

Elements of Finnish Syntax

Manuscript

Revised July 2020

Pauli Brattico

## Abbreviations

$\Lambda$	Proto-verb, i.e. eventive head that becomes a full verb when combined with further elements
$\varphi$	$\varphi$ -features (e.g., number, person)
$\varphi_*$	Unvalued/uninterpretable $\varphi$ -feature
0	Agreementless form
A	Adjective
all	Allative case
A/inf	A-infinitival
acc	Accusative case, any of the three forms
acc/gen	Genitive-looking accusative case
acc/nom	Nominative-looking accusative case (or nominative case)
acc(t)	Accusative case, the t-form
C	Complementizer
C/op	Operator field
cau	Causative morpheme or head
cond	Conditional mood
D	Determiner
$D_*$	Unvalued/uninterpretable D-feature
EPP	Extended projection principle
$EPP_{D,\varphi}$	Generalized EPP, relativized to probe features $D_*$ , $\varphi_*$
$EPP_{wh}$	Generalized EPP, relativized to probe feature $wh_*$
ela	Erelative case
expl	Expletive (e.g. <i>sitä</i> )
foc	Contrastive/corrective focus feature (often expressed by prosodic stress)
Force	Force head
gen	Genitive case
hAn	Second position discourse suffix/clitic <i>-hAn</i>
kO	Yes/no question suffix/clitic <i>-kO</i>
KSE/inf	KSE-infinitival
impass	Impersonal passive verb form
imp	Imperative form

ine	Inessive case
N	Noun
Neg	Negative head, negative word <i>e-</i>
nom	Nominative case
Num	Numeral head
MA/inf	MA-infinitival
OP	Operator
P	Preposition, or more generally an adposition
pA	Second position discourse suffix/clitic <i>-pA</i>
par	Partitive case
past	Past tense
pl	Plural
PRO	Controlled null subject pronoun
pro	Null subject pronoun (small pro)
prtcl	Participle form
Q	yes/no question feature
sg	singular
T	Tense (past, present)
T/fin	Finite tense (tense head with feature FIN)
T/past	Past tense
T/pres	Present tense
TUA/inf	TUA-infinitival
v	The small verb <i>v</i> ; transitivizer
V	Verb
VA/inf	VA-infinitival (both present and past forms)
VAR	Variable (for an operator)
wh	interrogative feature
wh_	Uninterpretable <i>wh</i> -feature

1	Theoretical background.....	10
1.1	Aims and scope .....	10
1.2	Language and cognition .....	11
1.2.1	The human language faculty.....	11
1.2.2	Merge.....	14
1.2.3	Move.....	21
1.2.4	Agree and the lexicon.....	26
1.2.5	An overview of Finnish clause structure .....	30
2	Overview .....	38
2.1	Finnish as a Finno-Ugric language.....	38
2.2	Finnish in a crosslinguistic perspective .....	38
3	CP-cartography and operator movement (K-field).....	50
3.1	Introduction .....	50
3.2	The high complementizer <i>että</i> ‘that’ .....	50
3.3	The operator field <i>C/op</i> .....	58
3.3.1	The operator position hosts several types of phrases.....	58
3.3.2	Head movement into the operator field .....	70
3.3.3	Relative clause lacks Force and discourse-features.....	72
3.3.4	The operator position and semantic interpretation .....	73

3.3.5	Optionality, the high complementizer, and the head/feature distinction .....	77
3.3.6	The syntactic status of the operator field features and criterial scanning .....	79
3.4	Operator movement .....	83
3.4.1	Introduction and terminology .....	83
3.4.2	Pied-piping and the edge generalization.....	84
3.4.3	The edge generalization is satisfied by operator movement.....	90
3.4.4	Pied-piping and adjunction.....	90
3.4.5	Islands and optional pied-piping.....	91
3.5	Horizontal and vertical modularity.....	96
3.6	Summary .....	100
4	The preverbal subject (T-field).....	101
4.1	Introduction .....	101
4.2	The preverbal subject field and its syntax .....	102
4.2.1	Distinguishing preverbal subjects from grammatical subjects .....	102
4.2.2	Disentangling the preverbal field from the tense projection .....	104
4.2.3	The preverbal subject position is regulated by an EPP principle .....	108
4.2.4	Discourse-configurationality and the topic.....	110
4.2.5	An analysis of the Finnish finite clause EPP .....	115
4.2.6	Feature inheritance or Fin projection?.....	121
4.2.7	Finnish expletive <i>sitä</i> : syntax and pragmatics .....	125
4.3	Nonsubject topicalization as a grammatical operation .....	128
4.3.1	The problem.....	128
4.3.2	Nonsubject topicalization as A-bar movement.....	129

4.3.3	Free word order as a post-syntactic stylistic movement.....	134
4.3.4	Free word order as adjunction, not movement.....	136
4.3.5	What is adjunction and how it applies to arguments?.....	141
4.4	Verb-initial clauses.....	144
4.4.1	Introduction.....	144
4.4.2	Null subject pronoun pro.....	144
4.4.3	Coming into existence.....	145
4.4.4	Generic sentences.....	148
4.4.5	Impersonals.....	153
5	The postverbal field.....	154
5.1	Introduction.....	154
5.2	Lexical roots and the argument structure.....	154
5.2.1	Lexical roots and labels.....	154
5.2.2	Derivational morphology.....	158
5.3	Infinitivals.....	164
5.3.1	Introduction.....	164
5.3.2	The A-infinitival.....	171
5.3.3	VA-infinitival.....	183
5.3.4	MA-infinitivals and MA-adverbials.....	191
5.3.5	E- and ESSA adverbs.....	197
5.3.6	TUA-adverbial.....	200
5.3.7	KSE-adverbial.....	202
5.4	Infinitival word order.....	204

5.5	Infinitival edge positions .....	210
5.6	Impersonal passive .....	214
5.6.1	Basic properties and classification .....	214
5.6.2	The impersonal agent .....	217
5.6.3	Voice as a verbal feature (Manninen and Nelson 2004).....	219
5.6.4	The impersonal passive and the EPP.....	222
5.6.5	Implicit argument or no argument at all? .....	223
5.6.6	Negated passives .....	226
5.6.7	The passive morpheme -tAAn: one or two affixes? .....	227
5.6.8	Impersonal passive and the direct object case .....	229
5.6.9	The “active” impersonal construction .....	230
5.7	Modal construction.....	235
6	Morphosyntax: agreement and case assignment .....	239
6.1	Morphosyntax and the d-structure.....	239
6.1.1	Morphosyntax is computed before argument float and A-bar movement.....	239
6.1.2	Morphosyntax and A-movement.....	243
6.1.3	The morphosyntactic hierarchy .....	247
6.2	Structural case assignment .....	254
6.2.1	Vainikka’s theory of Finnish structural case assignment .....	254
6.2.2	The analysis: an overview .....	263
6.2.3	Partitive (Vainikka’s partitive generalization).....	264
6.2.4	Genitive (non-finite subject/specifier case).....	269
6.2.5	The possessive-noun construction.....	271

6.2.6	Nominative (finite subject/specifier case) .....	276
6.2.7	Accusative and aspect.....	278
6.2.8	Accusative and the partitive of negation .....	299
6.2.9	Accusative, $\phi$ -agreement and singular full DPs .....	303
6.2.10	The numeral-noun construction.....	309
6.2.11	Spec-head agreement? .....	317
6.2.12	Alternative analyses.....	321
6.2.13	Summary.....	323
6.3	Phi-agreement: finite and non-finite.....	324
6.3.1	Finite agreement .....	324
6.3.2	Infinitival agreement .....	325
6.3.3	Agreement and locality.....	336
6.4	Nonlocal case assignment.....	340
6.4.1	Introduction .....	340
6.4.2	Agree is made of two separate grammatical processes .....	341
6.4.3	Case checking is initiated by the goal, not the probe .....	341
6.4.4	Principles of nonlocal case assignment .....	343
6.4.5	Deriving the data .....	346
6.4.6	Nonlocal case checking and the problem of the upper boundary condition.....	349
7	Special issues.....	355
7.1	Null subjects .....	355
7.1.1	Null subject pro .....	355
7.1.2	Control and PRO .....	362

7.2	Word order .....	366
7.2.1	Ordering of arguments.....	366
7.2.2	Ordering of heads .....	371
7.2.3	Ordering of adverbs, adverb scope and d-structure .....	372
7.2.4	In what sense is Finnish word order “free”?.....	376
7.3	Noun phrase.....	379
7.3.1	Determiners, definiteness and the status of D .....	379
7.3.2	Head marking .....	381
7.3.3	Word order inside the noun phrase.....	383
7.3.4	Assignment of the genitive inside the noun phrase .....	388
7.3.5	A note on quantifiers and quantifier phrases .....	389
7.4	Adpositions, the preposition phrase (PP) .....	394

## 1 Theoretical background

### 1.1 Aims and scope

This book covers basic syntactic mechanisms and structural properties of Finnish together with the major analysis present in the literature. It first tackles the Finnish left periphery and A-bar/operator movement (Chapter 3), the preverbal subject position and finiteness (Chapter 4), then moves on to discuss the syntactic construction of infinitival structures (Chapter 5). Several exceptional finite clauses are also discussed, such as the modal construction and the impersonal passive. These chapters cover what we could label as the basic ‘syntactic template’ of Finnish. Morphosyntax is addressed in a separate chapter (Chapter 6), followed by special topics that would merit a book of their own but which must be discussed, no matter how briefly, in order to understand the rest of the material. These include null subjects, pro-drop and control (7.1), word order (7.2), and the structure of the Finnish noun phrase (7.3) and the preposition/adposition phrase (7.4). Many advanced topics, such as control, quantifier scope, long head movement, and binding, are excluded.

The volume is neither a textbook nor a completely impartial review of the literature. It has been composed within a narrowly defined research agenda, much of which was defined by Anne Vainikka in her seminal thesis on Finnish syntax. This work focuses on syntax with the exclusion of phonology, phonological side of derivational morphology, semantics and pragmatics, the latter which are only occasionally referred to. On the other hand, within that narrow research agenda the book aims for a relative broad and comprehensive coverage, with emphasis on more recent results. Thus, a reader who comes to this book without any prior knowledge of Finnish syntax should, to the extent that I have succeeded in fulfilling the intended goal of the book, find a balanced presentation that also attempts at drawing lines between what is considered uncontroversial and what is relatively well-established. Most of the key claims presented are argued for

empirically, at least the key empirical arguments are provided, thus a reader can form an educated opinion concerning the force of each argument.

The chapter discussing morphosyntax merits a special mention. Morphosyntax is a contested area of Finnish. The literature offers several radically different approaches to the topic that are difficult to compare. My presentation of the topic can only be described as idiosyncratic, in that the material is organized against the background of Vainikka's theory of Finnish structural case that the chapter then expands and modifies in the face of empirical evidence Vainikka's earlier work did not take into account. This way of organizing the material has the advantage that it allows the reader to gather the basic empirical generalizations before moving into more controversial and complex topics, but a disadvantage is that several other approaches meriting detailed discussion can only be mentioned.

The natural-scientific study of syntax is unpopular in Finland and has been at least since the 1970s due to various sociology of science related factors. The result is that there are very few natural scientific studies of Finnish syntax if compared to, say, the situation in Hungarian or Swedish. This has been a blessing in the sense that a treasure trove of unexplored data is something any author can dream of, but the negative side is that for most phenomena the range of alternative hypotheses is narrow; in some cases, zero. This should be kept in mind when considering how I cite the literature. This also explains why I sometimes discuss analyses and claims that are perhaps not that well justified in the sources and which the author(s) mention only passing or perhaps only in a footnote; again, if cursory comments or speculative ideas are all there is, then that's what we must build on. My point was not to pick up and specifically focus on points that the author perhaps did not want to be interpreted as fully fledged analyses, but instead to use them as a starting point.

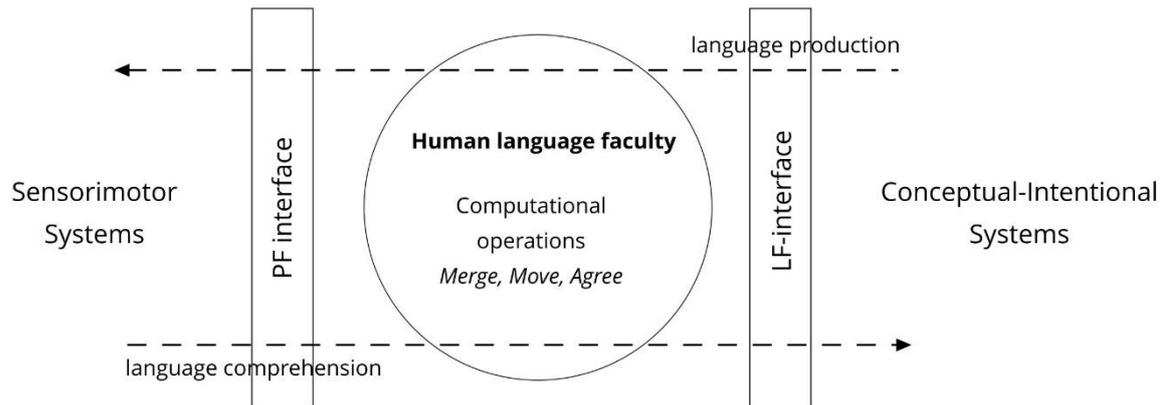
## 1.2 Language and cognition

### 1.2.1 *The human language faculty*

The human language faculty interacts with the sensorimotoric systems of the brain to produce and interpret overt forms of language, such as its spoken, written, or gestured forms. At the outer edge of these processes are the final, concrete motoric programs or early sensory processing systems involved in concrete sensorimotoric activity, not only language, and which are then connected, via neuronal wiring, with the low-level systems specialized in processing and representing language-specific information, such as its phonetic

and phonological forms. The latter two can be thought of as constituting language's sensorimotor interfaces, levels of representations and processing systems that mediate or sustain communication between low-level non-linguistic sensorimotoric processing and language-specific processing. It is customary to refer to this interface as the *phonetic(-phonological) form* (also PF-interface).

In addition to the sensorimotoric properties, language must exchange information with cognitive and perceptual systems responsible for thinking, decision making, emotions, communication and systems of conceptual-intentional representations. We can use language to express thoughts, to influence others and perhaps also to clarify and explicate our own internal thoughts and thinking processes. The semantic systems can, of course, be accessed without the intervention to language. Playing a musical instrument, seeing a smiling face, or hearing the baby crying impact our mental, emotional lives through the sensory systems in a way that does not necessarily involve the use of any linguistic forms. But when these semantic systems are mobilized by means of linguistic forms, a specialized system, the human language faculty, comes into play that recognizes, represents, processes and repackages linguistic information, specifically, into a form that is understood by the semantic systems. This requires that there is a connection between language and the *conceptual-intentional* (C-I) processes. This interface is called *Logical Form* or LF. It constitutes a level of representation, a system or perhaps several, that is connected causally (and logically, representationally) both to the linguistic, language-specific computations and the processes that take place in the conceptual-intentional systems. The LF has no a priori connection to logic or any logical system; it denotes a neuronal wiring configuration that connects linguistic representations and its internal computations with meaning. Its properties are a matter of empirical inquiry. The whole architecture, comprising of both the PF-interface and LF-interface and assumed as a background in this study, is depicted in Figure 4.



*Figure 4. The human language faculty as embedded within other cognitive systems, with which it interacts through the PF-interface and LF-interface. All the components depicted abstractly in the figure are biological components of the human brain and its neuronal architecture. The computational operations of the language faculty (Merge, Move, Agree) will be discussed in the subsequent sections of this chapter.*

The three systems (sensorimotor systems, conceptual-intentional systems and the language faculty) and their interactions (PF, LF) are described in terms of information processing. The sensory stimulus is represented as a physiological activation pattern at the sensory organ(s), after which it is processed by upstream systems that represent its properties in an increasingly abstract way. If the brain interprets sensory stimulus as consisting of linguistic or linguistically relevant input, it will arrive at and is interpreted by the PF-interface system(s), and is then further processed by the higher-level modules, connecting physiological stimuli with abstract linguistic features, such as syllables, stress, prosody, words, lexical categories, inflectional forms, word order and lexical, phrasal meanings and communicative intentions. Similarly, when we generate linguistic utterances by free will, the process begins in an abstract system responsible for representing an idea (e.g., a mental representation, image, dream) we hope to communicate, often together with a larger intention or purpose, which is then transduced, perhaps in several phases and in several relatively independent parallel streams, into concrete representational forms until a sequence of motoric actions is orchestrated and transmitted through some medium to the hearer. These processing steps are necessary because there is no linguistic meaning or linguistic intention present in any raw stimulus (e.g., in an auditory waveform). Its linguistic properties must be hallucinated, hence represented in an increasingly abstract way

before the hearer can understand what might have prompted the speaker to generate the auditory waveforms in the first place.

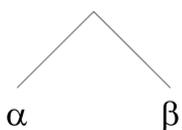
### 1.2.2 Merge

Information processing in the human language faculty makes use of computational operations. These operations manipulate (generate, store, retrieve and transform) linguistic representations for the purpose of linguistic communication and other forms of linguistic behavior. The basic properties of the few of the most fundamental operations will be elucidated next. The following presentation does not aim for completeness, and it does not discuss the matter exhaustively with the aim of a broad or deep theoretical coverage; rather, my purpose is to provide a minimal background for the rest of the book. A detailed theoretical model of Finnish clause structure, build on the assumptions laid down here, will emerge as we proceed to later chapters.

A core property of the human language system is our ability to control some of its representations creatively by free will, hence independently of both its sensory and semantic properties. Once we have acquired basic words and rules that can be used to put them together, we can generate new linguistic expressions without limit. The resulting word combinations might have coherent meaning, or they might not—whether they have does not limit our ability to manipulate them. Hence the linguistic system is in principle autonomous with respect to both the sensory systems and the conceptual-intentional systems.

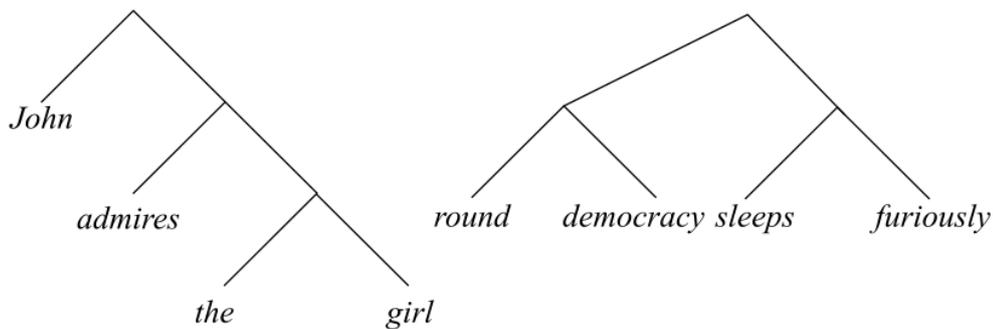
Linguistic theories capture creativity by assuming that there is a cognitive operation that produces complex linguistic elements by combining simpler elements. Thus, language “is a system of discrete infinity consisting of hierarchically organized objects” (Chomsky 2008:137). Suppose we take two linguistic elements  $\alpha$  and  $\beta$  and combine them. The combinatorial operation, call it Merge, yields  $[\alpha, \beta]$ , a combination of the two (1).

(1)



The operation is recursive: after applying to  $\alpha$  and  $\beta$ , the same operation can apply to its own output  $[\alpha, \beta]$  to yield  $[\chi [\alpha, \beta]]$ . If we add one more element, the result could already constitute a rudimentary sentence, such as [*John [admires [the girl]]*] or [[*round democracy*] [*sleeps furiously*]]. The phrase structures corresponding to expressions, created by Merge, are shown in (2).

(2)



All rules of Finnish examined in this book make use of Merge and, vice versa, we cannot make any sense of the syntactic regularities of Finnish without assuming at least (1).<sup>1</sup> The operation is implemented by the neuronal systems of the human brain: a matter that is investigated in neurolinguistics.

Prior to the theory of Merge, linguistic phrase structures were described as language-specific representations, generated and computed by using language-specific rules such as ‘ $S \rightarrow NP + VP$ ’, ‘generate a sentence by combining a noun phrase and a verb phrase’. In subsequent work, it was found that the combinatorial principles involved in language were more abstract, a development that led to the theory of Merge (Chomsky 1992; Collins 2002; Fukui 1986; Speas 1986). This theory is so abstract that it can, in principle at least, apply to other creative modalities, such as music, planning, tool making or mathematics. We can combine musical elements, mental plans, or create complex tools by combining simpler ones. It is possible, in other words, that the operation is not restricted to language. Perhaps, as Chomsky has speculated, it emerged as a critical mutation associated with the emergence of the exceptional creative-cultural abilities of *Homo sapiens* (Chomsky 2005). The issue is more complex because some amount of the language-specific properties we see in connection with the application of Merge to linguistic material, as will be evident in the material

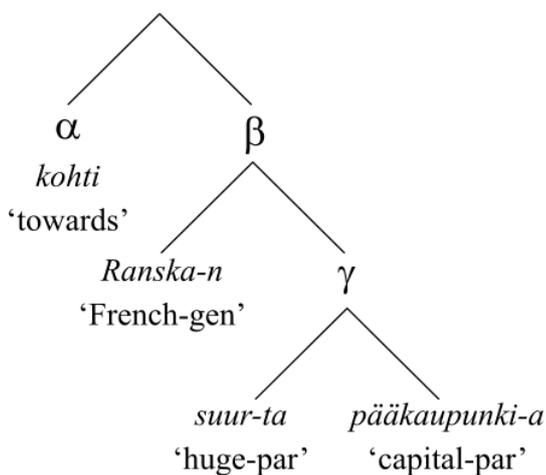
<sup>1</sup> See (Sammallahti 2002, 2003), among many others, for claims to the contrary. I have discussed such alternatives in (Brattico 2019e).

discussed in this book as well, might arise from the non-recursive components of language, such as its lexicon, especially from its functional lexicon (Brattico 2010b; Brattico and Liikkanen 2009).<sup>2</sup> The theory of linguistic creativity need not be thought of as an exclusively language-specific computational operation. For the purposes of this work it suffices to assume that Merge exists and that it can combine linguistic objects recursively.

Elements that are merged are related to each other structurally. In the simplest possible theory, the operation creates a set  $\{\alpha, \beta\}$  (Chomsky 1995, 2008; Collins 2002). This could be called “simplest MERGE,” in that no simpler operation is imaginable (Chomsky, Gallego, and Ott 2019:236); for discussion, see (Collins 2017). An alternative, assumed here, is that Merge creates an asymmetric structure  $[\alpha, \beta]$  with a left and right constituent. Thus, in the example (3), the left constituent is the preposition *kohti* ‘towards’, while the right constituent is the complex phrase *Ranska-n suur-ta pääkaupunki-a* ‘French-gen huge-par capital-par’.

- (3)  $[\alpha$  kohti  $[\beta$  Ranska-n [suur-ta pääkaupunki-a ]]]
- towards France-gen huge-par capital-par
- ‘Towards the capital of France’

(4)



<sup>2</sup> A functional word has the defining property that it requires a complement, which could go some way towards explaining the language-specific properties of the head-complement configuration, the “edge feature” of (Chomsky 2008). They also possess a curious affinity to D,  $\phi$  features that is visible in  $\phi$ -agreement, control and specifier filling, which might account for some of the remaining properties. The matter remains controversial.

The same logic applies to the constituents of  $\beta$ . Here, the possessor *Ranska-n* ‘French-gen’ constitutes the left constituents,  $\gamma$  the right. The assumption that Merge is asymmetric should be considered as a theoretical principle that makes the analysis of Finnish simpler. Working with the “simplest” set-theoretical Merge, on the other hand, creates complications in the data analysis that we want to avoid in this book in order to render the data more transparent. Finally, the terms “left constituent” and “right constituent” separate constituents from each other based on their left-right ordering in phrase structure illustrations; we do not know how the asymmetry is ultimately implemented in the brain.

Phrase structure objects created by simple asymmetric union do not have labels. Thus, there are no verb phrases or noun phrases. Yet such notions play a role in any descriptively adequate grammar. The label of any complex constituent can be defined by rule (5), which suffices for the purposes of this book.

(5) *Labeling of  $\gamma$*

Suppose  $\gamma = [\alpha, \beta]$ , then

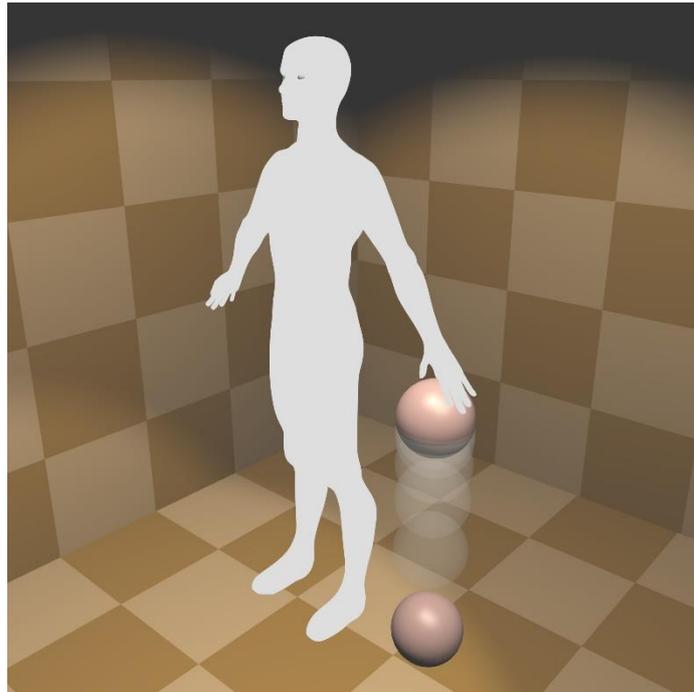
- a. If  $\alpha$  is primitive, it will be the label; else,
- b. if  $\beta$  is primitive, it will be the label; else,
- c. apply (5) to  $\beta$ .

The intuitive function of (5) is to find the highest primitive head from  $\gamma$  and designate it as the most dominant category, providing the label for the whole phrase. To illustrate, consider again the preposition phrase *kohti Ranskan suurta pääkaupunkia* ‘towards France.gen huge.par capital.par’ and its structure [towards [Franch.gen [huge.par capital.par]]]. The highest unit consists of [P NP] structure, hence the label of the whole phrase is P according to (5)a and the whole phrase will be a preposition phrase. This is because the preposition is the highest primitive constituent to the left. If ‘France.gen’ and ‘huge.par’ are DPs and APs, respectively, then (5) will provide N as the label of the complement (P-NP). Labels will often be marked explicitly into the analyses, as shown in (6), but they are determined by an algorithm and are not intrinsic properties of the complex phrases.

(6) [PP towards [NP [DP France’s] [NP [AP really huge] capital]]]

Rule (5) is simpler than what is commonly found in the primary literature, but this simplicity is due to the fact that I decided to treat Merge as an asymmetric operation.

Representations created by Merge are associated with meaning through the LF-interface (see Figure 4). Consider a prototypical finite sentence describing a situation in which somebody does something to someone, such as when Jari drops the ball (see Figure 5 below).



*Figure 5. The meaning of the sentence “Jari dropped the ball,” represented in an image format. The event of Jari’s dropping the ball begins in some manner from within Jari (in his conscious decision, for example), flows from his hand and other actions into the ball, that will then participate in the event by falling down; the event terminates when the ball touches the ground.*

The sentence *Jari drops the ball* consists of a predicate (a finite verb, for example) and number of arguments (e.g., ‘Jari’, ‘the ball’). The verb denotes an event of dropping, while the arguments denote the participants of that event. The participants have asymmetric roles in that event. One participant is the ‘agent’, the participants whose actions cause the event to occur, or who participates in the event first, and another, the ‘patient’, whose participation in the event comes next and/or who is affected by the event without having originated it. In this case, the patient is the ball. These roles are called *thematic roles*. Thus, the sentence *Jari drops the ball* is mapped systematically to things that are part of its meaning, here conceived as a conscious (nondiscursive)

representation, precept, or a mental image. Arguments ('Pekka', 'the ball') are merged close to lexical items representing the predicates ('dropped'). The complement of the lexical verb (intransitive verbal stem) represents the patient, and the specifier of a functional head, small verb *v*, encoding transitivity (here *pudo-* 'fall' → *pudo-ttaa* 'drop'), represents the agent (7).

- (7) [Jari                    [ *v*(-tta-) [ *pudo-*            *pallo-n.*]] = *Jari pudo-tta-a pallon*  
 Jari.nom                *v*            fall-            ball-acc/gen  
 'agent'                'cau'        'event'        'patient'  
 'The agent (Jari) caused the event, i.e. the ball to drop.'

Empirical data suggests that structural configurations are mapped systematically to various things in the mental representation of meaning. This arrangement is illustrated in Figure 7, following standard assumptions in current linguistic theorizing.

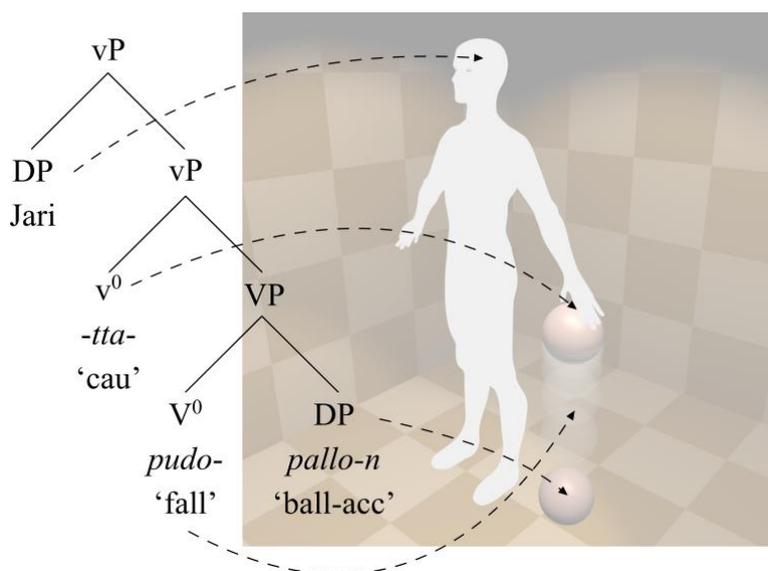


Figure 7. Relationship between clause structure (at the LF-interface) and semantics as assumed in this study. The constituents of the clause and their mutual configurational relations are mapped systematically to meaning (here conscious, nondiscursive and image-like representation).

Let us examine these assumptions further by using a more interesting example. The partitive-accusative alteration in Finnish direct object case is known to correlate with aspects of semantic interpretation. When the accusative is used, the event corresponding to the verb phrase is perceived as being terminated or

completed (i.e. it is perceived as being telic), whereas the partitive signifies that it might still be ongoing (being atelic). We therefore have two systems of representation that we correlate, a semantic representation, shown in Figure 9, and case assignment, illustrated in (8).

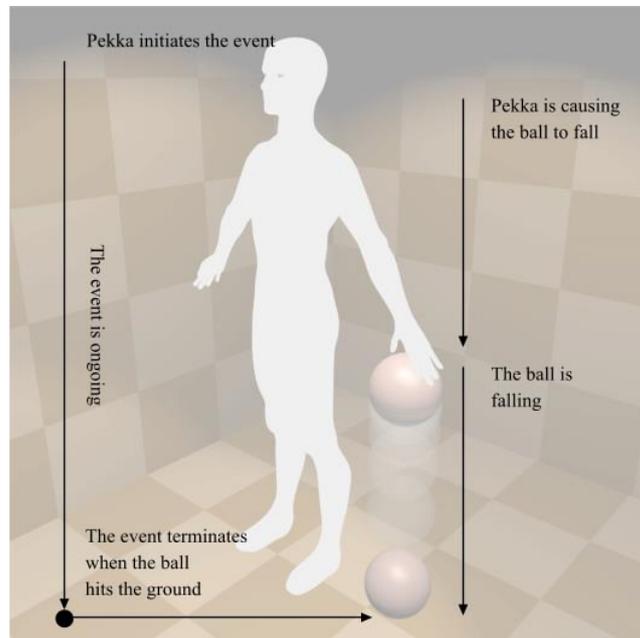


Figure 9. An anatomy of human experience as dissected under a particular linguistic perspective.

(8)

- a. Pekka pudotti pallo-a.  
 Pekka dropped ball-par  
 ‘Pekka was dropping the ball.’
- b. Pekka pudotti pallo-n  
 Pekka dropped ball-acc  
 ‘Pekka dropped the ball.’

The case alteration at the level of language and linguistic representations correlates with something at the level of semantic representation, in this case it is sensitive to whether the sentence includes or excludes the termination point. Language can therefore be said to provide “perspectives” to meanings in the sense that the same semantic representation could be described by an unlimited number of different sentences, each which provides a different perspective to it (e.g., the ball is falling, the ball is being dropped, there is an event of

ball's falling). By using the accusative case, specifically, the speaker is specifically emphasizing the termination point of the event.

In sum, linguistic representations created by Merge are mapped, systematically, to semantic representations, the latter which consists ultimately of human experience viewed under particular perspectives provided by both linguistic and extralinguistic prisms.

### 1.2.3 *Move*

In Finnish, standard canonical finite clause follows the SVO pattern, in which the subject (S) is followed by the verb (V), which is followed by the direct object (O)(9)a. It is possible to reverse the order without changing the propositional meaning (9)b.

(9)

a. Jari                ihailee                Merja-a.                (SVO)

Jari.nom                admires                Merja-par

'Jari admires Merja.'

b. Merja-a                ihailee                Jari.                (OVS)

Merja-par                admires                Jari.nom

'Jari admires Merja.'

A native-speaker would not treat the two sentences as exactly identical in meaning: example (b) carries a stylistic component, often difficult to elucidate without linguistic training. The example shows, however, that it is possible to use word order to convey some aspects of semantic interpretation. The operation is subject grammatical restrictions. Reversal of the subject and object positions is impossible when performed inside an infinitival clause, for example (10).

(10)

a. Pekka                käski                [Jari-n                hakea    heidä-t.]                (SV<sub>inf</sub>O)

Pekka.nom                ordered                Jari.gen                to.pick    they-acc

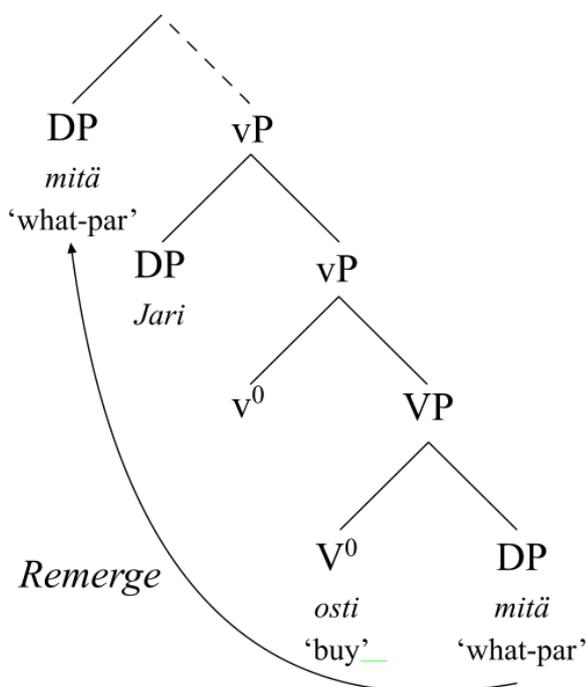
'Pekka ordered Jari to pick up them.'

- b. \*Pekka käksi [heidä-t hakea Jari-n.] (OV<sub>inf</sub>S)  
 Pekka.nom ordered they-acc to.get Jari-gen

Since noncanonical word orders are subject to grammatical constraints, the operation cannot be assumed to literally “free.” Instead, it is based on a computational operation of the language faculty. To capture it, we assume that Merge can combine  $\alpha$  and  $\beta$  even when  $\alpha$  is taken from within  $\beta$  (Chomsky 2008). A constituent that has already been merged to the structure is effectively re-merged into another position. I will use the term *movement* or operation Move in this book to refer to this special form of Merge. A constituent  $\alpha$  that is merged with  $\beta$  will be copied and remerged from within  $\beta$ . The result is  $[\alpha [\beta. . \alpha. . ]]$ . Interrogative clause provides one much-studied example of movement. The interrogative word is first-merged to its canonical position and is then merged again to a higher position to signal the scope of the interrogative, as illustrated by the example (11).

- (11) [Mitä<sub>i</sub> [Jari osti \_\_\_<sub>i</sub>?]  
 what.par Jari.nom bought  
 ‘What did Jari buy?’

(12)



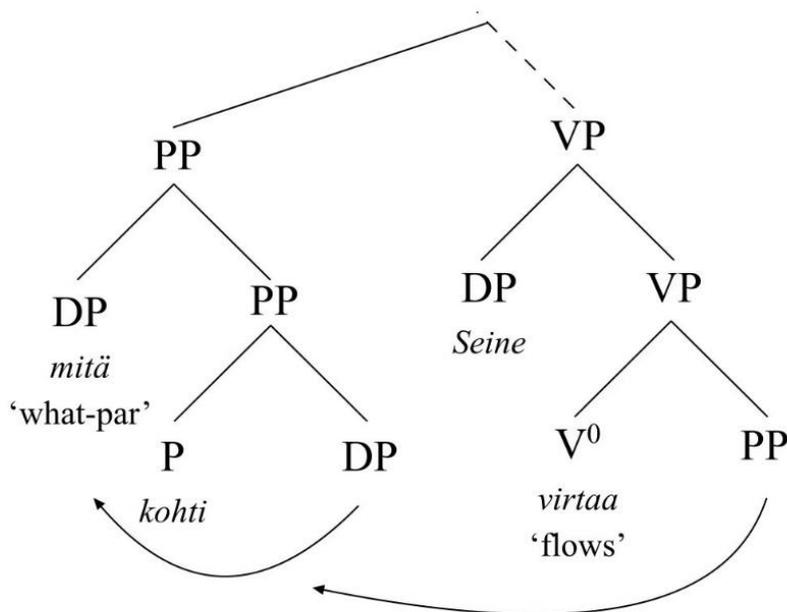
Overt *wh*-movement does not always involve scope marking or checking a criterial force *wh*-feature, as shown in (13). In this example the DP complement of the preposition moves to Spec,PP. Similar properties are true of core examples of A-movement, such as raising (14), in which the thematic subject of the complement clause is dislocated to the position of grammatical subject (to the specifier of  $\alpha$ ) for reasons that have to do with morphosyntax.

- (13) [Mitä<sub>1</sub> kohti \_\_<sub>1</sub>]<sub>2</sub> Seine virtaa \_\_<sub>2</sub>?  
 what.par towards Seine.nom flows  
 ‘Towards what does Seine flow?’

- (14) [ <sub>$\alpha$ P</sub> Jari<sub>1</sub> näyttää( $\alpha$ ) [<sub>\_\_<sub>1</sub></sub> voittavan.]]  
 Jari.nom seem.3sg to.win  
 ‘Jari seems to be winning.’

The following phase structure illustrates (13).

(15)



The standard theory holds that functional heads may have a feature or diacritic, call it EPP, that causes something to move (get remerged) to its specifier position. The EPP is like “specifier selection feature” that

is satisfied by remerge (“second edge feature” in (Chomsky 2008)). The functional head imposing the requirement and the phrase that is moved are related to each other by means of grammatical features. It is of course not a coincidence that an interrogative clause involves movement of a phrase that contains an interrogative feature. To describe dependencies of this type, I assume that the functional head has a feature  $EPP_{F...}$ , where  $F...$  denotes the relevant features, in this case the interrogative feature  $wh$ . We can think of  $EPP_F$  as a specifier selection targeting feature  $F$ . Example (16) provides several examples of this mechanism.  $C$  denotes a higher functional head hosting the interrogative or relative feature discussed later in this book.

(16)

- a. Kenet  $C(wh)$  Pekka tapasi \_?  
 who  $EPP_{wh}$  Pekka met  
 ‘Who did Pekka meet?’
- b. [Mitä kohti \_]  $C(wh)$  Seine virtaa \_?  
 what towards Seine flows  
 $EPP_{wh}$   $EPP_{wh}$   
 ‘Towards what does Seine run?’
- c. Arkku, jonka  $C(rel)$  Pekka rikkoi \_  
 chest which  $EPP_{rel}$  Pekka break  
 ‘A chest that Pekka broke.’

Existence of movement has constituted something of an enigma in linguistic theorizing, including generative theory. Finnish facts, examined in detail later in this book, suggest that we will have to live with an irreducible EPP property, and the facts are important enough to mention already here. In Finnish, many prepositions (adpositions) can exhibit two forms, one in which the argument noun phrase occurs in the complement position (17)a and another in which it exhibits the EPP-behavior (17)b.

(17)

- a. lähellä minu-a  
 near I-par  
 ‘near me’

- b.       minu-n        lähellä-ni  
           I-gen        near-1sg  
           'near me'

The two forms do not differ in meaning. Furthermore, some prepositions only exhibit (a), others exhibit only (b). Given there is no difference in meaning, there seems to be no other alternative than to describe the variation by means of a lexical feature that is present in some prepositions and not in others, triggers the EPP-profile either optionally, as in the above example (a-b), obligatorily, when the preposition only accepts a genitive argument at its Spec,PP position (b). That lexical feature will correspond to the  $EPP_F$  diacritic assumed in this study, forcing some phrase to move to the specifier position in a way that is often associated with morphosyntactic effects, such as phi-agreement and case assignment, as is the case here.

*Head movement* refers to a similar operation, but it targets a primitive head and merges it with another head:  $[\alpha^0 + H^0 [ \dots \alpha^0 \dots ]]$ . Properties of the complex head  $\alpha + H$  and the nature of the combinatorial operation that creates it are controversial, but the outcome is not: the operation produces complex words such as tensed verbs and deverbal adjectives. The operation is particularly transparent in an agglutinative language such as Finnish, in which words can be generated productively by applying head movement. Thus, *pudo-tta-utu-isi-n-ko-han* is a complex but not in any way marginal word in Finnish that is composed of several morphemes representing (in this order) the root stem 'to fall' (*pudo-*), transitivization/causativization (*-tta-*), reflexivization (*-utu-*), conditional (*-isi-*), first person singular (*-n-*), yes/no interrogativization (*-ko-*) and topicalization (*-han*) and is interpreted to mean 'would I perhaps cause myself to fall?'. Phrasal movement and head movement can be denoted by the common term *movement* or the operation *Move*. The morpheme structure of a complex word reflects its syntactic derivation in a transparent fashion (Baker 1985; Julien 2002). The assumption is justified empirically and has been assumed to my knowledge in virtually all generative studies of Finnish at least since (Holmberg et al. 1993). The basic generalization underlying this hypothesis is that as the morphological richness (or the range of expressive options) of a verb increases, so does its position both in the hierarchical order and in the left-to-right order, as shown in (18).

(18)

- a.       *Only one form, A-infinitival (=grammatical information available at the word)*

Jari            ei        halun-nut        **maala-ta**        talo-a.  
 Jari.nom        not.3sg want-past        paint-INF        house-par

'Jari did not want to paint the house.'

b. *Lower tense, number, voice*

Jari            ei        ollut        **maalan-nut**        talo-a.  
 Jari.nom        not.3sg be.past paint-past.sg        house-par

'Jari had not painted the house.'

c. *Tense, number, mood, voice*

Jari            ei        **maalan-nut**        talo-a.  
 Jari.nom        not.3sg paint-past        house-par

'Jari did not paint the house.'

d. *Tense, number, person, mood, voice, finiteness, positive polarity*

Jari            **maala-si**        talo-a.  
 Jari.nom        paint-3sg.past        house-par

'Jari painted the house.'

e. *Tense, number, person, finiteness, mood, voice, positive polarity, operator, force (interrogative)*

**Maala-si-ko**        Jari        talo-a.  
 paint-3sg.past-Q        Jari.nom        house-par

'Did Jari paint the house?'

The grammatical features exist as independent heads and are then picked up as the verb moves higher by head movement, or are expressed by independent morphemes (e.g., negation, auxiliary) which then blocks verb movement. For an early and still correct analysis of the full finite clause in Finnish along these lines, see (Holmberg et al. 1993). An analysis of finite clause, based on this analysis of complex words, will be sketched in Section 1.2.5 of the present chapter.

#### 1.2.4 Agree and the lexicon

The lexicon is often conceived as the storage of language's words. This starting point is insufficient for linguistic analysis. Many segments that speakers classify intuitively as words are composed of further parts

that cannot be ignored in linguistic analysis. A verb may be composed out of a verbal stem, a tense suffix, and an agreement cluster (*juoks-e-n* ‘ran-present-1sg’ v. *juoks-i-n* ‘ran-past-1sg’).

To understand the meaning of a complex word such as *juoksin* ‘run-past-1sg’, it is not necessary for the speaker to decompose the word consciously into its parts. Instead, the surface form /juoksin/ is mapped directly into its constituents ‘run’, ‘past tense’ and ‘1sg’ by an automatized table-lookup memory. This memory system will be called a *surface vocabulary* or *surface lexicon* in this study. It is a lexical storage medium that makes word recognition faster and facilitates language use. The surface lexicon and the information processing associated with the automatic retrieval of complex words must, however, make use of some ultimate, primitive building blocks of words. We can think of the surface lexicon as a fast mapping between surface forms and a linear ordering of its ultimate linguistic constituents (i.e. /juoksin/ → ⟨‘run’, ‘past’, ‘1sg’⟩). The (ultimate) *lexicon* holds the ultimate lexical elements. This lexicon will contain elements corresponding to the verbal stem ‘run’, tense information ‘present/past’, and an element corresponding to the first person singular inflection ‘1sg’ (or possibly their decomposition into separate number and person features, and possibly others). These elements have the property that they can no longer be decomposed lexically. They might receive a further decomposition in the semantic system, however.

The lexicon provides the syntactically primitive elements of language that have therefore no further structure generated by Merge. The elements in the lexicon have their own internal structure, however, minimally a *set* of lexical features. Most lexical items have *phonological features*, representing its form at the PF-interface, and *semantic features*, linking the word with meaning. The verbal stem ‘run’, for example, is connected to a specific phonological form /juokse-/ and to a semantic concept RUN that expresses the meaning ‘to run’ that we conceptualize and can link with conscious experience. Furthermore, the tense morpheme T has a property that allows it to occur in connection with verbs, not nouns, which means that the lexical item for *run* must also contain a lexical category feature V. Selection properties of lexical items must likewise be represented as features. When two elements are merged to form  $[\alpha, \beta]$ , the well-formedness of the operation depends on the lexical category features of the heads of the two elements  $\alpha$  and  $\beta$  (determined by (5)). For example, it is possible to combine a preposition P such as *kohti* ‘towards’ with a noun phrase NP, not with an infinitival verb phrase (19).

(19)

- a. kohti [Pariisi-a] P + NP  
towards Paris-par  
'towards Paris'
- b. \*kohti [Peka-n hakea heidä-t] P + SVO  
towards Pekka-gen to.pick they-acc

We therefore assume that the preposition *kohti* 'towards' comes with a *complement selection feature* allowing it to be merged with a NP (analysed later in this book as a DP); example (b) is ungrammatical because the expression violates the complement selection feature of the preposition. What the exhaustive, complete list of lexical selection features are is an empirical issue that we have to decide on the basis of empirical inquiry.

When a preposition *kohti* 'towards' is merged with a noun phrase, the noun phrase must be marked with the *partitive case*, as shown in (19)a. I will say that the preposition *assigns* the partitive case to its sister constituent. In some cases, the preposition (adposition) also phi-agrees with its argument. These two alternatives, one with the partitive case and another with the genitive case + phi-agreement, are again shown in (20).

(20)

- a. lähellä minu-a, lähellä sinu-a  
near I-par near you-par  
'near me' 'near you'
- b. minun lähellä-ni, sinun lähellä-si  
I.gen near-1sg you.gen near-2sg  
'near me' 'near you'

The phi-features '1sg' at the adposition *lähellä-ni* 'near-1sg' are due to the existence of the local genitive argument *minun* 'I.gen'. I say that the adposition *phi-agrees* with the argument NP. Whether and how a lexical item assigns a case to something else or phi-agrees with something else, or whether it receives a case

feature, must be represented in some manner in the corresponding lexical item. Thus compare the behavior of the adposition *lähellä* ‘near’ with that of *kohti* ‘towards’: the latter does not exhibit the behavior profile (b), only (a) (*kohti minu-a* ‘towards I-par’ vs. \**minu-n kohti-ni* ‘I-gen towards-1sg’). Thus, in addition to lexical category features (e.g., P, A, N), selection features, the lexical items must contain formal features representing its ‘morphosyntactic signature’. I will argue in Chapter 6 that we need two lexical features to cover these properties for Finnish:  $\pm$ ARG, which determines whether or not the lexical item comes with an unspecific/unvalued pronominal argument features (denoted as  $\varphi$ \_, D\_ in this book) and  $\pm$ VAL which determines whether or not the lexical element can value these feature by morphosyntactic operations. This process in which a lexical head assigns case to something and/or phi-agrees with an element is called *Agree*. An element that acquires new features and initiates the operation (here the preposition and the tensed verb) is called the *probe*, and the element that donates the features (D) the *goal* (Chomsky 2008). The operations and terms are illustrated in the example (21).

(21)

a.      *lähellä*            *minu-a*.

near      →      I-par

Partitive case assignment by Agree(P, NP)

b.      *minu-n*            *lähellä-ni*

I.gen      ↔      near-1sg

Genitive case assignment and phi-agreement by Agree(P, NP)

c.      *Minä*              *nuku-n*.

I.nom      ↔      sleep-1sg

Nominative case assignment and phi-agreement by Agree(T, NP)

d.      *Minä uskon nukkuva-ni hyvin*.

I.nom    believe to.sleep-1sg    well

Phi-agreement by Agree(infinitival, antecedent)

Agree and its underlying explanation is controversial. It is hoped that by systematically examining data from various languages, here Finnish, we could learn more than in order to guide linguists towards a more comprehensive and truthful explanation.

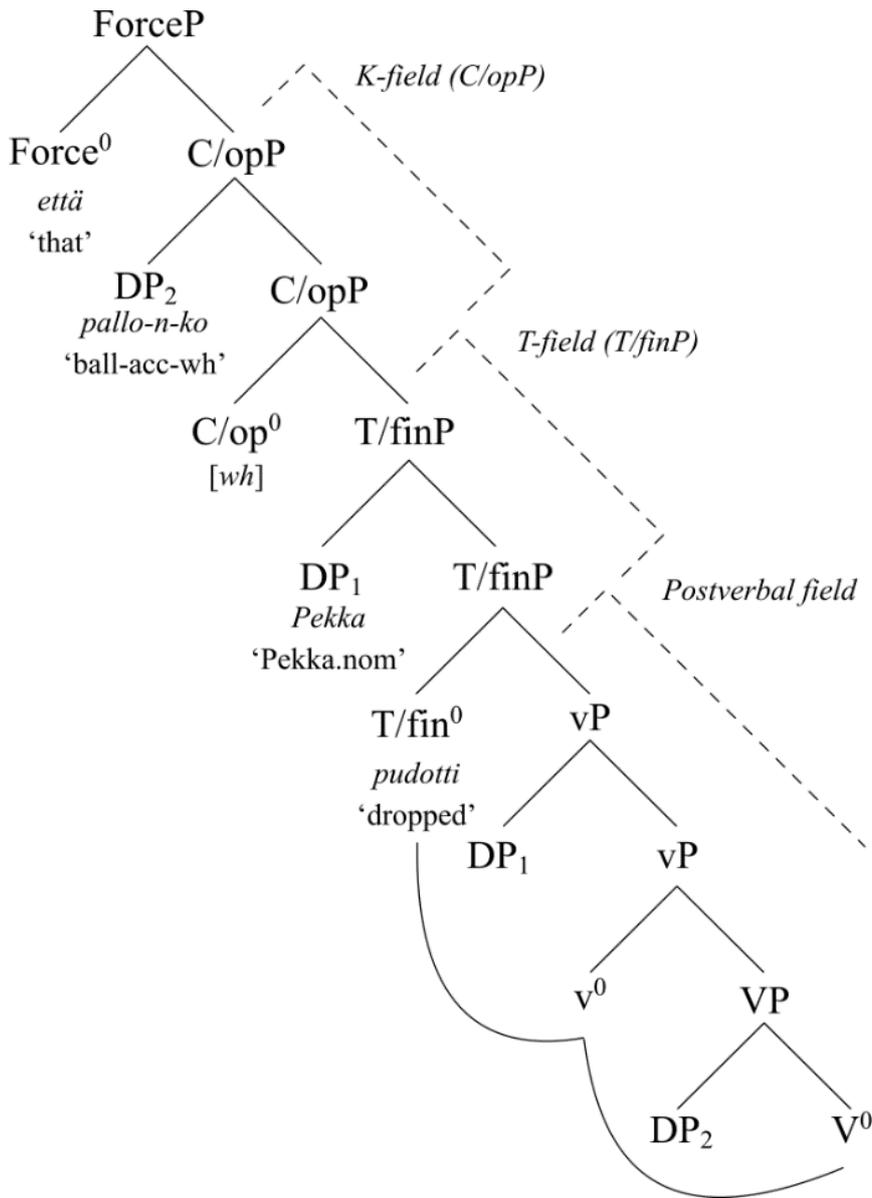
The fact that a preposition selects for a noun phrase complement applies to most prepositions. A theory in which the selection feature is added to each preposition individually fails to capture a generalization. The theory would be unnecessarily complex. The theory is also wrong because all new words in category L that are introduced to language are automatically regulated by many of the same principles as the existing words in the same category. Lexical items are therefore subject to *lexical redundancy rules*. Among the redundancy rules is a rule which says that any element with lexical category label P will have a feature determining that it selects for a noun phrase. Should a feature of a specific lexical entry and a lexical redundancy rule conflict, specific lexical information wins.

#### *1.2.5 An overview of Finnish clause structure*

(Vilkuna 1989, 1995) divided Finnish finite clause syntax into three fields she named the K-field, T-field and the postverbal field. This distinction has proven useful in analyzing Finnish and will be assumed in this book as a starting point (properties of the K-field are examined in Chapter 3, T-field in Chapter 4 and the postverbal field in Chapter 5). Here I provide a brief summary of their properties.

The simplest possible analysis of the Finnish finite clause that can arguably handle the main empirical patterns of Finnish is (22). This analysis constitutes the null hypothesis one should assume as a starting point when working with Finnish material. Possible extensions of this structure are discussed further below.

(22)



The postverbal field contains the verb phrase denoting the core propositional content of the sentence, representing who did what to whom, when, and in what manner. This layer is minimally followed by a finite tense T/fin corresponding to a bare finite verb. Finnish finite clause exhibits systematic EPP behavior, thus a subject or subject-like element must occur in the preverbal subject position. This condition is here captured by assuming that T/Fin has  $EPP_F$ , targeting some feature F. There is disagreement in the literature on the nature of F. The two main lines of research are that it is either a discourse feature, such as topic or non-focus (Holmberg and Nikanne 2002; Huhmarniemi 2019a; Vilkkuna 1989, 1995), or a formal feature (Brattico

2019d; Brattico and Huhmarniemi 2006; Vainikka 1989; Vainikka and Levy 1999) with the discourse effects arising as secondary effects. Almost any phrase with any label (subject, object, indirect object, adverbial) can move to Spec,T/finP in Finnish. The  $EPP_F$  can be checked also by an expletive, generated in situ. It is uncontroversial that whatever satisfies the  $EPP_F$  tends to constitute the topic of the clause. It is also uncontroversial that the Finnish EPP, although it does constitute a specifier requirement of some sort, is not identical to an English-type EPP that must be satisfied by the grammatical subject. A good starting point is to view the finite clause  $EPP_F$  through the lens of a “topic-criterion” of some type, but without making strong assumptions concerning the explanation (e.g., grammatical or syntactic). Finally, suffixes indicating conditional (*pudo-tta-isi* ‘drop-cond’) or potential (*pudotta-nee* ‘drop-pot’) mood are part of the feature content at T/fin and never co-occur with tense. Mood is not part of overtly tensed infinitivals (Section 5.3.3) but they do occur in relative clauses, so they are connected to finiteness.

Finnish negation constitutes its own head  $Neg^0$  between C/op and T/fin. When Neg is part of the clause, finiteness properties such as agreement and EPP shift from T into Neg. Tense remains at the lower main verb; EPP and full phi-agreement are at  $Neg^0$ . This has led to two main analyses of the Finnish T-field.

According one, going back to (Holmberg et al. 1993), there is a separate finiteness head  $Fin^0$  between C/op and T hosting either Neg or T, depending on the presence of Neg. This explains why finiteness (EPP, full phi-agreement, topic criterion) cluster to the highest head. Another possibility is that finiteness cluster to a head selected by C/op and are is inherited from C (Brattico and Huhmarniemi 2006). There is currently no clear evidence in favor or against either of these hypotheses, hence (22) can be assumed as the null hypothesis. The important point is that the finiteness diacritic “fin” marked on the tense node is not a lexical feature of T: it is either inherited from C/op or there is an additional finiteness projection into which T has raised.

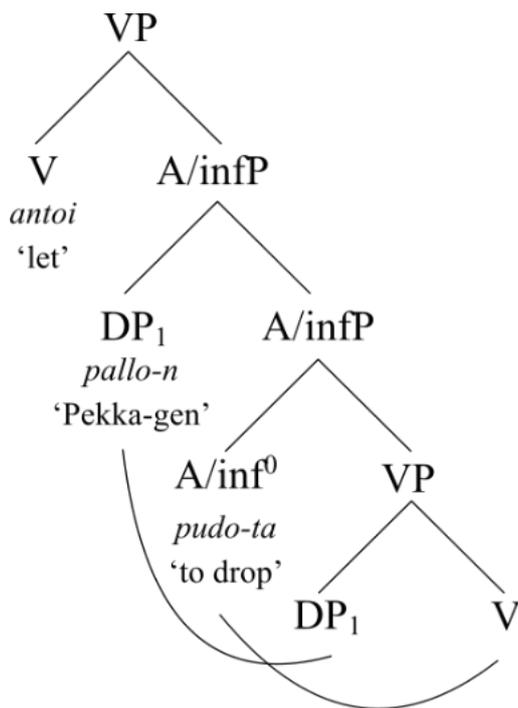
Some Finnish modals, illustrated in (23), have special properties that seem at first to require projection of a modal head.

- (23) Meidän            täytyy            pudottaa            pallo.  
        we.gen            must(3sg)        to.drop            ball  
        ‘We must drop the ball.’

The modal verb is finite, takes a genitive subject, possibly exhibits default third person agreement, exhibits past-present tense alteration, inflects for mood and selects for the A-infinitival, the latter which is discussed in Section 5.3.2. Yet none of these properties force us to assume that there is a special finite modal projection above T/fin. Instead, modals are like auxiliary verbs in the sense that they do not project independent thematic roles or event structures but like ordinary verbs in having special lexical content, occurring together with tensed auxiliary verbs (*Pekka-n oli täyty-nyt unohta-a avaimet autoon* ‘Pekka-gen be-past must.prtcpl forget.A keys to.car’). They are analysed as quasi-functional elements in the sense of (Cardinaletti and Shlonsky 2004) in this book (Section 5.7). In the sentence (23), then, we can assume that this modal lexical content is at T/fin and maintain the null hypothesis.

Deverbal infinitivals are created by combining the verb phrase introducing the argument structure with an *infinitival head* that then corresponds with an overt infinitival morpheme(s). The infinitival phrase, headed by an infinitival head that comes from a large repertoire of infinitivals, is merged to the structure either as an adjunct or as a selected argument. Infinitival heads behave like finite T in that they exhibit strict EPP behavior, often trigger full agreement with the thematic argument, may contain pro subjects, but their tense-aspect properties are impoverished in comparison to finite tense. Mood is absent. Due to the lack of Force-C/op layer most infinitivals are unable to articulate information structure. Most Finnish infinitivals behave like English finite clauses with respect to their EPP behavior, forcing closest argument to SPEC, hence I will argue that they exhibit  $EPP_F$ , F identified in this work with nominal features D and  $\varphi$ . A typical Finnish infinitival structure is illustrated in (24), which represents the postverbal field of the sentence *Pekka antoi [A/infP pallo-n pudota]* ‘Pekka let ball-gen to.drop’ containing an A-infinitival complement (A/infP). Few examples are provided in (25). The infinitival morpheme, corresponding to an infinitival head, is emphasized.

(24) General syntactic template for Finnish infinitivals



(25)

- a. Pekka antoi [pallo-n puto-ta.] (A-infinitival)  
 Pekka let ball-gen fall-A/inf  
 'Pekka let the ball to fall.'
- b. Pekka uskoi [pallo-n putoa-van.] (VA-infinitival)  
 Pekka believed ball-gen fall-VA/inf  
 'Pekka believed that the ball will fall.'
- c. Pekka heräsi [pallo-n puto-ttua.] (TUA-infinitival)  
 Pekka wake-up ball-gen fall-TUA/inf  
 'Pekka wake up after the ball drop.'

*C/op* introduces a position for a head *C/op*<sup>0</sup> or a phrase *Spec,C/opP*, the latter which is moved to the phrasal position by A-bar movement and only by A-bar movement (Huhmarniemi 2012). It hosts several types of elements, among them relative pronouns, interrogative pronouns and phrases; phrases hosting second position discourse clitics, such as *-han*, *-pa*, *-s*, and their combinations; contrastively focused phrases; corrective phrases and contrastive topic phrases; and imperative verbs. The reason such a wide variety of

phrases occur at the operator field is because there is no left peripheral cartography; instead, one syntactic position hosts all A-bar moved phrases, while the different semantic interpretations created by moving something to the operator field are generated by using suffixes and their combinations. This means that the syntax of Finnish left periphery is relatively simple, while the semantic repertoire, conveyed by means of morphology, is considerable. The standard analysis of these patterns is one in which the C functions as a probe (in the sense of Section 1.2.4), looks for features corresponding to the second position clitics and other features associated with phrases and triggers A-bar movement to Spec,C/opP. The phenomenon instantiates  $EPP_F$ , F being selected from the list of C-features (e.g., *wh*). It is also possible to move a head to C/op<sup>0</sup>, but phrasal movement is then barred. The head must contain one of the C-features. An interesting and currently little studied matter of Finnish grammar is the observation that it instantiates long head movement to C/op.<sup>3</sup>

In Vilkuna's original analysis the semantics of the K-field were captured in terms of "corrective focus," hence the explanation was rooted in discourse. The operator field is, indeed, often used to express corrective or contrastive focus. Vilkuna did notice, however, that corrective focus semantics does not exhaust the semantic interpretation associated with the operator field. Relative pronouns, for example, use the same position, but are not associated with discourse properties. The unifying factor is that anything that goes to the C/op field creates an operator-variable interpretation, in which variable abstraction is applied to some constituent. I will call the K-field as the *operator field* in this book. Looking at all the evidence we currently have on the Finnish operator field supports in my view an approach in which the operator-variable mechanism – the position and the associated grammatical mechanisms – are implemented by a different "module" that those involved with the T- and the postverbal fields.

The evidence for Vilkuna's two-position theory positing C/op-T/fin structure is compelling. Again, the null hypothesis can be maintained and is supported by a wealth of evidence. There are several issues that remain open. The exact semantic contribution of the discourse clitics such as -hAn and -pA remains to be worked out. A possible problem for the analysis are noncanonical clauses in which several phrases appear to occur between C/op and T/fin, as in (26).

---

<sup>3</sup> I have collected all the data I currently have on Finnish long head movement into an unpublished manuscript (Brattico 2020d).

- (26) Milloin Pekka auton Merjalta osti?  
 when Pekka car-acc from.Merja bought  
 C/op T/fin  
 ‘When did Pekka buy the car from Merja?’

T/fin may contain an auxiliary, while C/op can be filled in independently, thus this sentence hosts three arguments in the T-field, or at the very least between C/op and T/fin. The sentence is noncanonical and presupposes that the three arguments are topics. I will return to this phenomenon later in this book, where I argue this is a consequence of an independent property of Finnish, its free word order signature. There is no evidence suggesting that these elements are hosted by a rigid cartography of heads between C/op and T/fin; they behave syntactically like adjuncts or adverbials. If so, we can maintain the null hypothesis.

The operator-variable system relies on internal snowball movement discovered by (Huhmarniemi 2012). When the C-probe targets features inside a phrase that it also pied-pipes to Spec,C/opP, a word containing the goal feature must occur at the left edge of the pied-piped phrase. If the element carrying the feature is not merged initially to the edge position, it is moved there. The rule applies iteratively, creating cascades of internal movement (27).

- (27) [[[Mitä kaupunki-a<sub>1</sub> kohti <sub>-1</sub>]<sub>2</sub> virtaamalla <sub>-2</sub>]<sub>3</sub> Seine saavuttaa valtameren <sub>-3</sub>?  
 what city-par towards by.flowing Seine reaches ocean  
 ‘By flowing towards which city does Seine reach the ocean?’

The internal movement satisfies properties of A-bar/operator movement (Huhmarniemi and Brattico 2013a). In the latter paper, we assumed that pied-piped phrases must be headed by an EPP<sub>C/op</sub> features, C/op being a probe feature in the C/op-system (*rel*, *wh*, etc). EPP<sub>C/op</sub> must be a formal trigger, as internal *wh*-movement does not create secondary operator-variable constructions. The same assumption was argued for in (Brattico and Chesi 2020), which presents a fully formal, computational theory of operator-variable constructions and pied-piping, including Finnish.

Finnish morphosyntax is particularly contested area due its complex properties, but some of the simpler areas can be captured straightforwardly. This can be done by assuming, following (Chomsky 2001), that structural

cases are assigned under probe-goal by the core functional heads T/fin (nominative), H/inf (genitive) and v (accusative, partitive). To this we add Vainikka's partitive generalization, according to which the partitive constitutes a default complement case in Finnish (Vainikka 1993, 2003) and is, therefore, assigned by several functional heads (e.g., v, P, A, N) to its sister under probe-goal. Vainikka has argued that the genitive constitutes a specifier case in Finnish, which can be captured by assuming the above probe-goal mechanism combined with  $EPP_{D,\varphi}$ , which will bring genitive arguments to local specifier positions. Some residuum problems remains, discussed in (Vainikka 2011).

$Force^0-C/op^0-T/fin^0-VP$  constitutes a minimal structural analysis of Finnish finite clause that can, I believe, capture all the data;  $H/inf^0-vP$  for infinitivals. The literature contains several analyses positing more heads, such as  $Fin^0$  (finiteness head above TP, (Holmberg et al. 1993; Holmberg and Nikanne 2002; Mitchell 1991)), modal heads or two (Thomas 2012), Polarity (Culicover 1991; Laka 1990; Mitchell 1991; Pollock 1989),  $Asp^0$  (aspect head inside the VP, (Megerdoomian 2000)),  $Mood^0$  (mood below or above T),  $Pass^0$  (impersonal passive head), QP (Thomas 2003) and  $Agr^0$  (agreement head). (Thomas 2003) proposes to split the CP-layer into a complex cartography proposed by (Rizzi 1997). The postverbal field could contain category neutral roots and functional material corresponding to the major lexical categories (Brattico 2005; Brattico and Leinonen 2009; Pylkkänen 2002). The alternative is that the semantic and syntactic effects of these heads result from features residing inside the core functional heads. For example, we could assume that  $Agr^0$  constitutes not a head but uninterpretable  $\varphi$ -set at T, or that the aspectual properties of the  $Asp^0$  head are in fact features of v-V. Critical data is hard to find. A separate head could be justified based on an extra specifier position, extra head position, word order, or scope ambiguities. Adverbial order is difficult to use in Finnish due to relatively free word order, and there is to my knowledge no compelling demonstration for the extra specifiers or head positions generated by the extra heads. Scope ambiguities can support syntactic analyses but only if such phenomena are ultimately syntactic in nature, not an uncontroversial assumption. I assume the null hypothesis and will maintain it until compelling evidence emerges for more complex structure that cannot be explained by assuming the feature theory. This analysis should be considered the minimal null hypothesis; there is no convincing evidence against the extra heads either.

## 2 Overview

### 2.1 Finnish as a Finno-Ugric language

Finnish belongs to the Finno-Ugric group of the Uralic languages. It is spoken by approximately 5,5 million people living the Finnish peninsula, where it constitutes one of the national languages of Finland (the other is Swedish). Finnish is related to two larger Finno-Ugric languages, Estonian (1,1 million speakers) and Hungarian (14,5 million), both which also function as national languages. Of these, Estonian is closely related to Finnish, as they both belong to the larger Baltic-Finnish languages including languages such as Karelian, Veps, Ingrian, Votian, Ludian and Livonian. Other Finno-Ugric languages include Mari, Mordvin, Komi, Udmurt, Saami, Khanty and Mansi. Finno-Ugric languages are thought of being derived from an ancestral Proto-Uralic language, spoken about 5,000 years ago in what is today western Russia and Ukraine.

### 2.2 Finnish in a crosslinguistic perspective

Finnish can be characterized as a configurational SVO language that nevertheless exhibits certain ‘free word order’ traits due to the employment of a wide variety of noncanonical word orders. These noncanonical words orders correlate with discourse properties such as topic and focus, a fact that has led several authors to characterize Finnish as ‘discourse configurational’ language, a language in which word order articulates discourse properties rather than grammatical roles. Example (28)a represents a canonical Finnish finite clause, with (b-d) illustrating some of the possible word order variations and the discourse properties correlated with these variations. Propositional meaning remains the same, and each clause constitutes an active, declarative finite clause. Morphosyntactic properties, likewise, remain unaltered between the examples.

(28)

a. Jari antoi kirja-n Merja-lle. (S-V-O-IO)

Jari.nom          gave.3sg          book-acc          Merja-all

‘Jari gave a book to Merja.’

b. Kirja-n          Jari          antoi          Merja-lle          (O-S-V-IO)

book-acc          Jari.nom          gave.3sg          Merja-all

‘It was the book (and not the magazine) that Jari gave to Merja.’

c. Kirja-n          antoi          Merja-lle          Jari.          (O-V-IO-S)

book-acc          gave.3sg          Merja-all          Jari.nom

‘It was Jari who gave a book to Merja.’

d. Kirja-n          antoi          Jari          Merja-lle.          (O-V-S-IO)

book-acc          gave.3sg          Jari.nom          Merja-all

‘When it comes to the book, it was to Merja that Jari gave it.’

The translations approximate the meanings underlying these clauses. They should give the reader some idea of how word order in Finnish affects the articulation of the sentence’s information structure.

The free word order property is related to the fact that Finnish morphosyntax is rich both in terms of the number of morphosyntactic devices it uses and in the range of constructions in which such devices are applied. Finnish has fifteen case suffixes and it exhibits productive and systematic phi-agreement (in person and number) in both finite and several non-finite domains. Thus, not only finite verbs but also noun heads, prepositions, adverbs and non-finite verbs exhibit full phi-agreement with a local argument. Furthermore, case suffixes are attached to words, not phrases, hence Finnish constitutes a prototypical head-marking language, often marking every word inside the noun phrase. It is generally true that when an argument is moved to a noncanonical position, its thematic role is reconstructed from overt case suffixes and/or agreement patterns. Similarly, while word order is flexible, case assignment and phi-agreement are not: changes in case assignment or phi-agreement result in ungrammaticality, whereas word order changes only result in degrees of marginality. Finite agreement is illustrated in (28); example (29) shows few examples of infinitival agreement. The infinitival agreement morpheme is called “possessive suffix” in the traditional Finnish scholarship, but it has no necessary relationship to possession (see, in particular, (Kanerva 1987) and (Nelson 1998, §5)).

(29)

- a. Jari<sub>1</sub> väitti pro<sub>1</sub> ihaile-va-nsa Merja-a.  
Jari.nom claimed admire-VA/inf-3sg Merja-par  
'Jari claimed to admire Merja.'
- b. Jari<sub>1</sub> rikkoi pro<sub>1</sub> auto-nsa./ Jari rikkoi (minun) auto-ni.  
Jari.nom broke car-3sg Jari.nom broke (my) car-1sg  
'Jari broke his car.' 'Jari broke my car.'
- c. Jari asui (minun) lähellä-ni.  
Jari.nom lived (my) near-1sg  
'Jari lived near me.'

The combination of relatively free word order with rich morphosyntax has led some authors to view Finnish from a nonconfigurational perspective (e.g., Sammallahti 2002, 2003; Välimaa-Blum 1988; Vilkuna 1989). Hungarian, a distantly related Finno-Ugric language, has been argued to be partially nonconfigurational, having no phrase structure in its postverbal field (Kiss 1987). There is no independent evidence, however, for radical nonconfigurationality. Binding, adverb scope, morphosyntax, A-bar movement, control, quantifier scope, ordering between grammatical heads, ellipsis and virtually all infinitival constructions behave as if Finnish would be fully configurational (Brattico 2019e; Manninen 2003b: §1.4; van Steenbergen 1989). To my knowledge all generative approaches to Finnish view Finnish as configurational, and the same is true of this work as well. The free word order property, which must still be accounted for, will be analysed in this work as a form of argument adjunction, similar to that previously proposed by (Baker 1996; Chomsky 1995: 4.7.3; Jelinek 1984) for other languages.

Finnish morphology and morphosyntax is productive and agglutinative, meaning that many verbal and nominal words are composed out of several distinct and easily distinguishable morphemes, each contributing some semantic property. Words can therefore become complex, in fact, derivational morphology is recursive, although seldom used recursively. Derivational morphology is followed by functional layer responsible for lexical categorization and the concomitant introduction of syntactic arguments. Word formation can then continue by further addition of verbal material encoding impersonal passivization, mood,

conditionals and tense, followed by inflectional suffixes and finite forms and, finally, left peripheral second position clitics.

Rich phi-agreement is connected with the fact that Finnish is a pro-drop language: subject pronouns can remain phonologically null in the presence of agreement. Finnish pro-drop profile has three special properties. First, Finnish is not an VSO language, and most verb-initial constructions are ungrammatical (Vainikka 1989; Vainikka and Levy 1999; Vilkkuna 1989). In fact, there is a productive expletive strategy that is used to avoid verb-initial clauses (30) (Holmberg 2005; Holmberg and Nikanne 2002). In this Finnish resembles languages like English and departs from the typical pro-drop profile.

(30) Sitä on Jari taas valvonut koko yön.  
expl has Jari.nom again stayed.awake all night  
'Jari has again stayed awake whole night.'

The second interesting aspect is that the pro-drop phenomenon extends to all those infinitival constructions that exhibit productive and overt phi-agreement (Brattico and Huhmarniemi 2016; Huhmarniemi and Brattico 2015). Hence also adpositions, several non-finite verbs and several adverbs project a subject-specifier position that can be filled in by a null pronoun *pro* (see (29)). And third, Finnish belongs to the group of partial pro-drop languages: the third person null pronoun cannot be dropped freely, in opposition to a language such as Italian, but requires the presence of an (often c-commanding) antecedent (Vainikka and Levy 1999). Thus, the third person null pronoun is perhaps not fully 'pronominal' but behaves like a variable (31).

(31)

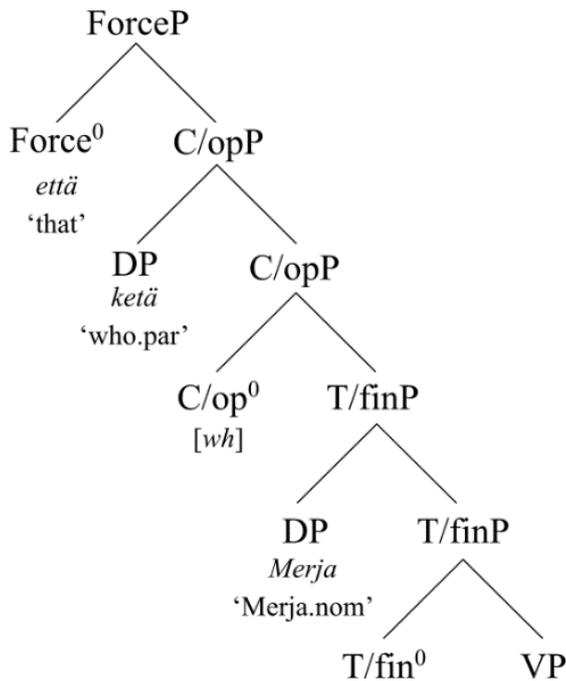
a. \*Ei ihaile Merja-a. En ihaile Merja-a.  
not.3sg admire Merja-par not.1sg admire Merja-par  
Intended: 'He does not admire Merja.' 'I don't admire Merja.'

b. Jari<sub>1</sub> väitti että [pro<sub>1</sub> ei ihaile Merja-a.]  
Jari claimed that not.3sg admire Merja-par  
'Jari claimed that he (=Jari) does not admire Merja.'

The syntax of Finnish left periphery or CP-cartography is impoverished in terms of syntactic positions, but rich in terms of the grammatical features that it can express by means of separate morphemes. Between the highest element in the Finnish finite clause, the high complementizer, and the preverbal subject, there exists only one syntactic position that is filled in (and filled in only) by means of A-bar/operator movement. Thus, interrogative pronouns, relative pronouns, contrastively focused or topicalized phrases, and phrases suffixed with discourse-motivated second position clitics target the same position sandwiched between the high complementizer and the preverbal subject. Some of the features characterizing these items can be combined, for example, it is possible to combine an interrogative pronoun with a second position clitic -hAn (as in *kuka-han siellä on?* ‘who-hAn there is’). If a phrase is not moved to the C/operator position, then a head can be moved there by standard T-to-C movement. I will call this position as C/operator or C/op position in this book; the term will be motivated later. The high complementizer, C/op and the preverbal subject field provide that the basic structure of the Finnish finite clause is ‘that + Spec + C/op + Spec + T/fin’, with the lower Spec constituting the preverbal subject position. Example (32) shows a sentence in which each field is filled in. Notice that the interrogative pronoun can co-occur with the high complementizer.

- (32) Jari kysyi että ketä<sub>1</sub> Merja ihailee \_\_<sub>1</sub>.  
 Jari asked that who.par Merja.nom admires  
 ‘Jari asked who Merja admires?’

(33)



In short, the left peripheral cartography is simple, but the morphological features associated with it more complex, as would perhaps be expected from a language that is agglutinative.

One special property of Finnish is that fact that in a negated clause, the negation agrees with the grammatical subject and exhibits the EPP profile, requiring that a phrase occur in its specifier (thus, \*Neg-X orders are ungrammatical unless Neg occurs at C/op). Finnish negation is therefore like a ‘tenseless auxiliary’ that exhibits finiteness but no tense. A negative sentence is illustrated in (34).

(34) Jari            e-i    ihaile   Merja-a.  
      Jari.nom       not-3sg admire Merja-par  
      ‘Jari does not admire Merja.’

Because the negation constitutes a finite element, it is unable to appear in non-finite contexts. The tensed auxiliaries and/or verb occur in the presence of the negation in lower positions and often exhibit partial phi-agreement in number. It follows from this that properties of finiteness, phi-agreement, grammatical subject assignment and finite pro-drop are not properties of tense: they emerge to a head selected by C/op.

Finnish structural case assignment is for the most part relatively straightforward matter. The grammatical subject of a finite clause occurs in the nominative (35)a, whereas the thematic subject of most infinitival constructions is in the genitive (35)b (when the argument is overt).

(35)

- a.      Jari                ihaile-e            Merja-a.  
           Jari.nom          admire-3sg        Merja-par  
           'Jari admires Merja.'
- b.      Jari                käski              Merja-n            lähteä.  
           Jari.nom          ordered.3sg      Merja-gen        to.leave  
           'Jari ordered Merja to leave.'

The nominative and the genitive therefore constitute the default subject-specifier cases in Finnish (Vainikka 1989, 1993). In a finite clause, only the nominative argument phi-agrees with the finite element: thematic agents that are in any other cases, such as in the genitive or partitive, do not. When the grammatical subject is missing, the finite element is either agreementless or exhibits third person default agreement. Furthermore, there are no situations in which the thematic object would be promoted to the subject and would agree with the finite element.

The partitive case is assigned by default when an argument occurs in a complement position of some head and does not engage in agreement with that head. For example, many Finnish prepositions assigns the partitive case to their complement (*lähellä minu-a* 'near I-par'). The partitive is also assigned to arguments in the direct object position (*Jari ihailee Merja-a* 'Jari.nom admires.3sg Merja-par', ex. (35) above). The Finnish partitive case expresses a 'partitive part of relation' only in exceptional cases, thus it constitutes a structural complement case in Finnish (Vainikka 1989, 1993, 2003).

Direct object case assignment as a whole is much less straightforward. There are four direct object cases. The partitive complement case alternate with various accusative forms, and this alteration encodes aspect among other associated semantic notions discussed later: when the direct object is in the partitive, the clause is interpreted as being aspectually unbounded and/or incomplete; when an accusative case (again, any

accusative form) occurs, it is interpreted as being aspectually bounded and complete. The system is illustrated in example (36).

(36)

- a. Jari pesi hevos-ta.  
Jari.nom washed horse-par  
'Jari washed the horse (but did not necessarily complete the washing).'
- b. Jari pesi hevos-en.  
Jari.nom washed horse-acc  
'Jari washed the horse (and completed the task).'

When the direct object is in the plural, the accusative further encodes quantificational determinacy: *Pekka pesi hevoset* 'Pekka washed horses.acc' means that Pekka washed all horses from a given, determinate set, whereas the use of the partitive means that only some horses were washed, but not necessarily them all.

When the clause is interpreted as being aspectually bounded (all else being equal), one of the three accusative forms are used. These forms are the *t-accusative*, assigned to pronouns and full DPs in plural, the *nominative-looking accusative* (acc/nom), assigned to singular full DPs under certain conditions, and the *genitive-looking accusative* (acc/gen), assigned to singular full DPs under certain condition. The inanimate pronoun *se* 'it' inflects like full DPs. These forms are illustrated in (37).

(37)

- a. Me löysi-mme hän-et. (t-accusative)  
we.nom found-1pl he-acc(t)  
'We found him.'
- b. Me löysi-mme avaimet. (t-accusative)  
we.nom found-1pl key-acc(t)  
'We found the keys.'
- c. Me löysi-mme avaimet. (genitive-looking accusative)  
we.nom found-1pl key-acc/gen

‘We found the key.’

d. Me löyde-ttiin avain. (nominative-looking accusative)  
we.nom found.impass key.nom

‘We found the key.’

e. Me löysi-mme se-n./ Me löyde-ttiin se.  
we.nom found-1pl it-acc/gen we.nom find-past.impass it.nom

‘We found it.’

‘We found it.’

The paradigm has been subject to much controversy, with several different approaches in the current literature. The problem is not made easier by the fact that the case assignment paradigm illustrated above exhibits nonlocal effects, in which the form of the direct object case depends on the grammatical properties of one or even several nonlocal heads. One example is provided by the negation. It assigns the partitive case to all direct objects within its scope, here to the direct object inside a non-finite complement clause (38)a-b.

(38)

a. Jari kannusti Merjaa voittamaan kilpailu-n/ \*kilpailu-a.  
Jari encourage Merja to.win competition-acc competition-par  
‘Jari encouraged Merja to win the competition.’

b. Jari e-i kannustanut Merja-a voittamaan \*?kilpalu-n kilpailu-a.  
Jari not-3sg encourage Merja-par to.win competition-acc competition-par  
‘Jari did not encourage Merja to win the competition.’

Finnish has no grammaticalized determiners or articles. The indefinite-definite distinction is expressed by means of demonstratives, demonstrative-like elements, word order, and by means of context. Two demonstrative-like elements in particular behave like determiners: the plural numeral expressing ‘one’ and the inanimate pronoun/demonstrative *se* ‘it’. The former expresses indefiniteness (39)a, the latter definiteness (b).

(39)

a. Ostin yhde-t suka-t.

bought.1sg one-pl sock-pl

‘I bought (some indefinite) socks.’

b. Ota se sukka mukaan.

take that/it sock.nom with

‘Take that sock with you.’

It is hard to establish conclusively if these (and other demonstrative-like elements) represent the category of D or Dem. If they are D-elements, their occurrence is optional. With respect to the ordering of other nominal elements within the Finnish noun phrase, the D/Dem layer is followed by quantifiers (Q), numerals (Num) and the noun head (N), in this order, with both adjectives and possessive DPs being located more freely (40)a. Both adjectives and the possessive DPs are adjoined to the projectional spine of the DP, hence their order is free and occurrence is optional. The postnominal position, however, is reserved for PPs and partitive complements, the latter under certain special constructions: neither adjectives nor the possessive DPs can occur postnominally (b).

(40)

a. ne kaikki Cruisen kolme täyspitkää elokuvaa salaisesta agentista

those all Cruise.gen three full.length movies of.secret agent

Dem Q DP.gen Num AP N PP

‘All those three full length movies by Cruise about a secret agent’

b. \*ne kaikki kolme täyspitkää elokuvaa Cruisen

those all three full.length movies Cruise.gen

Furthermore, Finnish DP can host two genitive possessive DPs, in which case the higher DP expresses the ‘agent’ while the lower the ‘patient’ (41). When they occur in this configuration, their order is not free with respect to each other.

(41) Isä-n syntymäpäivi-en juhliminen kesti pitkään.

father-gen birthday-gen celebration lasted long

‘The celebration of the birthday by the father lasted long.’

Finnish surface word order cannot be said to be head-initial or head-final. The best way to capture Finnish surface word order is to say that it is canonically ‘head second’. This is due to the extensive use of the EPP condition, which states that most heads must be preceded by a phrase of some type. Head second configurations are generated from an underlying head-initial structure by moving or inserting something to the prehead position. We know this from several facts, among them the observation that often it does not matter much which element is fronted (and indeed sometimes it does not have to move). In addition, the fronting operation is often not linked or even correlated with specific discourse interpretation. On the other hand, it does not follow that everything must be head second: this is only a tendency (furthermore, Finnish is also not a V2-language, because, although it is often required that one element be fronted, the operation is not limited to one element). Examples of the head second phenomenon are shown in (42). In each example, the expression becomes grammatical if one of the post head elements occurs in the prehead position. Example (43) illustrates an expression in which several phrases are fronted, creating a head-final configuration that is nevertheless noncanonical and associated with a specific discourse interpretation.

(42)

- |    |               |                  |                               |           |
|----|---------------|------------------|-------------------------------|-----------|
| a. | *Ihaile-e     | Pekka            | Merja-a.                      | (V-X)     |
|    | admire-3sg    | Pekka.nom        | Merja-par                     |           |
| b. | *Pekka käski  | ___ <sub>i</sub> | lähteä Merja-n <sub>i</sub> . | (V/inf-X) |
|    | Pekka ordered |                  | to.leave Merja-gen            |           |
| c. | *lähellä      | minun            |                               | (P-DP)    |
|    | near          | I.gen            |                               |           |
| d. | *auto         | Peka-n           |                               | (N-DP)    |
|    | car           | Pekka-gen        |                               |           |

- (43) Milloin Pekka kirja-n Merjalle antoi!?
- when Pekka.nom book-acc to.Merja gave
- ‘When did Pekka give the book to Merja!?’

The head second property surfaces in surprising contexts. As shown by (Huhmarniemi 2012), formation of an interrogative clause triggers a cascade of internal *wh*-movement operations that satisfy the head second property by moving the *wh*-element to the left edge of its containing phrase, thus leaving the head for the second position (44).

- (44) [[Mitä kaupunki-a<sub>1</sub> kohti \_<sub>1</sub>]<sub>2</sub> virtaamalla \_<sub>2</sub>]<sub>3</sub> Seine saavuttaa valtameren \_<sub>3</sub>?  
 which city-par towards by.flowing Seine reaches ocean-acc  
 'By flowing towards which city does Seine reach the ocean?'

Leaving any of the internal *wh*-movement steps undone in an example such as (44) would make the sentence unable to function as a regular interrogative, making only the echo-question possible. This behavior, like the one illustrated by the previous example (43), exhibits the systematic feature of Finnish that often requires the specifier position of a grammatical head to be filled. The behavior is connected to the observation that many grammatical heads are 'morphosyntactically active' and exhibit also phi-agreement (Brattico, Chesi, and Suranyi 2019). Thus, both the preposition and the adverbial in the example (44) belong to a lexical class in Finnish that exhibits phi-agreement. It is not related to the fact that some phrases are islands to *wh*-movement: the preposition phrase, when it occurs alone, does not constitute an island, yet *wh*-movement to its edge is still required.