Merja-n. (#181, 184).
	Pekka.nom	not-3sg	win-pst.prtcpl	Merja-par	Merja-acc(n)
	’Pekka did not win/beat Merja.’				

The analysis calculated for (20) is (21).

(21)



When the direct object is marked for the accusative (any form, #182-185), its positive polarity feature –NEG ([*POL:NEG]) matches with the negative polarity feature at Neg and causes the derivation to crash. No reconstruction rescues the solution. Notice that {+ASP, +ARG} and {–NEG} are checked separately: the former against v-V, the latter against Neg. Reconstruction was tested by #186-194, both for grammatical and ungrammatical case configurations. (22) lists the grammatical sentences that were tested.

(22) a.	Merja-a	e-i	voittanut	Pekka. (#186)
	Merja-par	not-3sg	win-pst.prtcpl	Pekka.nom
	O	Neg	V	S
b.	Merja-a	ei	Pekka	voittanut. (#187)

	Merja-par	not	Pekka	won
	O	Neg	S	V
c.	Pekka	ei	Merja-a	voittanut. (#188)
	Pekka	not	Merja-par	won
	S	Neg	O	V
d.	Merja-a	Pekka	ei	voittanut. (#189)
	Merja-par	Pekka	not	won
	O	S	Neg	V
e.	Pekka	Merja-a	ei	voittanut. (#190)
	Pekka	Merja-par	not	won
	S	O	Neg	V

The analyses calculated for these sentences are in Figure 13.

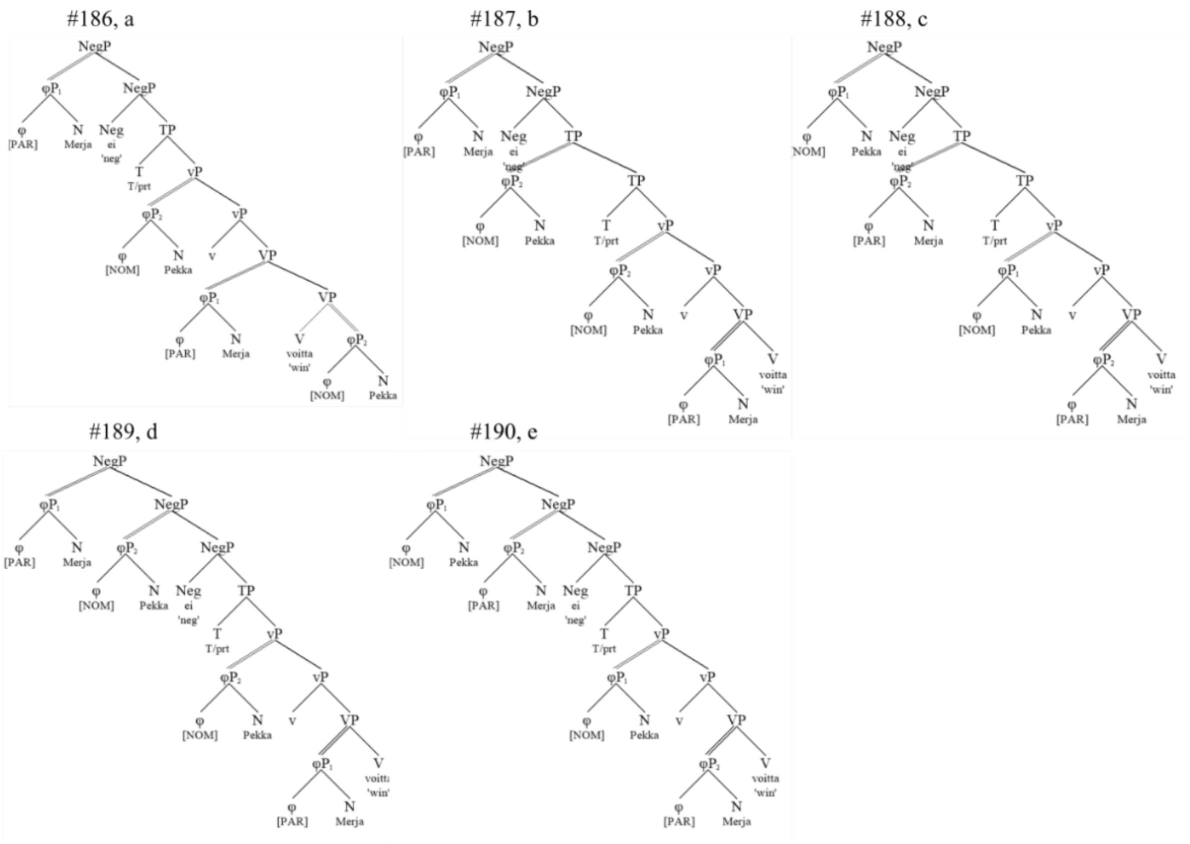


Figure 13. Calculated analyses for (22).

All accusative forms constitute positive polarity cases, matched with −NEG.[28] There is an alternative analysis which maintains that a negated event cannot be aspectually complete and hence the accusative is ungrammatical (Leino, 1991; Thomas, 2003; Vainikka, 1989). This analysis looks problematic. Consider the behavior of time-frame and time-span adverbs that are only compatible with aspectually incomplete or aspectually complete events, respectively (23).

(23)

a.	Jari	pes-i	hevos-en	tunnissa/	*tunnin.
	Jari.nom	washed-pst.3sg	horse-acc(n)	in.hour	for.hour
	‘Jari washed the horse (accomplished the washing) in an hour/ *for an hour.’				
b.	Jari	pes-i	hevos-ta	*tunnissa/	tunnin.
	Jari.nom	washed-pst.3sg	horse-par	in.hour	for.hour
	‘Jari washed the horse (but did not accomplish it) *in an hour/ for an hour.’				

If the partitive of negation were based on aspect, then negated sentences should not be compatible with time-span adverbs. This prediction is not borne out.

(24) Jari	e-i	pes-syt	hevos-ta	tunnin/	tunnissa.
Jari	not-3sg	wash-pst.ptcp	horse-par	for.hour	in.hour
‘Jari did not wash the horse in an hour/ for a hour.’					

What is negated is either a completed event or a non-completed event. Use of the time-frame adverbial provides a reading in which Jari did not wash the horse for an hour, but some other time period; use of a time-span adverbial means (among other possible readings) that completing the task did not take an hour (the most likely reading being one in which it took more time), but it was nevertheless

[28] There are certain cases in which a negative polarity element partitivizes a direct object optionally. Two examples of this phenomenon are hesitative adverbials such as ‘hardly’ and yes/no questions (Kaiser, 2002, 2003; Kiparsky, 1998), both arguably falling under some form of “sensitivity to (non)veridicality” phenomenon in the sense of (Giannakidou, 1998). This phenomenon was left for future research. The fact that nonveridical expressions license the partitive raises the question of whether the obligatory partitive in connection with a clause-mate sentential negation is due to negative polarity/strong antiveridicality or due to the presence of the negation as a formal item. Reime (1986) shows that the former is the case: when the negative word is used in a ‘non-negative’ way, as in the sentence *ei-kö-hän jouda vielä oluet?* ‘not-q-fam drink still beer-acc(t)’, meaning ‘should we still drink a few beer’, the partitive phenomenon disappears.

accomplished. In addition, negating an achievement is only compatible with aspectually bounded time-span adverbials, showing that the underlying aspect interpretation remains intact:

(25)	Jari	e-i	voitta-nut	kilpailu-a	*tunnin/	tunnissa.
	Jari	not-3sg	win-past.ptcp	competition.par	for.hour	in.hour

‘Jari did not win the competition *for an hour/ in a hour.’

The third and most important reason to keep the partitive of negation distinct from the partitive of aspect is that the partitive of negation, unlike the partitive of aspect, is a nonlocal phenomenon. The negation in a main clause affects the direct object(s) inside non-finite complement clauses. I will return to the long-distance effects later in this section.

5.2.4 Agreement

Pronouns and plural DPs are marked by the t-accusative form, and are only sensitive to aspect and polarity. Singular full DPs exhibit the n-accusative/zero-accusative alteration that is sensitive to agreement in addition. Descriptively the n-accusative appears under the scope of predicates that exhibit overt agreement (with another argument); the zero-accusative occurs under the scope of a non-agreeing predicate. This phenomenon is one of the most controversial domains in this arena. I return to the controversy at the end of this section. The feature for “overt agreement” in the present analysis is ±PHI: +PHi = overt agreement present, –PHI = no overt agreement. Phi checking was added to polarity and aspects rules for singular full DPs. Because the relevant behavior is controlled by *overt* agreement, not ‘possible agreement’ and hence ±VAL, +PHI is part of the lexical entries of inflectional phi-features. Negative condition –PHI is checked like polarity in this study, thus matching with +PHI causes rejection (technically the specification is therefore [*PHI] = reject if +PHI is matched). It is be possible to stipulate a negative –PHI feature, allowing for the third option “no feature,” but this was not needed to calculate the present dataset. The impersonal construction, which does not exhibit systematic inflectional agreement with an argument, was used in this study to verify the agreement rule (#195-203). The basic cases are (26), but also other case combinations were tested.

(26) a.	Me	löys-i-mme	*avain/	avaim-en. (#202, 195)	
	we.nom	find-pst.1pl	key.acc(0)	key-acc(n)	
		(+PHI)			

‘We found the key.’

b.	Me	löyde-ttiin	avain/	*avaim-en. (#196, 199)	
	we.nom	find-pst.impass	key.acc(0)	key-acc(n)	

‘We found the key.’

Structurally these are standard transitive clauses. Impersonal passive construction (b) is discussed in Section 5.7. Interaction between negation and agreement was tested by sentences #204-211.

The zero accusative, which is identical to the nominative in the case of singular full DPs, appears in many constructions in which the nominative grammatical subject is missing. This has led number of authors to develop analyses where the nominative case is assigned or shifted, in some manner depending on the particulars of the analysis, to the direct object (Anttila & Kim, 2011; Jahnsson, 1871; Maling, 1993, 2004; Poole, 2015, 2022; Toivainen, 1993; Vainikka, 1989; Wiik, 1972).[29] The problem is that there are sentences which contain several nominative-looking and nominative arguments. See (27).

(27)

a.	Me	ostet-tiin	uusi auto.			
	we.nom	buy-pst.impass.0	new car.acc/nom			
	‘We bought a new car.’					
b.	Me	näh-tiin	Pekka	ostamassa	uutta	auto-a.
	we.nom	see-pst.impass.0	Pekka.acc/nom	to.buy	new.par	car-par
	‘We saw Pekka buying a new car.’					
c.	Me	näh-tiin	Pekka	ostamassa	uusi	auto.
	we.nom	see-pst.impass.0	Pekka.acc/nom	to.buy	new.acc/nom	car.acc/nom

(28) (Examples from Holmberg & Nikanne (2008))

a.	Se	on Jari	lopetta-nut	tupakoinn-in.	
	it.nom	has Jari.nom	quit-pst.prtcpl	smoking-acc/gen	
	‘Jari has quit smoking.’ (=ex. 1a in H&N, p. 326)				
b.	Ne	sai kaikki	lapset	samat	oiree-t.
	they.nom	got all	children.nom	same	sympton-pl.acc(t)
	’All the children got the same symptons.’ (=ex. 1b in H&N, p. 326)				
c.	Me	ol-laan	mekin	lopetettu	tupakointi.
	we.nom	be-prs.impass.0	we.nom-kin	quit	smoking.acc/nom

[29] Many of these analyses are inspired by the Case-in-Tiers approach of (Yip et al., 1987), according to which case forms are assigned to arguments from a case tier list. When the external nominative argument is missing, first item in the case tier list, the nominative, goes to the next argument in the syntactic prominence hierarchy, the direct object. Hence we get nominative case assignment to the direct object in the absence of other nominative argument.

‘We have quit smoking, too.’ (=ex. 4g in H&N, p. 328)

d.	Se	se	on	Tarja-kin	lopettanut	tupakoinn-in.
	It.nom	it.nom	has	Tarja.nom-kin	quit	smoking-acc/gen

‘Also Tarja has quit smoking.’ (=ex. 30a in H&N, p. 338).

e.	Me	on	mekin	nähty	Pekka	ostamassa	uusi	auto.
	we.nom	be	we.nom	saw.impass	Pekka.nom	to.buy	new.nom	car.nom

‘We have seen Pekka buying a new car.’

Furthermore, in order to maintain case shifting these proposals are forced to ignore or exclude the partitive, plural and pronouns objects, all which do not exhibit the proposed shifting behavior. Anttila & Kim (2011), for example, claim on such grounds that the partitive is a lexical case. The pronouns, which exhibit the regular accusative case, are usually ignored. The matter remains debated.

Another controversial property is the observation that there are complex cases where two direct object case forms are possible (example (a) from Anttila & Kim, 2017, p. 593, ex. 30).

(29)

a.	Mati-n	luul-tiin	[$_{VA/infP}$ampuneen	karhu/	karhu-n.]
	Matti-gen	thought-impass.0	to.shoot	bear.acc(0)	bear-acc(n)

‘It was thought that Matti had shot a/the bear.’

b.	Me	tehtiin	[$_{DP}$ sopimus	ampua	karhu/	?karhu-n.]
	we.nom	made.impass.0	agreement	to.shoot	bear.acc(0)	bear.acc(n)

‘We made an agreement to shot a/the bear.’

Anttila & Kim (2017), who examine this phenomenon in considerable detail, propose an optimality theoretical analysis where case configurations compete with each other in such a way that the competition can end up in a stalemate, leading into labile case forms. This phenomenon was ignored in the present study.[30] I think the approach by A&K has considerable merit: some type of competition could well be at stake.

[30] Including labile case constructions into the present study would require modelling very complex constructions where these case phenomena occur. Furthermore, the resulting analysis would be credible only to the extent that the underlying basic case signature is first calculated correctly and in a plausible way. I decided on such grounds to put the matter aside.

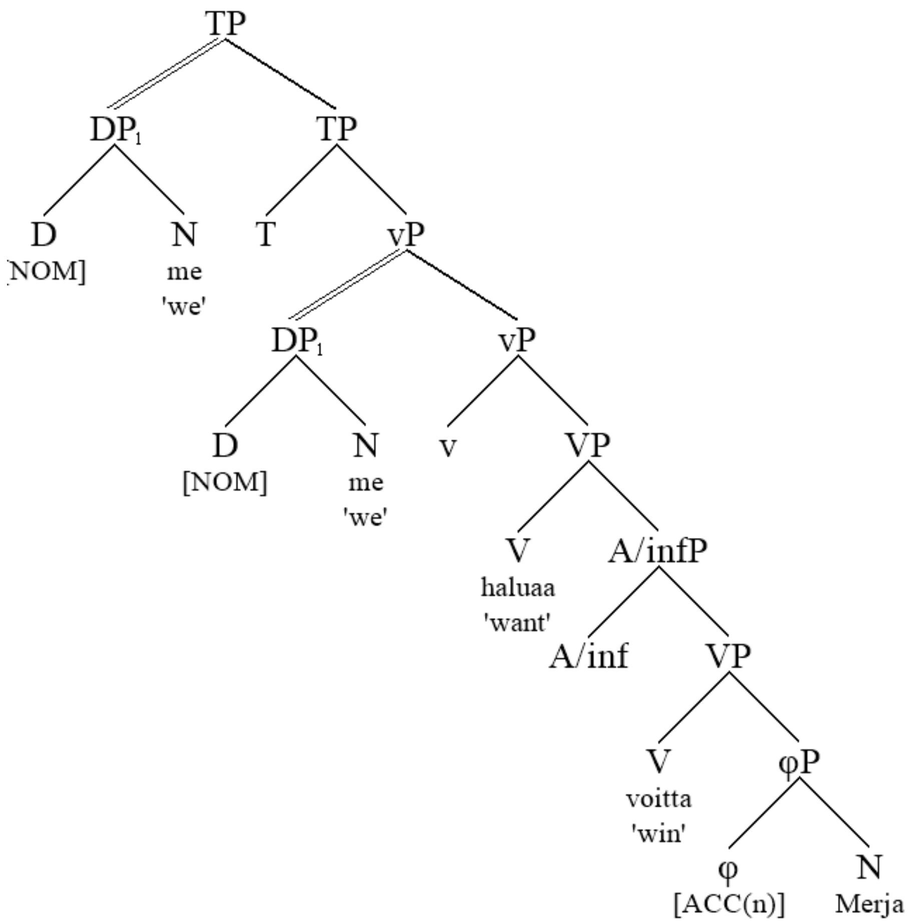
5.2.5 Nonlocal effects

Finnish case assignment is nonlocal. The phenomenon has been known since the 1950s; perhaps the first rigorous study was (Toivonen, 1995). The agreement and negative rules span arbitrary number of infinitival boundaries, both complements and adjuncts; and the negative rule seems to cross even the finite clause boundary (although the effect is weakened). The data involved in this domain is complex, and the case forms can under some circumstances become labile, unable to decide between two forms. Since the literature was cited in the main article, I don't review it here. See (Anttila & Kim, 2017; Brattico, 2014; Poole, 2022; Vainikka & Brattico, 2014 for recent work). Only the basic cases were included into this study to verify that the analysis does cover the effects, and that the core cases are calculated correctly. Let us consider how the hypothesis captures the long-distance effects. We compare (30)a-b.

(30) a.	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.'				
b.	Me	halut-tiin	voitta-a	Merja/	*Merja-n. (#216, 226)
	we.nom	want-impass	win-a/inf	Merja.acc(0)	Merja-acc(n)
	'We wanted to win Merja.'				

They are calculated as (31), with the impersonal verb replacing the active verb in (b).

(31)



The relevant agreement features (hence ±PHI) occur at the main clause T. Similarly, polarity effects are based on a dependency between Neg in the main clause and the direct object inside the embedded A-infinitival (#228-229). The accusative form that responds to these features is at the direct object inside the A-infinitival. Both the nature of the infinitival and the number of embedded infinitivals are in principle irrelevant[31]; and it does not matter whether the direct object occurs inside an adjunct or complement. These dependencies are therefore not local. The reason these sentences are calculated correctly by the model is because if a case is checked against one feature, in this case separately against {±NEG} and {±PHI}, the case checking relation scans the structure until either there is match or there is no more structure. As a consequence, the dependency is in principle unbounded. In the phrase structure

[31] One can demonstrate long-distance effects through any infinitival.

above, ACC(N) at the direct object searches through the whole structure for +PHI until it encounters a match at main clause T. The real problem then becomes how to explain case patterns that are local. Locality follows when cases are checked against several features (e.g., AGR, –VAL). This triggers intervention.[32]

Negation and phi-features can affect several direct objects.

(32)

a.	Me	e-mme	nähneet	Jari-a	syömässä	leipä-ä.
	we	not-1pl	saw-pst.ptcp	Jari-par	to.eat	bread-par
	‘We didn’t see Jari eating the bread.’					
b.	Me	näi-mme	Jari-n	syömässä	leivä-n.	
	we	saw-1pl	Jari-acc(n)	to.eat	bread-acc(n)	
	‘We saw Jari eating the bread.’					

(33)

a.	Me	nähtiin	Jari	syömässä	leipä.
	we	saw.impass.0	Jari.acc(0)	to.eat	bread.acc(0)
	‘We saw Jari eating the bread.’				
b.	Me	näi-mme	Jari-n	syömässä	leivä-n.
	We	saw-1pl	Jari-acc(n)	to.eat	bread-acc(n)
	‘We saw Jari eating the bread.’				

The fact that several objects can be affected nonlocally by one and the same lexical item is captured since an upward path can extend from each relevant case assignee to the one and the same case assigner.

[32] Previous literature has established that restrictions on nonlocal case assignment are not identical to restrictions on *wh*-movement, a fact first reported by (Toivonen, 1995). The hypothesis that nonlocal Case assignment is based on an Ā-dependency (Brattico, 2012) must be rejected on such grounds. (Anttila & Kim, 2017; Toivonen, 1995; Vainikka & Brattico, 2014) all argue against it. The model proposed here suggests that they are based on feature intervention, following (Brattico, 2014).

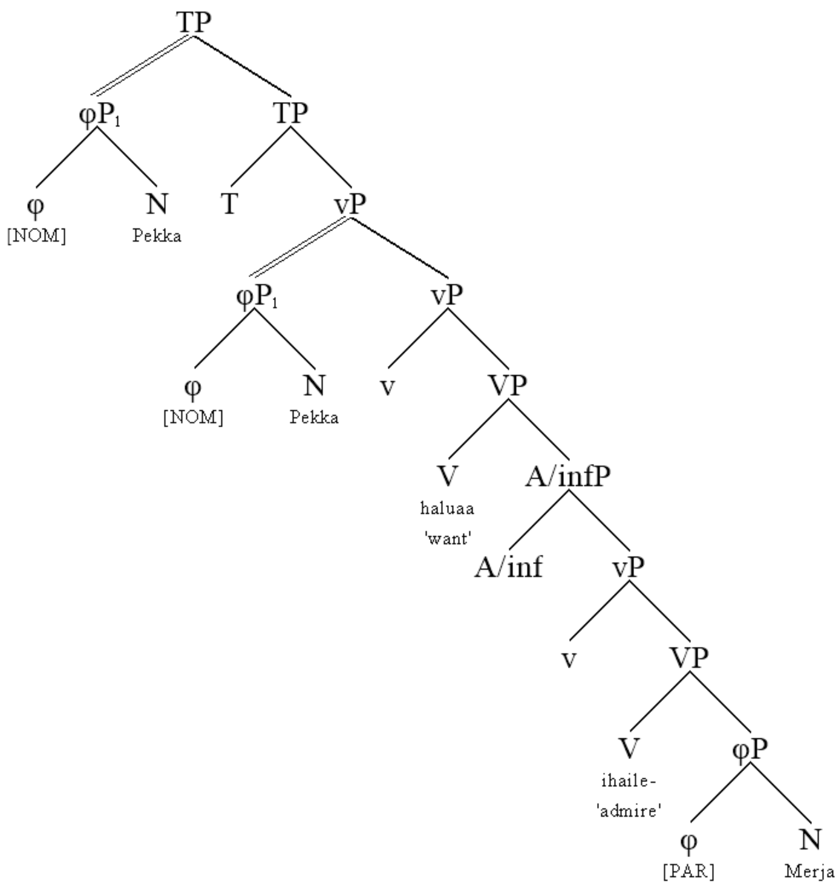
5.3 Infinitivals

Finnish infinitival clauses and their syntactic analysis constitute a controversial topic. The syntactic background theory analyses them as projecting structure overtly available in the input. Consider, therefore, what is available in the input (34):

(34)	Pekka	halu-si	ihail-la	Merja-a. (#19, 86)
	Pekka.nom	want-pst.3sg	admire-A/inf	Merja-par
	‘Pekka wanted to admire Merja.’			

The sentence contains two arguments plus the finite main verb and an infinitival verb that is composed of two elements, the verb stem *ihail(e)-* and the infinitival suffix *-lA* glossed as A/inf from “A-infinitival.” We can compare it to the English ‘to’: it is a morphological element that signals the presence of an infinitival form, but is a suffix in Finnish. The lexicon creates mapping /ihail-la/ ~ V#A/inf, which the syntactic component extracts into [A/inf…V] we see in the analysis:

(35)



Both the finite transitive verb and the infinitival verb have been decomposed, and by the same lexical decomposition + syntactic extraction mechanism. Case checking proceeds as in all previous examples: the grammatical subject checks the nominative case against T while the partitive direct object against v-V. The A-infinitival head A/inf^0 represents reduced or completely neutralized TAM (tense-aspect-mood) properties. All infinitivals are calculated in the same way: morphemes of the surface verb are decomposed and then project corresponding syntactic heads. We therefore generate VA-infinitivals (VA/inf substituting for A/inf, sentence #231), MA-infinitivals (#87), and so on. This analysis follows the style of analysis suggested in (Koskinen, 1998; Vainikka, 1989). The test corpus contains sentences where the case-marked arguments occur in noncanonical positions to verify that case checking and reconstruction still works correctly. The following symbols will be used for glossing word orders: S = main clause subject, O = main clause object, V = main clause verb, s = infinitival subject, o = infinitival object, v = infinitival verb.

(36)	a.	Merja-a	halu-si	ihail-la	Pekka. (#14)
		Merja-par	want-pst.3sg	admire-a/inf	Pekka.nom
		o	V	v	S
	b.	Merja-a	halu-si	Pekka	ihail-la. (#15)
		Merja-par	want-pst.3sg	Pekka.nom	admire-A/inf
		o	V	S	v
	c.	Merja-a	Pekka	halu-si	ihail-la. (#16)
		Merja-par	Pekka.nom	want-pst.3sg	admire-A/inf
		o	S	V	v
	d.	Pekka	Merja-a	halu-si	ihail-la. (#17)
		Pekka.nom	Merja-par	want-pst.3sg	admire-A/inf
		S	o	V	v
	e.	Pekka	halu-si	Merja-a	ihail-la. (#18)
		Pekka.nom	want-pst.3sg	Merja-par	admire-A/inf
		S	V	o	v

The corresponding syntactic analysis are collected into Figure 9.

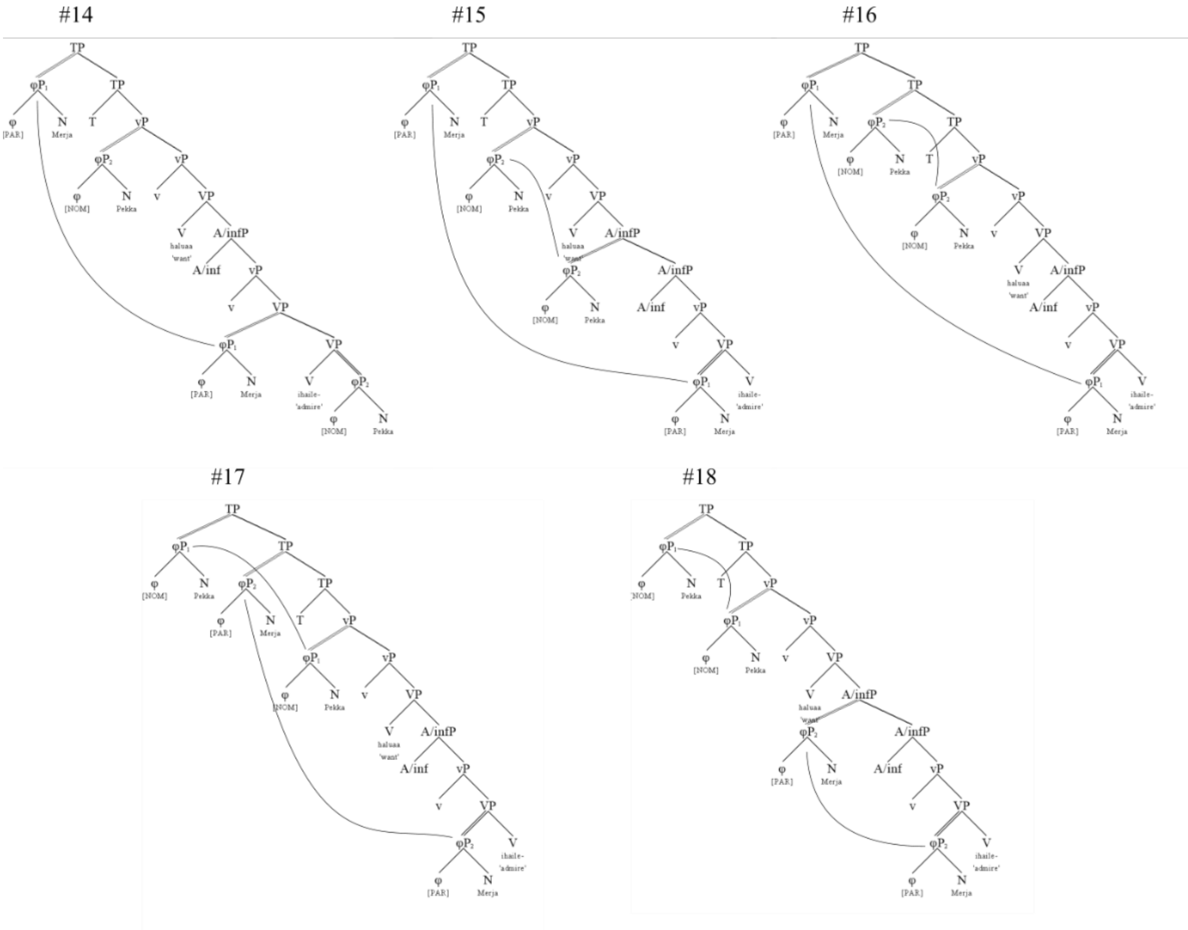


Figure 9. Analyses of infinitival words orders (sentences #14-18)

Preverbal arguments are interpreted as topics, arguments that appear too late are interpreted as representing new information focus.

Not all infinitival word order permutations are possible. As observed by (Manninen, 2003b), the ordering between verbal heads is much more limited than the ordering of referential arguments, and many such orders are correctly ruled out because head reconstruction is a limited operation. The relevant test sentences are #32-37. I am not completely sure if the model is making the correct judgments here; the sentences are hard to judge.[33] Example (e) is clearly ungrammatical, while (f) has a poetic quality to me.

[33] Crucially, these sentences must be interpreted so that nothing is moved into the CP-domain. For example, (a) cannot be interpreted so that the infinitival verb is focused or contrasted (it is grammatical under such interpretation); rather, we are looking at an analysis where it constitutes some kind of *topic* of the clause. The

(37)	a.	*?Ihail-la	halusi	Merja-a	Pekka. (#32)
		admire-A/inf	want-pst.3sg	Merja-par	Pekka.nom
		v	V	O	S
	b.	*?ihail-la	halusi	Pekka	Merja-a. (#33)
		admire-A/inf	want-pst.3sg	Pekka.nom	Merja-par
		v	V	S	O
	c.	*?ihail-la	Pekka	Merja-a	halu-si. (#34)
		admire-A/inf	Pekka.nom	Merja-par	want-pst.3sg
		v	S	O	V
	d.	*?ihail-la	Pekka	halusi	Merja-a. (#35)
		admire-A/inf	Pekka.nom	want-pst.3sg	Merja-par
		v	S	V	O
	e.	*Pekka	ihail-la	halu-si	Merja-a. (#36)
		Pekka.nom	admire-A/inf	want-pst.3sg	Merja-par
		S	v	V	o
	f.	*?Merja-a	Pekka	ihail-la	halu-si. (#37)
		Merja-par	Pekka.nom	admire-A/inf	want-pst.3sg
		o	S	v	V

These sentences are judged ungrammatical by the model because there are no parsing solutions + reconstruction that would yield well-formed representations. To being with, the model does not allow one to topicalize heads into specifier positions,[34] which means that sentence (a) for example will always have an interpretation $[_{\text{A/infP}}$ A/inf [V TP]] that cannot generate a well-formed output.

5.3.1 VP-fronting

Some infinitival word orders give rise to VP-fronting analyses, as already discussed in the main article. These sentences are #20-23, 30-31 in the test corpus. A VP-fronting (here, A/infP fronting) is possible when the infinitival predicate is adjacent to its argument and the two occur in a noncanonical position,

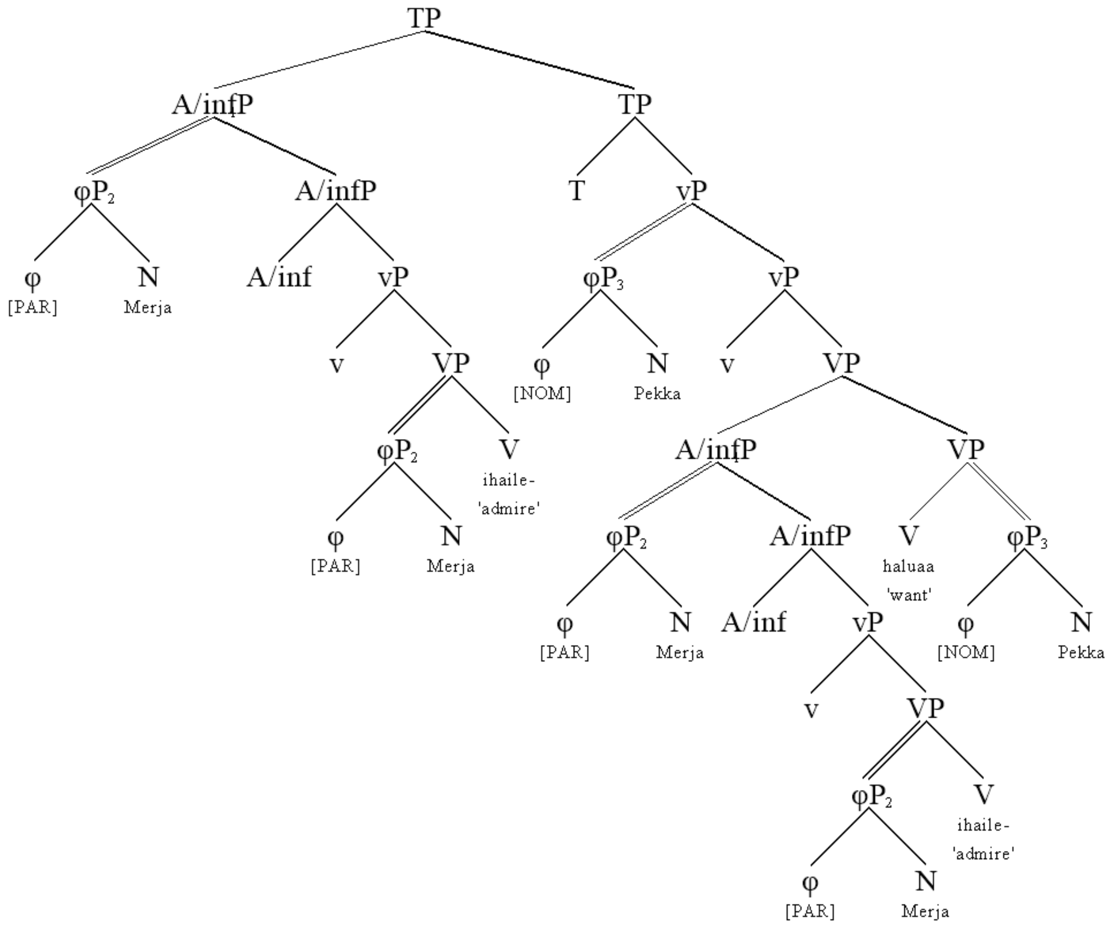
---

following could provide a possible context: *Tuollainen ihailu on lapsellista. Ihailla halusi Merjaa Pekka* 'Admiration like that is stupid. Admiring was what Pekka wanted to do to Merja'. The sentence feels very odd; ungrammatical to me.

[34] In other words, $[_{XP}$ H^0 $[_{XP}$ X YP]] is not a possible configuration. This is a fundamental property of the underlying syntactic theory.

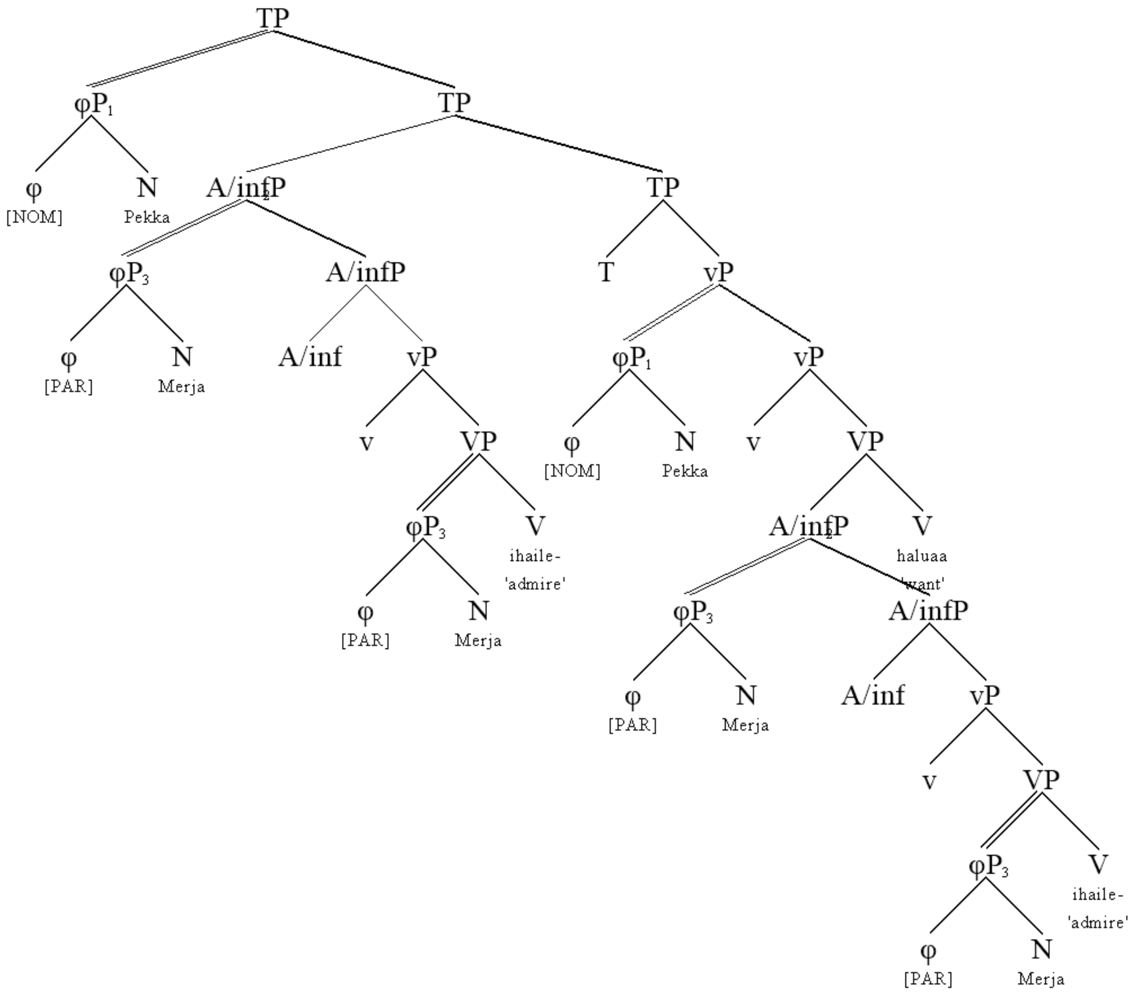
typically at the front of the clause. Then the parsing component can, typically after considerable amount of computations, access a VP-fronting construction. Whether Finnish exhibits VP-fronting is controversial. In my Finnish the construction exists, but it doesn't seem to have real use (i.e., contexts of use). Therefore, the model was programmed in such a way that a VP could fill in the SpecTP, satisfy EPP checking, but was not interpreted as the topic, making it an expletive-type of EPP filler. In addition, I wrote the code in such a way that this stipulating occurs as a special condition that can be removed to experiment with Finnish that does not have this option. A VP-fronting construction *Merja-a ihail-la halusi Pekka* 'Merja-par admire-A/inf want-pst.3sg Pekka.nom' (#20) was analysed as (38).

(38)



The analysis reconstructs A/infP from SpecTP into VP, which corresponds to VP-fronting. Notice that the object of the infinitival is reconstructed inside the A/infP, producing embedded chains. Sentence #21 is analysed in the same way. If both the grammatical subject and the A-infinitival phrase have been fronted (#22), they are both reconstructed:

(39)



The grammatical subject is reconstructed from Spec_2TP into SpecvP, and the infinitival from Spec_1TP into VP. They both occurs inside the TP projection, creating a double-spec construction but without multitopic interpretation. A/infP is not case marked, so in this case reconstruction cannot use case information. The algorithm creates an A-chain (inverse of A-movement), which does not employ any case information. Thus, the grammatical subject and the A-infinitival are reconstructed by different operations. This is visible also in the derivational log file: A/infP reconstruction occurs earlier, line

16239, A-reconstruction later, line 16267.[35] Ungrammatical case forms are sentences #93-102, where all combinations (nom/nom, nom/acc(0)…) were again tested and correctly ruled ungrammatical.

5.3.2 Adjunct infinitival (MA-infinitival)

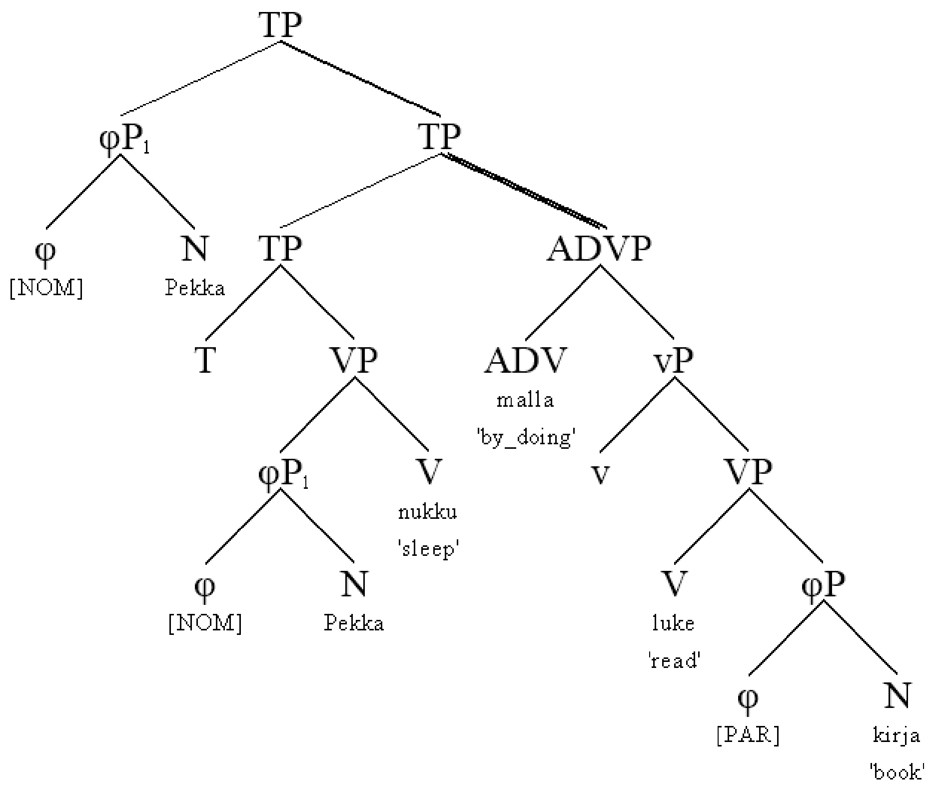
Sentence (40) illustrates the Finnish MA-infinitival that was part of the study. This form was included because, contrary to the A-infinitival, it behaves like an adjunct adverbial and therefore has a different syntactic analysis. Our main interest in this specimen is in verifying that case checking and reconstruction still works correctly. This MA-infinitival corresponds roughly to the English 'by doing something' adverbial, as shown by the translation.

(40)	Pekka	nukahti	[luke-malla	kirja-a.]
	Pekka.nom	fell.sleep	read-ma	book-par
	'Pekka fell asleep by reading a book.'			

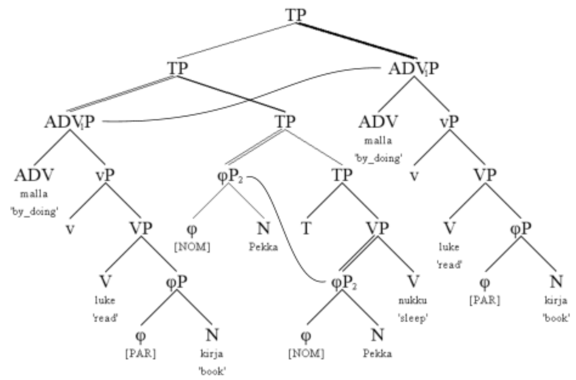
The analysis calculated by the model is (41).

[35] One problem is that the model interprets the A-infinitival as the 'agent' of T. This problem can be fixed by stipulation, but the problem is more general: phrasal nonreferential elements that appear on SpecTP, such as expletives, impact semantic interpretation (e.g., definiteness effects, generic interpretations). Similar questions arise in connection with the A/inf head, which represents reduced TAM properties and corresponds to T in some ways. A phrase at SpecA/inf will be interpreted as its argument. The issue was left for future.

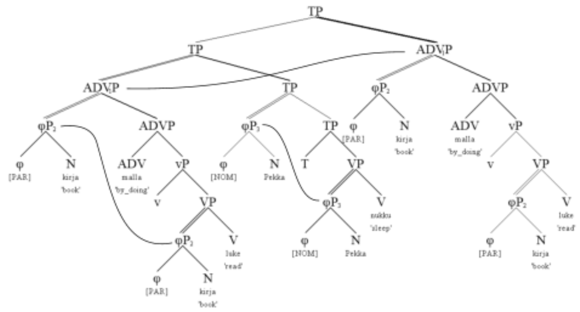
(41)



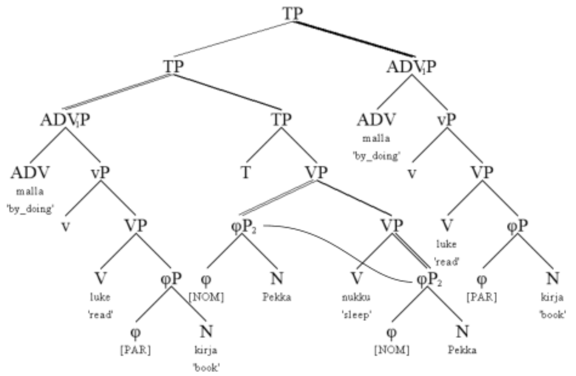
The MA-infinitival is merged as a right-adjoined to TP, where it is interpreted as modifying the event. Word order tests are sentences (#88-92), with the calculated analyses collected to Figure 10.



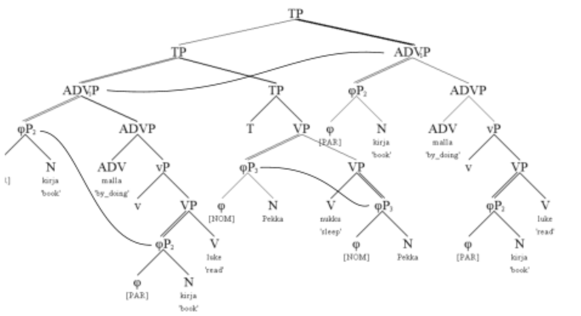
luke-malla kirja-a Pekka nukahti (#89)
Adv o S V
'read-ma book-acc Pekka.nom fell.sleep'



kirja-a luke-malla Pekka nukahti (#90)
o Adv S V
'read-ma book-acc Pekka.nom fell.sleep'



luke-malla kirja-a nukahti Pekka (#91)
Adv o V S
'read-ma book-acc Pekka.nom fell.sleep'



kirja-a luke-malla nukahti Pekka (#92)
o Adv V S
'read-ma book-acc Pekka.nom fell.sleep'

Figure 10. Analyses for word order variations in a sentence containing an adverbial and its direct object.

The fronted adverbial is reconstructed from SpecTP into [TP AdvP] position, thus in absolute terms it moves higher in the phrase structure. This is possible due to the nature of case-based reconstruction, which must likewise reconstruct postverbal arguments into higher SpecvP positions (42). Although not standard, this is required by empirical evidence and should not be rejected on theoretical grounds alone.

(42) Merja-a T __$_1$ halusi ihail-la Pekka$_1$.
Merja-par want-prs.3sg admire-A/inf Pekka.nom
'Pekka (information focus) wanted to admire Merja (topic).'

Ungrammatical case forms were tested by sentences #110-124, with all combinations correctly ruled out.

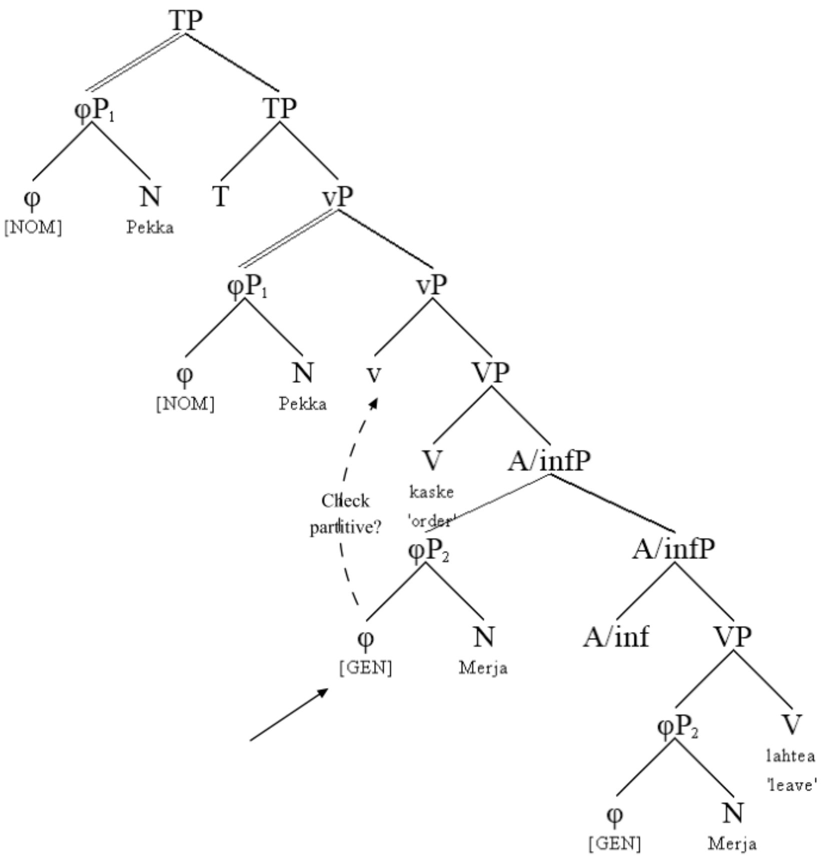
5.3.3 *Genitive subjects of infinitivals*

Most Finnish infinitival clauses have genitive subjects that represent the thematic agent of the underlying infinitival predicate (43).

(43)	Pekka	käsk-i	Merja-n/	hän-en	lähte-ä.
	Pekka.nom	order-pst.3sg	Merja-gen/	she-gen	leave-A/inf
	‘Pekka ordered Merja/her to leave.’				

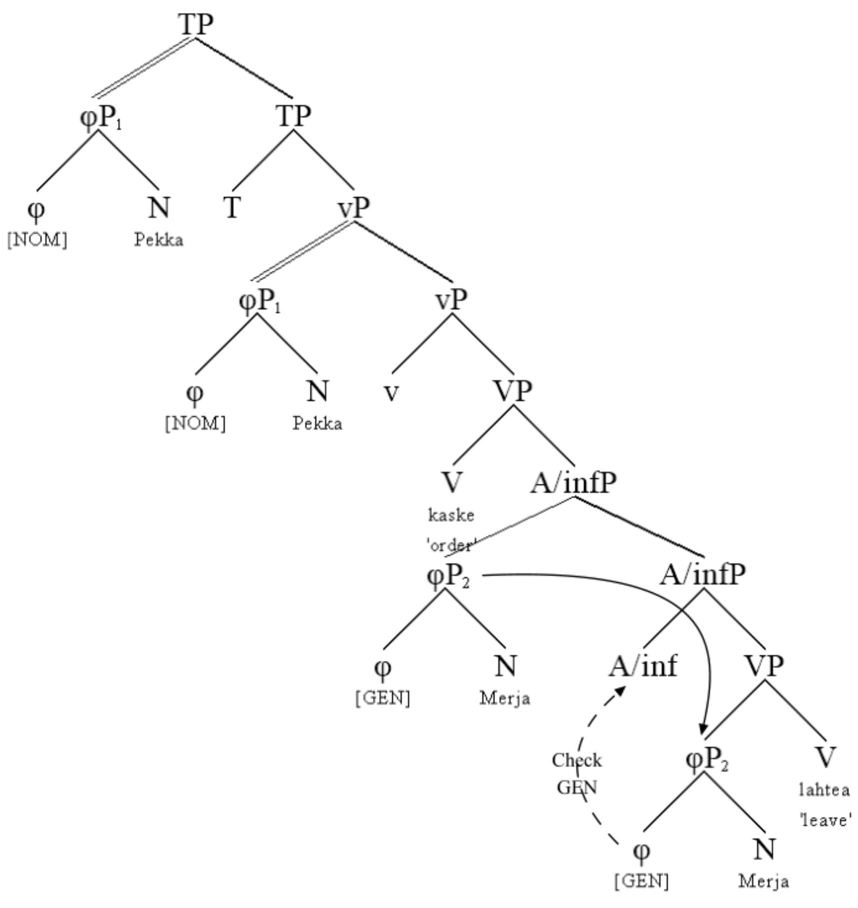
Here the structure differs from English, where the embedded subject has accusative case, as shown by the translation, and it behaves in several ways as if it were part of the main clause. Notice, in particular, that the embedded subject pronoun takes the genitive case, not the accusative that would be *hän-et* ‘he-acc(t)’. The parser puts the genitive subject at SpecA/infP right before the predicate, following the surface order *Merja-n* + *lähte-ä* ‘Merja-gen + leave-A/inf’, but the problem is that *at that position* it would check the accusative against lexical items in the main clause. We can see this if we consider the analysis calculated by the model and consider only the first pass parse position of the genitive, marked by the arrow in the example below:

(44)



This is the surface position, where the argument checks the partitive (or accusative) from the small verb v. To solve this issue, (Brattico, 2020) assumed that the genitive argument is reconstructed from SpecA/infP into SpecVP so that it checks the genitive case against A/inf. This assumption was carried over to the present model. The genitive argument was reconstructed to SpecVP in the analysis above, and it checks the genitive case against {+ARG, –FIN} at A/inf. All genitive cases are checked in the same way. Example (45) shows how the genitive is checked.

(45)



The same analysis applies to the VA-infinitival (#231). Several tests were performed where the genitive subject was replaced with nominative, partitive, accusative and zero-accusative subjects, all correctly judged ungrammatical (#235-238). The same tests were performed also in connection with noncanonical word orders (#239-242, #239-242).

5.3.4 Problem: the postverbal genitive

Finnish genitive involves a problem I have not been able to solve satisfactorily.[36] (Brattico, 2016, 2018) noted that the distribution of genitive arguments is more limited than those of nominative, partitive and accusative arguments. The basic generalization is that while nominative, partitive and accusative arguments can move rightwards, genitive arguments can't.

(46) a.	*Pekka	sanoi	__$_1$	lähte-vän	Merja-n$_1$. (#234)
	Pekka.nom	said		leave-VA/inf	Merja-gen
	Intended: 'Pekka said that Merja will leave.'				
v.	*Pekka	käski	__$_1$	lähte-ä	Merja-n$_1$.
	Pekka.nom	ordered		leave-A/inf	Merja-gen
	'Pekka ordered Merja to leave.'				

(Brattico, 2020) captured these limitations by assuming that Finnish genitive arguments cannot undergo case-based reconstruction; instead, they behave like English case-marked pronouns that are limited to the standard cases of A/Ā-reconstruction. This captures (46), because standard A/Ā-reconstruction cannot reconstruct rightward elements. Working with a comprehensive case corpus containing noncanonical genitive arguments revealed, however, that this assumption is problematic. There are situations where noncanonical leftward genitive movement cannot be reconstructed via standard A/Ā chains. While it was possible to make some cases to work, all my solutions generated problems elsewhere. In addition, the assumption that the genitive is excluded from case-based reconstruction is a stipulation that has no motivation. I therefore rejected this assumption in this study and allowed the genitive argument to reconstruct like arguments marked by other structural cases, which removed the stipulation from the system but allows (46)(#234 in the test corpus) and therefore judges it wrongly as grammatical. This means that part of the problem with the noncanonically positioned genitive subjects remains unsolved.[37]

---

[36] As this article goes into the print, one new solution has been considered and is presented in a manuscript "Agreement and chains" written together with Cristiano Chesi and Balász Surányi, submitted for publication.

[37] (Brattico, 2016, 2018) considers another solution, but without rigorous testing. The essence of that proposal is that due to the lack of verbal agreement at the infinitival, the genitive argument, carrying the phi-features, must occupy the SpecA/infP position. The analysis claims that verbal phi-agreement and the presence of the argument itself can both satisfy the EPP.

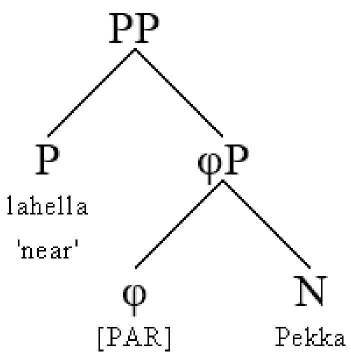
5.4 Adpositions

Finnish adpositions can be grouped into three classes, with a residuum of exceptional cases. The first (call it Class I) contains prepositions checking the partitive and never showing agreement under any argument, independent of its case (47)(c-d). The partitive can appear both before (a) and after (b) the adposition.

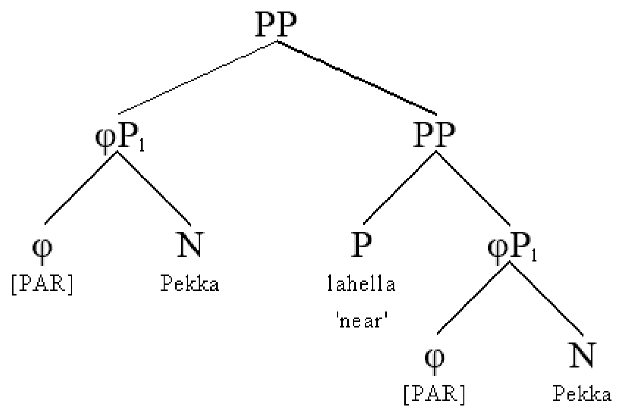
(47) a.	minu-a	kohti	b.	kohti	minu-a	c.*minu-a/minu-n	kohti-ni
	I-par	towards		towards	I-par	I-par /I-gen	towards-px/1sg
	’towards me’			’towards me’			

As far as I can see, both word orders (a-b) are always possible, and there seems to be no difference in semantic interpretation (i.e., no natural contexts of use, no expressions that could not occur at both positions). As explained in the main article, prepositions in this class do not exhibit agreement and have –VAL (c) which checks the partitive together with +ARG; if the partitive is in the prehead position (a), it is reconstructed before case checking. These analyses are shown in (48) for *lähellä Pekka-a* ‘near Pekka-par’ (#125) and (49) for *Pekka-a lähellä* ‘Pekka-par near’ (#126), the latter which involves reconstruction.

(48)



(49)



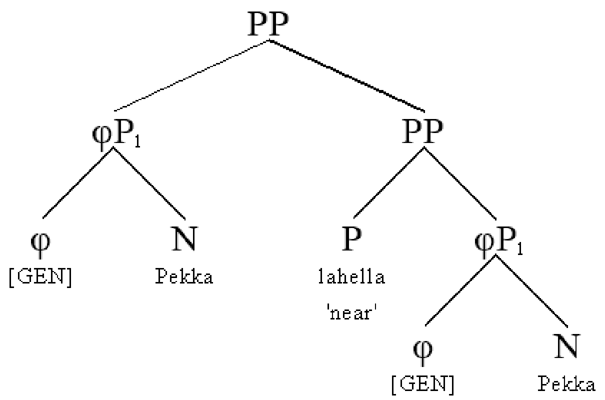
The partitive argument is reconstructed from SpecPP into CompPP, by A-reconstruction. Ungrammatical case configurations are sentences #128-133. This type of analysis was first proposed by (Manninen, 2003a).

Class II contains adpositions which accept only genitive arguments, such as *kanssa* ‘with’ that show optional phi-agreement with the genitive argument (*minun kanssa(-ni)* ‘I.gen with(-px/1sg)’). The genitive argument must occur at the specifier position (**kanssa-ni minun* ‘with-px/1sg I.gen’). I return to this class below. Class III has adpositions exhibit both behaviours. The preposition *lähellä* ‘near’ belongs to this class, thus it exhibit the partitive behavior shown above but also the genitive behavior (50).

(50) a.	Minun	lähellä(-ni)	b.	*lähellä(-ni)	minun
	I.gen	near(-px/1sg)		near(-px/1sg)	I.gen
	’near me’				

I have not been able to differentiate Class II from the genitive variants in Class III, so it is possible to assume Class III = I + II, thus adpositions in Class III are ambiguous, and only profiles I and II exist as primitive classes. Class II is derived by reconstructing the genitive to the complement position and checking the case against the preposition. It must have +VAL due to the agreement, therefore this form cannot check the partitive. The analysis for *Peka-n lähellä* ‘Pekka-gen near’ is (51)(#127).

(51)



The ungrammaticality of **lähellä Peka-n* ‘near Pekka-gen’ (#138) is explained by assuming that Class II adpositions have the EPP feature in addition of triggering agreement, requiring that a suitable phrase appears at its specifier position (Brattico, 2011b; Manninen, 2003a). The two classes plus one combined class leaves few exceptional cases which license the genitive argument at post-head positions (52).

(52) a.	silla-n	yli	b.	yli	silla-n
	birdge-gen	over		over	bridge-gen
	’over the bridge’				

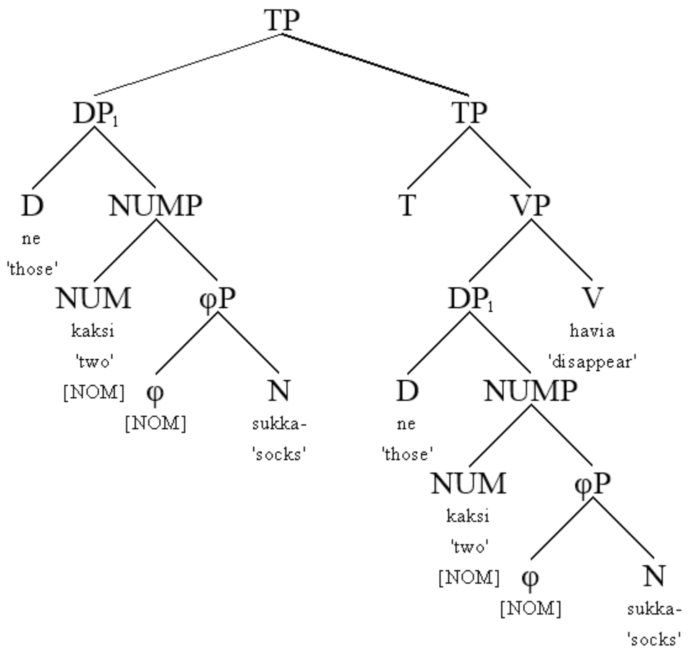
These were not included into the present study.

5.5 Numerals and case concord

5.5.1 Type I adjectival numeral

Case concord is handled in model by linking each phonological word independently with a case checker element, following (Brattico, 2010, 2011a). This can be seen from the analysis of sentence #145 = *ne kahdet sukat hävisivät* ‘those two-pl.nom sock-pl.nom disappeared’ which is shown in (53).

(53)

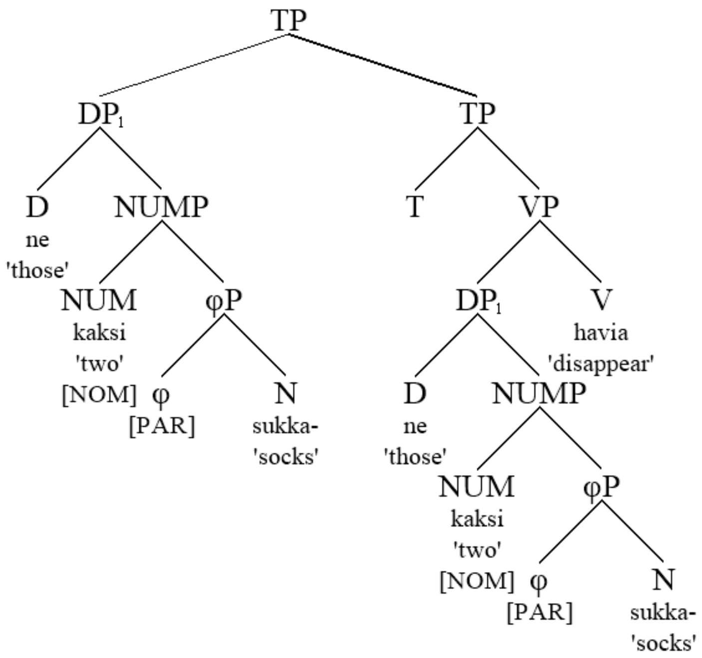


Here, both the numeral *kahde-t* 'two-pl.nom' and the φ (or the noun head, in other theories) check the nominative case independently against finite T at the reconstructed canonical position inside the VP. The model does not attempt to project abstract case features over and above what it sees at the phonological words (and thus at the corresponding lexical items). There is no separate case concord. Notice in particular that since the plural numeral *kahde-t* 'two-pl.nom' does not contain +ARG or –VAL, it allows the lower φ (or noun head) to check its features against T.

5.5.2 Type II QCC numeral

Finnish lexicon has another form of the numeral, however, which appear in a bare uninflected form and assigns the partitive case to the elements below. I call it quantificational case construction (QCC). It is illustrated by sentence #144 = *ne kaksi sukka-a hävisi* 'those.0 two.0 sock-sg.par disappeared', where I have glossed the two bare forms by 0.

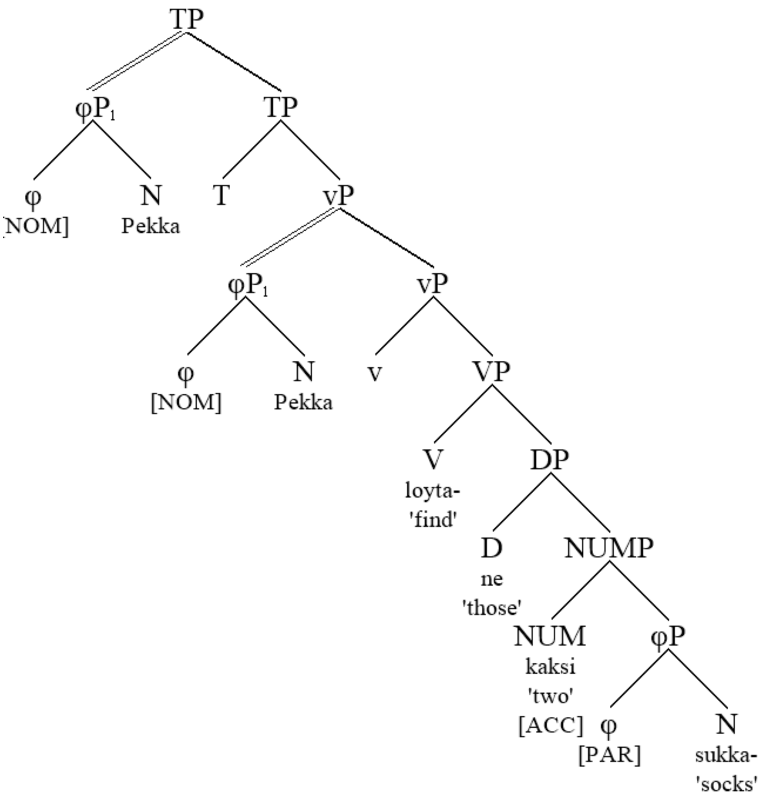
(54)



The partitive at φ is checked by the numeral (NUM) which has +AGR, –VAL, while the nominative case at Num is checked by finite T. The noun phrase comes to exhibit discontinuous case concord. Notice that the bare numeral takes the singular form and does not agree with the demonstrative that is in plural, justifying –VAL, which prevents agreement. The system creates layered pluralities where the singular form of *kaksi* ‘two.sg’ denotes two, while the plural form *suka-t* ‘sock-pl’ denotes two pairs of socks, i.e. four socks.

One problem with the analysis above is that the same *ne kaksi sukka-a* ‘those two sock-sg.par’ can occur also in the direct object position. If *kaksi* has nominative case, as shown above, then this phrase cannot occur in a direct object position. This issue was solved by assuming that *kaksi* is ambiguous between nominative and accusative forms, or that this form expresses both. Thus the analysis calculated by the model is (55)(for #147).

(55)

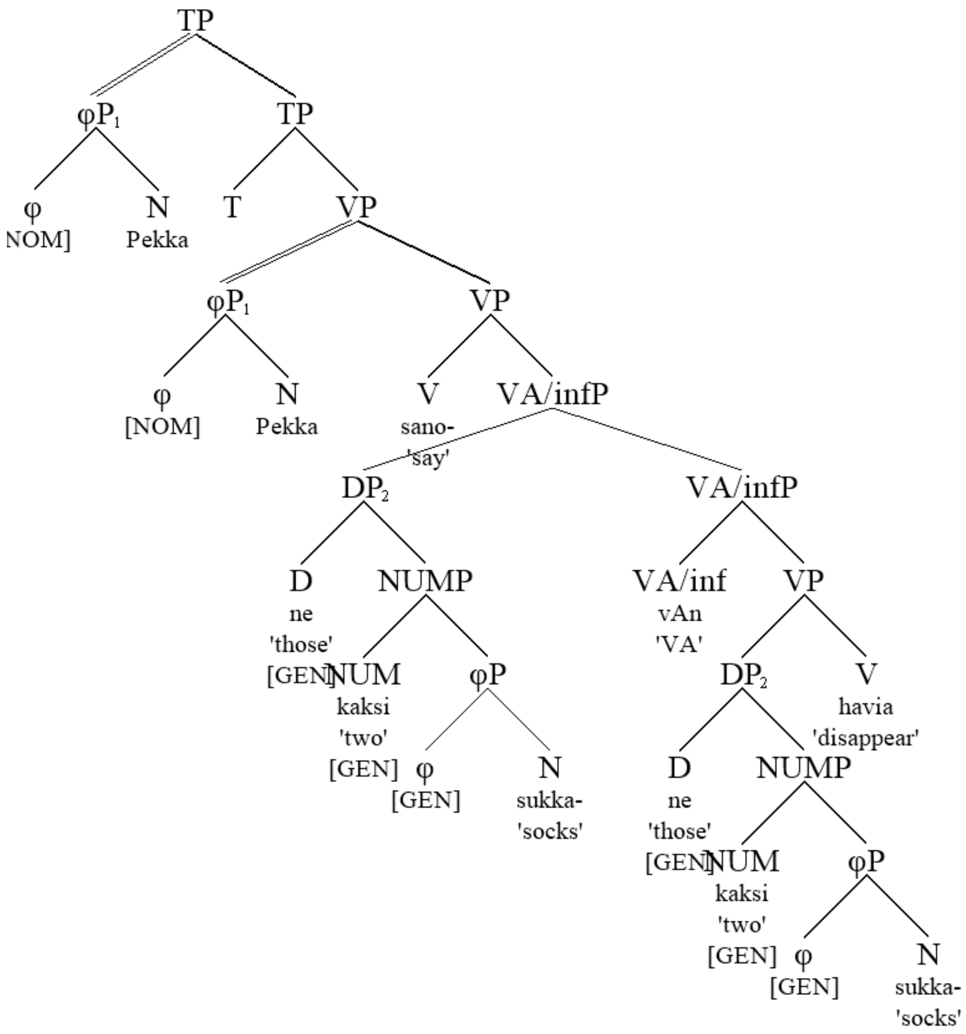


Notice the change in case marking. This assumption, although at first stipulative, explains why QCC occurs only in nominative and accusative contexts. When the noun phrase is assigned any other case, such as the genitive (#148), the quantificational case construction disappears (56).

(56) a.	Pekka	sanoi	niiden	kahden	sukan	häviävän. (#148)
	Pekka.nom	said	those-gen	two-gen	sock-gen	disappear
	‘Pekka said that those two socks will disappear.’					
b.	*Pekka	sanoi	niiden	kaksi	sukka-a	häviävän. (#158)
	Pekka.nom	said	those-gen	two.0	sock-par.sg	disappear

Phrase structure (57) illustrates the analysis for (56)a.

(57)

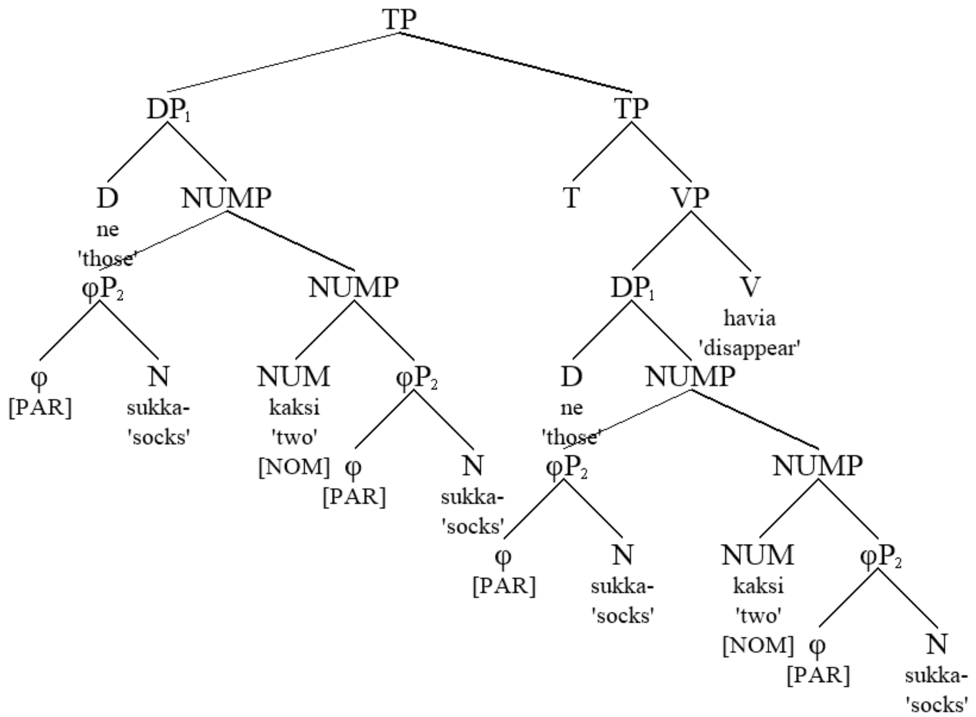


Sentence (b) is ungrammatical because the bare numeral cannot check nominative or accusative in this context. Various ungrammatical case forms were tested in this context (#150-157), which were correctly rejected, together with four noncanonical word orders (#159-162).

5.5.3 Problem: spurious reconstruction

I encountered one problem that was left unsolved. The model analysed the ungrammatical word order **ne sukkaa kaksi hävisi* ‘those sock-sg.par two disappeared’ (D + N + Num + V, #162) by reconstructing φN:

(58)

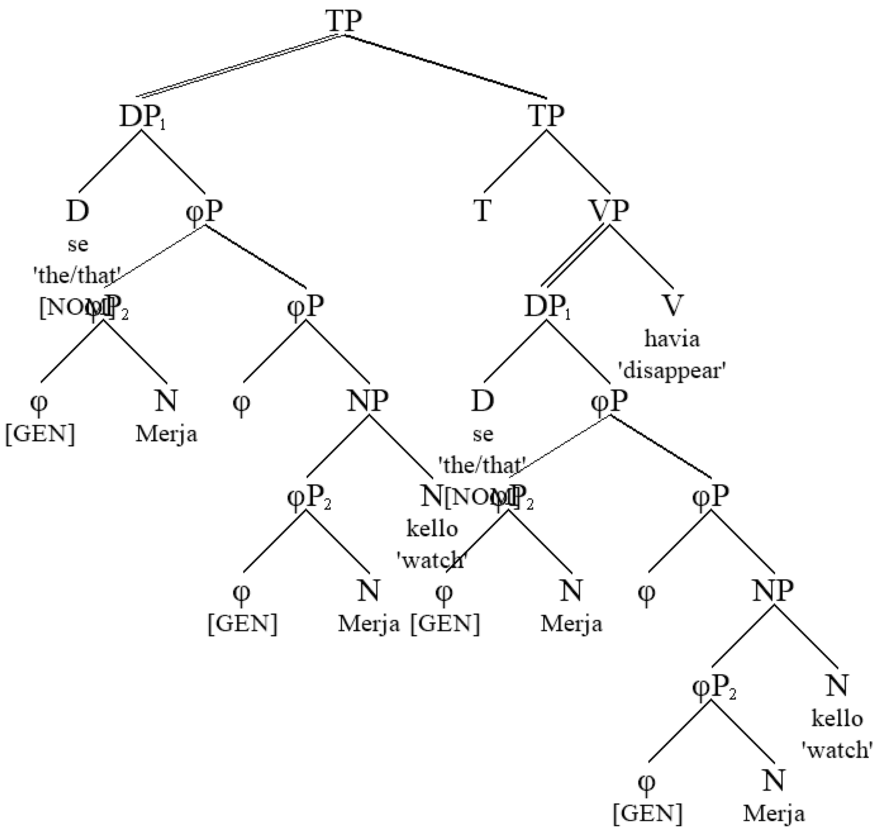


Since the result is well-formed, the model judges this sentence grammatical (#162). It is, however, ungrammatical. This type of noun-internal reconstruction operation is not possible in Finnish. One could rule it out trivially by stipulation, but some general principle must be at stake. It implements a local Spec-to-Comp reconstruction that we could rule out as “too local,” but the problem is that exactly this operation is required to calculate the properties of adpositions (Section 5.4). One problem in this arena is that there are very few rigorous studies of Finnish noun phrase, and certainly there is no consensus on how to best approach the issue.

5.6 Noun phrase and possessor genitives

The internal syntactic structure of Finnish noun phrase has received relatively little attention in previous literature. Only few examples were added to the present study, and their purpose was restricted to verifying that the genitive case checking works correctly inside this context. The basic approach is to assume that the genitive possessor is checked by an extra functional head that appears inside the noun phrase, in our case φ which contains phi-features and creates a minimal domain that can functions as a fully referential argument in syntax and semantics, taking into account the fact that D-elements are not obligatory in Finnish. The analysis for *se Merja-n kello hävisi* ‘that Merja-gen watch disappeared’ (#243) is calculated as (59).

(59)



The first parse structure is $[_{DP}$ D $[[_{\varphi P}$ φ Merja] $[_{\varphi P}$ φ N]], based on the input string *that + Merja's + watch*, which leads into reconstruction $[_{DP}$ D $[[_{\varphi P}$ φ Merja$]_1$ $[_{\varphi P}$ φ $[\_\_{}_1$ N]] above. The genitive is checked by φ. An alternative, discussed in the main article, is to project *n* instead of φ. Several ungrammatical variants were correctly ruled out, listed below.

(60) *Word order violations*

a.	*se kello	Merja-n	hävisi. (#244)
	that watch	Merja-gen	disappeared
b.	*se kello	hävisi	Merja-n. (#245)
	that watch	disappeared	Merja-gen
c.	Merja-n	se kello	hävisi. (#246)
	Merja-gen	that watch	disappeared

(61) *Case violations*

a.	*se Merja	kello	hävisi. (#247)

	that	Merja.nom	watch	disappeared.
b.	*se	Merja-a	kello	hävisi. (#248)
	that	Merja-par	watch	disappeared.
c.	*se	Merja-n	kello	hävisi. (#249)
	that	Merja-acc(n)	watch	disappeared.
d.	*se	Merja	kello	hävisi. (#250)
	that	Merja.acc(0)	watch	disappeared.

5.7 Impersonal passive

Finnish impersonal passive construction has generated substantial literature. For case checking purposes its most salient property is the occurrence of two nominative-looking arguments, the second which turns out to be the 0-accusative (62).

(62)	Me	löydet-tiin	avain/	hän-et.
	we.nom	find-pst.impass	key.0	he-acc(t)
	‘We found the key/him.’			

I will assume from now own that the singular DP direct object has the zero-accusative case, glossed as ACC(0). If the sentence lacks bounded aspect, partitive direct object surfaces. The first interesting thing about (62) is the fact that, as shown by the translation, this is a regular transitive clause in active voice. It is interpreted as ‘we found the key’. The reason the impersonal passive verb form appears in this construction is because it has replaced the correct third person plural active verb form *me löys-i-mme avaimen* -‘we.nom found-pst-1pl key-acc(n)’ virtually everywhere in spoken language, the later which now feels hypercorrect and is only used in written Finnish. Only the first person plural subject can be used in connection with the impersonal passive form, however (e.g., **te löydet-tiin avain* ‘you.pl found-pst.impass key-acc(n)’). Still, notice that once the verb form changes, also the case pattern changes: the n-accusative form of the regular active form (*avaim-en* ‘key-acc(n)’) is replaced by the zero-accusative ant t-accusative forms (*avain* ‘key.acc(0)’)(63)a-b.

(63) a.	*Me	löydet-tiin	avaimen. (#264)
	we.nom	find-pst.impass	key-acc(n)/gen
b.	Me	löys-i-mme	avain. (#265)
	we.nom	find-pst-1pl	key.acc(0)/nom

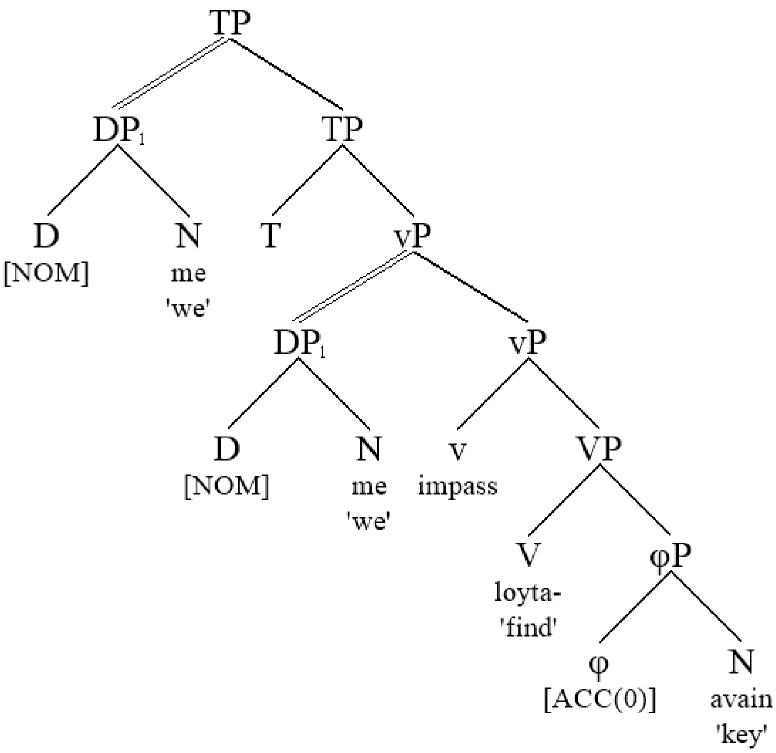
The verb form in (62) is called impersonal because once we remove the subject, an interpretation emerges where an unnamed plural sentient agency, representing the agent, caused the action on the patient (64). Both word orders VO (a) and OV (b) are possible.

(64) a.	Löydet-tiin	avain.	b.	Avain	löydet-tiin.
	find-pst.impass	key.acc(0)		key.acc(0)	find-pst.impass
	‘The key was found, by a collective agent.’				

The fact that the agent must by a (typically human) collective is a strong tendency, not strictly necessary. Construction (64) is regarded as genuine syntactic and semantic impersonal construction, since the agent remain unmentioned. The object is still accusative, as can be verified by using the pronoun in this context (*hän-et löydet-tiin* ‘he-acc(t) find-pst.impass’). The construction does not resemble English personal passive, which involves syntactic promotion of the object to the subject.

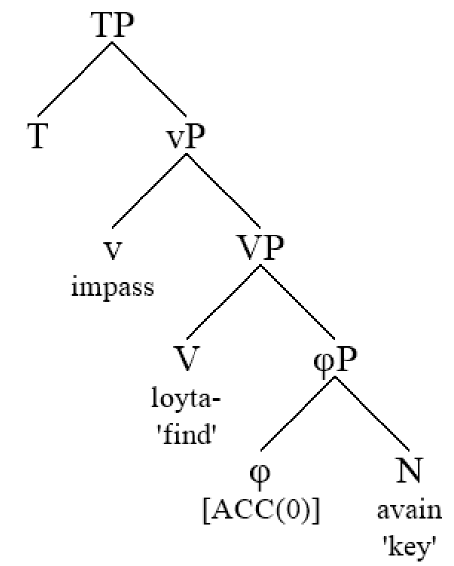
The analysis for (62) calculated by the model is (65), which build on (Manninen & Nelson, 2004).

(65)



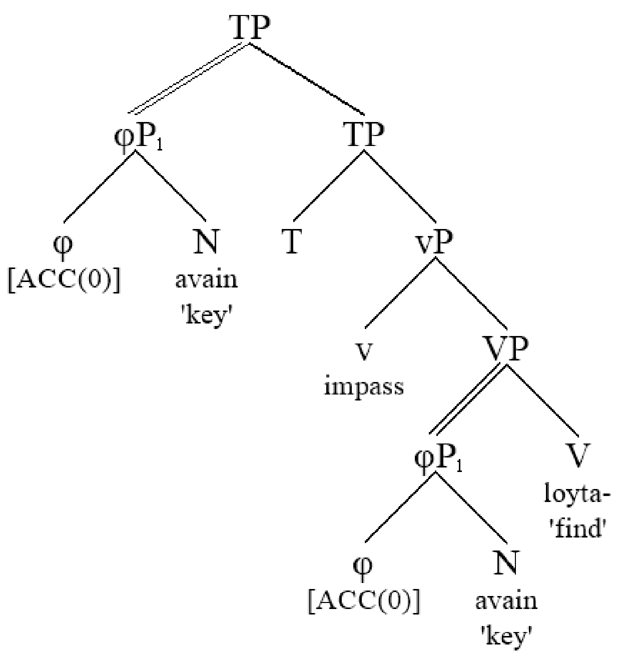
M&N proposed in essence that the impersonal passive verb form is generated by a special impersonal transitivizer v glossed in the above analysis as v/impass which, in their view, “supresses” the agent in the syntactic impersonal passive construction (64). Since this element does not exhibit productive phi-agreement with any element in the sentence, it checks the zero-accusative (–PHI) and calculates the data correctly. If the grammatical subject is missing (64)a-b, the model calculates an subjectless analysis (66) that again follows closely M&N’s analysis.

(66)



If the direct object is fronted, as in (a), it will be reconstructed, with its case checked at the canonical thematic position against the agreementless v/impass (67).

(67)



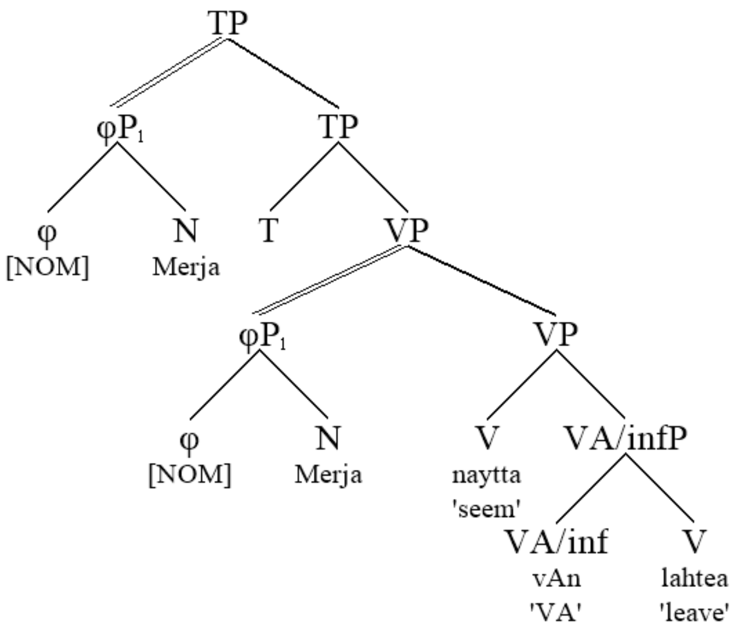
5.8 Raising construction and A-chains

Finnish has at least one potential personal raising construction that involves movement of an embedded subject into the position of a grammatical subject. The construction is illustrated in (68)a. Example (b) shows a similar construction where the embedded subject occurs in the canonical thematic position inside the VA-infinitival and the main verb is a regular active verb.

(68)	a.	Merja$_1$	näyttä-ä	[__$_1$	lähte-vän.]
		Merja.nom	seem-prs.3sg		leave-VA/inf
		‘Merja seems to be leaving.’			
	b.	Pekka	näki	[Merja-n	lähte-vän.]
		Pekka.nom	saw	Merja-gen	leave-VA/inf
		‘Pekka saw Merja leaving.’			

One way to analyse (68)a is to assume that the grammatical subject has been moved from the infinitival into the grammatical subject position, and is assigned the nominative case and agrees with the finite main verb at the surface position. It receives its thematic role from the embedded infinitival. An analysis along these lines was presented in (Brattico, 2018). Reversing this process for present purposes suggests that the nominative grammatical subject is reconstructed into the canonical position. This assumption turned out to be very problematic. The reason is that the canonical thematic position of the reconstructed subject, namely VP, can only check the genitive case (thus, see (b)), while the surface subject is case-marked by the nominative. Thus, while the model is able to reconstruct the subject into VP, as verified by simulation experiments, it cannot check the nominative case at that position. If reconstruction into the embedded VP is blocked, on the other hand, then the sentence will be analysed as an obligatory control (OC) construction. Since that assumption provides correct grammaticality judgments and semantic interpretations, the sentence was handled in this way, although the solution is certainly not compelling.

(69)



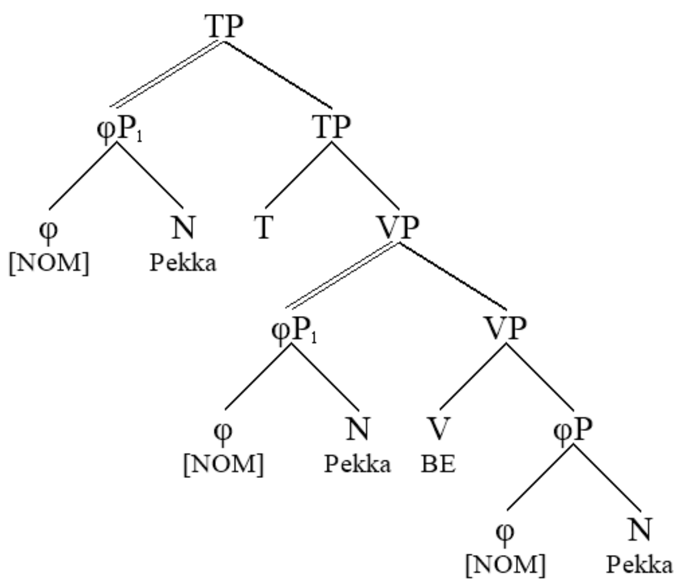
A control dependency is established between the verb 'leave' and the subject. (Brattico, 2018) considers a control analysis of this type by rejects it on the grounds that the raising verb does not seem to assign any thematic role on the subject; (69) implies that a thematic role is assigned, and indeed the current model provides one. The problem with case shifting was also noted in this study, but ignored on the basis that case shifting occurs also in English (p. 79). This does not solve the issue, however: an enumerative grammar requires that the case is checked at the surface configuration, not at the canonical base construction, which would make the construction still exceptional and a challenge to work with.

Raising constructions are usually explained by relying on local A-movement, which would here translate into A-reconstruction. Since Finnish case forms trigger case-based reconstruction, this cannot be the case here: nominative arguments are reconstructed into SpecVP where they check their case against finite T. How, then, could the nominative case "guide" reconstruction into the embedded VP if the case form is wrong and requires changing? On the other hand, if the sentence is linked with a syntactic analysis that cannot check the nominative, say a "raising" T that has slightly different feature composition (e.g., lack of ARG), then the failure of case-based reconstruction will still allow the model to try A-reconstruction. This is because A-reconstruction is modelled as a last resort option when both case-based reconstruction and Ā-reconstruction fail. There are at least two problems with this approach. First, I have never been able to formulate a convincing algorithm for A-reconstruction that handles a large nontrivial dataset of the relevant constructions, and second, no detailed studies of Finnish A-movement/A-reconstruction exists. The current model, which handles several basic cases in Finnish, English and Italian, attempts local A-reconstruction if and only if both case-based reconstruction and Ā-reconstruction fail. The algorithm finds the first empty specifier position for the reconstructed element and merges it there. The result is successive-cyclic A-reconstruction when the projectional spine of the finite clause contains several verbal elements such as negation, auxiliary, modals or other verbal components.

5.9 Copular construction

Copular constructions were analysed by assuming that the copular verb BE is invisible for the case mechanism, lacking –ARG and allowing both arguments to check nominative case against T. Sentence *Pekka on Pekka* 'Pekka.nom is Pekka.nom' (#272) was analysed as (70).

(70)



Impossible case configurations are sentences #273-280.

5.10 Adverbials and case-marking

Finnish DP-adverbials are assigned direct object cases much like DPs in direct object positions. This dataset, together with long-distance case assignment, provides a compelling argument for the conclusion that case assignment or case checking must rely on some notion of syntactic scope, namely because adverbial case forms depend on the presence of governing negation and overt phi-agreement. In fact, the present analysis is specifically tailored to account for these cases since an upward path can be established between a DP-adverbial and negation as easily as it can be established between a direct object and negation; the two cases are very similar when looked through the lens of syntactic scope and feature checking. Despite its prima facie plausibility, the data presented significant empirical and technical difficulties and, as pointed out in the main article, the model cannot be regarded as complete or even successful. Consider (71).

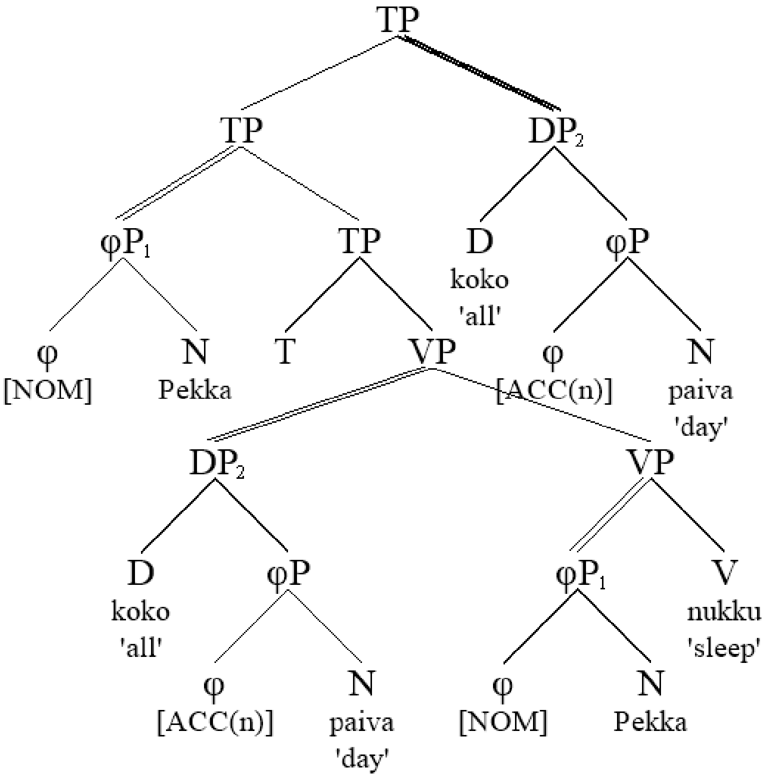
(71)	Pekka	nukku-i	[$_{DP}$	koko	päivä-n.] (#281)
	Pekka.nom	sleep-pst.3sg		all	day-acc(n)
	'Pekka slept all day.'				

The first problem is to separate direct object DPs from adverbial DPs, so that *koko päivän* 'all day' is not analysed as an illegitimate direct object of the verb 'sleep'. One solution is to allow any DP to function as an argument and adverbial, but this cannot be correct since the model will then create sentences like **Pekka nukkui Merjan* 'Pekka slept Merja-acc', *Merja* being interpreted as an unattested "proper name adverbial." Not all DPs can function as adverbs. A separate lexical feature was created for this purpose to license adverbial uses shown in (71). DPs that were licensed in this way, such as *koko päivä* 'all day', can still be used as arguments, the feature only licenses adverbial uses.

The next issue is that the accusative case occurs also in connection with intransitive verbs. We cannot therefore say that the accusative is literally a direct object case. This implies that the telic interpretation cannot be restricted to transitive verbs and, indeed, in this study the aspectual feature was part of both v and V. Furthermore, the accusative adverbial creates, all by itself, a bounded reading for the event. The sentence means that sleeping terminated after one day elapsed.

The third problem is the syntactic analysis of the DP adverbial. After several rounds of unsuccessful simulations and hypotheses, I found a solution by assuming that (i) the DP quantifier *koko* 'all' must be checked against feature [T], so that the adverbial comes to modify the tensed event, and that (ii) T must have the aspectual features as well. The link between aspect at T and the case form of the accusative then capture the way adverbial case marking affects the aspectual properties of the whole sentence. I return to these assumptions and the data that motivated them below; (i-ii) generates (72) for (71).

(72)



The DP adverbial is reconstructed from [TP DP] into a left adjoined position inside VP where both the quantifier D and φ can check their features against T, per (i-ii). This analysis reveals how the model connects the accusative form of the DP adverbial with polarity (73)a,b and agreement (c, d): both features are present in upward paths from the reconstructed adverbial DP.

(73) a.	Pekka	e-i	nukku-nut	koko	päivä-ä. (#283)
	Pekka	not-3sg	sleep-pst.prtcpl	all	day-par
b.	*Pekka	e-i	nukku-nut	koko	päivä-n. (#287)
	Pekka	not-3sg	sleep-pst.prtcpl	all	day-acc(n)
c.	Me	nuku-ttiin	koko	päivä. (#284)	
	we	sleep-pst.impass	all	day.acc(0)	
d.	??Me	nuku-ttiin	koko	päivä-n. (#288)	
	we	sleep-pst.impass	all	day-acc(n)	

Why not to assume that the DP adverbial checks V, so that V could contain the aspectual feature as well? If both DPs are inside the VP, then they should always be marked by the same case, which is not the case (74).

(74) Pekka	auttoi	Merja-a	koko	päivä-n. (#285)
Pekka.nom	help-pst.3sg	Merja-par	all	day-acc(n)

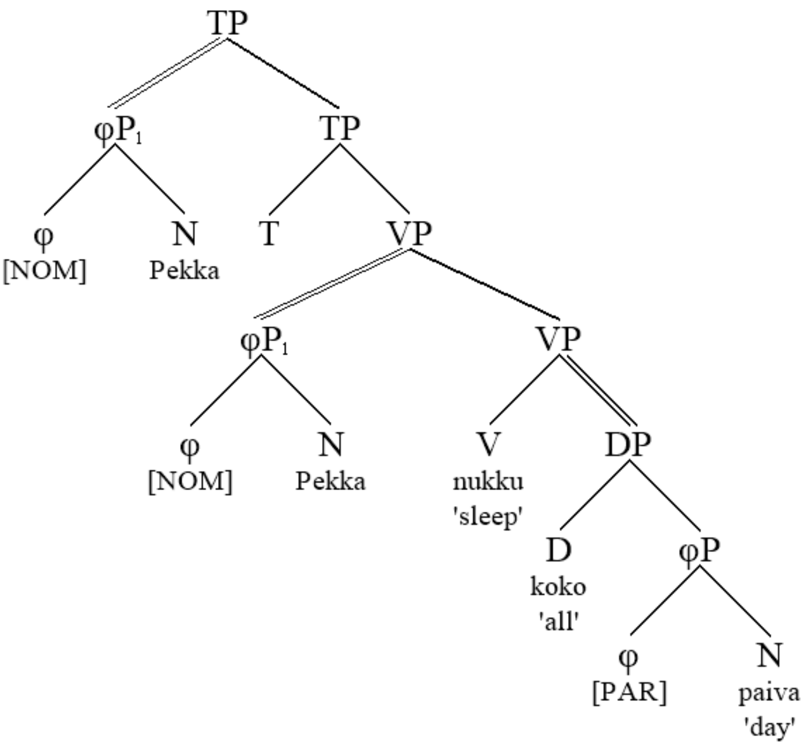
‘Pekka helped Merja all day.’

This data consisted one of the most puzzling aspects of the whole phenomenon. The analysis calculated by the model solves this problem by assuming, as shown by the analysis in (72), that the adverbial DP is reconstructed into a higher position that the direct object. Perhaps all the heads in the whole verbal spread T-v-V must carry the same semantic verbal features. Ungrammatical case configurations were items #286-292. Sentence (75) remains a problem:

(75) *Pekka	nukku-i	koko	päivä-ä. (#293)
Pekka	sleep-pst.3sg	all	day-par

The sentence is ungrammatical, but it is judged grammatical by the model. The problem is that nothing prevents it from adjoining the DP adverbial to the complement position of the verb, as shown in the analysis (76), where it can check the partitive case against v-V.

(76)



Since the DP is adjoined to V, it does not count as an illegitimate complement. D can still check [T] from this position. Further exploration of this problem and examination of additional data led me to believe that the adverbial case marking – although technically sufficient to capture the phenomenon present in the present dataset – needs rethinking.

Much of the current understanding concerning adjunct case-marking in Finnish comes from research which has accounted for their properties by using some variation of the case tier hypothesis (Anttila & Kim, 2011; Maling, 1993, 2004; Poole, 2022). The main argument comes from expressions such as (77).

(77)	Liisan	täytyy	muistaa	matka	koko	viikon.
	Liisa.gen	must.0	remember-a/inf	trip.acc(0)	whole	week-acc(n)

‘(we) remembered the trip for the whole week.ä

The subject is assigned the genitive case, the direct object gets the nominative, while the adverbial is assigned the accusative (LEX > NOM > ACC). The NOM > ACC pattern is ‘shifted’ into the object-adverb complex when the subject is assigned (by assumption) the lexical case. This can be described elegantly by assuming that the nominative and accusative cases are distributed to the unmarked arguments from the case tier hierarchy. Under the present system, the case-marking at the adverbial differs from that of the direct object because the former is merged (right-adjoined) outside of the VP, causing its upward path to differ from elements inside the verb phrase. It then reconstructs into the second/outer specifier position of the verb phrase, and then becomes sensitive only to the agreement and polarity effects (reconstruction is not necessary, however). As a consequence, its upward path differs from that of the direct object. It checks the aspectual feature from T, which then explains how the accusative adverbial affects the aspectual properties of the whole event and how the mechanism operates in the case of intransitives.[38] The second problem with the case tier model is that it does not capture equally well the distribution of the partitive, genitive, personal pronouns and plurals. The third issue is the existence of constructions with several nominative arguments, including those with a nominative subject. In addition, the accusative in (77) is not obligatory: also the zero-accusative/nominative adverbial is possible (*Liisan täytyy muistaa matka koko viikko* ‘Liisa.GEN must remember trip whole week.NOM’).

---

[38] This assumption overgeneralizes, because it allows the researcher to assign aspectual features to any heads in the input. This, however, is a pseudo-problem in and itself, because it cannot be true that the aspectual ambiguity relies on the lexicon; some more general system must handle aspectual computations. The lexical strategy was used in this study only to test the behavior of the aspect features under various conditions.

6 Some residuum issues

6.1 Properties of case-based reconstruction

The analysis relies on the notion of case-based reconstruction that reconstructs arguments from positions where they cannot check their cases into positions where they can.

Overtly case-marked arguments (in Finnish) are first merged into the structure as adjuncts to account for the fact that their ordering is free. This means that the syntax of case-marked arguments in Finnish is unified with the syntax of adverbs, following (Jelinek, 1984). Adjuncts are attached to the primary phrase structure like any other constituent but processed inside a secondary processing pipeline and become invisible inside the primary plane. For example, they are invisible for labelling and sisterhood. We can imagine that they are "pulled" from the primary phrase structure and processed independently. Adjuncts are marked in the phrase structure trees by double lines, and in the output of the algorithm by using <,> instead of [,].

Case-based reconstruction is adjunct reconstruction or "adjunct float" operation that is applied to both case-marked arguments and adverbials. It locates the local finite boundary and searches for a suitable position downstream by using a minimal search developed on the basis of (Chomsky, 2008). Suitable positions are defined by relying on the case checking mechanism which demands checking of F for [TAIL:F]. The first suitable position found is always selected. The same mechanism is used for adverbials, where the relevant features are [TAIL:V] for 'V-adverbials', [TAIL:C] to 'T-adverbials', and so on. Thus, case forms are nominal versions of features which guide adverbials into their canonical positions where they can be interpreted. The same applies to PP-adverbials, and thus also to Finnish semantics cases. Notice that these operations makes it possible to reconstruct arguments and adverbials both leftwards and rightwards.

Interestingly, I have not been able to accomplish complete unification: nominal tail-features require matching via an upward path, while adverbial tail-features must satisfy a stronger requirement and occur inside the projection from the lexical head that contains the matched feature. It is easy to see what the difficulty is: the checking of the –NEG for the accusative, for example, is so nonlocal and liberal that adverbial licensing cannot be based on the same relation. We could try to use intervention to restrict the distribution of adverbials at the syntax-semantics interface, but so far I have not been able to achieve good results.[39]

Adjunct reconstruction differs from A/Ā-reconstruction. It resembles "scrambling" in the generative theory, assuming, of course, that scrambling is not A/Ā-movement.

6.2 Are genitive and nominative assigned to direct objects?

The n-accusative is homophonous with the genitive, zero-accusative with the nominative. This is true of singular full DPs, but breaks down in plural and for pronouns which have their own t-accusative forms. We can still ask if the n-accusative and zero-accusative forms are the *sui generis* genitive and nominative cases. The theory could possibly be more elegant if two case forms could be eliminated from the theory. I have been unable to accomplish such unification, and would like to present few remarks concerning the reasons.

The problem is that the syntactic contexts for the hypothetical unified genitive/n-accusative and nominative/zero-accusative cases are be different. To calculate the properties of the genuine nominative case correctly, it has to be tied in some manner to both agreement and finiteness. If the direct object can take the same form, so that we would have a "direct object nominative", it follows that it, too, must be related to the same properties. Yet, it is not obvious how direct objects could depend on agreement and finiteness, ostensible they have nothing to with either one: direct objects never agree, and they occur in syntactic positions that are not directly related to finiteness. Similarly, if the properties of the genuine genitive case are related to –FIN, how does this apply to the hypothetical direct object genitive? The contexts of genitive/n-accusative and nominative/zero-accusative are vastly different. In addition, recall that plural DPs and pronouns take the t-accusative form in these same contexts.

Those who have advanced unified models have typically claimed that the genitive case assignment can "target" the direct object if the right conditions are met, and by doing so produce the n-accusative form. Vainikka (1989), for example, claimed that the genitive percolates to the direct object when the relevant subject argument is missing. Similar analyses have been proposed for the nominative, which can also be said to target direct objects in the absence of the subject. Thus, although the analysis could reduce two case forms, the checking mechanism now requires that we detect more complex properties, such as "absence of subject."

[39] The upward path cannot capture specifier-head relations that are required, or so it seems, to model adverb placement.

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.

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. (2011a). Case assignment, case concord, and the quantificational case constructions. *Lingua*, *121*, 1042–1066.

Brattico, P. (2011b). The diehard extended projection principle. *SKY Journal of Linguistics*, *24*, 181–188.

Brattico, P. (2012). Long distance case assignment in Finnish. *Studia Linguistica*, *66*(3), 245–284.

Brattico, P. (2014). Long distance Case assignment and intervention. *Lingua*, *148*, 309–336.

Brattico, P. (2016). Is Finnish topic prominent? *Acta Linguistica Hungarica*, *63*, 299–330.

Brattico, P. (2018). *Word Order and Adjunction in Finnish*. Aguila & Celik.

Brattico, P. (2019). *Computational implementation of a linear phase parser. Framework and technical documentation (version 14.0)*.

Brattico, P. (2020). Finnish word order: does comprehension matter? *Nordic Journal of Linguistics*, *44*(1), 38–70.

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. (2021c). Null arguments and the inverse problem. *Glossa: A Journal of General Linguistics*, *6*(1), 1–29.

Brattico, P. (2022a). Computational analysis of Finnish nonfinite clauses. *Submitted*.

Brattico, P. (2022b). Predicate clefting and long head movement in Finnish. *Linguistic Inquiry*, *54*(4), 663–692.

Brattico, P. (2022c). Structural case assignment, thematic roles and information structure. *Studia Linguistica*. In press. https://doi.org/DOI: 10.1111/STUL.12206

Brattico, P. (2022d). Towards an information processing model of binding and coreference in Finnish. *Submitted for publication*.

Brattico, P., & Chesi, C. (2020). A top-down, parser-friendly approach to operator movement and pied-piping. *Lingua*, *233*, 102760.

Brattico, P., Chesi, C., & Suranyi, B. (2022). Agreement and chains. *Submitted for publication.,*

Chesi, C. (2012). *Competence and Computation: toward a processing friendly minimalist Grammar*. Unipress.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. MIT Press.

Chomsky, N. (1995). *The Minimalist Program*. MIT Press.

Chomsky, N. (2008). On Phases. In C. Otero, R. Freidin, & M.-L. Zubizarreta (Eds.), *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud* (pp. 133–166). MIT Press.

Csirmaz, A. (2005). *Semantics and phonology in syntax*. MIT.

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*. John Benjamins.

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

Holmberg, A., & Nikanne, U. (2008). Subject doubling in Finnish: the role of deficient pronouns. In S. Barbiers, O. Koeneman, & M. Lekakou (Eds.), *Microvariations in Syntactic Doubling* (pp. 325–349). Emerald.

Huumo, T. (2013). Incompleteness:, On the many faces of Object, Hide-and-seek with the Finnish partitive. *Folia Linguistica*, *47*(1), 89–111.

Jahnsson, A. W. (1871). *Finska Språkets Satslära: För läroverkens behof*. Finska Litteratursällskapets tryckeri.

Jelinek, E. (1984). Empty categories, case and configurationality. *Natural Language & Linguistic Theory*, *2*, 39–76.

Kaiser, E. (2002). Case Alternation and NPIs in Questions in Finnish. In L. Mikkelsen & C. Potts (Eds.), *WCCFL 21 Proceedings of the 21st West Coast Conference on Formal Linguistics* (pp. 194–207). Cascadilla Press.

Kaiser, E. (2003). Surprises at the Left Pheriphery. *GLOW Workshop Talk*.

Kayne, R. (1983). Connectedness. *Linguistic Inquiry*, *14.2*, 223–249.

Kayne, R. (1984). *Connectedness and Binary Branching*. Foris.

Kiparsky, P. (1998). Partitive Case and Aspect. In M. Butt & W. Geuder (Eds.), *The Projection of Arguments: Lexical and Compositional Factors* (pp. 265–307). CSLI Publications.

Kiparsky, P. (2001). Structural case in Finnish. *Lingua*, *111*, 315–376.

Koskinen, P. (1998). *Features and categories: Non-finite constructions in Finnish: Vol. Ph.D* [PhD. dissertation]. University of Toronto.

Krifka, M. (1989). Nominal Reference, Temporal Constitution and Quantification in Event Semantics. In R. Bartsch, J. van Benthem, & P. van Emde-Boas (Eds.), *Semantics and Contextual Expressions* (pp. 75–115). Foris.

Leino, P. (1991). *Lauseet ja tilanteet: suomen objektin ongelmia*. SKS.

Link, G. (1998). *Algebraic semantics in language and philosophy*. CSLI Publications.

Maling, J. (1993). Of Nominative and Accusative: the Hierarchical Assignment of Grammatical Cases in Finnish. In A. Holmberg & U. Nikanne (Eds.), *Case and Other Functional Categories in Finnish Syntax* (pp. 49–74). Mouton.

Maling, J. (2004). Case and grammatical functions: structural vs. hierarchical approaches to case assignment. *Proceedings of the 2004 Linguistic Society of Korea International Conference*.

Manninen, S. (2003a). Finnish PPs and the Phase Impenetrability Condition. In D. Nelson & S. Manninen (Eds.), *Generative approaches to Finnic and Saami linguistics* (pp. 295–320). CSLI Publications.

Manninen, S. (2003b). *Small Phrase Layers: A study of Finnish Manner Adverbials*. John Benjamins.

Manninen, S., & Nelson, D. (2004). What is a passive? The case of Finnish. *Studia Linguistica*, *58*(3), 212–251.

Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. MIT Press.

Megerdoomian, K. (2000). Aspect and Partitive objects in Finnish. In Billerey & Lillehaugen (Eds.), *WCCFL 19: Proceedings of the 19th West Coast Conference on Formal Linguistics* (Vol. 19, pp. 316–328). Cascadilla Press.

Megerdoomian, K. (2008). Partitive Case and Aspect. In R. Freidin, C. P. Otero, & M. L. Zubizarreta (Eds.), *Foundational Issues in Linguistic Theory* (pp. 73–104). MIT Press.

Nikanne, U. (1993). On assigning semantic cases in Finnish. In A. Holmberg & U. Nikanne (Eds.), *Case and other functional categories in Finnish syntax* (pp. 75–87). Mouton de Gruyter.

Phillips, C. (1996). Order and structure. *Dissertation, MIT*.

Pollock, J. Y. (1989). Verb Movement, Universal Grammar and the structure of IP. *Linguistic Inquiry*, *20*, 365–424.

Poole, E. (2015). A Configurational Account of Finnish Case. *University of Pennsylvania Working Papers in Linguistics*, *21*(1), 1–10.

Poole, E. (2022). Improper case. *Natural Language and Linguistic Theory*. https://doi.org/10.1007/s11049-022-09541-6

Reime, H. (1986). *On the notion of "Subject of" and its role in the theory of grammar*.

Rothstein, S. (2004). *Structuring events: A Study in the Semantics of Lexical Aspect*. Blackwell.

Tenny, C. (1994). *Aspectual roles and the syntax-semantics interface*. Kluwer.

Thomas, R. (2003). *The Partitive in Finnish and its relation to the weak quantifiers*. University of Westminster.

Toivainen, J. (1993). The nature of the accusative in Finnish. In A. Holmberg & U. Nikanne (Eds.), *Case and Other Functional Categories in Finnish Syntax* (pp. 111–128). Mouton de Gruyter.

Toivonen, I. (1995). *A study of Finnish infinitives*. BA thesis.

Vainikka, A. (1989). *Deriving Syntactic Representations in Finnish*. University of Massachusetts Amherst.

Vainikka, A. (1993). The three structural cases in Finnish. In A. Holmberg & U. Nikanne (Eds.), *Case and other functional categories in Finnish syntax* (pp. 129–159). Mouton de Gruyter.

Vainikka, A. (2003). Postverbal case Realization in Finnish. In D. Nelson & S. Manninen (Eds.), *Generative approaches to Finnic and Saami linguistics* (pp. 235–266). CSLI Publications.

Vainikka, A., & Brattico, P. (2014). The Finnish accusative: long distance case by agreement. *Linguistics*, *52*(1), 73–124.

Verkuyl, H. J. (1972). *On the Compositional Nature of the Aspects*. Reidel.

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

Yip, M., Maling, J., & Jackendoff, R. (1987). Case in Tiers. *Language*, *63*, 217–250.