

Editorial

The editors are pleased to welcome you to the fifth volume of FULL, an open-access international journal that is meant to provide a platform for linguistic research on modern and older Finno-Ugric or other Uralic languages and dialects. FULL publishes comparative research, as well as research on single languages; with comparison of just Uralic languages or comparison across family lines; soliciting both formal linguistic accounts and empirically oriented contributions.

Issue 1 of the present volume is comprised by a selection of papers from a workshop on the Syntactic Structure of Uralic Languages, organized as part of the 12th International Congress for Finno-Ugric Studies (17–21 August 2015, Oulu, Finland). The theme of this special issue is the morphosyntax of agreement and case, with a special focus on Finnish and Estonian.

The first article in the collection, ‘Unifying subject agreement across clause types in Estonian’ by Mark Norris, looks at the realization of agreement in negated and non-negated indicative and imperative clauses in Estonian. Arguing in favour of a unified syntax for agreement across these different clause types, centering around a single functional head element in clause structure, Norris provides support for a general theory of agreement according to which morphological agreement and the syntactic Agree(ment) relation do not track each other directly. From a typological perspective, a repercussion of the paper is that the syntax of agreement in Estonian need not be radically different from its neighbours.

The article by Phil Crone, titled ‘Finnish first conjunct agreement and the direction of Agree’, also addresses the syntax of agreement, this time, based on data from Finnish, in particular, Finnish first conjunct agreement. In first conjunct agreement, a cross-linguistically widely attested agreement pattern, some agreement-bearing element stands in agreement with the first conjunct of a coordinated nominal expression, rather than with the full conjoined nominal phrase. Showing that none of the previous analyses are able to adequately explain the Finnish data, a new analysis is proposed, which is crucially based on the notion that the syntactic Agree(ment) process operates not only “downward” within constituent structure, as traditionally assumed, but also “upward.” In addition to bearing on general theoretical questions related to the directionality and timing of syntactic Agree(ment), Crone’s paper demonstrates that an analysis drawing on bidirectional Agree(ment) is a promising candidate for the explanation of first conjunct agreement cross-linguistically.

Saara Huhmarniemi and Merilin Miljan’s article titled ‘The partitive split in Finnish and Estonian’ is a study of discontinuous noun phrases in Finnish and Estonian. Both languages have several types of constructions in which a noun is separated from the modifying quantifier or numeral. The paper focuses on what is referred to as the partitive split, a construction in which the nominal appears in partitive case, and the two parts of the split nominal phrase exhibit morphological mismatch. It is proposed that partitive split is derived by a syntactic displacement of a subpart of what is originally a single nominal phrase. Unearthing intricate differences between Finnish and Estonian, the authors consider two alternative syntactic derivations. The first analysis is based on a classifier head which facilitates the selectional requirements of the quantifier/numeral

and which directly accounts for the apparent morphological mismatch. This analysis is shown to be preferable for the Finnish data, but it is only partly supported in Estonian. For Estonian a second analysis is contemplated, based on morphological repair.

Issue 1 of Vol. 5 concludes with a report by Tommi A. Pirinen, Eszter Simon, Francis M. Tyers, and Veronika Vincze on the Second International Workshop on Computational Linguistics for Uralic Languages (SIWCLUL), held in Szeged in January earlier this year. The aims of the SIWCLUL conference series, meeting with the general goals of FULL, include increased co-operation between the researchers, universities and research centres working on Uralic languages. As the One of the specific objectives is to avoid unnecessary duplicated work in the field of Uralistics by establishing connections and interoperability standards between researchers and research groups working at different sites. This year the conference also marked a start of an Association for Computational Linguistics' Special Interest Group for Uralic Languages (ACL SIGUR).

We take this opportunity to thank the anonymous reviewers who generously lent their time and expertise to FULL.

Our publications can be freely accessed and downloaded without any need for prior registration. At the same time, those who register, or have already registered, are provided with the benefit of getting notified of new issues, calls, etc. via the occasional email.

FULL welcomes manuscripts from all the main branches of linguistics, including phonology, morphology, syntax, semantics and pragmatics, employing a diachronic or synchronic perspective, as well as from first language acquisition and psycholinguistics. Whatever the theoretical or empirical orientation of the contributions may be, our leading principle is to maintain the highest international standards.

The Editors

Unifying subject agreement across clause types in Estonian^{*}

Mark Norris

Estonian negated indicative clauses show no agreement, whereas Estonian negated imperative clauses show agreement twice: once on the main verb and once on the negation word *ära*. This contrasts with affirmative clauses, where agreement appears only once. I propose a unified syntax for agreement across these clausal types, arguing that there is one head which bears a ϕ -feature probe in all Estonian sentences. There is no agreement in negated indicatives because this head has only one suitable vocabulary item in this context: *ei*. Doubled agreement arises due to a rule of post-syntactic Feature Copying in imperative contexts. I argue that this analysis is superior to an analysis making use of multiple ϕ -feature probes in the syntax, as such analyses struggle to account for the optionality of doubling in first-person plural contexts. The proposed analysis makes predictions about the kinds of marking possible in negated imperatives, which appear to be borne out in related Uralic languages. This investigation supports a view of the morphosyntax of agreement whereby the syntax and morphology of agreement overlap but do not coincide.

Keywords: *agreement, Estonian, imperatives, negation*

1 Introduction

In the canonical case of subject-verb agreement, there is one instance of subject agreement in person and number (hereafter: ϕ -features) on the finite verb. This is shown for Estonian in (1) and (2).¹

^{*} I started thinking about these patterns in collaboration with Anie Thompson in 2013, and her contributions to the project are gratefully acknowledged. Thanks are also due to Claire Halpert, Jorge Hankamer, Boris Harizanov, Ruth Kramer, Ethan Poole, Virve Vihman, and audiences at UCSC's Morphology Reading Group, the 88th Annual Meeting of the Linguistic Society of America, and the Workshop on Syntactic Structures of Uralic Languages at the XII International Congress for Finno-Ugric Studies in Oulu, Finland. Thanks to András Bárány for technical support. Thanks to two anonymous reviewers, whose critique and suggestions have improved this paper in terms of argumentation and rhetorical flow. Finally, I thank the following speakers of Estonian for discussing their language with me: Leelo Kask, Kärt Lazić, and Katrin Jänese. I am responsible for any remaining errors.

¹ Glossing abbreviations are as follows: 1 first person, 2 second person, 3 third person, ACC accusative case, ADE adessive case, CNG connegative, COND conditional, DA.INF *da*-infinitive, DU dual number, FRQ frequentative, HOR hortative, IMP imperative, INE inessive case, NEG negation, PAR partitive case, PASS passive/impersonal, PCPL participle, PL plural number, PST past tense, SG singular number.

Some examples come from online resources for the Estonian language. The first (BALANCED) is a balanced literary corpus containing equal parts fiction, journalism, and academic writing. The second (EKSS) is an online dictionary of standard Estonian (*Eesti keele seletav sõnaraamat*). Both are available online at <http://www.keeleveeb.ee/>.

- (1) *Sa vaata-d filmi.*
 you watch-2SG movie.PAR
 ‘You are watching a movie.’
- (2) *Te vaata-te filmi.*
 you.PL watch-2PL movie.PAR
 ‘You (pl/formal) are watching a movie.’

In (1), the verb bears the suffix *-d* indicating agreement with a second-person singular subject, and in (2) the verb bears *-te*, indicating agreement with a second-person plural subject. Under standard Minimalist conceptions, this kind of agreement is formalized as an Agree relation (between the subject and the verb, loosely speaking) in the syntax correlating with one agreement marker in the morphology.

However, this one-to-one relationship between syntactic agreement relation and morphological exponence of agreement does not always obtain in Estonian. In addition to contexts where there is one morphological exponent of agreement, there are situations in Estonian where there is no agreement and situations where agreement is doubled. These complex patterns require a closer look at the relationship between the syntax and morphology of agreement in the language. In this paper, I investigate and analyze the agreement patterns in negated imperatives, which show doubling, and negated indicatives, where agreement disappears, and I argue that the most successful analysis is one in which syntactic agreement (i.e., the Agree relation) and morphological agreement (i.e., the presence of agreement markers) overlap but are not isomorphic. Under the analysis I propose, the differing morphological situations disguise a system that is syntactically uniform. Before getting to the analysis, I present the morphological patterns in more detail.

Like indicative clauses, affirmative imperative clauses also exhibit one instance of subject-verb agreement. The paradigm of agreement markers for imperatives is different than the paradigm used in indicatives, as is visible in (3) and (4).

- (3) *Vaata filmi!*
 watch.IMP.2SG movie.PAR
 ‘Watch a movie!’
- (4) *Vaada-ke filmi!*
 watch-IMP.2PL movie.PAR
 ‘Watch a movie!’

The second-person singular imperative (in (3)) has no ending (I represent this in the table as a null morpheme $-\emptyset$), and the second-person plural (in (4)) uses the morpheme *-ge/ke*.²

Full paradigms for indicative and imperative clauses are presented in Table 1 on the following page.³ Examples such as affirmative clauses in Estonian are the canonical case, a one-to-one correspondence between syntactic and morphological agreement.

The situations where the agreement patterns diverge are found in negated clauses. Negated indicative clauses exhibit no morphological agreement. In the present tense, the main verb appears in a form that resembles an inflected stem without the inflection. This

² The choice between *-ge* and *-ke* is part of Estonian’s complex morphophonological system of stem alternation known as *gradation*. It is largely irrelevant to the issues I address here, so I will not discuss it

INDICATIVE			IMPERATIVE		
	SG	PL		SG	PL
1	-n	-me	1	-gu/ku	-me
2	-d	-te			-gem/kem
3	-b/-s	-vad/-d	2	-∅	-ge/ke
			3	-gu/ku	-gu/ku

Table 1: *Estonian agreement paradigms: indicative (left) and imperative (right)*

form is sometimes called the CONNEGATIVE form, as it is a verb form that appears with negation. Examples are given in (5).

- (5) a. *Sa ei vaata(*-d) filmi.*
 You NEG watch-2SG movie.PAR
 ‘You are not watching a/the movie.’
- b. *Te ei vaata(*-te) filmi.*
 You.PL NEG watch-2PL movie.PAR
 ‘You (PL/formal) are not watching a/the movie.’

As shown in (5), negated indicative clauses with overt agreement morphemes are ungrammatical.

In contrast, negated imperatives exhibit agreement twice: once on the main verb and once on the word *ära*, which is traditionally identified as an imperative negative auxiliary.⁴

- (6) a. *Ära vaata filmi!*
 IMP.NEG.2SG watch.IMP.2SG movie.PAR
 ‘Don’t watch a/the movie!’
- b. *Är*(-ge) vaada*(-ke) filmi!*
 IMP.NEG-2PL watch-IMP.2PL movie.PAR
 ‘Don’t watch a/the movie!’

The imperative negative auxiliary’s agreement paradigm is clearly the imperative paradigm; compare the forms of *ära* shown in Table 2 with the forms in Table 1.

Thus, both *ära* and the main verb morphologically express imperativity and agreement. I refer to this pattern as DOUBLING. There are thus three types of exponence in Estonian: standard (single) agreement, no agreement, and doubling.

I argue that these three patterns of exponence arise from a unified syntax of agreement: in Estonian clauses, there is always exactly one head bearing $[u\phi]$. I propose this head is the polarity head Pol^0 . This straightforwardly predicts one instance of agreement, as in affirmative clauses. The patterns of doubled agreement and no agreement emerge

in detail. I assume for concreteness that the alternation amounts to contextual allomorphy determined by particular verbal roots, but see Blevins (2007) for further discussion.

³ The *-gu/ku* forms are not always included as part of the imperative paradigm (see, e.g., Ereht 2003), as that is the ending associated with the jussive. I have included them in the paradigm in the interest of completeness.

⁴ This pattern is also observable in negated jussive clauses (see, e.g., Tamm 2015, 408–410). I focus here on imperatives, but the patterns seen in jussive clauses are fully compatible with my analysis, as far as I can tell.

	SG	PL
1	<i>är-gu</i>	<i>är-me</i> <i>är-gem</i>
2	<i>ära</i>	<i>är-ge</i>
3	<i>är-gu</i>	<i>är-gu</i>

Table 2: *Agreement paradigm for the imperative negative auxiliary ära*

due to the morphology of Estonian, in ways that are made precise below. The end result is a system with a unified syntax of agreement across clausal types but a complex mapping between the syntax of agreement and its morphological exponence. This is consistent with a theory where morphological and syntactic agreement overlap but are not isomorphic (Chung 2013). This analysis also has support from patterns seen in other Uralic languages: while doubling occurs in some languages, the most common pattern is for imperative-marking and agreement to appear only on negation, consistent with the idea that doubling is idiosyncratic morphology rather than evidence of a syntactic relation.

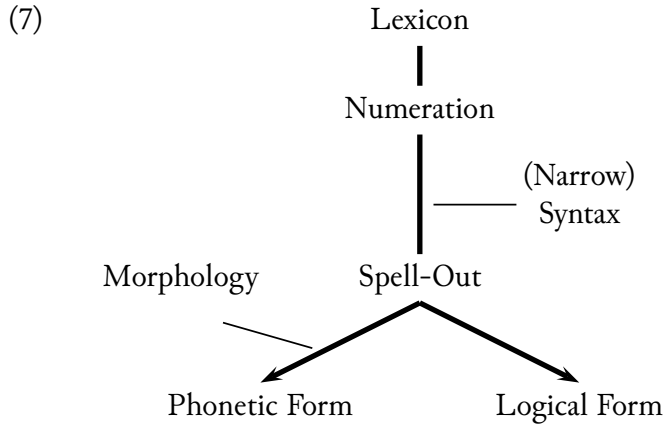
An alternative viewpoint that I consider is that every instance of morphological agreement corresponds to an agreement relationship in syntax, and vice versa. This alternative view is tacitly assumed in some research on agreement, especially in analyses of multiple agreement in the clausal domain by Baker and Willie (2010) and Carstens (2001), and Henderson (2006). For Estonian, negated indicatives would have no agreement—neither syntactic, nor morphological—and negated imperatives would have two instances of syntactic agreement that correspond to two morphological exponents of agreement. While this simplifies the mapping between the syntax and morphology of agreement, it precludes the possibility of a unified syntax of agreement across clause types in Estonian. I believe it is desirable to seek a unified syntax of agreement in the language, as it extends more readily to systems in other Uralic languages (and beyond). More strongly, I show that it is difficult for this type of analysis to generate all of the correct results in Estonian, casting doubt on its general viability in the language.

I begin by providing additional background information in Section 2 before proposing my analysis in Section 3. I consider and reject the alternative analysis mentioned above in Section 4. In Section 5, I consider how my analysis helps us understand the patterns of agreement and imperative-marking in negated imperatives in other Uralic languages, and I discuss a type of unattested imperative-marking within Uralic that is predicted to be nonexistent by my analysis. I offer some directions for future research in Section 6.

2 Background Assumptions

This analysis is formalized within a Minimalist approach to syntax (Chomsky 1995, *et seq.*) combined with the proposals of Distributed Morphology (Halle 1990; Halle and Marantz 1993). In this framework, the syntax manipulates sets of abstract morphological features with no phonological content. Morphological forms are supplied after syntactic operations at a step known as Vocabulary Insertion, which is taken to occur as part of the interface responsible for phonetic and phonological form (PF). Importantly, within this approach,

there are postsyntactic morphological operations that can alter the syntactic representation before Vocabulary Insertion takes place (see, e.g., Arregi and Nevins 2012; Embick 2010; Harley and Noyer 1999). This system is schematized below.



One particular syntactic operation figures prominently in the analysis. It is commonly known as Agree (Chomsky 2000, 2001; Preminger 2014, *a.o.*). The formalization of Agree is subject to some variation, but the core aspects are fairly consistent. The formalization I adopt is given in (8).

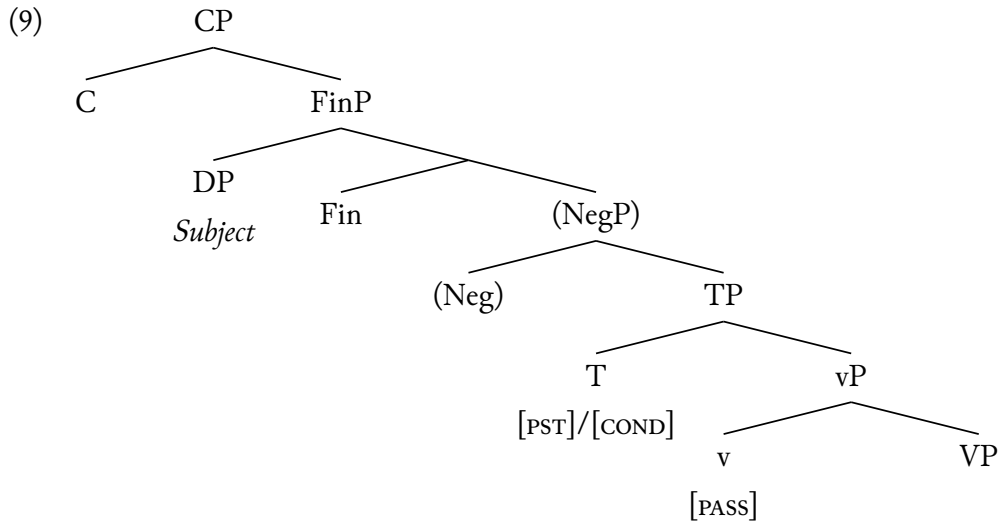
- (8) Agree:
- a. A syntactic head with an unvalued feature or set of features [μ FEAT] (the PROBE) searches within its c-command domain for a node with a corresponding valued feature or feature set (the GOAL).
 - b. If the probe finds a suitable goal, the goal's relevant features are copied to the probe, and Agree is complete.
 - c. If the probe fails to find a suitable goal, Agree fails, and no values are supplied to the probe.⁵

In the situations considered here, the probe has a set of unvalued ϕ -features ($[\mu\phi]$), and it finds a suitable set of ϕ -features on the subject DP, which I assume (for concreteness) to be generated in Spec,vP. Those ϕ -features are copied to the probe and they are canonically realized (at PF) as an agreement marker. In what follows, the precise question that I investigate is how closely morphological agreement tracks syntactic Agree.

2.1 Basic proposals regarding Estonian clause structure

While there are no significant proposals regarding the functional structure of the Estonian IP within a generative framework, there are a number of proposals for Finnish, and they more or less converge on the same structure (Brattico and Huhmarniemi 2006; Brattico et al. 2014; Holmberg and Nikanne 2002; Holmberg et al. 1993; Huhmarniemi 2012; Koskinen 1998; Mitchell 1991). A synthesis of these views is presented below.

⁵ Under some formalizations of the operation, failure of Agree leads to ungrammaticality. See Preminger (2014) for a thorough rebuttal against these kinds of approaches.



Two particular aspects of this proposal are worth mentioning at the moment. First, Fin(ite)^0 is the head that is responsible for finiteness, as visible by subject-verb agreement. In addition, the neutral position of the subject (or more properly, topic) is taken to be Spec,FinP . Second, if Negation is present, it is in between Fin^0 and T^0 . This reflects the fact that, in negated sentences, negation bears agreement, but tense is still reflected only on the main verb. This is visible for Finnish in (10) and (11).

- (10) *pubu-n* / *pubu-i-n*
 speak-1SG / speak-PST-1SG
 'I speak' / 'I spoke'
- (11) *e-n pubu* / *e-n pubu-nut*
 NEG-1SG speak / NEG-1SG speak-PST.PCPL(SG)
 'I don't speak' / 'I didn't speak'

The examples in (10) establish that, in affirmative clauses, the verb reflects both tense and agreement. In the negated sentences in (11), the expression of agreement and tense is split, with negation bearing agreement and the main verb expressing the tense of the clause.

Turning to Estonian, the facts regarding negation and tense are the same: the main verb bears tense in negated clauses, and negation does not (see (12) and (13)).

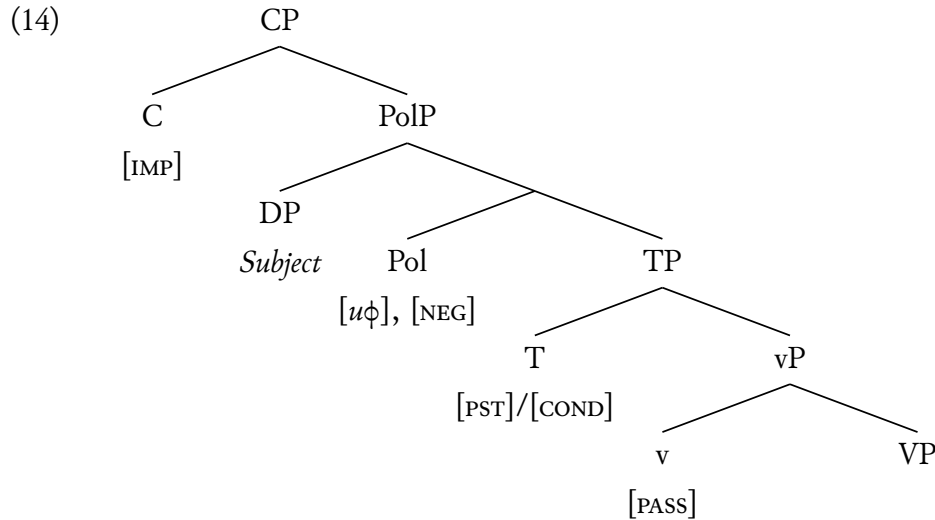
- (12) *Ma ei maga.* (13) *Ma ei maga-nud.*
 1SG NEG sleep / 1SG NEG sleep-PST.PCPL
 'I don't sleep.' / 'I didn't sleep.'

These facts follow the same pattern as Finnish with respect to tense-marking: compare *maga* 'sleep' in (12) to *maganud* 'sleep.PST.PCPL' in (13). As in Finnish, the main verb does not bear agreement in negated indicatives in Estonian, whether past or present. I take this as evidence that the ϕ -feature probe in Estonian clauses is also higher than negation (as in Finnish), as the introduction of negation blocks ϕ -feature agreement on the main verb. However, unlike negation in Finnish, Estonian indicative negation does not show any agreement: it is always *ei* (see also (5)). I set this aside for the moment, returning to an analysis of indicative negation in Section 3.2.

Imperative clauses are not discussed in detail in much of the literature on Finnic clausal architecture; some important exceptions are Brattico et al. (2014), Mitchell (1991),

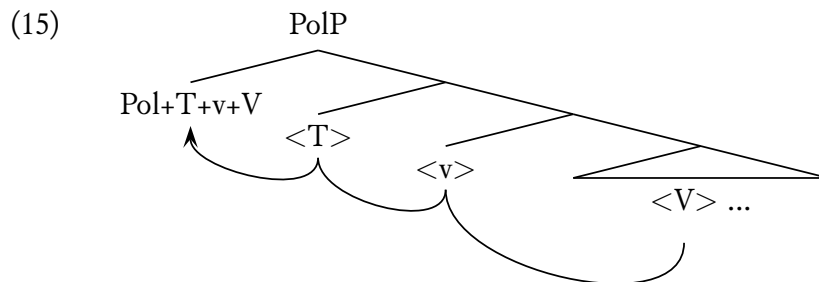
and Nelson (1998). Nelson (1998) proposes that the morphosyntactic feature(s) unique to imperative clauses are lower (on T^0 , below negation and Fin^0), while Brattico et al. (2014) and Mitchell (1991) propose that the morphosyntactic features unique to imperative clauses are located on C^0 , which is above negation and finiteness (see also Han 1999; Rivero and Terzi 1995). I propose the same for Estonian, i.e., that morphosyntactic imperative features are located on C^0 , but I will come back to Nelson's (1998) proposal in Section 5.

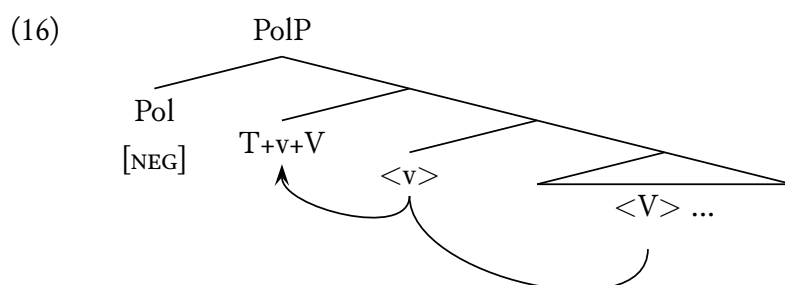
Taking these proposals together, the structure of the Estonian clause that serves as the basis for my analysis is presented in (14).



Some aspects of the structure deserve comment. First, though I have not discussed it here, I assume that v^0 is the location of the impersonal/passive suffix. Second, I have collapsed Neg^0 and Fin^0 into one head: Pol^0 . As far as I can tell, Neg^0 serves no purpose other than the introduction of the negation head, which always moves to Fin^0 . Instead of this, I propose a head Pol^0 (for polarity) that comes in two flavors: one with $[\text{NEG}]$ and one without. $\text{Pol}_{[\text{NEG}]}^0$ spells out as negation, and plain Pol^0 spells out as ϕ -feature agreement. The proposals I make in what follows are compatible with an analysis that does not collapse Fin^0 and Neg^0 , but I will do so in the interest of simplicity. This means that whatever information besides $[u\phi]$ is assumed to be present on Fin^0 must be present on Pol^0 instead, although the discussion that follows focuses only on $[u\phi]$.

Within this structure, I assume that the verb undergoes head movement up the clausal spine. I assume it moves as high as T^0 in all clauses, as has been proposed for Finnish (Holmberg and Nikanne 2002, *a.o.*). In affirmative clauses, the verb continues to Pol^0 . In negated clauses, head movement to $\text{Pol}_{[\text{NEG}]}^0$ is blocked, and the verb stays in T^0 . This is shown for affirmative clauses in (15) and negated clauses in (16).





The interaction of Pol^0 and $\text{C}_{[\text{IMP}]}^0$ is part of the core of the analysis, which is discussed in the next section. Let us turn to it now.

3 *ära* as a conglomeration of negation and imperativity

As we have seen, negated imperative clauses in Estonian are not formed with *ei*, but with a special negative auxiliary *ära*.⁶ Unlike Estonian *ei*, *ära* morphologically agrees with the subject, and the set of agreement markers that appear is the same imperative paradigm as that which appears on the main verb in an affirmative imperative.⁷ Additional examples are in (17).

- (17) a. *Är-ge kõndi-ge muru-!*
 IMP.NEG-2PL walk-IMP.2PL grass-ADE
 ‘Don’t walk on the grass!’ (Erelt et al. 1993, 155)
- b. *Ära kõnni muru-!*
 IMP.NEG.2SG walk.IMP.2SG grass-ADE
 ‘Don’t walk on the grass!’ (Erelt et al. 1993, 155)

Recall as well that the main verb bears imperative agreement in negated imperatives (e.g., *kõndi-ge* ‘walk-IMP.2PL’ in (17a)). This means that agreement and imperativity are real-

⁶ An anonymous reviewer asks about the etymology of *ära* and *ei*, and in particular, whether their relationship can be illuminated by historical evidence. I have not found anything particularly interesting as of yet. The etymological dictionary of Estonian (*Eesti etümoloogiasõnaraamat*, available online at <http://www.eki.ee/dict/ety/>) collapses the entries for *ära* and *ei* into one entry. The entry provides the cognate forms in a variety of languages, including both the indicative and imperative negative auxiliary. Some of these forms are given in Section 5 of the present article. However, I will refrain from speculating and simply note that this is a question worth investigating. Thanks to Anne Tamm for helpful discussion of this issue.

⁷ I will continue to refer to imperative agreement as agreement with the subject, and I will represent it formally as such. However, the source of this agreement is a matter of some debate within the literature. Kiparsky (2001) proposes that agreement in imperatives is not the same as agreement in non-imperative clauses in Finnish (pp. 335–6), and Miljan and Cann (2013) claim something similar for Estonian, proposing the agreement is “not syntactic but pragmatic” (p. 360). I note first that the observations Kiparsky and Miljan & Cann make do not preclude a syntactic account: see, e.g., Zanuttini (2008) and Zanuttini et al. (2012). However, even if the pragmatic approach is superior, the choice between the pragmatic account and the syntactic account is ultimately about the proper *controller* of agreement, whereas the phenomenon I am investigating is about the *exponence* of agreement. It seems to me that the puzzle of multiple agreement exponence in negated imperatives remains even if the controller is pragmatic in nature. Thus, in what follows, I assume a syntactic approach for concreteness.

ized twice: once on *ära* and once on the main verb.⁸ It is ungrammatical to leave either agreement exponent out, whether on negation (as in (18a)) or on the main verb (as in (18b)).

- (18) a. **Ära söö-ge seda kooki!*
 IMP.NEG eat-IMP.2PL that.PAR cake.PAR
 Intended: ‘Don’t eat that cake!’
- b. **Är-ge söö seda kooki!*
 IMP.NEG-2PL eat.IMP that.PAR cake.PAR
 Intended: ‘Don’t eat that cake!’

However, there is one exception to the general pattern of obligatory imperative agreement doubling in Estonian negated imperatives. For [1PL] imperatives using the ending *-me*, agreement on the main verb is optional.⁹ When agreement does not appear on the main verb, the main verb instead surfaces in the ordinary present connegative form. This is illustrated in (19) and (20).

- (19) *Är-me vaata(-me) filmi!*
 IMP.NEG-1PL watch(-1PL) film.PAR
 ‘Let’s not watch a/the movie!’
- (20) *är-me tee(-me)*
 IMP.NEG-1PL do(-1PL)
 ‘Let’s not do’ (Tamm 2015, 407)

As Tamm (2015) shows, both forms—with doubling and without it—exist in modern spoken and written registers of varying levels of formality in the modern language, though she suggests that the standardization of the non-agreeing form is a somewhat recent development.

There are two primary questions that I address with my analysis. First, why are both negation and the main verb inflected for imperative features? Assuming that imperative morphology is connected to a morphosyntactic feature (e.g., [IMP]), it is worth considering which of the imperative inflections (i.e., that on negation or that on the main verb) is connected to this feature. Second, what is the source of the doubled ϕ -feature marking, and how can we formalize it such that it can be optional in the case of [1PL] agreement?

In response to the first question, I argue that the imperative marking on negation is connected to “true” (i.e., syntactic) imperativity. In response to the second question, I argue that agreement on *ära*, i.e., negation, is the reflex of an Agree relationship. Thus, the imperative negator is directly connected to both syntactic features, [IMP] and ϕ -features.

⁸ It is difficult to know whether the main verb is inflecting for person and number in the context of second-person singular subjects, as it is in a form that is homophonous with the present connegative form for nearly every verb in the language. The only verb for which this is not true is *minema* ‘go’, where the 2SG imperative form is *mine* but the present connegative form is *lähe*. Because of this, I believe the verb is bearing agreement in this context.

⁹ Another ending is possible for [1PL] imperatives: *-gem/-kem*. As Tamm (2015) observes, there is a register difference between *-me* and *-gem*, such that the *-gem* form is quite formal. Unlike *-me*, *-gem* obligatorily doubles in negated imperatives. I assume this difference can be attributed to a morphosyntactic feature related to formality or to different dialects, which is independently necessary in order to capture the fact that *-me* and *-gem* are both possible exponents for [1PL] in imperatives.

In contrast, I propose that the agreement and imperative-marking on the main verb is the result of a different operation. Furthermore, I argue that the optionality of [1PL] doubling suggests that this operation is morphological rather than syntactic in nature. I propose that it is an instance of Feature Copying (Kramer 2010; Norris 2014), an idiosyncratic rule of Estonian triggered by the presence [IMP] and [NEG] on the same head.

The primary focus of this section is to lay out my analysis and show how it works. The two sections following this one show in more detail why the analysis I propose is preferable. In Section 4, I consider an alternative that is more along the lines of the work that proposes multiple ϕ -feature probes (e.g., Baker and Willie 2010; Carstens 2001; Henderson 2006), and I show that it is less plausible than the analysis I advocate for. Then in Section 5, I show that my analysis makes interesting predictions with respect to negated imperatives in the Uralic family, lending further support to the analysis proposed in this section.

3.1 Analysis: morphology and syntax of negated imperatives

The analysis is broken up into two parts. Beginning with the assumption that imperative clauses are connected to a feature (or set of features) in the syntax, I propose that Estonian has a C^0 with an imperative feature, which I call [IMP] for concreteness. In an affirmative imperative clause, the finite verb is high, indicated by the fact that it can readily appear before the subject when the subject is overt.

- (21) *(Sa) kasta lilli!*
 you water.IMP.2SG flower.PL.PAR
 ‘Water (some) flowers!’ (Erelt et al. 1993, 175)

- (22) *Kasta SINA lilli!*
 water.IMP.2SG you flower.PL.PAR
 ‘YOU water (some) flowers!’ (Erelt et al. 1993, 175)

In (21), the pronoun *sa* ‘you’ can appear before the verb. Importantly, the subject pronoun can also appear after the verb, as in (22).¹⁰ I attribute this to head movement: the verb moves all the way to C^0 in imperative clauses. For concreteness, I suggest the variable positioning of the subject (before or after the verb) is due to an optional movement of the subject to Spec,CP, though the position of the subject will not affect the analysis I propose.¹¹ If the subject remains in Spec,PolP, then it is post-verbal.¹²

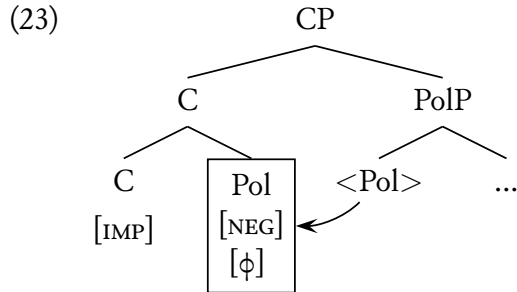
Turning to negated imperatives, recall that head movement of the verb stops at T^0 when $\text{Pol}_{[\text{NEG}]^0}$ is present. This accounts for the fact that the verb inflects for tense but not

¹⁰ The form, *sina* or *sa*, does not affect the possible positions, as far as I have been able to tell.

¹¹ Based on the research conducted on Finnish, I believe this movement to Spec,CP is likely driven by information-structural considerations. For example, Holmberg and Nikanne (2002) note that Spec,CP is generally reserved for contrastive topics. I have not yet uncovered any significant contrasts in the interpretation of imperatives with preverbal or postverbal subjects in fieldwork, though a reviewer notes that postverbal subjects in imperatives are normally interpreted as contrastive, suggesting the default position for the subject in an imperative is preverbal. I must leave this issue unresolved here, although I note that the exact positioning of the subject is tangential to the point that I am trying to make about agreement exponence.

¹² In Finnish, the imperative subject, when present, must be post-verbal. I have no explanation for why Finnish and Estonian differ in this respect.

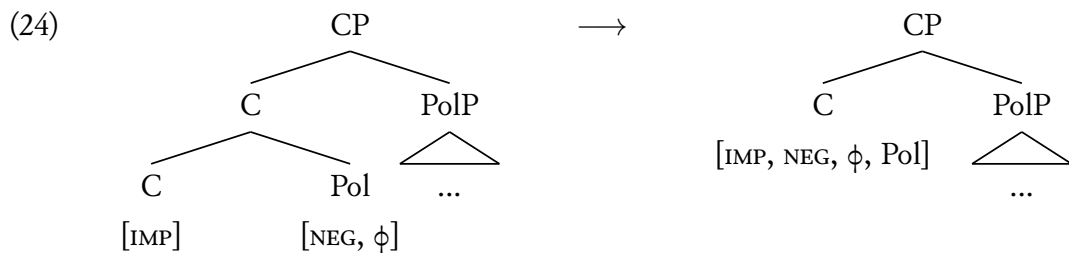
agreement in negated clauses. This also precludes the verb from reaching C^0 as it does in affirmative imperatives. However, Pol^0 is able to move to C^0 as normal—it just does not contain the verbal complex. This movement yields the complex head in (23).¹³



Thus, the only difference between negated imperatives and affirmative imperatives is whether or not the verbal complex raises to Pol^0 . In either case, Pol^0 always raises to $C_{[\text{IMP}]}^0$.

As indicated in (23), $\text{Pol}_{[\text{NEG}]}^0$ has already established Agree with the subject DP by the time it raises to C^0 , and the subject's ϕ -features, along with negation and imperativity, now form a complex head in C^0 . This is the essence of the special negative imperative auxiliary *ära*: it is the morphological form supplied to the complex C^0 head, expressing negation, imperativity, and the subject's ϕ -features. Imperativity is visible at least in the agreement paradigm used, which is the language's imperative agreement paradigm. It is reasonable to suggest *ära* reflects negation given (i) its restriction to negative imperatives, and (ii) the fact that no other sentential negation is possible.¹⁴

At this point, the representation is sent to the PF interface where it can be interpreted by the morphology. To begin, I propose that $C_{[\text{IMP}]}^0$ and $\text{Pol}_{[\text{NEG}]}^0$ undergo Fusion (Halle 1990; Halle and Marantz 1993), which takes two terminal nodes which are sisters and combines their feature bundles to form a single node as represented in (24). The new fused head is ultimately spelled out as *ära*.



¹³ An issue that I do not address here is how such a complex head would be interpreted, especially as far as scope is concerned. This issue could be avoided under an account where head movement does not occur until PF (e.g., Schoorlemmer and Temmerman 2012) or where *ära* is not derived via Head Movement, but Spanning (Merchant 2015; Svenonius 2012). I will not attempt to construct such an account here.

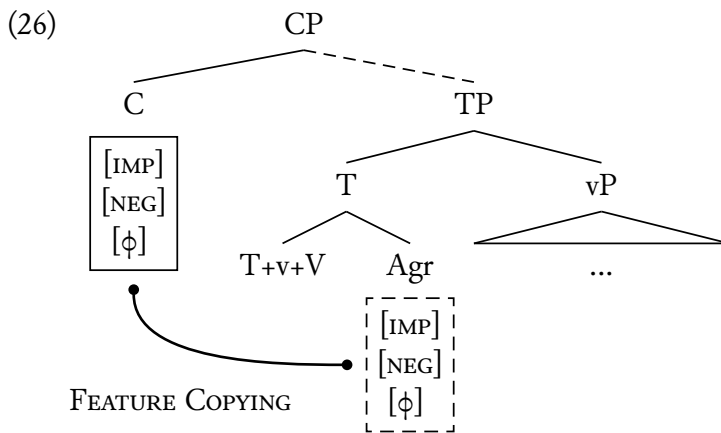
¹⁴ However, I note here that the alternative analysis that I consider in Section 4 does not adopt this characterization for *ära*.

- (25) Imperative Negative Auxiliary Vocabulary Items (simplified):¹⁵
- a. C, [IMP, NEG, 2SG] ↔ *ära*
 - b. C, [IMP, NEG, 1PL] ↔ *ärme*
 - c. C, [IMP, NEG, 2PL] ↔ *ärge*
 - d. C, [IMP, NEG] ↔ *ärgu*

Thus, under this analysis, the term “negative imperative auxiliary” refers to a C^0 that contains ϕ -features, the imperative feature [IMP], and the negation feature [NEG].

Turning now to the main verb, recall that it expresses imperative agreement as well in negated imperatives. Thus, as far as the morphology is concerned, the main verb needs to acquire the imperative feature [IMP] as well as the subject’s ϕ -features. Since those features are bundled up in C^0 , we have to say something more. It will not do to propose a second ϕ -feature probe on a lower head, e.g., T^0 . We have already seen evidence against such a proposal from negated indicatives, where the main verb bears tense but no agreement. If T^0 were a probe, this would be unexpected. Thus, while two ϕ -feature probes might simplify negated imperatives, they complicate negated indicatives, as well as affirmative clauses in general.

I propose instead that doubling is the result of a morphological rule of Feature Copying, triggered by C^0 specified as [IMP, NEG]. First, an Agr node is adjoined to T^0 , and subsequently, C^0 ’s features are copied onto that Agr node. This is represented schematically in (26), where the dashed branch indicates a slight abbreviation in the structure.



This copies all of the features of C^0 , including [NEG], which is essentially ignored at Vocabulary Insertion. I do this in the interest of simplicity: it is simpler to have C^0 ’s entire feature set copied rather than allowing Feature Copying to choose which features are copied. See

¹⁵ These Vocabulary Items are simplified in that I abstract away from whatever operation (e.g., Fission), separates the agreement features from the negative and imperative features. It would be empirically identical to suggest that there is no Fusion, but instead, mutual contextual allomorphy. We could propose that *är* is an expression of $C_{[IMP]}^0$ in the context of [NEG], and $\text{Pol}_{[NEG]}^0$ with ϕ -features is realized as the imperative agreement paradigm in the context of $C_{[IMP]}^0$. I adopt the Fusion account here because it allows for a more straightforward account of the doubling of imperative and agreement features, as it joins them as a single node.

also Kramer (2010) for a similar conclusion regarding Feature Copying in Amharic adjectival agreement.¹⁶

Because Feature Copying applies postsyntactically, it necessarily applies after Agree has taken place. This means that the ϕ -feature values of the subject are known when Feature Copying applies, which opens the door for an understanding of the optionality in the case of [1PL] *-me*. Concretely, I propose that the operation of Feature Copying is optional when C^0 is specified as [1PL]. This is similar to the analysis of Basque auxiliaries proposed by Arregi and Nevins (2012), where certain PF operations may apply (or not apply) based on ϕ -feature specifications. The final formalization I adopt for Feature Copying Estonian Imperatives is presented in (27).

- (27) Feature Copying (Estonian Imperatives): Copy C^0 's features to an Agr node adjoined to T^0 if C^0 is specified for [IMP, NEG]. This is optional if C^0 is specified as [1PL].

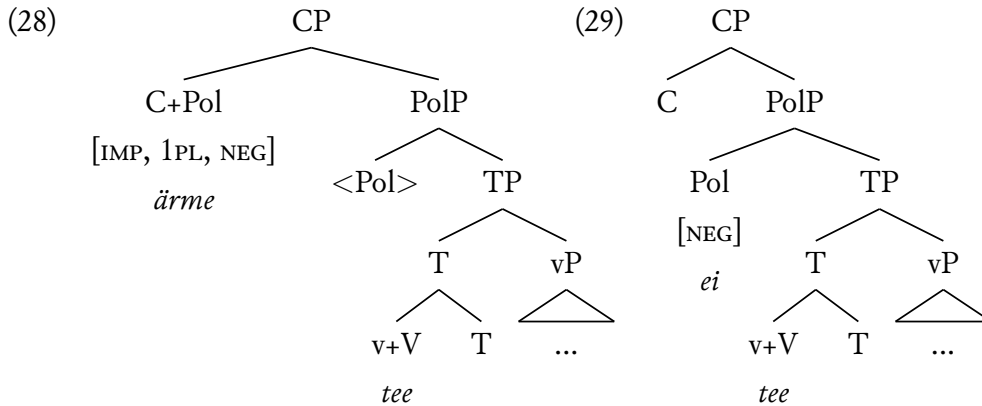
A reviewer has two related questions about the arbitrariness of rules of Feature Copying. First, the reviewer observes that the optionality of first-person plural doubling is rather arbitrary as stated here. The issue of arbitrariness comes up again in Section 5, where some differences in patterns of doubling in other Finno-Ugric languages are investigated. There is some variation in what features are copied in languages with doubling, suggesting some degree of arbitrariness may be necessary.

Second, the reviewer asks whether the heads and features involved in Feature Copying rules like (27) are completely arbitrary. While I cannot offer a definitive answer to this question, I comment briefly on this question here. The question is both empirical and theoretical, and it has not yet been systematically investigated. However, the rule presented here shares similarities with other uses of Feature Copying in the literature (Kramer 2010; Norris 2014). First, Feature Copying copies features to an Agr node adjoined to some head postsyntactically. Second, the origin of the features is higher in the structure than the target; in other words, features are copied downward, as opposed to Agree, which can only transfer features upward, in at least some conceptualizations.¹⁷ Third, there is a sense in which Feature Copying takes place within a single domain reminiscent of Grimshaw's (1991/2005) Extended Projection: for Kramer (2010) and Norris (2014), features are copied from some nominal head to attributive adjectives, and for my analysis, features are copied from a verbal head to another verbal head. Whether these properties can hold for all plausible instances of Feature Copying is an issue that I must leave to future work.

Returning to the main point, when Feature Copying does not apply in these [1PL] contexts, the main verb is in a structural position that is identical (or very similar) to its position in negated indicative clauses: it is in T^0 with no ϕ -features. This is visible in the imperative structure in (28) and the indicative structure in (29).

¹⁶ Feature Copying of the kind proposed here is also similar to the operation of Enrichment proposed by Müller (2007), although Enrichment operates more locally, only allowing features to be duplicated inside a single complex head.

¹⁷ There are many existing proposals concerning the direction of Agree, including that it only transmits feature values downward and that it can transfer feature values in any direction. For recent discussion, see Bjorkman and Zeijlstra (2015) and Preminger and Polinsky (2015).



In (28), Pol^0 has raised to C^0 , but Feature Copying has not applied, leaving the main verb in T^0 with no [IMP] feature nor ϕ -features. This is the same as the main verb in a negated indicative clause: it has no ϕ -features (nor the [IMP] feature, for that matter).

This analysis thus predicts that, when Feature Copying does not apply, the main verb should surface in its ordinary connegative form. It is actually somewhat difficult to tell, because for most verbs in the language, the connegative form is homophonous with the second-person singular imperative form. However, as noted in footnote 8, there is one verb with distinct forms (*minema* ‘go’), and it must be in the connegative form in first-person plural imperatives without doubling. This is shown in (30) and (31).

- (30) *är-me läbe*
 NEG.IMP-1PL go.CNG
 ‘Let’s not go’
- (31) **är-me mine*
 NEG.IMP-1PL go.IMP.2SG
 Intended: ‘Let’s not go.’

The connegative form of *minema* ‘go’ is *läbe*, while its second-person singular imperative form is *mine*. Only *läbe* can be used in negated [1PL] imperatives without Feature Copying, which is exactly what the Feature Copying analysis predicts.

Under this analysis, the relationship between the syntax of agreement (i.e., Agree) and its morphological expression is imperfect for negated imperatives: there is only one instance of Agree, but potentially two exponents of ϕ -feature agreement. In fact, the analysis also requires an imperfect relationship for negated indicatives: one instance of Agree in the syntax, but no morphological expression. This is a complication, but it is warranted. Let us now discuss indicative clauses for some independent evidence for that claim.

3.2 *ei* as a morphologically deficient negative auxiliary

As we have seen, negated indicative clauses in Estonian do not show morphological agreement, regardless of tense.

- (32) *Sa ei vaata filmi.*
 2SG NEG watch movie.PAR
 ‘You are not watching a/the movie.’
- (33) *Sa ei vaada-nud filmi.*
 2SG NEG watch-PST.PCPL movie.PAR
 ‘You did not watch a/the movie.’

If we take the term *negative auxiliary* to mean that the negation word inflects like a verb, then Estonian *ei* does not qualify. It does not show verb-like inflection of any kind. If we take the canonical example of a *negation particle* to be a negation form that does not affect the form of the main verb (SYMMETRIC NEGATION in the terms of Miestamo 2005), Estonian *ei* does not seem to qualify with this either, as its presence prevents the main verb from showing agreement. Under the analysis I proposed above, Estonian *ei* is the realization of $\text{Pol}_{[\text{NEG}]}$ ⁰, a head which bears $[u\phi]$ and establishes Agree in the syntax. In other words, it is syntactically a negative auxiliary. I believe that this is the right approach for the syntax of Estonian for two reasons.

First, discussion by Erelt (2003) and Tamm (2015) suggests that the negative auxiliary bore inflection in older forms of the language. Tamm provides the example in (34).

- (34) *et e-b se mei-lle voi mitte kuria teb-da.*
 that NEG-3SG this we-ALL can.CNG NEG harm.PAR do-DA.INF
 ‘That this cannot harm us.’ (Tamm 2015, 419)

Here, the form of negation is *eb*, indicated as bearing third person singular agreement. Tamm also notes that, although *ei* does not show inflection in Standard Estonian, it does agree in some Southern varieties of the language. There is thus historical precedent and potential dialectal evidence for a negative auxiliary in Estonian.

Second, Estonian is the only language among its closest relatives where the negative auxiliary fails to inflect (Miestamo et al. 2015; Mitchell 2006). As we have seen, it inflects in Finnish, but it also inflects in the Finnic languages Ingrian, Karelian, Livonian, Veps, and Vod.¹⁸

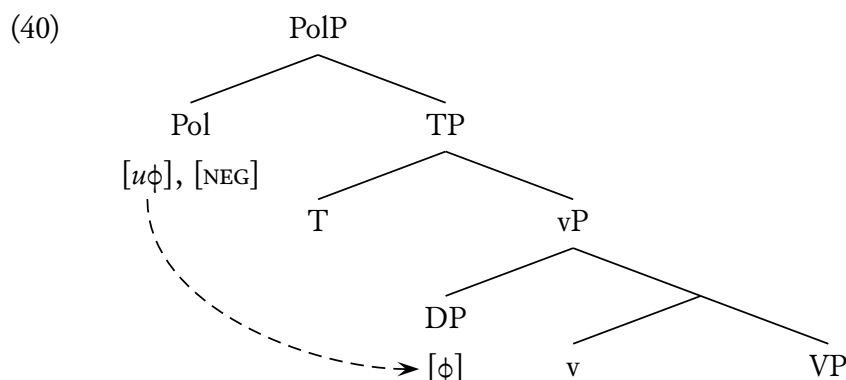
- (35) Ingrian (Mitchell 2006, 235)
e-n ompēle
 NEG-1SG sew
 ‘I don’t sew’
- (36) Karelian (Mitchell 2006, 235)
e-n šano
 NEG-1SG say
 ‘I don’t say’
- (37) Livonian (Mitchell 2006, 230)
Miná ä-b uo
 1SG NEG-1SG be-PRS.PCPL
 ‘I am not.’
- (38) Veps (Mitchell 2006, 230)
e-n to
 NEG-1SG bring
 ‘I don’t bring’

¹⁸ Some of the Finnic languages’ names are rendered in English in a variety of ways. Veps is sometimes called Vepsian, and Vod is sometimes called Votic or Votian. Here, I have used the main entry name according to Ethnologue.

- (39) Vod (Mitchell 2006, 236)
e-n jō
 NEG-1SG drink
 ‘I don’t drink.’

In addition, though not shown here, all of the closely related Saami languages have inflecting auxiliaries (Toivonen and Nelson 2007, 8–9).¹⁹ Proposing that the Estonian *ei* is still at some level a negative auxiliary thus puts it in line with the other members of its family and the closely related Saami languages.

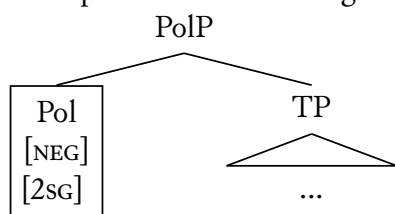
Speaking more concretely, I propose that Estonian *ei* (or more accurately, the syntactic head Pol^0) still has the syntax of a negative auxiliary. This means it is generated with $[\mu\phi]$ in the syntax, and it undergoes Agree with the subject, obtaining its ϕ -features. Estonian has this in common with the languages where this agreement is overt. This is depicted in (40).



In (40), Pol^0 establishes Agree with the subject DP in Spec,vP. Thus, $\text{Pol}_{[\text{NEG}]}^0$ acquires a ϕ -feature set in the narrow syntax, and this is the representation that undergoes Vocabulary Insertion in indicatives.

This is where Standard Estonian differs from the rest of the members of its family. Whereas a language like Finnish has multiple vocabulary items for $\text{Pol}_{[\text{NEG}]}^0$, Estonian only has one: *ei*. Thus, when $\text{Pol}_{[\text{NEG}]}^0$ undergoes Vocabulary Insertion, *ei* is inserted regardless of the ϕ -features. An example terminal node is presented in (41), and the competitions for Finnish and Estonian are represented in (42) and (43), respectively.

- (41) Example terminal node targeted for insertion:



¹⁹ Toivonen and Nelson (2007) provide examples from South Saami and Inari Saami, but their discussion suggests this is the case in all of the Saami languages. The literature also contains examples from Finnmark Saami (Mitchell 2006), North Saami (Nickel 1990), Pite Saami (Wilbur 2014), and Skolt Saami (Feist 2010), all of which have agreeing negative auxiliaries.

- | | | | |
|------|-----------------------------------|------|-----------------------------|
| (42) | Finnish Vocabulary Items | (43) | Estonian Vocabulary Items |
| | a. Pol, [NEG, 1SG] ↔ <i>en</i> | | a. Pol, [NEG] ↔ <i>ei</i> ⇐ |
| | b. Pol, [NEG, 2SG] ↔ <i>et</i> ⇐ | | |
| | c. Pol, [NEG, 1PL] ↔ <i>emme</i> | | |
| | d. Pol, [NEG, 2PL] ↔ <i>ette</i> | | |
| | e. Pol, [NEG, 3PL] ↔ <i>eivät</i> | | |
| | f. Pol, [NEG] ↔ <i>ei</i> | | |

Thus, the Estonian *ei* amounts to a highly underspecified vocabulary item for negation. Following standard DM assumptions, the vocabulary item matching the largest subset of features of the terminal node is inserted (on the Subset Principle, see Halle 1997; Hankamer and Mikkelsen 2005, *a.o.*). For Finnish, this is *et*, which matches the entire set. For Estonian, *ei* matches a subset of the features of the terminal node, and there are no Vocabulary Items matching a greater set. Regardless of the ϕ -feature set, *ei* is inserted. Thus, the indicative negative auxiliary in Estonian establishes a syntactic Agree relationship, but that relationship is not realized by morphological agreement.

As an alternative, it could be proposed that the difference between Estonian and other Finno-Ugric languages is that $\text{Pol}_{[\text{NEG}]^0}$ in Estonian has lost its status as a ϕ -feature probe. There would be thus no agreement in the syntax, and the spell-out would be the same. This provides a more transparent mapping between the syntax and morphology of Estonian *ei*, but it also sets Estonian apart from the rest of its relatives in terms of the syntax. I believe this alternative analysis and the analysis I propose are indistinguishable in terms of empirical coverage for indicatives. Conceptually, my analysis is preferable for two reasons. First, it fits within a general theory of language variation where variation is located in the morphology of languages rather than their syntax (Chung 2013, 2014).²⁰

Second, the analysis whereby Estonian *ei* establishes Agree in the syntax despite its lack of demonstrable agreement is easier to incorporate with an analysis of the pattern of negated imperatives. In that analysis, the imperative negative auxiliary *ära* is the spell-out of a complex head involving [NEG], [IMP], and [ϕ]. If $\text{Pol}_{[\text{NEG}]^0}$ is simply not a ϕ -feature probe (i.e., if it lacks [$u\phi$]), then the only way for $\text{C}_{[\text{IMP}]^0}$ to have ϕ -features in its feature bundle would be if $\text{C}_{[\text{IMP}]^0}$ could idiosyncratically select $\text{Pol}_{[\text{NEG}]^0}$ with ϕ -features, which seems difficult to motivate independently. It is also not obviously superior to propose that $\text{C}_{[\text{IMP}]^0}$ Agrees directly with the subject, as this would mean there are two ϕ -feature probes in affirmative imperatives— $\text{C}_{[\text{IMP}]^0}$ and Pol^0 , necessary for affirmative indicatives—even though there is only one instance of agreement. It also leads to a system that is less uniform, as the analysis I propose holds that there is always only one ϕ -feature probe per finite CP, and it is located on Pol^0 . For these reasons, I do not adopt an analysis where $\text{Pol}_{[\text{NEG}]^0}$ lacks [$u\phi$] in Estonian.

²⁰ It is also worth noting here that the account whereby $\text{Pol}_{[\text{NEG}]^0}$ is simply not a probe in Estonian is incompatible with the proposal that some clausal features are born on C^0 but inherited by the head of C^0 's complement (see Brattico et al. 2014 for a discussion of this in the context of Finnish). Under this analysis, it is C^0 that determines the features relevant for the clause. Thus, whether C^0 's complement is $\text{Pol}_{[\text{NEG}]^0}$ or simply Pol^0 would be irrelevant, formally speaking.

3.3 Negated Imperatives: analysis summary

In this section, I presented an analysis of the agreement patterns in negated imperatives in Estonian as resulting from syntactic and morphological agreement. On the syntactic side, I proposed that $\text{Pol}_{[\text{NEG}]^0}$ Agrees with the subject before moving to $\text{C}_{[\text{IMP}]^0}$, ultimately spelling out as the imperative negative auxiliary *ära*. In the morphology, the combination of [IMP] and [NEG] triggers Feature Copying from C^0 to the verbal complex in T^0 . In Estonian, Feature Copying is obligatory for all combinations of person and number except first-person plural, where agreement is only obligatory on the negative auxiliary. As a consequence of this proposal, I also discussed the indicative negative auxiliary *ei*, proposing that it establishes Agree in the syntax even though that agreement is not reflected morphologically as it is in all of Estonian's closest relatives.

The upshot of the analysis is that the complicated morphology of agreement exponence in Estonian disguises a system that is straightforward syntactically. There is one head that bears $[\mu\phi]$ — Pol^0 —and that head always establishes Agree in the syntax. This opens the door for the possibility of a more uniform analysis of the syntax of agreement in the Finnic and Saamic languages. This line of research ties language-specific stipulations about Estonian (e.g., doubling, *ei* showing no inflection) to the morphology as opposed to the syntax, following the argumentation of Chung (2014).

This analysis also involves a maximally transparent syntax for negated imperatives in Estonian, as the syntax of imperativity is simply laid on top of the syntax of negation. While I believe this is a welcome result, it is not necessarily an argument in favor of the current proposal, as it is well-known that negation and imperativity are not readily combined in the world's languages. In the WALS chapter on the prohibitive (van der Auwera et al. 2013), only 22.8% of languages in the sample form negative imperatives with a morphologically-imperative verb and the same negation as declaratives, and we would expect that number to be higher if there were no issues combining imperative and negation. Furthermore, based on surface forms alone, Estonian does not fall into that group. Rather, Estonian is a member of the largest group of languages in the sample (36.7%), whose negative imperatives utilize a special form of negation. It is thus reasonable to question whether it is right to ascribe the transparent syntax I have proposed for negated imperatives in the language.

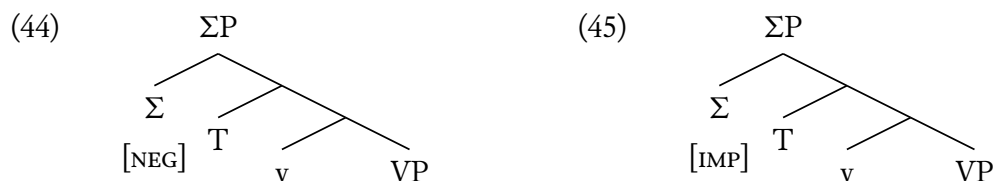
In the next section, I consider the prospects for an account of Estonian that takes a different tack, treating negation and imperativity as syntactically incompatible. The alternative hypothesis I will consider also differs from the analysis presented in this section in that it accounts for doubling in imperatives by adding a second ϕ -feature probe to the clause. As I have mentioned, this is in line with some previous work on multiple agreement for ϕ -features (Baker and Willie 2010; Carstens 2001; Henderson 2006; Polinsky 2016). I show that the account struggles to account for the optionality of agreement in first-person plural imperatives. I also discuss some conceptual reasons why the alternative developed in the next section is inferior to the analysis I argued for in this section.

4 Alternative: Multiple ϕ -feature probes

The morphological analysis presented in Section 3 requires a morphological method of feature transfer (Feature Copying), and thus it is worth considering the prospects for an account that does not utilize additional mechanisms. There are no previous accounts that can be compared, so far as I know. Instead, I present the best account I have been able to construct, building on the idea that imperative syntax and ordinary negation syntax are incompatible (Laka 1990; Zanuttini 1994). This alternative analysis accounts for doubling through the use of multiple ϕ -feature probes, as has been proposed for other instances of multiple ϕ -feature exponence in Archi (Polinsky 2016), Ibibio (Baker and Willie 2010), and Swahili (Carstens 2001; Henderson 2006). We will see that it cannot straightforwardly account for the optionality of doubling in [1PL] contexts. Furthermore, it does not generalize as easily to the morphological patterns seen in other Uralic languages, which is the focus of Section 5.

4.1 Laka (1990): Negation and imperative realize the same head

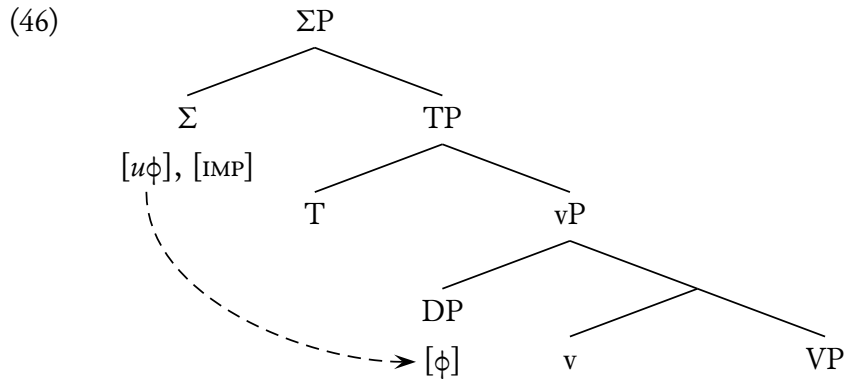
To account for the fact that negation and imperativity are incompatible in some languages, Laka (1990) proposes that negation and imperative realize the same head. Laka calls this head Σ^0 , and Σ^0 selects a TP complement. Thus, a clause can either have Σ^0 with an imperative feature [IMP] or Σ^0 with a negation feature [NEG], but not both at once. This is shown in (44) and (45).



Laka's discussion about imperativity focused on Spanish, where the morphological imperative form cannot be combined with the standard negation marker. Instead, the language uses a subjunctive verbal form. To adapt this for Estonian, we would further propose that it is Σ^0 that bears unvalued ϕ -features in imperative clauses in Estonian. I leave open the question of whether all Σ^0 heads bear [$u\phi$] (that is, both Σ^0 with [NEG] and Σ^0 with [IMP]) or if it is just the imperative Σ^0 . This differs from the analysis proposed in Section 3, as imperative and negation are realized by separate heads in that analysis— C^0 and Pol^0 , respectively—and are thus fully compatible.

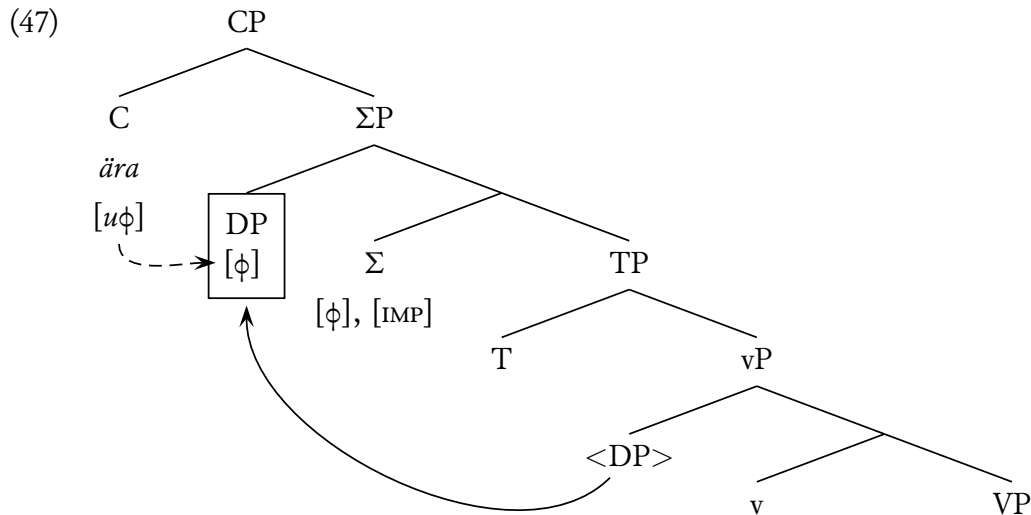
Since negation and imperative features would be syntactically incompatible under this analysis, Estonian would require an alternative structure to express prohibitive semantics. Namely, Estonian makes use of *ära*, a word traditionally called the imperative negative auxiliary. To concretize this traditional definition, I suggest that *ära* is a C^0 that selects a ΣP with [IMP] features. This constrains *ära* such that it would only surface in imperative clauses. Because *ära* shows agreement, it must be idiosyncratically associated with its own set of unvalued ϕ -features [$u\phi$]. Let us see how this alternative hypothesis works by walking through a sample derivation.

In the first step, Σ^0 (with imperative features) establishes Agree with the subject DP as normal, schematized by the dashed arrow in (46).



Then, assuming that the verb moves up the clausal spine to $\Sigma^0_{[IMP]}$ —just as I proposed in Section 3—the imperative verb comes to reflect the ϕ -features of the imperative subject. Under this account, this much would be identical for both affirmative and negated imperatives.

In the second step, *ära* (more accurately, the C^0 that ultimately spells out as *ära*) is merged with the structure, and it too has unvalued ϕ -features. It searches its c-command domain for a suitable set of ϕ -features, which I assume it finds on the subject—which has moved to Spec, ΣP —for concreteness.²¹



We thus have two instances of ϕ -features in the same clause: on C^0 (*ära*) and Σ^0 (main verb).²²

This analysis succeeds in certain respects. First, it captures the fact that *ära* is restricted to imperatives, as *ära* selects only ΣP headed by $\Sigma^0_{[IMP]}$. Second, it captures the fact that *ära* and the main verb must bear the same set of ϕ -features, as they acquire their ϕ -features from the same source. However, what is not immediately clear from this analysis is why *ära*'s ϕ -feature marking is from the imperative paradigm rather than the indicative paradigm (or some other paradigm entirely). It cannot be because *ära* is string-adjacent

²¹ The particulars of this analysis would not change if the goal was Σ^0 or ΣP , as in the analyses of Baker and Willie (2010) and Henderson (2006).

²² At this point, this alternative analysis is quite similar to the analyses of multiple ϕ -feature agreement in some Niger-Congo languages: see Carstens (2001) and Henderson (2006) for discussion of Swahili (Bantu) and Baker and Willie (2010) for Ibibio (non-Bantu).

to $\Sigma^0_{[IMP]}$ (e.g., as a kind of contextual allomorphy) because *ära* and the main verb can be separated, as in (48) and (49).

- (48) *Ja är-ge te öue-s kukku-ge, libe on.*
 and NEG.IMP-2PL you.PL outside-INE fall-IMP.2PL slippery be.3.PRS
 ‘And don’t you fall down outside. It’s slippery.’ (BALANCED)
- (49) *Är-ge te seda arva-ke ega loot-ke.*
 NEG.IMP-2PL you.PL that.PAR think-IMP.2PL or hope-IMP.2PL
 ‘Don’t you think or hope that.’ (BALANCED)

In these examples *ärge* is separated from the main verb(s), suggesting that an analysis based on linear adjacency would run into trouble with a broader range of data.

A more promising possibility for *ära*’s imperative marking is that *ära* is associated not only with $[u\phi]$, but with its own imperative feature $[IMP]$. With this additional stipulation, there would be two components that are used to account for the imperative characteristics of *ära* under this analysis. It would be restricted to imperative clauses because it only selects ΣP with imperative features. Its inflection would be from the imperative set because it is associated with its own imperative feature.

In this respect, the analysis proposed in Section 3 is superior, as the imperative aspects of *ära* are tied to the same claim. In that analysis, *ära* is the spell-out of a complex head containing $\text{Pol}_{[NEG]}^0$ (with ϕ -features) and $C_{[IMP]}^0$. Because this $C_{[IMP]}^0$ is the source of imperative-marking under this analysis, we capture the fact that *ära* only surfaces in imperative clauses. It is that same feature that is responsible for the fact that the ϕ -feature inflection is from the imperative paradigm. Thus, though *ära*’s restriction to imperatives must be stipulated in either account, the alternative account currently under discussion would require two separate stipulations to capture all of *ära*’s imperative properties, but the account proposed in Section 3 requires only one.

However, a bigger concern for the account currently under consideration is that there is no clear path to generate the optionality of $[1PL]$ doubling. Recall that $[1PL]$ can be dropped on the main verb, but importantly, this is only true in negated imperatives. In affirmative imperatives, the verb must bear agreement, even in the case of $[1PL]$ agreement. Leaving off $[1PL]$ agreement in affirmative imperatives is ungrammatical, as shown in (50).

- (50) * *(Me) läbe sinna!*
 we go.CNG there
 Intended: ‘Let’s go there!’

Under this account, both the main verb and *ära* establish Agree relationships and acquire ϕ -features, and there is no formal link between *ära*’s agreement and the main verb’s (i.e., Σ^0 ’s) agreement. The main verb has to be able to see that *ära* is present, and when it is, the verb can drop its $[1PL]$ inflection, informally speaking.

This hypothetical operation is reminiscent of the Distributed Morphology operation Impoverishment, so it is worth taking a moment to consider such an approach. As we will see, the operation required would be significantly more powerful than existing formalisms for Impoverishment.

4.2 Optional first-person plural agreement is not Impoverishment

Broadly speaking, Impoverishment removes features that were present in the syntactic representation before Vocabulary Insertion can take place. Informally, Impoverishment in this case would remove [1PL] features from $\Sigma^0_{[IMP]}$ just in case it is in a structure with a C^0 with [IMP], i.e., *ära*.²³ This way of looking at [1PL] optionality is essentially the reverse of how I framed it in Section 3. Here, agreement doubling always happens, but agreement on the main verb can be Impoverished if it is [1PL]. In the analysis I proposed, doubling is the norm, but the doubling operation can be ignored or might not apply in the context of [1PL].

At first glance, the Feature Copying operation that I propose seems just as arbitrary as the Impoverishment operation described here. Unlike Feature Copying, Impoverishment has received a fair amount of attention in the literature, largely focused on how exactly it should be constrained. Based on the existing literature on Impoverishment, it seems to me that the Impoverishment-like operation that this alternative account requires would be outside the bounds of our current understanding of Impoverishment.

It has been proposed that rules of Impoverishment are not arbitrary, but emerge from cross-linguistically supported scales of markedness (see, e.g., Arregi and Nevins 2012; Keine and Müller 2014). It is not clear how [1PL] could be considered more marked than other sets of ϕ -features, even in the context of imperatives. For example, in the system developed by Harley and Ritter (2002), first-person plural must be less-marked than second-person plural: it contains the default [SPEAKER] feature whereas second-person plural comes pre-specified with the feature [ADDRESSEE].²⁴ Unlike the first-person plural, second-person plural must always be marked on the main verb in a negated imperative (see (18b)). Thus, it seems unlikely that the optionality of [1PL] agreement on the main verb in Estonian negated imperatives has anything to do with cross-linguistic markedness.

It has also been proposed that Impoverishment rules refer only to isolated feature bundles, not to the contexts in which those feature bundles appear (see, e.g., Keine 2010; Müller 2007). Thus, for example, a gender feature might be deleted just in case it is in a feature bundle with [PL], capturing the fact that gender features are neutralized in the plural for some languages (Corbett 1991; Kramer 2015). Along these lines, one might suggest that [1PL] is deleted in Estonian in the context of [IMP]. However, this would be too strong: in affirmative imperatives, where [1PL] and [IMP] can readily appear in the same feature bundle, agreement is obligatory, as we just saw in (50). The [1PL] *-me* can only be deleted when the higher C^0 that ultimately spells out as *ära* is present—in other words, it is not triggered by the mere presence of [1PL] and [IMP] in the same feature bundle.

Finally, Arregi and Nevins (2012) propose that Impoverishment rules may refer to more than one feature bundle, but those bundles must be at least in the same M-word (roughly, the same complex head). In order for this to be applicable in Estonian, *ära* and the main verb would have to be one M-word. However, this seems unlikely, as *ära* and the

²³ Recall that, under this alternative analysis, *ära* is not associated with the morphosyntactic feature [NEG].

²⁴ The feature [SPEAKER] is not always a default. Harley and Ritter (2002) propose that [SPEAKER] is always fully specified in languages with an inclusive/exclusive distinction for first-person plural. Estonian does not have this distinction, and so [SPEAKER] must be supplied as a default in Estonian. As a result, it does not count for Harley and Ritter's node-counting metric for markedness.

main verb can be separated, as seen in (48)–(49). Additional examples are shown in (51) and (52).

- (51) *Ära homme tule!*
 IMP.NEG tomorrow come.IMP.2SG
 ‘Don’t come tomorrow!’ (Erelt 2003, 111)
- (52) *Är-me selle-st rohkem räägi!*
 IMP.NEG-1PL this-ELA more speak
 ‘Let’s not speak about this anymore!’ (EKSS, entry for *ära*)

In (51), *ära* is separated from the main verb *tule* ‘come’ by *homme* ‘tomorrow’, and in (52), they are separated by two words. Significantly, (52) involves a main verb with no agreement exponence, which is exactly the environment where the putative rule of Impoverishment would have to apply.

Thus, it seems unlikely that the optionality of [1PL] agreement on the main verb in negated imperatives is to be explained by Impoverishment. It would involve deletion of features that are not driven by markedness in any straightforward way, and it would require a greater amount of syntactic context than has previously been included in the domain of Impoverishment.²⁵

4.3 Alternative: prospects & summary

The syntactic account presented in this section eschews DM postsyntactic operations in favor of syntactic Agree relations. This fits with a general theory of clausal agreement wherein more than one head in a clause can serve as a ϕ -feature probe (Baker and Willie 2010; Carstens 2001; Henderson 2006; Kalin and van Urk 2015; Polinsky 2016). An even stronger version of this theory is one in which Agree underlies all forms of morphological agreement, and all instances of syntactic Agree result in morphological agreement.²⁶ This would, of course, provide a transparent mapping between the syntax and morphology, at least as far as agreement is concerned, and this would be a positive result.

However, the analysis struggles to account for the full range of patterns in negated imperatives. Because the groundwork for extended exponence is laid at the start of the derivation, it is difficult to produce the examples where [1PL] agreement is not doubled. There is no formal connection made between the ϕ -feature probe of *ära* and the ϕ -feature probe of the main verb. I conclude that the account that makes use of morphological operations is the more promising approach of the two. At a minimum, it is able to produce the attested forms: first-person plural agreement on the main verb is optional only in the context of *ära* because (i) it is $C_{[IMP]}^0$ (*ära*) that passes the features to the main verb,

²⁵ To be sure, the Impoverishment approach is not inconceivable. It may be that future research on Impoverishment finds reason for it to be less constrained than its current form. In that case, it would be worth revisiting the alternative analysis presented here to see if it can be subsumed under the umbrella of Impoverishment. I thank an anonymous reviewer for helpful discussion of this point.

²⁶ Polinsky (2016) does not adopt the strong version of Agree, wherein Agree underlies all forms of morphological agreement. In particular, she argues that nominal concord in Archi, which is unquestionably a form of morphological agreement, does not arise from Agree. Instead, she advocates for the view that concord is a fundamentally different operation from argument-predicate agreement (see also den Dikken 2006; Kramer 2009; Norris 2014).

and (ii) the operation is postsyntactic, meaning it can be value-sensitive. Furthermore, my analysis locates plausibly general facts about (Uralic) imperatives in the syntax, and it locates Estonian-specific facts in the morphology, which I take to be a desirable result (see also Chung 2013, 2014).

The alternative account also does not fare as well in another domain: the patterns of agreement and imperative-marking in other Uralic languages. In this alternative account, both instances of imperative-marking and agreement are syntactically real— that is, they are both connected to features in the syntax. However, there are a number of Uralic languages where the main verb bears no imperative marking nor inflection in negated imperatives, suggesting the imperative-marking on the negative auxiliary is in some sense more robust than that on the main verb. From the perspective of the analysis I argue for, that is because it is only the imperative-marking on the negative auxiliary that is connected to the syntactic imperative feature. Let us now turn to a discussion of the patterns in other languages.

5 Negative Imperatives across Uralic

The analysis I argue for contains both Estonian-specific and more general proposals. On the general side, the analysis holds that the syntactic locus of imperativity is C^0 . Thus, in a language like Estonian, where imperativity is expressed twice in negated imperatives, it is the imperativity on the negative auxiliary that corresponds to real (that is, syntactic) imperative features. From the perspective of this analysis, the doubled imperative marking on the verb is special. Also on the general side, the analysis holds that heads other than T^0 may be the main ϕ -feature probe in a language, though this much is already assumed in much research on Finnish (Holmberg and Nikanne 2002; Holmberg et al. 1993; Mitchell 1991).

On the Estonian-specific side, the syntax of Estonian involves head movement of $\text{Pol}_{[\text{NEG}]^0}$ to $\text{C}_{[\text{IMP}]^0}$ followed by Fusion, which results in a negative auxiliary that is particular to imperative clauses. This much is shared by some (not all) Uralic languages. Further, Estonian has a language-particular rule of Feature Copying, which copies features from C^0 to T^0 . Again, some (not all) Uralic languages share this property.

In this section, I discuss the picture from other languages within the Uralic family, focusing on languages that have negative auxiliaries at least some of the time. I show that peeling away some of the Estonian-specific operations predicts the existence of other patterns that are attested in other Uralic languages. I also show that, in virtue of the general proposal that imperativity is high, there is a pattern of imperativity and ϕ -feature marking that is predicted to be nonexistent. Based on the sample in Miestamo et al. (2015), this prediction is borne out for Uralic. Thus, the analysis presented here receives additional support from other Uralic languages.

5.1 Languages (almost) like Estonian

Vod, Livonian, Skolt Saami, and Finnish exhibit very similar morphology in their negated imperative clauses. They have a special negative auxiliary just for negated imperatives. In

addition, the main verb reflects imperativity in negated imperatives, and in Livonian and Vod, the main verb also reflects agreement.

- (53) Livonian: doubled imperative + agreement (Metslang et al. 2015, 440)
- a. *alā* *and-õ*
 IMP.NEG.2SG give-IMP.2SG
 ‘Don’t give’
- b. *al-gid* *anda-gid*
 NEG.IMP-2PL give-IMP.2PL
 ‘Don’t give’
- (54) Vod: doubled imperative + agreement (Rozhanskiy and Markus 2015, 494)
- a. *elä* *näe*
 IMP.NEG.2SG see.IMP.2SG
 ‘Don’t see!’
- b. *elka* *näbka*
 IMP.NEG.2PL see.IMP.2PL
 ‘Don’t see!’
- (55) Finnish: doubled imperative (+ 2SG agreement) (Sulkala and Karjalainen 1992)
- a. *älä* *tule*
 IMP.NEG.2SG come.IMP.2SG
 ‘don’t come’
- b. *äl-kää* *tul-ko*
 NEG.IMP-2PL come-IMP.CNG
 ‘don’t come’
- (56) Skolt Saami: doubled imperative only (Miestamo and Koponen 2015, 360)
- a. *jeä’lled* *porru/poor*
 IMP.NEG.2PL eat.IMP.CNG/CNG
 ‘Don’t eat!’
- b. *jeällap* *porru*
 IMP.NEG.1PL eat.IMP.CNG
 ‘Let’s not eat!’

In terms of the analysis proposed here, these languages would all have rules of Feature Copying, but they would have to be slightly different from Estonian’s. Doubling in Vod and Livonian is always obligatory. In Skolt Saami, Feature Copying copies only the imperative feature, but that copying is optional for [2PL].²⁷ When the imperative feature is not copied, then the main verb surfaces in its ordinary connegative form, just as in [1PL] contexts for Estonian. In Finnish, in most cases, the main verb only reflects imperativity (with the suffix *-ko*) and not ϕ -features— the one exception is second-person singular. Thus, it is like Skolt Saami in that most scenarios involve copying only the imperative feature, but slightly different in that it requires copying ϕ -features in second-person singular. These patterns suggest a degree of arbitrariness among the languages with doubling of some

²⁷ Alternatively, in line with the analysis of Estonian *ei*, we could propose that all the features are copied, but the ϕ -features do not affect vocabulary insertion in Skolt Saami.

kind, reminiscent of the kinds of seemingly arbitrary restrictions on PF operations for the varying dialects of Basque uncovered by Arregi and Nevins (2012).

5.2 Languages with no Feature Copying

There are also multiple Uralic languages which reflect imperativity only on the negative auxiliary and not on the main verb. Instead, the main verb in a negated imperative is indistinguishable from the form of a verb in a negated indicative clause. This is observable in Erzya, Forest Enets, Mari, North Saami, Pite Saami, South Saami, and Tundra Nenets. I present only the imperative examples.

(57) Erzya (Hamari and Aasmäe 2015, 299)

a. *il'a učo!*
NEG.IMP.2SG wait.CNG
'do not wait!'

b. *il'a-do učo!*
NEG.IMP-2PL wait.CNG
'do not wait!'

(58) Forest Enets²⁸ (Siegl 2015, 50)

iđ d'ori-r
NEG.IMP.2SG speak-FRQ.CNG
'don't speak'

(59) Eastern Mari (Saarinen 2015, 335)

a. *i-t tol*
IMP.NEG-2SG come.CNG
'don't come'

b. *i-da tol*
IMP.NEG-2PL come.CNG
'don't come'

(60) North Saami (Nickel 1990, 61)

a. *ale boađe*
IMP.NEG.2SG come.CNG
'don't come'

b. *allet boađe*
IMP.NEG.2PL come.CNG
'don't come'

²⁸ The imperative system of Forest Enets is complex, but Siegl (2015) explains it very carefully. The only person and number combination that has a distinct form of imperative negation is second-person singular. Other persons and numbers use the same negative auxiliary as indicatives and thus, negated imperatives and negated indicatives are formally indistinct for those persons and numbers.

- (61) Pite Saami (Wilbur 2014, 160, 181)²⁹
- a. *ele tsábme!*
NEG.IMP.2SG hit.CNG
'Don't hit!' (said to a child)
- b. *ellet/illut*
NEG.IMP.2PL
(no translation)
- (62) South Saami³⁰ (Blokland and Inaba 2015, 382)
- a. *aellieh bǎetieb*
IMP.NEG.2SG come.CNG
'don't come'
- b. *aellebe bǎetieb*
IMP.NEG.2PL come.CNG
'don't come'
- (63) Tundra Nenets (Eastern Dialect, Taymyr Subdialect) (Mus 2015, 80–81)
- a. *šimi ńo-n xaada-?!*
1SG.ACC NEG.IMP-2SG kill-CNG
'Do not kill me!'
- b. *tańa ńo-xo-ńi? xań-?!*
there NEG.IMP-HORT-2DU go-CNG
'Do not go there!'

In terms of my analysis, these languages all involve creation of a complex head comprising Pol⁰ and C⁰. This creates the negative auxiliary that is particular to imperative clauses. However, unlike Estonian, these languages do not have an additional rule of Feature Copying. Instead, the verb surfaces in its ordinary connegative form.

These languages also provide an argument in favor of the high origin of imperative features. In Section 2.1, I noted that Nelson (1998) proposed that imperative features in Finnish were generated lower (in T⁰). This was her analysis for the Finnish morpheme *-ko* that appears on the main verb in all negated imperatives except second-person singular. Nelson does state that there is a special negative auxiliary for imperatives (p. 171), but she does not specify how these morphological forms are generated. However, given that imperativity is otherwise lower in her analysis, it seems that the special negative imperative auxiliary is the unexpected part of the construction, with the main verb bearing the syntactic imperative feature. The languages of this class suggest that that is not right— it is the negative auxiliary that bears the main marker of imperativity, and imperative marking is copied onto the main verb in some languages.

²⁹ Wilbur provides only one complete example of a negated imperative. However, he does provide forms for the imperative negative auxiliary, and discussion surrounding negation (pp. 159–60) suggests that the main verb is identical in indicative and imperative negation, which means that it does not inflect for imperative features.

³⁰ South Saami has two negated imperative auxiliaries. One is used for prohibitives (the forms presented here), and one is used for warnings. Blokland and Inaba (2015) call the latter forms APPREHENSIVES, and they behave identically to prohibitives as far as the data given here are concerned.

5.3 A gap: No imperative marking on the negative auxiliary

The observations seen so far are summarized in Table 3 below.

<i>Negation</i>	<i>Main verb</i>	<i>Languages</i>
[IMP], [ϕ]	[IMP], [ϕ]	Votic, Estonian, Livonian
[IMP], [ϕ]	[IMP]	Finnish, Skolt Saami
[IMP], [ϕ]	\emptyset	Erzya, Forest Enets, Eastern Mari, North Saami, Pite Saami, South Saami
[ϕ]	[IMP], [ϕ]	Unattested
[ϕ]	[IMP]	Unattested

Table 3: *Imperative and ϕ -feature marking in Uralic imperative clauses*

The table focuses on just those Uralic languages where negation bears agreement in imperative clauses. The top three rows show the languages just discussed, where negation inflects for imperative features and ϕ -features. The bottom two rows present two kinds of languages that are unattested in Miestamo et al. 2015. In these patterns, imperative features are reflected only on the main verb, not on the negative auxiliary. A toy example based on Finnish is presented in (65), with the indicative in (64) for reference.

(64) *e-tte tule*
NEG-2PL come
'You (PL) do not come.'

(65) *e-tte tul-ko*
NEG-2PL come-IMP
'Don't come.' (Hypothetical Finnish)

In standard Finnish, the main verb does reflect imperativity in negated imperatives, but so does negation: it is *älkää* rather than the indicative *ette* as in the hypothetical Finnish example in (65).

This gap within the typology of Uralic negation (so far as I know) is predicted by my analysis, and it connects back to the question of which imperative markers correspond to imperative features in the syntax. In my analysis, the source of imperative-marking on the negative auxiliary is the syntactic head bearing imperative features, i.e., C^0 , located high in the clausal structure. The negative auxiliary comes to reflect imperative features because it undergoes head movement to C^0 . Thus, the imperative marking on the negative auxiliary is connected to the syntactic location of imperative in the clause.

In contrast, the imperative-marking on the main verb is not connected directly to the syntactic locus of imperativity. Rather, it is connected indirectly via the imperative features of the negative auxiliary. Because the negative auxiliary (i.e., negation) intervenes between the imperative C^0 and the main verb, there is no way for the verb to raise to C^0 without passing the negative auxiliary. Instead, it is the negative auxiliary that is associated with imperativity in the syntax, and these imperative features are sometimes passed to the main verb. This is essentially the same explanation for the languages whose indicative

connegative and imperative connegative forms are identical: the imperative marking on the negative auxiliary is syntactically real, and the marking on the main verb (when present) is redundant morphology.

At this point, I would like to note that the alternative analysis presented in Section 4 does not straightforwardly predict this same gap. Recall that, under that analysis, the negated imperative is formed from by merging the imperative negative auxiliary (C^0 / *ära*) with an affirmative imperative clause (ΣP headed by $\Sigma^0_{[IMP]}$). The imperative features and ϕ -feature probes associated with C^0 and Σ^0 are independent of each other; just as it is possible for both to be associated with $[IMP]$ and $[\mu\phi]$, it should be possible for one to be associated with $[IMP]$ and the other with $[\mu\phi]$, all else being equal. Thus, Σ^0 could be associated with $[IMP]$, and C^0 could be associated with $[\mu\phi]$, and the result would be very similar to the proposed gap: a negation word with non-imperative agreement, and a main verb with imperative-marking but no agreement.³¹

6 Conclusion

In this paper, I have argued for a straightforward syntax for agreement in Estonian negated clauses. $\text{Pol}_{[NEG]}^0$ is always the only ϕ -feature probe, even when morphological agreement is realized twice or not at all. The upshot of this claim is that the syntax of agreement in Estonian need not be radically different from its neighbors. The analysis presented here is also an argument in support of a general theory of agreement whereby morphological agreement and the syntactic relation Agree do not track each other directly (Chung 1998, 2013, 2014; Polinsky 2016; Sigurðsson 2004). Instead, there are cases of syntactic Agree that are not ultimately realized by morphological agreement, and there are instances of morphological agreement that are not (directly, at least) tied to a syntactic Agree relation.

There are a number of puzzles for agreement exponence inside Estonian and inside Uralic not investigated here that could be interesting domains for future research. First, in the Estonian conditional, the verb agrees only optionally, as shown in (66) and (67).

(66) *mina ela-ksi-n* / *ela-ks*
 I live-COND-1SG live-COND
 ‘I would live’ (Erelt et al. 2000, 282)

(67) *meie ela-ksi-me* / *ela-ks*
 we live-COND-1PL live-COND
 ‘we would live’ (Erelt et al. 2000, 282)

For example, in (66), the verb can either bear agreement with the first-person singular subject agreement (*elaksin*), or it can appear without agreement (*elaks*). This is not only part of the spoken language, but it is also acceptable in the written standard. Erelt et al. (2000) note that the forms without agreement are more common than the forms with endings in the spoken language. According to the framework laid out here, agreeing and non-agreeing

³¹ In truth, the language predicted by the alternative proposal is not exactly the gap, because standard negation in that proposal would be the spell-out of Σ^0 , not C^0 . Thus, it predicts the presence of a special negator which nevertheless uses the standard indicative agreement. This kind of language is also not attested, and it cannot be generated by the analysis I proposed in Section 3.

conditionals would both involve an Agree relationship, but it would not always be realized morphologically.

Next, when we look at other Uralic languages, there are instances of multiple agreement exponence outside of imperatives. For example, in standard Finnish, the form known as the past participle inflects for number.³² This is true in both auxiliary constructions and negative constructions, where the past participle is used as a past indicative connegative form.

- (68) *Lapset o-vat luke-neet tämän kirjan.*
 children be-3PL read-PST.PCPL.PL this book
 ‘The children have read this book.’ (Holmberg et al. 1993, 200)
- (69) *Lapset ei-vät ol-leet luke-neet tätä kirjaa.*
 children NEG-3PL be-PST.PCPL.PL read-PST.PCPL.PL this book
 ‘The children had not read this book.’ (Holmberg et al. 1993, 200)

In (68), we see the past participle *lukeneet* ‘read’ in an auxiliary construction. In (69), there are two past participles: one that emerges as a result of the auxiliary, the other because of negation. In each case, the participle reflects the number of the subject: plural *-neet* as opposed to singular *-nut*. It is worth considering whether these forms could be amenable to the kind of analysis investigated here.

Outside of Finnic, there are languages with indicative connegatives that agree with the subject. For example, in Komi, the indicative connegative verb inflects in the plural.

- (70) Komi (Hamari 2015, 241)
- a. *o-g mun*
 NEG.PRS-1SG go.CNG
 ‘I don’t go’
- b. *o-gɛ(j) mun-ɛ(j) / o-g mun-ɛ(j)*
 NEG.PRS-1PL go.CNG-1/2PL / NEG.PRS-1SG go.CNG-1/2PL
 ‘We don’t go’
- c. *o-z mun*
 NEG.PRS-3 go.CNG
 ‘S/he doesn’t go’
- d. *o-z mun-nɨ*
 NEG.PRS-3 go.CNG-3PL
 ‘They don’t go’

The connegative verb clearly expresses number: compare first-person singular (70a) with plural (70b). In the first- and second-person plural, plurality can also optionally be reflected on negation, giving this the appearance of multiple exponence. This is certainly another domain that could raise interesting questions for the interaction between the syntax and morphology of agreement.³³

³² Holmberg et al. (1993) note that past participles do not inflect in most spoken varieties of the language. In these varieties, the singular form *-nut* is always used.

³³ Udmurt exhibits a very similar pattern to the Komi pattern in the present tense (Edygarova 2015: 267). It also exhibits a similar but simpler pattern in the future and 1st preterite, where the connegative varies

As I have mentioned, it is well-known that negation and imperativity are in some sense difficult to express in combination (van der Auwera et al. 2013). However, some languages express negated imperatives with morphosyntactic ease, suggesting this is only a tendency, not a universal. These languages make up 22.8% of the languages in the sample provided by van der Auwera et al. (2013). When the incompatibility of negation and imperativity manifests, its effects do not always have the same appearance. For example, some languages use a standard negation word combined with a non-imperative verbal form (e.g., Italian, Zanuttini 1994, 1997), others may use a special negation word but a normal imperative (like Estonian), and others may use both simultaneously. The analysis proposed here holds that negation and imperativity are in fact fully syntactically compatible in Estonian, but this is obscured by their surface forms. It is thus worth investigating other cases of apparent incompatibility to see if the incompatibility is deeply ingrained in the grammar of the language or, as I have argued for Estonian, a surface morphological effect.

References

- Arregi, Karlos and Andrew Nevins (2012). *Morphotactics: Basque auxiliaries and the structure of spellout*. Dordrecht: Springer.
- Baker, Mark and Willie Udo Willie (2010). "Agreement in Ibibio: From every head to every head". *Syntax* 13.2, 99–132.
- Bjorkman, Bronwyn and Hedde Zeijlstra (2015). *Upward Agree is superior*. Presented at the 33rd West Coast Conference on Formal Linguistics (WCCFL 33).
- Blevins, James P. (2007). "Conjugation classes in Estonian". *Linguistica Uralica* XLIII, 250–267.
- Blokland, Rogier and Nobufumi Inaba (2015). "Negation in South Saami". In: *Negation in Uralic Languages*. Ed. by Miestamo, Matti, Tamm, Anne, and Wagner-Nagy, Beáta. Amsterdam/Philadelphia: John Benjamins Publishing Company, 377–398.
- Brattico, Pauli and Saara Huhmarniemi (2006). "Finnish negation, the EPP feature and the valuation theory of morphosyntax". *Nordic Journal of Linguistics* 29 (1), 5–44.
- Brattico, Pauli, Saara Huhmarniemi, Jukka Purma, and Anne Vainikka (2014). "The structure of Finnish CP and feature inheritance". *Finno-Ugric Languages and Linguistics* 2 (2), 66–109.
- Carstens, Vicki (2001). "Multiple agreement and case deletion: Against ϕ -incompleteness". *Syntax* 4.3, 147–163.
- Chomsky, Noam (1995). *The minimalist program*. Cambridge, MA: MIT press.
- Chomsky, Noam (2000). "Minimalist inquiries: The framework". In: *Step by step: Essays on minimalism in honor of Howard Lasnik*. Ed. by Martin, Roger, Michaels, David, and Uriagereka, Juan. Cambridge, MA: MIT Press, 89–156.
- Chomsky, Noam (2001). "Derivation by Phase". In: *Ken Hale: A Life in Language*. Ed. by Kenstowicz, Michael. Cambridge, MA: MIT Press, 1–52.
- Chung, Sandra (1998). *The design of agreement: Evidence from Chamorro*. Chicago: University of Chicago Press.

only for number but not for person. Another similar but simpler pattern exists in Western Mari, where the indicative connegative bears an additional suffix only in the third-person plural (Saarinen 2015: 330).

- Chung, Sandra (2013). "The syntactic relations behind agreement". In: *Diagnosing Syntax*. Ed. by Cheng, Lisa Lai-Shen and Corver, Norbert. Oxford: Oxford University Press.
- Chung, Sandra (2014). "On reaching agreement late". In: *CLS 48, Volume 1*. Ed. by Beltrama, Andrea, Chatzikonstantinou, Tasos, Lee, Jackson L., Pham, Mike, and Rak, Diane. Chicago: Chicago Linguistic Society, 169–190.
- Corbett, Greville (1991). *Gender*. Cambridge: CUP.
- den Dikken, Marcel (2006). "Agreement". In: *Encyclopedia of Cognitive Science*.
- Edygarova, Svetlana (2015). "Negation in Udmurt". In: *Negation in Uralic Languages*. Ed. by Miestamo, Matti, Tamm, Anne, and Wagner-Nagy, Beáta. Amsterdam/Philadelphia: John Benjamins Publishing Company, 265–291.
- Embick, David (2010). *Localism versus globalism in morphology and phonology*. Cambridge, MA: MIT Press.
- Erelt, Mati, ed. (2003). *Estonian language*. Tallinn: Estonian Academy Publishers.
- Erelt, Mati, Reet Kasik, Helle Metslang, Henno Rajandi, Kristiina Ross, Henn Saari, Kaja Tael, and Silvi Vare (1993). *Eesti Keele Grammatika II: Süntaks; Lisa: Kiri* [Estonian Grammar II: Syntax; Appendix: Written Language]. Tallinn, Estonia: Eest Teaduste Akadeemia Keele ja Kirjanduse Instituut.
- Erelt, Mati, Tiiu Erelt, and Kristiina Ross (2000). *Eesti keele käsiraamat* [Estonian language handbook]. Tallinn: Eesti keele sihtasutus.
- Feist, Timothy (2010). "A Grammar of Skolt Saami". PhD thesis. University of Manchester.
- Grimshaw, Jane (1991/2005). *Words and structure*. CSLI Publications.
- Halle, M. (1990). "An approach to morphology". In: *Proceedings of NELS 20*. University of Massachusetts, Amherst: GLSA, 150–184.
- Halle, Morris (1997). "Distributed Morphology: Impoverishment and Fission". *MITWPL* 30, 425–449.
- Halle, Morris and Alec Marantz (1993). "Distributed Morphology and the Pieces of Inflection". In: *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*. Ed. by Hale, K. and Keyser, S. J. Cambridge, MA: MIT Press, 111–176.
- Hamari, Arja (2015). "Negation in Komi". In: *Negation in Uralic Languages*. Ed. by Miestamo, Matti, Tamm, Anne, and Wagner-Nagy, Beáta. Amsterdam/Philadelphia: John Benjamins Publishing Company, 239–263.
- Hamari, Arja and Niina Aasmäe (2015). "Negation in Erzya". In: *Negation in Uralic Languages*. Ed. by Miestamo, Matti, Tamm, Anne, and Wagner-Nagy, Beáta. Amsterdam/Philadelphia: John Benjamins Publishing Company, 293–324.
- Han, Chung-Hye (1999). "Cross-linguistic variation in the compatibility of negation and imperatives". In: *Proceedings of West Coast Conference in Formal Linguistics*. Vol. 17, 265–279.
- Hankamer, Jorge and Line Mikkelsen (2005). "When movement must be blocked: A reply to Embick and Noyer". *Linguistic Inquiry* 36.1, 85–125.
- Harley, Heidi and Rolf Noyer (1999). "State-of-the-article: Distributed Morphology". *GLOT international* 4 (4), 3–9.
- Harley, Heidi and Elizabeth Ritter (2002). "Person and number in pronouns: A feature-geometric analysis". *Language* 78.3, 482–526.

- Henderson, Brent (2006). "Multiple agreement and inversion in Bantu". *Syntax* 9.3, 275–289.
- Holmberg, Anders and Urpo Nikanne (2002). "Expletives, subjects, and topics in Finnish". In: *Subjects, Expletives, and the EPP*. Ed. by Svenonius, Peter. Oxford: Oxford University Press, 71–105.
- Holmberg, Anders, Urpo Nikanne, Irmeli Oraviita, Hannu Reime, and Trond Trosterud (1993). "The structure of INFL and the finite clause in Finnish". In: *Case and other functional categories in Finnish syntax*. Ed. by Holmberg, A. and Nikanne, U. Berlin: Mouton de Gruyter, 177–206.
- Huhmarniemi, Saara (2012). "Finnish A'-movement: Edges and islands". PhD thesis. University of Helsinki.
- Kalin, Laura and Coppe van Urk (2015). "Aspect splits without ergativity". *Natural Language & Linguistic Theory* 33.2, 659–702.
- Keine, Stefan (2010). *Case and agreement from fringe to core: A minimalist approach*. Linguistische Arbeiten. Berlin: Walter de Gruyter.
- Keine, Stefan and Gereon Müller (2014). "Differential argument encoding by impoverishment". In: *Scales and Hierarchies*. Ed. by Bornkessel-Schlesewsky, I., Malchukov, A., and Richards, M. Trends in Linguistics. Berlin: Mouton de Gruyter, 75–130.
- Kiparsky, Paul (2001). "Structural case in Finnish". *Lingua* 111.4, 315–376.
- Koskinen, Päivi (1998). "Features and categories: Non-finite constructions in Finnish". PhD thesis. University of Toronto.
- Kramer, Ruth (2009). "Definite markers, phi-features, and agreement: A morphosyntactic investigation of the Amharic DP". PhD thesis. University of California, Santa Cruz.
- Kramer, Ruth (2010). "The Amharic definite marker and the syntax-morphology interface". *Syntax* 13, 196–240.
- Kramer, Ruth (2015). *The morphosyntax of gender: evidence from Amharic*. Oxford: Oxford University Press.
- Laka, Itziar (1990). "Negation in syntax: On the nature of functional categories and projections". PhD thesis. Massachusetts Institute of Technology.
- Merchant, Jason (2015). "How much context is enough? Two cases of span-conditioned stem allomorphy". *Linguistic Inquiry* 46 (2), 273–303.
- Metslang, Helle, Karl Pajusalu, and Tiit-Rein Viitso (2015). "Negation in Livonian". In: *Negation in Uralic Languages*. Ed. by Miestamo, Matti, Tamm, Anne, and Wagner-Nagy, Beáta. Amsterdam/Philadelphia: John Benjamins Publishing Company, 433–456.
- Miestamo, Matti (2005). *Standard negation: the negation of declarative verbal main clauses in a typological perspective*. Berlin: Mouton de Gruyter.
- Miestamo, Matti and Eino Koponen (2015). "Negation in Skolt Saami". In: *Negation in Uralic Languages*. Ed. by Miestamo, Matti, Tamm, Anne, and Wagner-Nagy, Beáta. Amsterdam/Philadelphia: John Benjamins Publishing Company, 353–376.
- Miestamo, Matti, Anne Tamm, and Beáta Wagner-Nagy, eds. (2015). *Negation in Uralic languages*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Miljan, Merilin and Ronnie Cann (2013). "Rethinking case marking and case alternation in Estonian". *Nordic Journal of Linguistics* 36 (03), 333–379.

- Mitchell, Erika (1991). "Evidence from Finnish for Pollock's theory of IP". *Linguistic Inquiry* 22 (2).
- Mitchell, Erika (2006). "The morpho-syntax of negation and the positions of NegP in the Finno-Ugric languages". *Lingua* 116.3, 228–244.
- Müller, Gereon (2007). "Extended exponence by enrichment: argument encoding in German, Archi, and Timucua". In: *Proceedings of the 30th Annual Penn Linguistics Colloquium*. Ed. by Scheffler, Tatjana, Tauberer, Joshua, Eilam, Aviad, and Mayol, Laia, 253–266.
- Mus, Nikolett (2015). "Negation in Tundra Nenets". In: *Negation in Uralic Languages*. Ed. by Miestamo, Matti, Tamm, Anne, and Wagner-Nagy, Beáta. Amsterdam/Philadelphia: John Benjamins Publishing Company, 75–102.
- Nelson, Diane (1998). *Grammatical case assignment in Finnish*. New York & London: Garland.
- Nickel, Klaus Peter (1990). *Samisk grammatikk*. Oslo: Universitetsforlaget.
- Norris, Mark (2014). "A theory of Nominal Concord". PhD thesis. University of California, Santa Cruz.
- Polinsky, Maria (2016). "Agreement in Archi from a minimalist perspective". In: *Archi: complexities of agreement in a cross-theoretical perspective*. Ed. by Bond, Oliver, Corbett, Greville, Chumakina, Marina, and Brown, Dunstan. Oxford: Oxford University Press.
- Preminger, Omer (2014). *Agreement and its failures*. Vol. 68. Linguistic Inquiry Monographs. Cambridge, MA: MIT Press.
- Preminger, Omer and Maria Polinsky (2015). *Upwards and onwards: On the direction of valuation in ϕ -feature agreement*. Handout of a talk presented at New York University, April 25, 2015.
- Rivero, Maria Luisa and Arhonto Terzi (1995). "Imperatives, V-movement and logical mood". *Journal of Linguistics* 31.02, 301–332.
- Rozhanskiy, Fedor and Elena Markus (2015). "Negation in contemporary Votic". In: *Negation in Uralic Languages*. Ed. by Miestamo, Matti, Tamm, Anne, and Wagner-Nagy, Beáta. Amsterdam/Philadelphia: John Benjamins Publishing Company, 487–515.
- Saarienen, Sirkka (2015). "Negation in Mari". In: *Negation in Uralic Languages*. Ed. by Miestamo, Matti, Tamm, Anne, and Wagner-Nagy, Beáta. Amsterdam/Philadelphia: John Benjamins Publishing Company, 325–352.
- Schoorlemmer, Erik and Tanja Temmerman (2012). "Head movement as a PF-phenomenon: Evidence from identity under ellipsis". In: *Proceedings of the 29th West Coast Conference on Formal Linguistics*. Ed. by Choi, Jaehoon, Hogue, E. Alan, Punske, Jeffrey, Tat, Deniz, Schertz, Jessamyn, and Trueman, Alex. Somerville, MA: Cascadia Proceedings Project, 232–240.
- Siegl, Florian (2015). "Negation in Forest Enets". In: *Negation in Uralic Languages*. Ed. by Miestamo, Matti, Tamm, Anne, and Wagner-Nagy, Beáta. Amsterdam/Philadelphia: John Benjamins Publishing Company, 45–73.
- Sigurðsson, Halldór Ármann (2004). "Agree in syntax, agreement in signs". *Working Papers in Scandinavian Syntax* 74, 1–42.
- Sulkala, Helena and Merja Karjalainen (1992). *Finnish*. London and New York, NY: Routledge.

- Svenonius, Peter (2012). *Spanning*. Ms., CASTL, University of Tromsø.
- Tamm, Anne (2015). "Negation in Estonian". In: *Negation in Uralic Languages*. Ed. by Miestamo, Matti, Tamm, Anne, and Wagner-Nagy, Beáta. Amsterdam/Philadelphia: John Benjamins Publishing Company, 399–431.
- Toivonen, Ida and Diane Nelson, eds. (2007). *Saami linguistics*. Amsterdam: John Benjamins.
- van der Auwera, Johan, Ludo Lejeune, and Valentin Goussev (2013). "The prohibitive". In: *The World Atlas of Language Structures Online*. Ed. by Dryer, Matthew S. and Haspelmath, Martin. Leipzig: Max Planck Institute for Evolutionary Anthropology. Chap. 71. URL: <http://wals.info/chapter/71>.
- Wilbur, Joshua (2014). *A grammar of Pite Saami*. Berlin: Language Science Press.
- Zanuttini, Raffaella (1994). "Speculations on negative imperatives". *Rivista di Linguistica* 6 (1), 67–89.
- Zanuttini, Raffaella (1997). *Negation and clausal structure*. Oxford: Oxford University Press.
- Zanuttini, Raffaella (2008). "Encoding the addressee in the syntax: evidence from english imperative subjects". *Natural Language & Linguistic Theory* 26.1, 185–218.
- Zanuttini, Raffaella, Miok Pak, and Paul Portner (2012). "A syntactic analysis of interpretive restrictions on imperative, promissive, and exhortative subjects". *Natural Language & Linguistic Theory* 30.4, 1231–1274.

Mark Norris
University of Oklahoma
mark-norris@ou.edu

The partitive split in Finnish and Estonian^{*}

Saara Huhmarniemi & Merilin Miljan

Finnish and Estonian have several types of *split noun phrases* where a noun phrase is separated from the modifying quantifier or numeral. This paper provides a preliminary classification of split noun phrases in both languages and proposes a syntactic analysis of a specific type of split NP, the *partitive split*, where the noun phrase is in the partitive case. We propose that the partitive split is derived by discourse-related movement of the partitive NP. Particular attention is paid on contexts where the partitive noun phrase does not reconstruct to its position prior to movement. For example, numerals higher than one induce morphological mismatches in partitive split in both languages. A solution is proposed, where the partitive split involves an optionally pronounced classifier head, which facilitates the semantic selection and morphology. This analysis is shown to apply to Finnish, but the evidence for Estonian is not conclusive.

Keywords: *syntax, split NP, partitive split, Finnish, Estonian*

1 Introduction

This paper examines constructions where a noun phrase is separated from the quantifier or numeral that modifies it. The following examples from Finnish and Estonian illustrate the phenomenon. In (1), the noun *miehiä* ‘men’ occupies the position at the front, while the numeral that modifies it is at the end of the clause. The word order in (2), where the quantifier is at the front is also available, but less common in both languages.

(1) Finnish

Miehiä saapui paikalle viisi.
man.PL.PAR arrived place.to five.NOM
‘Five men arrived to the place.’

^{*} We would like to thank the audience of the workshop Syntactic structure of Uralic languages at the XII International Congress for Finno-Ugric Studies in Oulu, Finland, August 2015. Thank you to Professor Gisbert Fanselow for sharing his expertise and insightful discussions and to Anja Arnhold for the original idea of the paper. We would like to thank the participants of the Biolinguistics seminar organised by the Cognitive Science unit in the University of Helsinki; Pauli Brattico, Tommi Gröndahl, Jukka Purma and Taija Saikkonen; Mark Norris for valuable comments, Helle Metslang and all the informants. Thank you also to anonymous reviewers for insightful comments. The work was funded by a grant from the Kone foundation for the first author.

(2) Estonian

Palju nägi Peeter kasse.
 many saw Peeter.NOM cat.PL.PAR
 ‘Peeter saw many cats.’

Constructions where the noun phrase is divided into two parts are here referred to as *split noun phrases* (following Fanselow 1988; van Riemsdijk 1989). Split NPs are frequent among languages and a subject of variation even within one language. The split noun phrases that involve an NP in the partitive case are sometimes referred to as *quantifier clauses* (kvanttorilauseet).¹ The quantifier clause has been considered as a special clause type in Finnish (e.g. Hakulinen and Karlsson 1979: 97–99, Hakulinen et al. 2004: §902).²

An analysis is proposed where the quantifying expression and the partitive NP are initially part of the same constituent, as in (3a),³ and the split construction is formed by moving the partitive NP out, as in (3b).⁴ This approach is referred to as *sub-extraction* account for split noun phrases and it was proposed by van Riemsdijk (1989) for German.

- (3) a. *Minä näin* [_{DP}[_{QP} *paljon* [_{NP} *lintuja*]]].
 I.NOM saw a lot bird.PL.PAR
 ‘I saw a lot of birds.’
- b. [_{NP} *Lintuja*] *minä näin* [_{DP}[_{QP} *paljon* ____]].
 bird.PL.PAR I.NOM saw a lot
 ‘I saw a lot of **birds**.’

We will show in this paper that the partitive split shares the basic properties with discourse-related movement to the left periphery: it has the same triggers and landing sites (e.g. contrastive focus) and it follows the basic constraints on movement, i.e. islands. For example, in (3b), the noun phrase *lintuja* ‘birds’ occupies a position where it receives a discourse interpretation of contrast. We propose that the movement is an instance of *A'-movement*, which displaces elements from their thematic positions and positions where the case and agreement properties are assigned.

¹ Finnish has two object cases, accusative and partitive, which are here glossed as ACC and PAR. The accusative form is realised with different suffixes depending on the type of the NP. Pronouns have accusative case suffix *-t*, plural NPs have suffix *-t*, which is the same as in the nominative case. In addition, Finnish has an unmarked object case, which is here glossed as NOM. The same convention is adopted for numerals with unmarked case. For Estonian, we use a different, albeit traditional convention whereby the object is glossed as GEN, since there is no unique morphology which can be identified as ACC. Other cases are glossed as ablative case = ABL, adessive = ADE, genitive = GEN, elative = ELA, illative = ILL, inessive = INE. Semantic cases are in many examples glossed with English prepositions. Singular/plural inflection on nominal present is either present in the English translation or marked explicitly with SG and PL. The person and number agreement on finite verbs is glossed only when needed for clarity, e.g. first person singular verb inflection = 1SG. Conditional = COND, infinitive = INF, passive = PASS, PTCPL = participial, possessive suffixes = PX, Q=question particle.

² The same convention has been adopted for Estonian in Erelt et al. (2016), where similar constructions are referred to as *kvanttorilause*, ‘quantifier clause’. Spoelman (2013: 65) uses the term *quantifying sentence*. However, these constructions may occur in smaller domains, such as in adverbial clauses.

³ The quantifier *paljon* ‘a lot’ appears in the object position in unmarked case form and is not sensitive to aspectual object case variation. Therefore, the accusative case marking of *paljon* is suppressed in the glosses.

⁴ The bold typeface indicates contrastive focus. The contrastive focus is marked only in sentences where the contrastive reading is strongly preferred for the constituent.

However, the analysis in terms of sub-extraction faces a problem in the morphological mismatch between the quantifying expression and the NP. In both Estonian and Finnish, numeral-noun constructions such as (4a) and (5a), require the NP in the singular, see (4b) and (5b). In the split construction, the partitive noun phrase is in the plural, as in (4c) and (5c). This means that the NP cannot be ‘returned’ to the complement of the numeral (see also Hakulinen and Karlsson 1979, Seppänen 1983: 165–169, Vilkuna 1996, Hakulinen et al. 2004: §903).

(4) Finnish

- a. *Minä näin viisi lintua.*
I.NOM saw five.NOM bird.SG.PAR
‘I saw five birds.’
- b. **Minä näin viisi lintuja.*
I.NOM saw five.NOM bird.PL.PAR
- c. *Lintuja minä näin viisi.*
bird.PL.PAR I.NOM saw five.NOM
‘I saw five **birds**.’

(5) Estonian

- a. *Peeter ostis kolm raamatut.*
Peeter.NOM bought three.NOM book.SG.PAR
‘Peeter bought three books.’
- b. **Peeter ostis kolm raamatuid.*
Peeter.NOM bought three.NOM book.PL.PAR
- c. *Raamatuid Peeter ostis kolm.*
book.PL.PAR Peeter.NOM bought three.NOM
‘Peeter bought three books.’

We propose that the morphological mismatch can be avoided by assuming that the structure contains a classifier that licenses the partitive NP. Finnish has a classifier *kappale*, ‘piece’, which is typically used for counting individuals in sentences such as (6a) (see also Alho 1992: 7). The classifier enables the partitive NP to escape the noun phrase, as in (6b), in which case the classifier is only optionally pronounced.⁵

(6) Finnish

- a. *Pekka osti [DP[N_{NumP} kolme [CIP kappaletta [NP kirjoja]]]].*
Pekka.NOM bought three.NOM piece.SG.PAR book.PL.PAR
‘Pekka bought three books.’
- b. *[NP Kirjoja] Pekka osti [DP[N_{NumP} kolme [CIP (kappaletta) ___]]].*
book.PL.PAR Pekka bought three.NOM piece.SG.PAR
‘Pekka bought three **books**.’

However, the analysis of Estonian is more complicated. As in Finnish, the classifier is optionally present in the split construction (7a), but ungrammatical in a continuous NP

⁵ It should be noted that sentence (6a) has an artificial tone; we will consider the style and variation of this construction in Section 5.

(7b). In a continuous NP, the noun *tükk* ‘piece’ takes only singular mass (or abstract) nouns as complement (7c). If a count noun occurs in this position, it is in the singular and coerced into a mass noun.⁶

(7) Estonian

- a. *Raamatuid Peeter ostis kolm (tükki).*
 book.PL.PAR Peeter.NOM bought three.NOM piece.SG.PAR
 ‘Peeter bought three books.’
- b. **Peeter ostis kolm tükki raamatuid.*
 Peeter.NOM bought three.NOM piece.SG.PAR book.PL.PAR
- c. *Peeter ostis kolm suurt tükki šokolaadi /*
 Peeter.NOM bought three.NOM big.SG.PAR piece.SG.PAR chocolate.SG.PAR
juustu.
 cheese.SG.PAR
 ‘Peeter bought three big pieces of chocolate/cheese.’

The classifier approach is compared to *morphological repair* account, where the partitive split is derived from the numeral-noun construction, but the plural number of the NP is assigned post-syntactically (Fanselow and Cavar 2002).

The paper is organised as follows: Section 2 provides a classification for the split NPs in Finnish and Estonian. Section 3 shows that the NP split is triggered by discourse and the landing sites are the same as in other types of discourse-related movement. Section 4 addresses the syntactic properties of partitive splits and section 5 provides an analysis. The paper is concluded in Section 6.

2 Introduction to split noun phrases in Finnish and Estonian

Finnish and Estonian have several types of expressions where a noun phrase is separated from the modifying quantifier or numeral. Some quantifiers permit splitting relatively freely, whereas others display a more restricted pattern. This section starts with an introduction to the general properties of noun phrases in both languages in section 2.1. We then outline the properties of three types of splits: the partitive split in Section 2.2, regular NP split in Section 2.3 and the relative split in Section 2.4.

2.1 Basic structure of the noun phrase in Finnish and Estonian

In Finnish and Estonian, noun phrases are composed of the noun head, adjectival modifiers, possessor, demonstrative/determiner and the quantifier, which all precede the noun head. Both languages display case concord within the noun phrase; adjectival modifiers, determiner/demonstratives and quantifiers generally inflect in the same case and number as the noun head, as illustrated in (8) (for Estonian, see example (10)).

⁶ The word *tükk* appears in split noun phrases mainly in spoken language (Erelt et al. 1993: 148). In written language, the meaning of ‘unit’ is implicit.

(8) Finnish

- a. *tämä pieni punainen talo*
 this.NOM small.NOM red.NOM house.NOM
 ‘this small red house.’
- b. *Nä-i-ssä pien-i-ssä punais-i-ssa talo-i-ssa*
 this-PL-INE small-PL-INE red-PL-INE house-PL-INE
 ‘in these small red houses’

However, numerals higher than one and certain quantifiers display a heterogeneous case assignment pattern, see e.g. Brattico (2008) and Nelson and Toivonen (2003). When the DP is in the nominative or accusative case, the numeral is in the nominative case and adjectives and noun head below the numeral are in the partitive singular (9a). The quantificational partitive case is absent when the DP appears in some other case, as in (9b). Same holds for Estonian (10a-b).

(9) Finnish

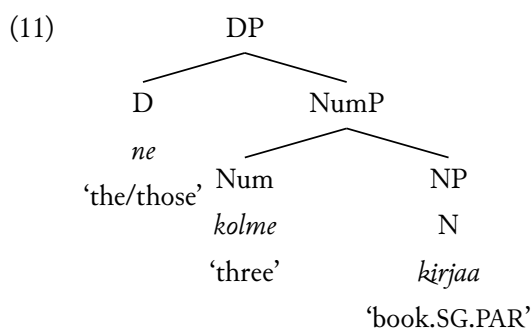
- a. *Nämä kaikki kolme pien-tä talo-a*
 this.PL.NOM all.NOM three.NOM small.SG-PAR house.SG-PAR
 ‘all these three small houses’
- b. *Nä-i-ssä kaikki-ssa kolme-ssa piene-ssä talo-ssa*
 this-PL-INE all-INE three-SG-INE small-SG-INE house-SG-INE
 ‘in all of these three small houses’

(10) Estonian

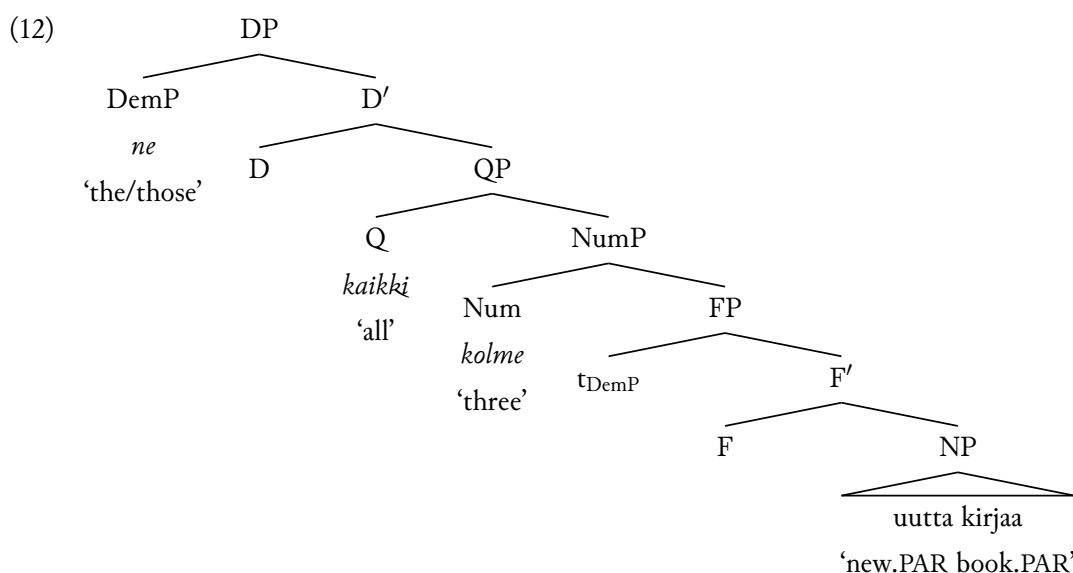
- a. *Ma leidsin [kaks pliiatsit].*
 1SG.NOM find.PST.1SG two.SG.NOM/(SG.ACC) pencil.SG.PAR
 ‘I found two pencils.’
 (Miljan and Cann 2013: 343)
- b. *kahelt teravalt pliiatsilt*
 two.SG.ABL sharp.SG.ABL pencil.SG.ABL
 ‘from two sharp pencils’
 (Miljan and Cann 2013: 343)

We assume here an analysis where the numeral is a head within the noun phrase (see Brattico 2008; Danon 2012; Ionin and Matushansky 2006; Nelson and Toivonen 2003; Norris 2014) and the NP occupies the complement of the numeral, see (19) below. In addition, the noun phrase may form a Determiner Phrase (DP) (see Gröndahl 2015; Norris 2014).⁷

⁷ In both languages, there are signs of development of indefinite and definite articles (Hiitam and Börjars 2003; Juvonen 2000; Laury 1991, 1997; Pajusalu 1997, 2000).



Finally, we will assume that demonstratives head their own phrase; see Norris (2014) for Estonian and Gröndahl (2015) for Finnish. The DemP moves to the specifier of DP from a specifier position of a lower functional projection. The basic structure of a noun phrase is presented in (12).⁸



We will adopt the following convention for the terminology: in an argument position, the noun phrase is referred to as a DP. In the split construction, the partitive noun phrase is referred to as ‘partitive NP’, even though it later turns out that the noun phrase may contain a D-projection.

2.2 The partitive split

Finnish and Estonian partitive splits can be divided into two classes: The first class does not display any morphological mismatches and the second class does. The former class involves Finnish quantifiers *paljon* ‘much, a lot’, *vähän* ‘little’, *hiukan* ‘a little’, *enemmän* ‘more’, see examples (13a-c). Among the Estonian quantifiers that belong to the first class are *palju*, ‘much, many’ and *vähe*, ‘little, few’, see examples (14a-c).

⁸ Another option would be to assume that the demonstrative occupies D⁰ (Brattico 2010).

(13) Finnish

- a. *Minä näin paljon lintuja.*
I.NOM saw a lot bird.PL.PAR
'I saw a lot of birds.'
- b. *Lintuja minä näin paljon.*
bird.PL.PAR I.NOM saw a lot
'I saw a lot of **birds**.'
- c. *Maitoa Merja osti vähän.*
milk.SG.PAR Merja bought little.NOM
'Merja bought a little bit of **milk**.'

(14) Estonian

- a. *Peeter nägi palju kasse.*
Peeter.NOM saw many cat.PL.PAR
'Peeter saw many cats.'
- b. *Palju nägi Peeter kasse.*
many saw Peeter.NOM cat.PL.PAR
'Peeter saw many cats.'
- c. *Kasse nägi Peeter palju.*
cat.PL.PAR saw Peeter.NOM many
'Peeter saw many cats.'

These quantifiers are often ambiguous between the reading where the quantifier modifies the NP and the clausal reading in Finnish and Estonian (see also Hakulinen et al. 2004: §657 and §994). For example, the quantifier *paljon* 'much, a lot' can modify an event, as in (15a). This means that the sentence (15b) can mean either that I read a lot of books in one day, or that I read books a lot in one day. We will ignore the clausal reading in this paper and concentrate on the split reading.

- (15) a. *Luen paljon kirjoja.*
read.1SG a lot book.PL.PAR
'I read a lot of books.' / 'I read books a lot.'
- b. *Kirjoja luen paljon.*
book.PL.PAR read.1SG a lot
'I read a lot of **books**.' / 'I read **books** a lot.'

Let us now turn to quantifying expressions that produce morphological mismatches in partitive split. In Estonian, mismatches are caused by numerals higher than one. In Finnish, all the numerals introduce mismatches, and, in addition, the quantifier *monta* 'many', and singular and plural forms of the quantifier *muutama* 'some'. A typical example is provided in (16a-c): in the complement of the numeral, the NP occurs in the singular, but in the split NP, the noun phrase is always in the partitive plural (16b). The plural NP cannot be "returned" to the complement (16c), and hence, there is a morphological mismatch.

(16) Estonian

- a. *Peeter ostis kolm raamatut.*
 Peeter.NOM bought three.NOM book.SG.PAR
 ‘Peeter bought three books.’
- b. *Raamatuid Peeter ostis kolm.*
 book.PL.PAR Peeter.NOM bought three.NOM
 ‘Peeter bought three books.’
- c. **Peeter ostis kolm raamatuid.*
 Peeter.NOM bought three.NOM book.PL.PAR

We intend to show in this paper that the morphological mismatches are restricted to quantifying expressions that can take only countable complements. We will return to this question in Section 4. In addition, we propose that Finnish and Estonian partitive splits can be analysed in terms of sub-extraction, despite of the mismatches.

2.3 Regular NP split

Another class of split noun phrases involves Finnish quantifiers *monet* ‘many’, *useat* ‘several’, *harvat* ‘rare’ and Estonian quantifiers *paljud* ‘many’, *mõned* ‘some’, *vähesed* ‘rare’, among others. These quantifiers can be separated from the NP in several contexts where the partitive split is not available. For example, the transitive clause subject can be split in (17a-b). Neither the case nor the number of the NP are altered during the split.⁹

(17) a. Finnish

Opiskelijat ovat monet ostaneet kirjan.
 student.PL.NOM be.PRES.3PL many.PL.NOM bought book.SG.ACC
 ‘Many students have bought a book.’

b. Estonian

Üliõpilased on paljud ostnud õpiku.
 student.PL.NOM be.PRES.3PL many.PL.NOM bought textbook.SG.GEN
 ‘Many students have bought a textbook.’

⁹ Interestingly, singular forms of quantifiers such as *moni* ‘many’, *usea* ‘several’ and *harva* ‘few’ do not permit NP-split at all. Both the quantifier and its complement inflect in the singular and in the same case. They disallow regular splitting (i.a-b) and the partitive split (i.c). These quantifiers cannot occur in the accusative case (i.d).

(i) Finnish

- a. *Minä ihailen harvaa opettajaa.*
 I.NOM admire few.SG.PAR teacher.SG.PAR
 ‘I admire few teachers.’
- b. **Opettaja minä ihailen harvaa ____.*
 teacher.SG.PAR I.NOM admire few.SG.PAR
- c. **Opettaja minä ihailen harvaa ____.*
 teacher.PL.PAR I.NOM admire few.SG.PAR
- d. **Minä näin harvan opettajan.*
 I.NOM saw few.SG.ACC teacher.SG.ACC

We will return to the regular NP split briefly in Section 4.2, which addresses the distribution of the split noun phrases.

2.4 The elative split

Third type of split noun phrase introduced here is the elative split, where an elative NP is separated from the modifying quantifier, as in (18a-b). Although the elative split is superficially similar to the partitive split (e.g. in targeting the same discourse positions), there are some fundamental differences between the two. First, the application of the elative split is almost unrestricted. All the numerals and most of the quantifiers enable the elative split in a variety of structural positions.

(18) a. Finnish

Oppilaista Pekka tuntee kaksi.
 student.PL.ELA Pekka.NOM knows two.NOM
 ‘Pekka knows two of the students.’

b. Estonian

Õpilastest Peeter kutsub kaks.
 student.PL.ELA Peeter.NOM invites two.NOM
 ‘Peeter invites two of the students.’

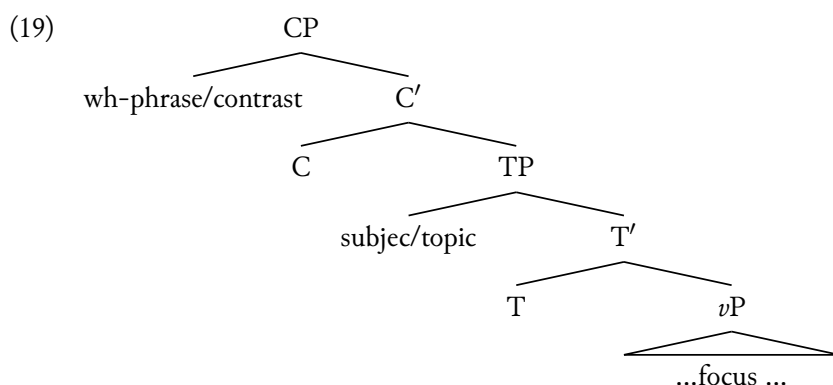
We will see later that the elative split does not obey island constraints (section 4.2) and permit noun doubling (section 4.3). This suggests that the elative split is not derived by movement. However, the analysis of the elative split is left for another occasion.

3 The partitive split and discourse

This section addresses discourse properties of the partitive split. It is proposed that in Finnish and Estonian, the partitive split is triggered by discourse features, such as topic, focus and contrast (see also Alho 1992; Arnhold 2009; Metslang 2016). In addition, wh-movement and relativization may induce splitting.

The basic word order in Estonian and Finnish is SVX, but the order is flexible. That is, the subject position can host also other elements in both languages. For example, it is typical that a non-subject occupies the subject position in sentences that do not contain a subject (Tael 1990; Vilkuna 1989, 1998). Nevertheless, the position of the wh-phrase is fixed to the beginning of the sentence (Erelt 2009; Vilkuna 1998).

We will follow the basic proposal by Vilkuna (1989, 1995) for Finnish, where the left periphery of a finite clause contains two discourse-related fields. The first one is able to host wh-phrases, relative pronouns and contrasted constituents and the second one is reserved for the subject or a topical element. These fields are represented structurally in (19) (Vainikka 1989). We will assume the same basic configuration for Estonian, although both positions have language-specific properties (see Henk 2010), such as the V2 constraint on Estonian word order (Tael 1990). Finally, new information focus occurs within the VP and is typically placed on a constituent at the end of the clause (Henk 2010; Tael 1990; Vilkuna 1989).



The partitive split may target both of these left-peripheral positions. Example (20a) shows that the split noun phrase can be a relative pronoun that occupies the Spec,CP of the relative clause. Example (20b) illustrates movement of a contrasted NP.

(20) Finnish

- a. *Merja näki leivokset, joita Pekka oli ostanut kolme.*
 Merja saw cake.PL.ACC which.PL.PAR Pekka be.PST.3SG bought three.NOM
 lit. 'Merja saw the cakes which Pekka had bought three.'
 'Merja saw the three cakes which Pekka had bought.'

Estonian

- b. *Raamatuid ta ostis palju.*
 book.PL.PAR s/he.NOM bought many
 'S/He bought many **books**.'

Similarly, the partitive NP can target the lower subject/topic position Spec,TP, as in examples (21a-b) (from Hakulinen et al. 2004: §902) and (22).

(21) Finnish

- a. *Vastauksia tuli vajaat 3000.*
 answer.PL.PAR came not.full 3000
 'Little less than 3000 answers arrived.'
- b. *Hakijoita kutsuttiin haastatteluun useita.*
 applicant.PL.PAR invited.PASS interview.to several.PL.PAR
 'Several applicants where invited to the interview.'

(22) Estonian

- Klaase purunes viis.*
 glass.PL.PAR broke five.NOM
 'Five glasses broke.'

According to Hakulinen and Karlsson (1979), the familiarity of discourse is not always required for the partitive NP to occur at the front of the sentence; this is illustrated with example (23a) (from Hakulinen and Karlsson 1979: p. 148). However, it is a general property of Finnish finite clauses that the element that occupies the subject position does not have to be familiar from the discourse, see (23b). For example, Holmberg and Nikanne (2002) propose that any element capable of functioning as a topic can occupy the subject position.

- (23) a. *Uistimia on kannettava mukana paljon.*
 spoon.bite.PL.PAR be.PRES.3SG carry.INF along a lot
 ‘You have to carry a lot of spoon bites with you.’
- b. *Uistimia on kannettava mukana.*
 spoon.bite.PL.PAR be.PRES.3SG carry.INF along
 ‘You have to carry spoon bites with you.’

It thus suffices to assume that the same properties that trigger movement of non-subjects to Spec,TP in other constructions trigger movement also in partitive splits.

In this paper, we concentrate on the order where the partitive NP occurs first, because it is more common in both languages. However, examples (24a-b) show that also the quantifying expression can move to the left periphery.

- (24) a. Finnish
Kuinka paljon Pekka kutsui vieraita?
 how much.NOM Pekka.NOM invited guest.PL.PAR
 ‘How many guests did Pekka invite?’
- b. Estonian
Kui palju Peeter sai toole?
 how many Peeter.NOM got chair.PL.PAR
 ‘How many chairs did Peeter get?’

This order is sometimes referred to as “inverted split” (Fanselow and Cavar 2002), since the linear order of the quantifier and the NP is opposite to the continuous NP. Hakulinen and Karlsson (1979: p. 149) provide examples such as (25a-b) of the order where the quantifier is at the front and write: “[...]it is possible to emphasise the quantifier by placing it to the beginning of the clause [...]” (Hakulinen and Karlsson 1979: p. 149) (author’s translation). This suggests that the quantifier has a discourse function at the left peripheral position. Metslang (2016) proposes that the movement of the quantifier in Estonian is triggered by focus. However, it is not always clear, whether the quantifier occupies the Spec,CP or the lower Spec,TP. The derivation of these type of splits is discussed in Section 4.5.

- (25) Finnish
- a. *Paljon olisi vielä kerrottavaa.*
 a lot be.COND still tellable.PAR
 ‘There is still a lot worth of telling’
- b. *Enemmän pitäisi siis olla opettajien, oppilaiden ja vanhempien
 keskinäistä kanssakäymistä.*
 more should therefore be teachers’ children’s and parents’
 mutual.SG.PAR interaction.SG.PAR
 ‘There should be more interaction between teachers, children and parents.’

In addition to the left peripheral positions, both languages permit other landing sites for the split constituents (26a-b) (see also Hakulinen et al. 2004: §902). These landing sites are often targeted by discourse-related movement, but their properties are not as well-known as the left periphery.

- (26) a. Finnish
Pekka on kirjoja ostanut kolme.
 Pekka.NOM be.PRES.3SG book.PL.PAR bought three.NOM
 ‘Pekka has bought three books.’
- b. Estonian
Mari sai seeni kolm.
 Mari.NOM got mushroom.PL.PAR three.NOM
 ‘Mari got three mushrooms.’

Finally, the constituent that is left behind (or, alternatively, displaced to the right edge of the clause) is typically interpreted as having new information focus or contrast.

In this paper, we propose that the partitive split is an instance of A' -movement. The fact that the split NP has the same triggers and the landing sites as regular discourse-related movement supports this proposal.

4 The sub-extraction analysis of the partitive split

This section examines the syntactic properties of the partitive split. It is proposed that the partitive split is derived by sub-extraction, where the partitive NP forms a constituent with the quantifying expression before the two parts are separated by movement. The proposed derivation is sketched in (27). We will leave the problem of morphological mismatch aside here; the analysis will be completed in Section 5.

- (27) **Sub-extraction of the partitive NP**
1. *Pekka osti [DP paljon [NP tuoleja]]*
 Pekka bought a lot chair.PL.PAR
 ‘Pekka bought a lot of chairs.’
 2. $[_{NP} Tuoleja]_i$ *Pekka osti [DP paljon ___i]*
 chair.PL.PAR Pekka bought a lot
 ‘Pekka bought a lot of **chairs**.’

The most important evidence for the sub-extraction analysis comes from islands; the partitive split is not possible in contexts that do not permit movement out of them. This data is discussed in Section 4.2. Other evidence in support for the movement account come from binding (section 4.1) and the absence of noun doubling (section 4.3). Finally, section 4.4 discusses controversial data from VP-fronting.

4.1 Binding

The evidence from reflexive binding is here used for ruling out the hypothesis, where the partitive NP would be base-generated (i.e. inserted directly) in the left-peripheral position. The binding data indicates that the partitive NP is base-generated in a low position, from where it moves to the left-peripheral position.

The Finnish third person possessive suffix is a reflexive anaphor that requires a correlate in a higher structural position, as exemplified in (28a-b) (Trosterud 1993; Vainikka 1989). The reflexive binding is not affected by movement of the NP (28c).¹⁰

- (28) a. *Pekka näki veljensä.*
 Pekka.NOM saw brother.SG.ACC.PX
 ‘Pekka_i saw his_i brother’
- b. **Veljensä näki Pekan.*
 brother.SG.NOM.PX saw Pekka.ACC
 Intended: ‘His_i brother saw Pekka_i.’
- c. [_{CP} [_{Veljensä}]_i [_C C [_{TP} *Pekka näki* ____i]]]!
 brother.SG.ACC.PX Pekka.NOM saw
 ‘Pekka_i saw **his_i brother!**’

Example (29) of a partitive split shows that, similarly as in the example (28c) above, the correlate of the partitive NP is the subject argument.

- (29) [_{CP} [_{Veljensä} *kirjoja*]_i [_C C [_{TP} *Pekka lukki* [_{viisi} ____i]]]].
 brother.GEN.PX book.PL.PAR Pekka.NOM read five.NOM
 ‘Pekka_i read five of **his_i brother’s** books’

We conclude that the partitive noun phrase is base-generated in a position below the subject, from where it moves to the left peripheral position. In the following section, we consider evidence from islands which suggest that this position is inside the same DP that contains the quantifying expression.

4.2 Islands

Islands are contexts that do not permit movement out of them. This section examines several types of islands and shows that the partitive split obeys island constraints. In addition, the comparison to regular NP split and relative split shows a clear contrast between the different types of split constructions.

However, let us first summon up the syntactic contexts where the partitive split commonly occurs. First, the direct object can be split, as we saw in examples (13)-(14). Second, subjects of unaccusative verbs and certain intransitive verbs permit the partitive split, as examples (30) and (31) illustrate.¹¹

¹⁰ The condition that the reflexive anaphor has to be bound by a c-commanding correlate is known as the Binding Condition A (Chomsky 1981).

¹¹ The subject of the unaccusative permits extraction in suitable contexts in Finnish (i.a-b).

- (i) a. *Kenestä saapui [kuva ___] toimistoon?*
 who.of arrived picture.SG.NOM office.to
 ‘Of whom did pictures arrive to the office?’
- b. *Mitä tuli [ajatus tehdä ___]?*
 what.PAR came idea.SG.NOM do.INF
 lit ‘What became an idea to do?’ ‘What was the idea to do?’

The movement to the left periphery is only possible when the subject occupies a low position.

(30) Estonian

- a. *Raamatuid ilmus palju.*
 book.PL.PAR appeared many
 ‘Many books appeared in print.’
- b. *Raamatuid ilmus kolm.*
 book.PL.PAR appeared three.NOM
 ‘Three books appeared in print.’

(31) Finnish

- Miehiä lähti kalaan viisi.*
 man.PL.PAR left.3SG fish.to five.NOM
 ‘Five men went fishing.’

Third, subjects of ECM (Exceptional Case Marking) constructions permit limited extraction and partitive split, as can be seen in examples (32) and (33). In the ECM-construction, the subject of the non-finite clause receives the case marking from the superordinate clause.¹²

(32) Finnish

- a. *Merja näki [INF lapsia leikkimässä].*
 Merja.NOM saw child.PL.PAR playing
 ‘Merja saw children playing.’
- b. *??Lapsia_i Merja näki [INF [kolme ___i] leikkimässä].*
 child.PL.PAR Merja.NOM saw three.NOM playing.
 ‘Merja saw three **children** playing.’

(33) Estonian

- a. *Mari nägi [INF lapsi mängimas].*
 Mari.NOM saw child.PL.PAR playing
 ‘Mari saw children playing.’
- b. *Lapsi nägi Mari mängimas kolm.*
 child.PL.PAR saw Mari.NOM playing three.NOM
 ‘Mari saw three children playing.’

In contrast, the examination of well-known islands such as Subject Condition and adjunct islands shows that the partitive split is not available in these contexts. First, the Subject Condition (Huang 1982; Ross 1967) states that extraction from the subject is more restricted than extraction from the object. For example, the nominative subject of a transitive verb does not permit the partitive split. Examples (34a-c) illustrate this for Finnish and examples (35a-b) for Estonian.¹³

¹² The word order in (33b), where the numeral occurs at the end of the sentence is preferred to the order where it occurs before the non-finite verb (32b). We propose that this preference is due to the information structure: in both languages, the element bearing the new information focus typically occurs at the end of the finite clause. If the numeral did not move, the new information focus would be placed on the non-finite verb. This is possible, but not a favored alternative. Note that in Estonian example (33b), also the V2 preference is in effect.

¹³ Note that the relative split is possible in all island contexts examined here. Thus, for example the sentence (i) is grammatical.

(34) Finnish

- a. *Kaksi miestä osti kirjan.*
two.NOM man.SG.PAR bought book.ACC
'Two men bought a book.'
- b. **Miehiä osti kaksi kirjan.*
man.PL.PAR bought two.NOM book.ACC
- c. **Miehiä kaksi osti kirjan.*
man.PL.PAR two.NOM bought book.ACC

(35) Estonian

- a. *Kaks meest ostsid raamatu.*
two.NOM man.SG.PAR bought book.SG.GEN
'Two men bought a book.'
- b. **Meest/mehi kaks ostis raamatu.*
man.SG.PAR/PL.PAR two.NOM bought.3SG/3PL book.SG.PAR
Intended: 'Two men bought a book.'

Another example is offered by *hiukan* 'a little', which does not trigger the morphological mismatch in (36a). Examples (b-c) show that splitting is not possible when the NP occupies the subject position. Example (36d) shows that the split is available in a non-island context.

(36) Finnish

- a. *Hiukan jauhoja korjaa taikinän rakenteen.*
little flour.PL.PAR fixes dough.GEN consistency.ACC
'A little bit of flour fixes the dough consistency'
- b. **Jauhoja korjaa hiukan taikinän rakenteen.*
flour.PL.PAR fixes little dough.GEN consistency.ACC
- c. **Jauhoja hiukan korjaa taikinän rakenteen.*
flour.PL.PAR little fixes dough.GEN consistency.ACC
- d. *Jauhoja Pekka osti hiukan.*
flour.PL.PAR Pekka.NOM bought little
'Pekka bought only a little bit of flour.'

Adjuncts offer another well-known context that resists movement out of them (Ross 1967). The following examples show that an adjunct cannot be split:

(37) Finnish

- a. *Pekka luki kirjaa kolme tuntia.*
Pekka.NOM read book.SG.PAR three.NOM hour.SG.PAR
'Pekka was reading a book for three hours.'
- b. **Kolme Pekka luki kirjaa tunteja.*
three.NOM Pekka.NOM read book.SG.PAR hour.PL.PAR

(i) *Miehistä osti kaksi kirjan.*
man.PL.ELA bought two book.ACC
'Two of the men bought a book.'

- c. **Tunteja Pekka luki kirjaa kolme.*
hour.PL.PAR Pekka.NOM read book.SG.PAR three.NOM

(38) Estonian

- a. *Peeter töötas kolm nädalavahetust.*
Peeter.NOM worked three.NOM weekend.SG.PAR
'Peeter worked for three weekends.'
- b. *?*Nädalavahetust Peeter töötas kolm.*
weekend.SG.PAR Peeter.NOM worked three.NOM

The final island context examined here is formed by DPs in semantic cases, which resist extraction. This is illustrated in example (39a-b), where the elative modifier cannot be moved out of a DP in the illative case (see also Huhmarniemi 2012). The partitive split is not permitted, even if the DP occupies the complement of the verb (39c).¹⁴

(39) Finnish

- a. *Pekka tarttui kirjaan presidentistä.*
Pekka grabbed book.ILL president.ELA
'Pekka grabbed the book about the president.'
- b. **Kenestä Pekka tarttui kirjaan?*
who.ELA Pekka.NOM grabbed book.ILL
- c. *?*Opiskelijoihin on Pekka tutustunut viiteen.*
student.PL.ILL be.PRES.3SG Pekka get.known five.ILL

To summarise, the partitive split is not available for transitive clause subjects, adjuncts or DPs in semantic cases. In contrast, the elative split is mostly not restricted by islands and the regular NP split may take place in at least some of these contexts. One of the quantifiers that enables NP split relatively freely is Finnish *monet*, 'many'. As can be seen in the following examples, subjects (40a), adjuncts (40b) and DPs in semantic cases (40c) all permit NP split in the presence of this quantifier.

(40) Finnish

- a. *Miehet ovat (monet) ostaneet kirjan.*
man.PL.NOM be.PRES.3PL many.PL.NOM bought book.SG.ACC
'Many men bought a book.'

¹⁴ Alho (1992: 8) notes that Finnish partitive verbs such as *ihailia* 'to admire' do not permit splitting, examples (i.a-b) are from Alho (1992). The example (i.b) is ungrammatical also when the classifier is not present. However, the object of a partitive verb appears to be an island also for other elements than split NPs (i.c).

(i) Finnish

- a. **Matti ihaillee kahta kappaletta lingvistejä.*
Matti admires two.PAR piece.SG.PAR linguist.PL.PAR
- b. **Lingvistejä Matti ihaillee kahta kappaletta.*
Linguist.PL.PAR Matti admires two.PAR piece.SG.PAR
- c. *?*Merjasta Pekka ihaillee kuvaa ____.*
Merja.of Pekka admires picture.SG.PAR
Intended: 'Pekka admires the picture of Merja.'

- b. *?Hotelleissa on Pekka työskennellyt monissa.*
 hotel.PL.INE be.PRES.3SG Pekka.NOM worked many.PL.INE
 ‘Pekka has worked in many hotels.’
- c. *?Moniin on Pekka vienyt Merjan kokouksiin.*
 many.PL.ILL be.PRES.3SG Pekka.NOM taken Merja.ACC meeting.PL.ILL
 ‘Pekka taken Merja to many meetings.’

The distribution of the elative split is even more widespread. For example, transitive clause subjects do not normally permit extraction, but the elative split is available (41a). Another example is offered by the DP in the illative case in (41b) (see also Alho 1992) and example (41c) shows that adverbial modifiers enable the elative split. However, relative clauses and adjective participials, among others, appear to be strong islands for the elative split.

- (41) a. *Miehistä osti kirjan viisi.*
 man.PL.ELA bought book.SG.ACC five.NOM
 ‘Five of the men bought the book’
- b. *Kirjoista Pekka tutustui viiteen.*
 book.PL.ELA Pekka.NOM explored five.ILL
 ‘Of the books, Pekka explored five of them.’
- c. *Näistä autoista Pekka on ajanut kolarin [viidellä ____].*
 these.PL.ELA cars.ELA Pekka.NOM be.PRES.3SG driven crash five.by
 ‘Pekka has caused a crash with five of these cars.’

Taken together, neither the elative split nor the regular NP split obey the island constraints typical for A'-movement.

4.3 Absence of noun doubling

The island data examined in the previous section indicates that the constituent containing the quantifier or the numeral and the partitive NP are syntactically related. However, the island data does not rule out the option that what we are seeing is not movement, as we will propose here, but some other type of A'-relation, which is sensitive to islands. This option can be excluded by investigating the phrase that contains the quantifier or the numeral. Movement typically leaves a *gap*, an empty position in the place of the moved element.¹⁵

It turns out that Finnish and Estonian partitive splits always involve a gap that cannot be filled by any other element. First, example (42a) shows that the noun head cannot be doubled to two locations. Second, example (42b) and (43) demonstrate that the quantifying phrase cannot contain any other noun head. We conclude that the structure contains a gap created by the movement of the partitive NP.

¹⁵ The copy theory of movement (Chomsky 1995) takes movement to be an instance of copying, where only one or some of the copies are pronounced. It is possible to adopt this hypothesis for A'-movement also for split NPs.

(42) Estonian

- a. **Linde ta tunneb ainult väikseid linde.*
 bird.PL.PAR s/he.NOM knows only small.PL.PAR bird.PL.PAR
 lit. 'Birds s/he only knows small birds.'
- b. **Juurvilju / *juurviljad Peeter sõi ainult kaks porgandit.*
 vegetable.PL.PAR vegetable.PL.NOM Peeter.NOM ate only two.NOM
 carrot.SG.PAR
 lit. 'Vegetables Peeter ate only two carrots.'

(43) Finnish

- *?*Lemmikkejä Pekka haluaa vain kolme koiraa.*
 pet.PL.PAR Pekka wants only three dog.SG.PAR
 lit. 'Pets Pekka only wants three dogs.'

Another example of how a filled complement position prevents splitting is offered by quantifying noun phrases. Consider example (44a), where a noun *pullon* 'a bottle' selects an NP-complement. When this complement is present, the partitive split cannot be formed (16b).¹⁶

(44) Finnish

- a. *Pekka osti pullon mehua.*
 Pekka bought bottle.SG.ACC juice.SG.PAR
 'Pekka bought a bottle of juice.'
- b. **Juomia Pekka osti viisi pulloa mehua.*
 drink.PL.PAR Pekka bought five.NOM bottle.SG.ACC juice.SG.PAR

The partitive NP thus appears to occupy a position in the complement domain of the NumP. Finally, the comparison to relative split shows that the two constructions involve a different syntactic derivation. The relative split permits noun doubling:

(45) a. Finnish

- Linnuista Pekka tuntee satakielen.*
 bird.PL.ELA Pekka.NOM knows nightingale.ACC
 'Of birds, Pekka knows the nightingale.'

¹⁶ The partitive split appears to escape this constraint in list contexts (i.a). In addition, the list context differs from the partitive split in other respects. For example, the adjective can be split in list context in (i.b), although this is not normally possible (i.c). A possible hypothesis is that list contexts enable elliptical constructions that are not available in the partitive split.

- (i) a. *Lemmikkejä Pekka haluaa kolme koiraa ja kaksi ponia.*
 pet.PL.PAR Pekka.NOM wants three.NOM dog.SG.PAR and two.NOM pony.SG.PAR
 'As for pets, Pekka wants three dogs and two ponies.'
- b. *Tuoleja Pekka osti punaisen ja sinisen.*
 chair.PL.PAR Pekka.NOM bought red.SG.ACC and blue.SG.ACC
 'As for chairs, Pekka bought a red one and a blue one.'
- c. **Tuoleja Pekka osti punaisen.*
 chair.PL.PAR Pekka.NOM bought red.SG.ACC

b. Estonian

Juurviljadest Peeter sõi ainult kaks porgandit.
 vegetables.PL.ELA Peeter.NOM ate only two.NOM carrot.SG.PAR
 ‘Of the vegetables, Pekka ate two carrots.’

To summarise, the presence of a gap indicates that the partitive NP has occupied a position in the complement domain of the numeral. Therefore, the two parts of the partitive split are structurally related and not independent phrases as has been proposed for some other languages (Fanselow 1988; Ott 2011). The absence of noun doubling thus supports the sub-extraction analysis for the partitive split.

4.4 Evidence from VP-fronting

The final diagnostic property of sub-extraction considered in this paper concerns the properties of VP-fronting, movement of the verb phrase to the left periphery. In Finnish, the verb phrase can move as a whole, but the construction is marked. This is indicated with ‘?’ in example (46b) below.

- (46) a. *Merja oli* [_{VP} *ostanut kirjaa*].
 Merja.NOM be.PST.3SG bought book.ACC
 ‘Merja had bought a book.’
 b. ?[_{VP} *Kirjaa ostanut*] *Merja oli* ____!
 book.ACC bought Merja.NOM be.PST.3SG
 ‘Merja had **bought a book**.’

In this paper, we defend an analysis where the partitive split is formed by moving the partitive NP, as in (47).

- (47) a. *Pekka on* [_{VP} *ostanut* [_{DP} *paljon* [_{NP} *tuoleja*]]]
 Pekka be.PRES.3SG bought a lot chair.PL.PAR
 b. [_{NP} *Tuoleja*]_i *Pekka on* [_{VP} *ostanut* [_{DP} *paljon* _____i]]
 chair.PL.PAR Pekka be.PRES.3SG bought a lot

Against this background, the example (48), where the moved VP contains the verb and the quantifier is expected to be possible. In this example, the partitive NP has moved out of the verb phrase, and after that, the verb phrase has been fronted.

- (48) ?[_{VP} *Ostanut paljon* _____i] *on* *Pekka* [_{NP} *tuoleja*]_i!
 bought a lot be.PRES.3SG Pekka.NOM chair.PL.PAR
 ‘Pekka has bought a lot of chairs.’

In contrast, the construction where the partitive NP moves together with the verb as in (49a-b) is expected to be ungrammatical. However, these type of sentences are accepted by some Finnish speakers. In a speaker experiment, test sentences such as (49a-b) were subject to a considerable amount of speaker variation. However, with the exception of one liberal speaker, none of our Finnish informants found splitting in VP fronting contexts completely acceptable.

- (49) a. ??[*Ostanut autoja*] *hän on kolme.*
 bought car.PL.PAR s/he be.PRES.3SG three.NOM
 lit. 'Bought **cars** she has three.'
- b. ??[*Autoja ostanut*] *hän on kolme.*
 car.PL.PAR bought s/he be.PRES.3SG three.NOM

In contrast, the VP-fronting appears to be possible in Estonian:

- (50) Estonian
- a. [*Raamatuid lugenud*] *on ta palju / kolm*
 book.PL.PAR read.PTCPL be.PRES.3SG s/he.NOM many / three.NOM
 'He has read many books.'
- b. [*Raamatuid lugeda*] *ta tahab kolm.*
 book.PL.PAR read.INF s/he.NOM want.3SG.PRES three.NOM
- c. [*Lugeda raamatuid*] *ta tahab kolm.*
 read.INF book.PL.PAR s/he.NOM want.3SG.PRES three.NOM

The fact that the Finnish speakers are reluctant to accept the VP-fronting where the quantifier/numeral has been stranded, points towards the sub-extraction account. However, more research is needed for determining the exact contexts that permit VP-fronting in both languages.

4.5 Movement of the quantifying expression

This far, we have provided evidence from islands, binding and noun doubling in support of the sub-extraction account of the partitive split. Before continuing with the analysis, we will briefly examine the movement of the quantifying expression to the left periphery. As we saw in (14b), repeated here as (51), quantifiers can occupy the left-peripheral position. Same holds for the numeral in example (52a) (see also Arnhold 2009; Metslang 2016). Example (52b) illustrates wh-movement.

- (51) Estonian
- Palju on Peeter näinud kasse.*
 many be.PRES.3SG Peeter.NOM seen cat.PL.PAR
 'Peeter has seen many cats.'
- (52) Finnish
- a. ?[_{CP} *Kolme* [_C *C* [_{TP} *Pekka on ostanut tuoleja*]]]
 three.NOM Pekka.NOM be.PRES.3SG bought chair.PL.PAR
 'Pekka has bought **three** chairs!'
- b. [_{CP} [*Kuinka monta*] [_C *C* [_{TP} *Pekka on ostanut tuoleja*]]]?
 how many Pekka.NOM be.PRES.3SG bought chair.PL.PAR
 'How many chairs has Pekka bought?'

For the derivation of these sentences, we propose an analysis where the partitive NP first moves out of the DP containing the quantifier/numeral. This is illustrated in steps 1-2 in example (53).¹⁷ In step 3, the rest of the DP moves to the left periphery.¹⁸

- (53) 1. *Pekka osti* [DP[QP *kuinka paljon* [NP *tuoleja*]]]
 Pekka bought how much chair.PL.PAR
2. *Pekka osti* [NP *tuoleja*]_i [DP[QP *kuinka paljon* ____i]]
 Pekka bought chair.PL.PAR how much
3. [DP[QP *Kuinka paljon* ____i]]_j C [TP *Pekka osti* [NP *tuoleja*]_i ____j]]?
 how much Pekka bought chair.PL.PAR
 ‘How many **chairs** did Pekka buy?’

Note that the quantifier *paljon* in above examples does not trigger morphological mismatches. However, we propose in Section 5 that the similar sub-extraction account applies also to numerals and quantifiers that trigger mismatches.

This analysis receives support from an analogous derivation of quantifying constructions, such as (54a) below, where the NP occupies the complement position of the measure expression. The measure phrase is able to move to the left-periphery, stranding the partitive NP, as in (54b).¹⁹

- (54) Finnish
- a. *Pekka osti* [*montako* [*pussia* [NP *jauhoja*]]].
 Pekka.NOM bought how many bag.PAR flour.PL.PAR
- b. [*Montako* *pussia* ____i]_j *Pekka osti* [NP *jauhoja*]_i ____j?
 how many bag.PAR Pekka.NOM bought flour.PL.PAR
 ‘How many bags did Pekka buy **flour**?’

It follows from this analysis that the movement of the quantifier/numeral is a more complex phenomenon than the movement of the partitive NP. This may be partially responsible for the fact that the order where the quantifier/numeral is at the front is less common than the order where the partitive NP is at the front.

¹⁷ Movement of the partitive DP could be an instance of object shift or similar phenomenon familiar among others from Finnish ditransitives (see Kaiser 2002).

¹⁸ Alternatively, if the partitive NP occurs at the end of the clause and is interpreted as focused, it is possible that the movement is rightward, targeting the right periphery of the finite clause (for examples of the position of the subject, see Brattico 2016).

¹⁹ In both constructions, the preferred means to form a wh-question is by moving the whole NP along with the wh-phrase, as in the examples below.

- (i) Finnish
- a. [CP [*Kuinka monta tuolia*] [C' C [TP *Pekka osti* ___]]]?
 how many chair.SG.PAR Pekka bought
 ‘How many chairs did Pekka buy?’
- b. [*Montako* *pussia* *jauhoja*] *Pekka osti* ___?
 how many.Q bag.SG.PAR flour.PL.PAR Pekka bought
 ‘How many bags of flour did Pekka buy?’

5 The structure of the partitive split in Finnish and Estonian

This section examines the syntactic derivation of the partitive split in Finnish and Estonian. First, the analysis of partitive splits that does not involve morphological mismatches is straightforward. The partitive NP is first-merged to the complement of the quantifier, as in (55a). In step (55b), the NP has moved to the left periphery of the finite clause and is interpreted as contrastively focused. The movement of the NP in this example is triggered by the discourse feature [+contrast].

(55) Sub-extraction of the partitive NP

- a. *Pekka osti* [DP *paljon* [NP *tuoleja*]]
 Pekka bought a lot chair.PL.PAR
- b. [NP *Tuoleja*]_i *Pekka osti* [DP *paljon* ____i]
 chair.PL.PAR Pekka bought a lot

However, not all partitive NPs *reconstruct* to the position below the quantifying expression. In partitive split, the NP is typically in the plural, as in (56a-b), but in a continuous NP, it has to be in the singular (56c). This morphological mismatch is problematic for the sub-extraction account.

(56) Finnish

- a. [NP *Lintuja*] *minä näin* [DP *kolme* ___].
 bird.PL.PAR I.NOM saw three.NOM
 ‘I saw three **birds**.’
- b. *[NP *Lintua*] *minä näin* [DP *kolme* ___].
 bird.SG.PAR I.NOM saw three.NOM
- c. *Minä näin* [DP [QP *kolme* [NP *lintua* / **lintuja*]]].
 I.NOM saw three.NOM bird.SG.PAR bird.PL.PAR

In this section, we consider two alternative approaches for the morphological mismatch. According to the first alternative, the split construction contains a classifier head that selects a partitive plural NP, as in (57a). In example (57b) the NP has moved to the left periphery of the finite clause and the classifier is only optionally present. This analysis for Finnish partitive splits has been previously presented by Alho (1992: 8).

(57) Finnish

- a. *Pekka osti* [DP [QP *kolme* [CI *kappaletta* [NP *kirjoja*]]]].
 Pekka bought three.NOM piece.SG.PAR book.PL.PAR
 ‘Pekka bought three books.’
- b. [NP *Kirjoja*] *Pekka osti* [DP [QP *kolme* [CI (*kappaletta*) ___]]].
 book.PL.PAR Pekka bought three.NOM piece.SG.PAR
 ‘Pekka bought three **books**.’

According to this proposal, the partitive split is thus not derivationally related to the numeral-noun construction, but has a different underlying syntactic structure. It follows that the morphological mismatch is only apparent. However, although this analysis accounts for the Finnish data, the same analysis cannot be applied directly to Estonian; we

will discuss the remaining problems in the end of this section and examine an alternative, *morphological repair* account in Section 5.4.

The following section points out the differences between the numeral-noun construction and the partitive split and motivates an analysis where the two constructions are not derivationally related. Section 5.2 introduces the classifier analysis and section 5.3 provides further evidence for this analysis. Finally, section 5.4 considers the morphological repair account.

5.1 Morphological mismatches in Finnish and Estonian

This section outlines the differences between the numeral-noun construction and the partitive split. We have already seen that in the partitive split, the partitive NP is often in the plural, whereas in the numeral-noun construction, it has to be in the singular. However, these constructions differ also in other respects. First, Finnish partitive split displays case mismatches, as is explained in Section 5.1.1. Second, in both languages, the partitive noun phrase can be a full DP, while this is not the case in the numeral-noun construction. This is discussed in Section 5.1.2.

5.1.1 Case mismatches

In addition to number mismatches, Finnish partitive split triggers case mismatches. Consider examples (58a-b). In a continuous NP, the noun phrase selected by the numeral *yksi*, ‘one’ appears always in the same case as the numeral. In example (a), the NP is in the accusative in the same context where the numeral *kaksi* ‘two’ requires a partitive NP. However, in the split construction, the NP is in the partitive plural (58b) and there is a morphological mismatch (58c).²⁰

(58) Finnish

- a. *Pekka löysi yhden kirjan / kaksi kirjaa.*
 Pekka found one.ACC book.SG.ACC two.NOM book.SG.PAR
 ‘Pekka found one book / two books.’

²⁰ Finnish numeral *yksi* ‘one’ offers another example of a case change, this time from zero-accusative to partitive in examples (i.a-c). In Finnish, the object argument appears in the zero-accusative (nominative) case in finite clauses that do not display subject agreement inflection (Vainikka and Brattico 2009). In (i.a), both the numeral and the NP are in the zero-accusative case. In (i.b), the fronted NP is in the partitive. Example (i.c) shows the mismatch.

(i) Finnish

- a. *Me ostettiin yksi talo.*
 we.NOM bought.PASS one.NOM house.SG.NOM
 b. *Taloja me ostettiin yksi.*
 house.PL.PAR we.NOM bought.PASS one.NOM
 ‘We bought one house.’
 c. **Me ostettiin yksi taloja.*
 we.NOM bought.PASS one.NOM house.PL.PAR

- b. *Kirjoja Pekka löysi [yhden ____].*
 book.PL.PAR Pekka found one.ACC
 ‘Pekka found one **book**.’
- c. **Pekka löysi [yhden kirjoja].*
 Pekka found one.ACC book.PL.PAR

Finnish quantifier *muutama* ‘a couple of’ displays the same pattern. In example (59a), the subject is in the nominative case, but in (59b), the moved NP is in the partitive. ²¹

(59) Finnish

- a. *Muutama lapsi lähti kotiin.*
 few.NOM child.NOM left home.to
 ‘A couple of children left for home.’
- b. *Lapsia lähti muutama kotiin.*
 child.PL.PAR left few.NOM home.to
 ‘Some children left for home.’
- c. **[Muutama lapsia] lähti kotiin.*
 few.NOM child.PL.PAR left home.to

In contrast, Estonian numeral *üks* ‘one’, with otherwise similar properties, does not permit the partitive split (60a-b). However, examples such as (60c), where the NP is in the singular and in the same case as the numeral, are marginally possible.

(60) Estonian

- a. *Ostsin ühe raamatu.*
 bought.1SG one.SG.GEN book.SG.GEN
 ‘I bought one book.’
- b. **?Raamatuid ostsin (ainult) ühe ____.*
 book.PL.PAR bought.1SG (only) one.SG.GEN
 ‘I bought (only) one book.’
- c. *(?)Raamatu ostsin (ainult) ühe ____.*
 book.SG.GEN bought.1SG (only) one.SG.GEN
 I bought (only) one book.’

Example (60c) thus seems to form a special case. However, since the numeral *üks* ‘one’ behaves like an adjective with regard to the case and number inflection, it may be proposed that Estonian numeral *üks* has an adjectival status. In Estonian, the adjective can be split, as in (61). Comparable example (62) from Finnish is impossible or very poetic.

(61) Estonian

- Püksid ostsin punased ____.*
 trousers.PL.NOM bought.1SG red.PL.NOM
 ‘I bought red trousers.’

²¹ In example (59b), the word order where the quantifier is later in the sentence is preferred to the word order where it would occupy the subject position. This might be due to the fact that the quantifier is in this context focused and focused phrases are not typically moved to the subject position (see Holmberg and Nikanne 2002). Therefore, the partitive NP moves to the subject position alone.

(62) Finnish

*?Housut ostin punaiset ____.
trousers.PL.ACC bought.1SG red.ACC.PL

It thus seems that the behavior of the numeral ‘one’ pairs up with adjectives in Estonian, but with numerals in Finnish.

Second class of morphological mismatches is formed by the plural forms of Finnish numerals. In example (63a), both the numeral and its complement are in the plural accusative form (which looks like the plural nominative). In the split construction (63b), the NP is nevertheless in the partitive. This produces a case mismatch, illustrated in (63c).²²

(63) Finnish

- a. *Olen kadottanut kahdet sukat.*
be.PRES.1SG lost two.PL.ACC sock.PL.ACC
‘I have lost two pairs of socks.’
- b. *Sukkia olen kadottanut kahdet ____.*
sock.PL.PAR be.PRES.1SG lost two.PL.ACC
‘I have lost two pairs of socks.’
- c. **Olen kadottanut kahdet sukia.*
be.PRES.1SG lost two.PL.ACC sock.PL.PAR

Again, Estonian behaves differently. Although numerals inflect in the plural, as in (64a), they disallow the partitive split (64b).

(64) Estonian

- a. *Ostin kolmed kõrvarõngad / püksid.*
bought.1SG three.PL.NOM earring.PL.NOM / trousers.PL.NOM
‘I bought three sets of earrings / pairs of trousers.’
- b. **Kõrvarõngaid / pükse ostsin kolmed.*
earring.PL.PAR / trouser.PL.PAR bought.1SG three.PL.NOM

In Estonian, the morphological mismatches are therefore restricted to the singular changing to plural when the noun phrase moves out of the scope of the numeral higher than one.

5.1.2 Demonstrative pronouns, determiners and pronouns in the partitive split

Another difference between partitive split and the numeral noun construction concerns the structure of the partitive noun phrase. In the partitive split, the partitive noun phrase may contain overt determiners or demonstrative pronouns or be replaced by a pronoun, as in (65a). In the numeral-noun construction, this is not possible (65b).²³ Example (65c) shows

²² The example in (63) is constructed for a noun that appears naturally in plural, as ‘the pair of socks’. However, similar examples are available, for instance, for the NP *kahdet kirjat*, which means ‘two sets of books’.

²³ The example in (65b) improves when the demonstrative is prosodically emphasised. However, we propose that the prosodic emphasis indicates contrastive focus that is associated with movement of the NP to the right edge of the clause.

that in this form, the numeral and the demonstrative pronoun do not form a constituent: they do not move as a whole.²⁴

(65) Finnish

- a. *Näitä Pekka osti viisi* ____.
 these.PAR Pekka.NOM bought five.NOM
 ‘Pekka bought five of these.’
- b. *?*Pekka osti viisi näitä*.
 Pekka.NOM bought five.NOM these.PAR
- c. *?*Viisi näitä*] *Pekka osti* ____.
 five.NOM these.PAR Pekka.NOM bought

Example (66a) illustrates that the partitive NP can contain a determiner/demonstrative, but this is not possible in the numeral-noun construction (66b).

(66) Finnish

- a. *Niitä kirjoja Pekka osti viisi* ____.
 that/the.PL.PAR book.PL.PAR Pekka.NOM bought five.NOM
 ‘Pekka bought five of those/the books.’
- b. *?[*Viisi niitä kirjoja*] *hävisi lomalla*.
 five.NOM that/the.PL.PAR book.PL.PAR disappeared vacation.in
 Intended: ‘Five of those/the books were lost during the vacation.’

Let us consider the above example in more detail. First, the word order where the partitive DP follows the numeral appears to be grammatical in (67a), especially if the DP is prosodically emphasized. However, as can be seen in (67b), the DP cannot move as a whole, which suggests that the partitive DP does not form a constituent with the numeral. Instead, in sentences such as (67a), the partitive DP is no longer inside the same constituent as the numeral. We thus propose that the DP *näitä kirjoja* has been moved. Note that overt demonstratives cannot normally occur in the complement of the numeral in Finnish (67c).

(67) Finnish

- a. *Pekka osti viisi näitä kirjoja, eikä noita!*
 Pekka.NOM bought five.NOM these.PAR book.PL.PAR not those.PAR
 ‘Pekka bought five of these books, not those!’
- b. *?[*Viisi näitä kirjoja*] *Pekka osti* ____!
 five.NOM these.PAR book.PL.PAR Pekka.NOM bought
- c. *?*Pekka osti viisi tätä kirjaa*.
 Pekka.NOM bought five this.SG.PAR book.SG.PAR

²⁴ Another example is offered by the split adjective phrase below (Hakulinen and Karlsson 1979: 98–99). Comparatives seem to form a special class of adjectives that can be split. We will leave them aside here.

(i) Finnish

- Näitä on pienempiä-kin.*
 these.PAR be.PRES.3SG smaller.PL.PAR-too
 ‘There are smaller of these too.’

The comparison to Estonian provides similar results. Overt demonstratives and determiners are commonplace in partitive splits, but they cannot occur in the complement of the numeral in plural (68a). However, when the DP is in the singular, the construction is marginally acceptable (68c).

(68) Estonian

- a. *Neid raamatuid Peeter ostis viis.*
 these.PAR book.PL.PAR Peeter.NOM bought five.NOM
 ‘Of these books, Peeter bought five.’
- b. **Peeter ostis viis neid raamatuid.*
 Peeter.NOM bought five.NOM these.PAR book.PL.PAR
- c. ?*Peeter ostis viis seda raamatut.*
 Peeter.NOM bought five.NOM this.SG.PAR book.SG.PAR

In conclusion, the partitive split differs from the numeral-noun construction in several respects: both the number and the case of the NP may be different in the two constructions. In addition, while numeral-noun constructions involve only ‘plain’ NPs, the partitive split targets full DPs and pronouns. This suggests that the two constructions have different syntactic analyses.

5.2 The classifier analysis of the split noun phrase

In this section, we provide an analysis for the partitive split, which embraces the sub-extraction account and explains the mismatches discussed in the previous section. The analysis is based on the observation that in both languages, the partitive split may contain an optionally pronounced classifier element, such as the word *tükk*, ‘piece’ (for inanimates) in Estonian and the word *kappale*, ‘piece’ in Finnish (69a-b).

(69) a. Finnish

Kirjoja Pekka osti kolme (kappaletta).
 book.PL.PAR Pekka.NOM bought three.NOM piece.SG.PAR
 ‘Pekka bought three books.’

b. Estonian

Raamatuid Peeter ostis kolm (tükki).
 book.PL.PAR Peeter.NOM bought three.NOM piece.SG.PAR
 ‘Peeter bought three books.’

In Finnish, the classifier can be present also in a continuous NP, as in (70). This sentence has an artificial tone, but it is well-formed. Estonian shows a different pattern; we will consider Estonian data at the end of this section.

- (70) *Pekka osti [NP kolme kappaletta kirjoja].*
 Pekka.NOM bought three.NOM piece.SG.PAR book.PL.PAR
 ‘Pekka bought three books.’

According to Alho (1992: 7), the classifier *kappale*, ‘piece’ is used for counting individuals.²⁵ This proposal is motivated by the fact that the morphological mismatch is restricted to quantifying expressions that require a countable NP-complement. For example, while quantifiers such as *paljon* ‘much, a lot’ take uncountable NP-complements (71a), numerals take only countable complements (71b). In addition, while it is effortless to insert a partitive plural NP to the complement of the quantifier *paljon* (71c), this is not possible for numerals (71d) unless the classifier is present, as in (71e). It thus seems that the partitive plural NP is interpreted as uncountable (or similarly to mass nouns) and counting requires support from a classifier.

(71) Finnish

- a. *paljon jauhoa*
a lot flour.SG.PAR
‘a lot of flour’
- b. **kolme jauhoa*
three flour.SG.PAR
- c. *paljon kirjoja*
a lot book.PL.PAR
‘a lot of books.’
- d. **kolme kirjoja*
three book.PL.PAR
- e. *kolme kappaletta kirjoja*
three piece.SG.PAR book.PL.PAR
‘three books’

Estonian displays a similar pattern, as can be seen in the following examples from Metslang (2013: 158). Mass nouns can appear in the singular in the complement of the quantifier *palju* (72a). However, a countable noun has to be in the plural (72b-c). Mass nouns cannot appear in the complement of the numeral (72d), and the same holds for plural countable nouns (72e).

(72) Estonian

- a. *palju liiva*
a lot sand.SG.PAR
‘a lot of sand’
- b. **palju poissi*
a lot boy.SG.PAR
- c. *palju poisse*
a lot boy.PL.PAR
‘a lot of boys’
- d. **kaks liiva*
two sand.SG.PAR

²⁵ In many languages, mass nouns require a support of a classifier or a measure phrase in order to be counted. However, in languages such as Mandarin, the classifier is also required for count nouns (see, e.g. Cheng and Sybesma 1999).

- e. **kaks poisse*
two sand.PL.PAR

The similarities between mass nouns and plural count nouns are pointed out by many authors (e.g. Quine 1960). We will not address the semantics of mass nouns and count nouns further in this paper, but merely point out that a related concept, *divisibility*, has been shown to have an effect on case marking in Finnish and Estonian existential clauses (for an overview, see Metslang 2013). Divisibility separates mass nouns and plural count nouns from singular count nouns and sets formed by individuals can be seen as being divisible in the same sense as mass nouns (Hakulinen et al. 2004: §555). A possible hypothesis is that a DP in the partitive plural does not enable counting directly, but requires a support of a classifier.

Consider Table 1, which summarises the properties of partitive splits in Finnish. The first row presents the characteristics of the reconstructing partitive split. Below the line are examples of elements that produce morphological mismatches.

The first column contains the quantifying expression and the next five columns the requirements that the quantifying expression normally poses to its complement. For example, the numeral *kaksi* ‘two’, takes only countable singular complements and assigns quantificational partitive case. With quantificational partitive case we refer to the case assignment that is described in Section 2.1, see examples (9) and (10). The characteristic property of this type of partitive case is that it is present only when the DP is assigned nominative or accusative/genitive object case.

The final two columns display the morphological mismatches in the partitive split. For example, with numeral *kaksi* ‘two’, the partitive split triggers a number mismatch. In contrast, with numeral *yksi* ‘one’, the partitive split displays both case and number mismatch, as we saw in the previous section (e.g. examples in (58)).

Quantifier	properties of the NP-complement					properties of the split NP	
	countable	uncountable	+sg	+pl	quantificational partitive case	case mismatch	number mismatch
paljon <i>much, a lot</i>	x	x	x	x			
kaksi <i>two</i>	x		x		x		x
monta <i>many</i>	x		x		x		x
yksi <i>one</i>	x		x			x	x
muutama <i>some+sg</i>	x		x			x	x
kahdet <i>two+pl</i>	x			x		x	
yhdet <i>one+pl</i>	x			x		x	
muutammat <i>some+pl</i>	x			x		x	

Table 1: Summary of the properties of the partitive split in Finnish

As can be seen in Table 1, the common denominator with the quantifying expressions that produce morphological mismatches is that they require a countable complement. In addition, none of the other factors correlate directly with the mismatches. This supports the hypothesis that the morphological mismatches are a side-effect of a presence of a silent classifier head that enables counting individuals.

The properties of Estonian partitive splits are summarised in Table 2.

Quantifier	properties of the NP-complement					properties of the split NP	
	countable	uncountable	+sg	+pl	quantificational partitive case	case mismatch	number mismatch
palju <i>much, many</i>	x	x	x	x			
kaks <i>two</i>	x		x		x		x

Table 2: Summary of the properties of the partitive split in Estonian

Thus, although Estonian does not display similar variation as Finnish, the same generalization holds: among the quantifying expressions that enable partitive split, only numerals, which do not take uncountable complements, produce morphological mismatches.

We thus propose an analysis, where the partitive NP is first base-generated to the complement of the classifier, and later moved to the left periphery of the finite clause, as illustrated with Finnish examples below.

- (73) a. *Pekka osti [kolme kappaletta [NP kirjoja]].*
 Pekka bought three.NOM piece.SG.PAR book.PL.PAR
 ‘Pekka bought three books.’
- b. *[NP Kirjoja] Pekka osti [kolme (kappaletta) ____].*
 book.PL.PAR Pekka bought three.NOM piece.SG.PAR
 ‘Pekka bought three books.’

This analysis solves the problems with morphological mismatches: the classifier selects a partitive plural NP in constructions such as (73a), and the NP retains its case and number when it is moved to the left periphery in (73b).

5.3 The classifier as a functional head

As we saw in the previous section, Finnish and Estonian partitive splits have language-specific properties. We will therefore investigate the analysis of Finnish first, and discuss Estonian at the end of the section.

We thus propose that Finnish has a classifier *kappale* ‘piece’. For example, the classifier may take pronouns and full DPs as complements, as in (74a-c). Example (74c) is from the Internet. In addition *kappale* has also other, lexical uses. As a noun head, it can mean ‘a piece of music’, ‘object’ (physics term), or ‘paragraph’. However, in the split NP, none of these meanings are available.

- (74) Finnish
- a. *Pekka hankki [kaksi kappaletta näitä sohvia].*
 Pekka got two.NOM piece.SG.PAR these.PAR couch.PL.PAR
 ‘Pekka got two of these couches.’
- b. *[Montako kappaletta näitä sohvia] hän haluaa ____?*
 how.many piece.SG.PAR these.PAR couch.PL.PAR s/he wants
 ‘How many of these couches does s/he want?’

- c. *Eli perjaatteessa tarvin* [*kaksi kappaletta niitä*
 so principle.in need.1SG two.NOM piece.SG.PAR those.PAR
ylihinnoiteltuja sieniä].
 over-priced.PL.PAR sponge.PL.PAR
 ‘So in principle, I need two of those over-priced sponges.’

The analysis as a functional head is motivated by the observation that the classifier does not permit adjectival modifiers when it occurs in the complement of a numeral (75a). Similarly, the construction does not permit splitting between the adjective and the noun (75b). Instead, the adjective has to move with the rest of the partitive NP, as in (75c).

(75) Finnish

- a. **Pekka näki* [*kolme suurta kappaletta kirjoja*].
 Pekka saw three.NOM big.SG.PAR piece.SG.PAR book.PL.PAR
- b. *?*Kirjoja Pekka näki* [*kolme suurta* ____].
 book.PL.PAR Pekka.NOM saw three.NOM big.SG.PAR
- c. [*NP Suuria kirjoja*] *Pekka näki* [*kolme* ____].
 big.PL.PAR book.PL.PAR Pekka.NOM saw three.NOM
 ‘Pekka saw three big books.’

Possessive modification provides similar results. In Finnish, the possessor can occur either below or above the numeral, as in (76a-b).²⁶ In the partitive split, the possessor moves with the partitive NP (76c)

- (76) a. *Pekka lainasi* [*kaksi Merjan levyä*].
 Pekka borrowed two.NOM Merja.GEN record.SG.PAR
 ‘Pekka borrowed two of Merja’s records’
- b. *Pekka lainasi* [*Merjan kaksi levyä*].
 Pekka borrowed Merja.GEN two.NOM record.SG.PAR
 ‘Pekka borrowed Merja’s two records.’
- c. [*Merjan levyjä*] *Pekka lainasi* [*kaksi* ____].
 Merja.GEN record.PL.PAR Pekka borrowed two.NOM
 ‘Pekka borrowed two of Merja’s records.’

However, the possessor cannot co-occur with the classifier *kappale* in a continuous NP (77a). If the partitive split is derived from (77), it should not be possible to strand the possessor. This prediction is borne out (b-c).

- (77) a. **Pekka lainasi kolme Merjan kappaletta levyjä*.
 Pekka.NOM borrowed three Merja.GEN piece.SG.PAR record.PL.PAR
- b. **Levyjä Pekka lainasi kaksi Merjan* ____.
 record.PL.PAR Pekka.NOM borrowed two.NOM Merja.GEN
- c. **Levyjä Pekka lainasi Merjan kaksi* ____.
 record.PL.PAR Pekka.NOM borrowed Merja.GEN two.NOM

²⁶ The two sentences differ in meaning; in example (a), Merja has more than two records, and in example (76b), Merja has only two records and Pekka borrowed both of them.

Finally, the classifier *kappale* should be kept separate from measure phrases which are required by mass nouns in order to be counted, for example *a glass of milk*. In Finnish and Estonian, measure phrases appear to be full noun phrases and enable different types of modifiers. However, it should be noted that in both languages, the measure phrases permit splitting, as can be seen in the following examples (e.g. Koptjevskaja-Tamm 2001; Seppänen 1983). The dislocation of the partitive NP in Finnish has been analysed as being an instance A'-movement by Brattico (2008: 145) and Huhmarniemi (2012).

(78) Finnish

- a. *Pekka lapioi [yhden [kasan [DP hiekkaa]]]*.
 Pekka.NOM shoveled one.GEN pile.SG.ACC sand.SG.PAR
 'Pekka shoveled one pile of sand.'
- b. [_{DP} *Hiekkaa*] *Pekka lapioi* [_{DP} *yhden [kasan ____]*].
 sand.SG.PAR Pekka.NOM shoveled one.ACC pile.SG.ACC
 lit. 'Of sand, Pekka shoveled one pile.'

(79) Estonian

- a. *Peeter ostis [_{DP} ühe [koti [_{DP} kartuleid]]]*.
 Peeter.NOM bought one.GEN bag.SG.GEN potato.PL.PAR
 'Peeter bought one bag of potatoes.'
- b. [_{DP} *Kartuleid*] *Peeter ostis* [_{DP} *ühe [koti ____]*].
 potato.PL.PAR Peeter.NOM bought one.GEN bag.SG.GEN
 lit. 'Of potatoes, Peeter bought one bag.'

In conclusion, *kappale* does not take any modifiers, which points towards to an analysis where it is a functional head which occurs between the numeral and the partitive NP. Functional elements do not identify objects, but rather contribute to the interpretation of their complements. Another property of functional heads is that they are phonetically minimal, and this accounts for the fact that the classifier can be unpronounced when the NP is split. The remaining problem is, why the classifier has to be pronounced in a continuous NP but is optional in the partitive split.

Let us now move to the analysis of Estonian, which is not as straightforward as Finnish. Unlike in Finnish, the partitive split does not reconstruct in the presence of the classifier:

(80) Estonian

- a. *Raamatuid Peeter nägi kolm (tükkj)*.
 book.PL.PAR Peeter saw three.NOM piece.SG.PAR
 'Peeter saw three books.'
- b. **Peeter nägi kolm suurt tükkj raamatut / raamatuid*.
 Peeter saw three.NOM big piece.SG.PAR book.SG book.PL.PAR

Estonian *tükk* requires a singular mass noun as complement (81).

(81) Estonian

Peeter nägi kolm suurt tükki šokolaadi /
 Peeter saw three.NOM big.SG.PAR piece.SG.PAR chocolate.SG.PAR
juustu.
 cheese.SG.PAR
 ‘Peeter saw three big pieces of chocolate/cheese.’

The Estonian *tükk* thus behaves like a measure expression, similar to *liter*, *some* and *slice*. It can occur with abstract nouns (82a) and its meaning is not restricted to counting.

(82) Estonian

- a. *tükk aega*
 piece.SG.NOM time.SG.PAR
 lit. piece of time, interpretation: ‘quite a while’
- b. *palju tükke graniiti*
 many piece.PL.PAR granite.SG.PAR
 ‘many pieces of granite’

It thus seems that Estonian *tükk* ‘piece’ has different properties in the split construction than inside a continuous NP. In order to nevertheless apply the classifier analysis to Estonian, we would have to assume that the word *tükk* ‘piece’ is ambiguous. In a continuous NP, *tükk* ‘piece’ is a measure expression with its own selectional properties. However, in the split construction, *tükk* ‘piece’ is similar to Finnish *kappale* ‘piece’: a functional head that does not take any modifiers. A piece of evidence in support for this hypothesis is offered by example (83). Whereas in a continuous NP, *tükk* ‘piece’ may take a possessor (83a-b), this is not possible in the split noun phrase (c-d).

(83) Estonian

- a. *Peeter laenas Marise kaks tükki šokolaadi.*
 Peeter.NOM borrowed Maris.GEN two.NOM piece.SG.PAR chocolate.SG.PAR
 ‘Peeter borrowed two pieces of Mari’s chocolate.’
- b. *Peeter laenas kaks tükki Marise šokolaadi.*
 Peeter.NOM borrowed two.NOM piece.SG.PAR Maris.GEN chocolate.SG.PAR
 ‘Peeter borrowed two pieces of Mari’s chocolate.’
- c. **Plaat laenas Peeter Marise kaks.*
 record.PL.PAR borrowed Peeter.NOM Maris.GEN two.NOM
- d. **Plaat laenas Peeter kaks Marise.*
 record.PL.PAR borrowed Peeter.NOM two.NOM Mari.GEN

To summarise, although some characteristics of Estonian provide support for the classifier analysis, the evidence is not conclusive.

5.4 Morphological repair

This section investigates an alternative analysis for the partitive split, referred to as *morphological repair*. This analysis accounts for the number mismatches and may therefore be a possible alternative for the analysis of Estonian partitive split. However, this alternative is

not an attractive solution for Finnish due to the fact that the morphological repair would need to account for both case and number changes.

Fanselow and Cavar (2002) discuss morphological mismatches in different languages and present the following example from German, which displays a similar phenomenon we have observed in Finnish and Estonian. In example (84a) from Fanselow and Cavar (2002), the left part bears plural marking, although in the continuous DP (84c), the noun is in the singular.

- (84) German
- a. *Zeitungen lese ich nur eine.*
newspapers read I only one
'I only read one newspaper.'
 - b. **Ich lese nur eine Zeitungen*
 - c. *Ich lese nur eine Zeitung.*
'I read only one newspaper.'

Fanselow and Cavar (2002) propose that since a singular countable noun cannot typically appear alone in a sentence in German, the split of a singular NP in (84c) would lead to ungrammaticality in the surface structure. This problem is solved by changing the singular number to plural post-syntactically.

Let us apply the repair strategy to Estonian and Finnish data. Under this analysis, the NP is first in the singular (85a), then moves to the left periphery, and receives plural inflection due to a post-syntactic repair rule, as in (85b).

- (85) a. *Peeter ostis [DP kolm [NP raamatut]].*
Peeter.NOM bought three.NOM book.SG.PAR
'Peeter bought three books.'
- b. *[NP Raamatuid] Peeter ostis [DP kolm ____].*
book.PL.PAR Peeter bought three.NOM
'Peeter bought three books.'

The repair strategy can be motivated analogously to the German example (84) above. In Finnish and Estonian, singular countable nouns do not generally appear in the partitive in any other than the complement position. For example, the singular count noun cannot appear in the partitive case in the subject position (86a), although the plural form is possible (86b).²⁷

²⁷ Nevertheless, A'-movement can target a full DPs in this form, as in the following examples (i.a-b). However, in the split construction, the moving element would be the complement of the numeral, which is not a DP, but an NP. This NP is moved out of the scope of the numeral, to the left periphery of the finite clause. In this position, the repair strategy assigns plural inflection to the NP.

- (i) a. *Pekka katseli valokuvaa.*
Pekka.NOM watched photo.SG.PAR
'Pekka was looking at a photo.'
- b. *Valokuva Pekka katseli ____.*
photo.SG.PAR Pekka.NOM watched
'Pekka was looking at a photo.'

- (86) a. **Lasta leikkii kadulla.*
 child.SG.PAR play.3SG street.at
 b. *Lapsia leikkii kadulla.*
 child.PL.PAR play.3SG street.at
 ‘Children are playing in the street.’

Assuming a repair strategy, the constituent is thus changed to plural in (86b). However, as noted by Fanselow (2012), the repair strategy is not a feasible alternative for all languages with morphological mismatches. Note that demonstrative pronouns do not pose a problem for the repair account in Estonian, because the demonstrative is marginally possible in the complement of the numeral, see examples in (87) repeated from (68). The constituent also moves as a whole, as in (87d).²⁸

- (87) a. *Neid raamatuid Peeter ostis viis.*
 these.PAR book.PL.PAR Peeter.NOM bought five.NOM
 ‘Of these books, Peeter bought five.’
 b. **Peeter ostis viis neid raamatuid.*
 Peeter.NOM bought five.NOM these.PAR book.PL.PAR
 c. ?*Peeter ostis viis seda raamatut.*
 Peeter.NOM bought five.NOM this.SG.PAR book.SG.PAR
 ‘Peeter bought five books of this type.’
 d. [*Viis seda raamatut*] *ostis Peeter.*
 five.NOM this.SG.PAR book.SG.PAR bought Peeter.NOM

In Finnish, the comparable example to (87c) would be ungrammatical, see example (67) above.²⁹

6 Conclusions

This paper addressed discontinuous noun phrases in Finnish and Estonian. It was proposed that both languages have at least three types of split noun phrases. While the general properties of the split noun phrases are similar in Finnish and Estonian, a more detailed examination reveals intricate differences between the languages.

The focus of this paper was on the partitive split and in particular, the morphological mismatch between the continuous NP and the split construction. It was argued that the partitive split is derived by sub-extraction of the NP, where the two parts are originally inside the same DP. The movement of the partitive NP was shown to have the general properties of A'-movement in both languages.

Two alternative analyses were examined: First, an account in terms of a classifier head which facilitates the selectional requirements of the quantifier/numeral and accounts for the morphological mismatch. The presence of the classifier was motivated by the observation

²⁸ In sentences (87c-d), the DP *seda raamatut* ‘this book’ appears to be coerced into type/kind reading, e.g. five books of this type/kind.

²⁹ The insertion of a determiner has been proposed for e.g. German as part of the repair strategy (Fanselow and Cavar 2002).

that morphological mismatches occur only with quantifying expressions that cannot take an uncountable NP-complements. Second, we investigated a morphological repair account, where the partitive NP receives morphological features only after movement.

The main advantage of the classifier analysis is that it deploys grammatical mechanisms that are already well-known and present in related constructions. In addition, the morphological mismatch is only apparent because the partitive NP is in the same form in the continuous NP and in the split construction. Finally, it provides a testable hypothesis for the analysis of split noun phrases in other languages with morphological mismatches.

It was shown that the classifier analysis accounts for the Finnish partitive splits, although the exact conditions for the pronunciation of the classifier were left open. However, the classifier analysis cannot be adopted as such to Estonian, and the morphological repair account was introduced as an alternative.

References

- Alho, Irja H. (1992). "Distinguishing kind and set in Finnish". *Studies in the Linguistic Sciences* 22.1, 1–16.
- Arnhold, Anja (2009). "Prosody of discontinuous noun phrases in Finnish". In: *Nordic Prosody: Proceedings of the Xth Conference*. Ed. by Martti Vainio, Reijo Aulanko, and Olli Aaltonen. Frankfurt am Main: Peter Lang, 37–48.
- Brattico, Pauli (2008). "Kayne's model of case and Finnish nominal phrases". *Nordic Journal of Linguistics* 31.2, 135–160.
- Brattico, Pauli (2010). "One-part and two-part models of nominal case: evidence from case distribution". *Journal of Linguistics* 46.1, 47–81.
- Brattico, Pauli (2016). "Finnish null subjects and their morphosyntax". *Manuscript*.
- Cheng, L. and R. Sybesma (1999). "Bare and not-so-bare nouns and the structure of NP". *Linguistic Inquiry* 30, 509–542.
- Chomsky, Noam (1981). *Lectures in government and binding: the Pisa lectures*. Dordrecht: Foris.
- Chomsky, Noam (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.
- Danon, Gabi (2012). "Two structures for numeral-noun constructions". *Lingua* 122, 1282–1307.
- Erelt, M., R. Kasik, H. Metslang, H. Rajandi, K. Ross, H. Saari, K. Tael, and S. Vare (1993). *Eesti keele grammatika II [The grammar of the Estonian language]*. Tallinn: Eesti Teaduste Akadeemia Eesti Keele Instituut.
- Erelt, Mati (2009). "Typological overview of Estonian syntax". In: *STUF, Language Typology and Universals*. Vol. 62. 1/2. Akademie Verlag, 6–28.
- Erelt, Mati, Helle Metslang, and Helen Plado (2016). "Alus". In: *Eesti keele lauseõpetus [The syntax of modern Estonian]*. Ed. by Mati Erelt and Helle Metslang. preprint June 2016. Tartu: Tartu Ülikooli kirjastus.
- Fanselow, Gisbert (1988). "Aufspaltung von NPn und das Problem der 'freien' Wortstellung". *Linguistische Berichte* 114, 91–113.

- Fanselow, Gisbert (2012). "Morphological mismatches in discontinuous noun phrases". In: *Interfaces of morphology. A Festschrift for Susan Olsen*. Ed. by Holden Härtl. *Studia grammatica* 74. Berlin, Boston: De Gruyter, 11–25.
- Fanselow, Gisbert and Damir Cavar (2002). "Distributed deletion". In: *Theoretical approaches to universals*. Ed. by Artemis Alexiadou. Amsterdam: John Benjamins, 65–107.
- Gröndahl, Tommi (2015). "Määräisyys funktionaalisenä pääsanana suomen kielen nominausekkeessa". MA thesis. University of Helsinki.
- Hakulinen, Auli and Fred Karlsson (1979). *Nykysuomen lauseoppi*. Helsinki: Finnish Literature Society.
- Hakulinen, Auli, Maria Vilkkuna, Riitta Korhonen, Vesa Koivisto, Tarja Riitta Heinonen, and Irja Alho (2004). *Iso suomen kielioppi*. Helsinki: Finnish Literature Society.
- Henk, Paula (2010). *Information Structure of Estonian – Compared to Finnish and Hungarian*. Bachelor's thesis. PICS Publications of the Institute of Cognitive Science. Osnabrück, Germany.
- Hiitema, Katrin and Kersti Börjars (2003). "The emergence of a definite article in Estonian". In: *Generative Approaches to Finnic and Saami Linguistics*. Ed. by Diane Nelson and Satu Manninen. Stanford: CSLI Publications.
- Holmberg, Anders and Urpo Nikanne (2002). "Expletives, subjects and topics in Finnish". In: *Subjects, Expletives, and the EPP*. Ed. by Peter Svenonius. Oxford: Oxford University Press, 71–106.
- Huang, Cheng-Teh James (1982). "Move wh in a language without wh movement". *Linguistic Review* 1, 369–416.
- Huhmarniemi, Saara (2012). *Finnish A'-movement: Edges and Islands*. Institute of Behavioural Sciences, Studies in Cognitive Science 2. Helsinki: University of Helsinki. URL: <http://urn.fi/URN:ISBN:978-952-10-7712-8>.
- Ionin, Tania and Ora Matushansky (2006). "The composition of complex cardinals". *Journal of Semantics* 23.4, 315–360.
- Juvonen, Päivi (2000). "Grammaticalizing the Definite Article. A study of definite adnominal determiners in a genre of spoken Finnish". PhD thesis. Stockholm University.
- Kaiser, Elsi (2002). "The syntax-pragmatics interface and Finnish ditransitive verbs". In: *Proceedings of the 9th Annual Conference of the Student Organization of Linguistics in Europe (ConSOLE IX)*. Ed. by M. van Koppen, E. Thrift, E. J. van der Torre, and M. Zimmermann. Lund.
- Koptjevskaja-Tamm, Maria (2001). "'a piece of the cake' and 'a cup of tea': partitive and pseudo-partitive nominal constructions in the circum-baltic languages." In: *The Circum-Baltic languages: Typology and contact*. Ed. by Östen Dahl and Maria Koptjevskaja-Tamm. Vol. 2, 523–568.
- Laury, Ritva (1991). "On the development of the definite article se in spoken Finnish". In: *The SKY Yearbook*, 93–121.
- Laury, Ritva (1997). *Demonstratives in interaction: the emergence of a definite article in Finnish*. Amsterdam: John Benjamins.
- Metslang, Helena (2013). "Grammatical relations in Estonian: subject, object and beyond". PhD thesis. University of Tartu.

- Metslang, Helle (2016). “Kvantorifraas [quantifier phrase]”. In: *Eesti keele lauseõpetus [The syntax of modern Estonian]*. Ed. by Mati Ereht and Helle Metslang. preprint June 2016. Tartu: Tartu Ülikooli kirjastus.
- Miljan, Merilin and Ronnie Cann (2013). “Rethinking case marking and case alternation in Estonian”. *Nordic Journal of Linguistics* 36.3, 333–379.
- Nelson, Diane and Ida Toivonen (2003). “Counting and the grammar: case and numerals in Inari Saami”. In: *Generative Approaches to Finnic and Saami Linguistics*. Ed. by Diane Nelson and Satu Manninen. Stanford, CA.: CSLI Publications, 321–340.
- Norris, Mark (2014). “A theory of nominal concord”. PhD thesis. Santa Cruz: University of California.
- Ott, Dennis (2011). “Local instability: The syntax of split topics.” PhD thesis. Harvard University.
- Pajusalu, Renate (1997). “Is there an article in (spoken) Estonian?” In: *Estonian: Typological Studies II*. Tartu: Tartu University Press.
- Pajusalu, Renate (2000). “Indefinite determiners üks and mingi in Estonian”. In: *Estonian: Typological Studies IV*. Ed. by Mati Ereht. Tartu: Tartu University Press.
- Quine, Willard van Orman (1960). *Word and object*. Cambridge, MA: MIT Press.
- Ross, John Robert (1967). “Constraints on variables in syntax”. PhD thesis. MIT.
- Seppänen, Aimo (1983). “Finnish “kaksi poikaa””. *Studia Linguistica* 37.2, 161–174.
- Spoelman, Marianne (2013). “Prior linguistic knowledge matters: The use of the partitive case in Finnish learner language”. PhD thesis. Oulu: University of Oulu.
- Tael, Kaja (1990). *An approach to word order problems in Estonian*.
- Trosterud, Trond (1993). “Anaphors and binding domains in Finnish”. In: *Case and other functional categories in Finnish syntax*. Ed. by Anders Holmberg and Urpo Nikanne. New York: Mouton de Gruyter, 225–243.
- Vainikka, Anne (1989). “Deriving Syntactic Representations in Finnish”. PhD thesis. University of Massachusetts Amherst.
- Vainikka, Anne and Pauli Brattico (2009). “The Finnish Accusative: Long Distance Case Assignment by phi-Agreement”. Manuscript submitted for publication.
- Vilkuna, Maria (1989). *Free word order in Finnish: its syntax and discourse functions*. Helsinki: Finnish Literature Society.
- Vilkuna, Maria (1995). “Discourse configurability in Finnish”. In: *Discourse Configurational Languages*. Ed. by Katalin É. Kiss. Oxford/New York: Oxford University Press, 244–268.
- Vilkuna, Maria (1996). *Suomen lauseopin perusteet*. Kotimaisten kielten tutkimuskeskuksen julkaisu 90. Helsinki: Edita.
- Vilkuna, Maria (1998). “Word order in European Uralic”. In: *Constituent Order in the Languages of Europe*. Ed. by Anna Siewierska. Berlin, New York: Mouton de Gruyter, 173–234.
- van Riemsdijk, Henk (1989). “Movement and regeneration”. In: *Dialect variation and the theory of grammar*. Ed. by Paola Benincà. Dordrecht: Foris, 105–136.

Saara Huhmarniemi

University of Helsinki, Institute of Behavioural Sciences, Cognitive Science

saara.huhmarniemi@helsinki.fi

Merilin Miljan

Tallinn University, School of Humanities & University of Tartu, Institute of Estonian and
General Linguistics

merilin.miljan@tlu.ee OR merilin.miljan@ut.ee

Report on the Second International Workshop on Computational Linguistics for Uralic Languages

Tommi A. Pirinen, Eszter Simon, Francis M. Tyers, Veronika Vincze

The Second International Workshop on Computational Linguistics for Uralic Languages (SIWCLUL) was held in Szeged in January 2016. The goals of the conference series include increased co-operation between the researchers, universities and research centres working on Uralic languages. The event gathered a number of participants from all over Eurasia, including Finland, Hungary, Estonia, Ireland, Germany, Austria and Norway among others. The conference also marked a start of an *Association for Computational Linguistics' Special Interest Group for Uralic Languages* (ACL SIGUR).

Keywords: *Finno-Ugric Languages and Linguistics, Computational linguistics*

1 Introduction

The Second International Workshop on Computational Linguistics for Uralic Languages was held in Szeged, Hungary, on 20 January 2016. The objective of the workshop was to bring together researchers working on computational approaches to working with the following languages: Finnish, Hungarian, Estonian, Voru, Setu, the Sámi languages, Komi (Zyrian, Permyak), Mordvin (Erzya, Moksha), Mari (Hill, Meadow), Udmurt, Nenets (Tundra, Forest), Enets (Tundra, Forest), Nganasan, Selkup, Mansi, Khanty, Veps, Karelian, Ingrian (Izhorian), Votic, Livonian, Ludic, Kven and other related languages.

The first edition of the workshop was held in Tromsø, Norway, in January 2015. This series of workshops is a new attempt to gather researchers of Uralic computational linguistics together, to ensure that they work towards common goals with a minimal amount of overlapping and redundant work. To that effect, the conference series has also formed a new special interest group under the guidance of the Association for Computational Linguistics (ACL).

Two organisers of the first event were also organisers of the second one, thus guaranteeing the continuity between the parts of the series. Local organisers were researchers from the University of Szeged and from the Research Institute for Linguistics of the Hungarian Academy of Sciences.

Original, substantial and unpublished papers were solicited that describe work-in-progress systems, frameworks, standards and evaluation schemes. Additionally, demos and tutorials were also invited which present systems and standards that pursue the goal of interoperability and unification of different projects, applications and research groups. The topics in which papers were expected are: parsers, analysers and processing pipelines of Uralic languages; lexical databases, electronic dictionaries; finished end-user applications

aimed at Uralic languages, such as spelling or grammar checkers, machine translation or speech processing; evaluation methods and gold standards, tagged corpora, treebanks; reports on language-independent or unsupervised methods as applied to Uralic languages; surveys and review articles on subjects related to computational linguistics for one or more Uralic languages; any work that aims at combining efforts and reducing duplication of work; and proposals concerning how to elicit activity from the language community, agitation campaigns, games with a purpose. To maximise the possibility of reproducibility, replication and reuse, submissions that present free/open-source language resources and make use of free/open-source software were particularly encouraged.

One of the aims of this gathering is to avoid unnecessary duplicated work in the field of Uralistics by establishing connections and interoperability standards between researchers and research groups working at different sites. It is now recognised as a serious problem that there is a lack of gold standards and evaluation metrics covering all Uralic languages including those with national support, thus any work towards better resources in these fields were greatly appreciated.

There were 10 accepted papers, 4 of which were presented as oral presentations in two sessions, while the others were poster presentations and/or interactive demonstrations. Additionally, two tutorials were included in a separate session. The topics and languages discussed in the workshop were wide and varied. This year the conference featured a state-of-the-art introduction to Estonian language technology resources. As one of the aims of the workshop series is to promote interoperability between the related Uralic languages, multiple presentations and two tutorials were held to highlight best common practices in the fields of computational linguistics intersecting software engineering.

The workshop gathered 28 scholars from 8 countries including Hungary, Estonia and Finland, where the national language belongs to the Uralic family, and countries such as Russia and Norway, where several Uralic languages are spoken as minority languages.

After a short opening, the first presentation was given by the invited lecturer, András Kornai. In a poster boaster session, each participant whose paper was accepted for poster presentation or demonstration had a few minutes to introduce his/her poster's topic. This was followed by the poster and demo session, and two sessions for oral presentations. In the afternoon, there was another poster and demo session, followed by two tutorials, while the event was closed with a SIGUR meeting and some closing remarks.

Below we report on the presentations and posters under thematic schemes: While Section 2 gives a brief overview of the presentations and discussions under the topic of best common practices, in Section 3 we introduce language-specific resources presented in the workshop. In Section 4 we describe our efforts to form a special interest group and possible related activities. Section 5 presents a sort of a desiderata for future revisions of the conference and pan-Uralic co-operation.

2 Best Common Practices in Uralic Computational Linguistics

The invited talk was given by András Kornai. He is full professor at the Budapest University of Technology, senior scientific advisor at the Computer and Automation Research Institute of the Hungarian Academy of Sciences, and the leader of the mathematical linguistics

research group in the Research Institute for Linguistics of the Hungarian Academy of Sciences. In his talk, entitled *Computational linguistics of borderline vital languages in the Uralic family*, he applied the methodology of Kornai (2013) to the Uralic family with the specific goal of *triage*, to help the community decide where the effort is best placed. As in battlefield triage, where the relatively lightly wounded and the very heavily wounded are treated last, he suggested to direct the very limited resources of the computational linguistics community towards the middle class of borderline languages where neither vital nor still/heritage status can be established.

Thierry Poibeau and Svetlana Toldova had a poster which presented some preliminary experiments concerning the automatic processing of Finno-Ugric languages. They presented symbolic methods as well as machine learning ones. Given the lack of corpora for some languages, they found that finite state transducers may sometimes be the best approach, even if machine learning techniques are supposed to outperform symbolic methods.

Kristian Kankainen demonstrated his tool, Minority Translate, which streamlines the process of creating, editing and saving new articles in any language edition of Wikipedia, also the new language editions starting out in the Incubator. Wikipedia can be treated as a language resource in itself for the lesser resourced languages, as well as a source of several other language technology tools.

Johannes Dellert introduced a new method for inducing a language contact model from lexical data. Based on automatically gathered and manually annotated sets of etymologically related words, the method analyses possible paths of borrowing in terms of lexical flow. In an evaluation on a large lexical database comprising 1,016 concepts across 26 Uralic languages and 18 neighbouring languages, the method detected and correctly inferred the directionality of many instances of cross-family language contact.

Francis Tyers and Tommi Pirinen reported their experience with regard to interoperability of the Uralic languages' practices and tagging standards when used in the context of rule-based machine translation. The Uralic languages exhibit certain resemblances: many of them have similar case inventories, word order and non-finite clause forms. However, current rule-based grammatical resources take many different approaches to encoding this information. In their presentation, Tyers and Pirinen provided some guidelines and suggestions to facilitate future work in the direction of interoperability.

In the tutorials session, first Trond Trosterud presented the language resource repository from the University of Tromsø *Giellatekno*¹ group with best common practices in rule-based open source natural language processing resources. Afterwards Veronika Vincze and Francis Tyers presented the *Universal Dependencies*² annotation scheme, which is on track to become an international de facto standard for part-of-speech tagging and dependency annotations.

¹ <http://giellatekno.uit.no>

² <http://universaldependencies.org/>

3 Language-specific Resources

Jeremy Bradley had a poster presentation in which he introduced his efforts to create a web-based automatic transcription and transliteration software for Uralic and non-Uralic languages. For four literary standards – Meadow Mari, Hill Mari, Russian, and Tatar – an operational interface can be found at transcribe.mari-language.com. His poster detailed many of the fine aspects of writing systems used for (Meadow) Mari that he had to take into consideration when creating transcription mechanisms for that language.

Trond Trosterud presented their common research with his colleagues: Lene Antonsen, Marja-Liisa Olthuis and Erika Sarivaara. Their poster, entitled *Modelling the Inari Sámi morphophonology as a finite state transducer*, presented a set of morphophonological problems coming up when they were working on a transducer for Inari Sámi, a language with a complex and not very well documented morphophonology. As they said: modelling the grammar as a finite state transducer gives more insight into the Inari Sámi morphophonology, and the resulting program will be the foundation of all future Inari Sámi language technology applications.

Tommi Pirinen and his colleagues, Antonio Toral and Raphael Rubino, reported on experiments with Finnish-English statistical machine translation. They jointly used rule-based and unsupervised approaches to segmentation. They found that in terms of automatic metrics, the best system is the one that combines both rule-based and unsupervised segmentations, while human evaluation shows that the outputs produced by a statistical machine translation system with rule-based segmentations are preferred over those of the system that uses unsupervised segmentations.

Axel Wisiolek and Zsófia Schön presented their poster on an Ob-Ugric database, a web-based framework for the storage and advanced retrieval of annotated corpora and corpus-based lexical databases of Khanty and Mansi dialects. The database building is a work in progress within the framework of the project titled *Ob-Ugric database: analysed text corpora and dictionaries for less described Ob-Ugric dialects (OUDB)*.

Peter Smit presented a generic model for automatic speech recognition applied to Northern Sámi as an example for a setup of lesser resourced languages. Since the lack of technology and applications may threaten the existence of these languages, it is important to study how to create speech recognizers with minimal effort and low resources.

Kadri Muischnek and her colleagues at the University of Tartu gave an overview of the state of the art of tools and resources for the syntactic analysis of Estonian. They presented a manually annotated dependency treebank containing 400,000 words. A morpho-syntactic disambiguator, a shallow parser and a dependency parser were also introduced, all of which are based on the Constraint Grammar formalism.

4 SIGUR

In the first workshop, there was a consensus that computational linguists working with Uralic languages should organise themselves in the form of an ACL-approved special interest group. The organisers then negotiated the founding of a group with the ACL secretary and the SIG officer. After affirmation from the ACL meeting, this second workshop's

business meeting was used as a co-ordinated founding meeting for the newly created SIG. The business meeting was attended by the participants who were also members of the SIG and decided upon the board and the founding papers of the new SIG. The details of the special interest group can also be found on the SIG website,³ which also includes the public minutes of the meeting.

5 Future plans and desiderata

In the formal meeting it was decided that the workshop series should carry on and plans were set for the forthcoming course of action, including the next workshop potentially to be organised in St. Petersburg. One of the rationales behind wishing to hold the next workshop in Russia is to increase co-operation with researchers in Russia. As the majority of Uralic languages are situated in Russia, there is a vast amount of ongoing research and resources that are of interest to workshop-goers and researchers.

The newly formed special interest group will take an active role in co-ordinating computational linguistics for Uralic languages, including forming best current practices and sharing information and resources in a centralised place.

References

Kornai, A. (2013). "Digital Language Death". *PLoS ONE* 8.10.

Tommi A. Pirinen
Hamburger Zentrum für Sprach Korpora, Universität Hamburg
tommi.antero.pirinen@uni-hamburg.de

Eszter Simon
Research Institute for Linguistics, Hungarian Academy of Sciences
simon.eszter@nytud.mta.hu

Francis M. Tyers
University of Tromsø
francis.tyers@uit.no

Veronika Vincze
MTA-SZTE Research Group for Artificial Intelligence
vinczev@inf.u-szeged.hu

³ <http://gtweb.uit.no/sigur>