Estonian and Finnish show that hybrid agreement is structural

Mark Norris

1. Introduction

This squib has two goals. The first goal is to describe a pattern of hybrid agreement in number in Estonian/Finnish numeral-noun constructions (hereafter, NNCs). The pattern is exemplified for Estonian in (1) and Finnish in (2). The numeral is boxed for clarity in each example.1

(1) nee-d (ilusa-d) viis (ilusa-t) maja
DEM-PL.NOM beautiful-PL.NOM 5.SG.NOM beautiful-SG.PAR house.SG.PAR
‘these beautiful five houses’ Estonian (Erelt et al. 1993: 143)

(2) ne (pilaantune-et) [kaksi] (pilaantunut-ta) leipä-ä
those.PL rotten-PL two.SG rotten-SG.PAR bread-SG.PAR
(no translation given) Finnish (Brattico 2010: 60-61)

In both examples, the noun, numeral, and everything in between are singular. Everything to the left of the demonstrative—in these examples, a demonstrative and an adjective—is plural. Note that adjective number tracks the adjective’s placement in the nominal phrase: plural on the left of the numeral, singular on the right. I take this to be an example of hybrid agreement, because there are two values for the same feature in one nominal phrase.2 Describing the pattern of hybrid agreement in Estonian/Finnish NNCs is one of the goals of this squib due to the fact that they have largely gone unnoticed in the literature on hybrid agreement, with Landau (2016) being the most prominent exception I am aware of (but see also Brattico (2010, 2011) and Norris (2014, 2018a)).3

The second goal of this squib is to use the Estonian/Finnish pattern to challenge a recent claim by Bruening (2020) that hybrid agreement can be analyzed without nominal functional structure. In that paper, Bruening discusses evidence that has been levied in favor of the so-called DP analysis and then reanalyzes those cases in an NP-only framework. He concludes that those case studies are not arguments in favor of the DP hypothesis insofar as they can be analyzed with only NP. Among the things he discusses is hybrid agreement in Bosnian/Croatian/Serbian (hereafter BCS), exemplified below (see also Puškar (2018) for detailed discussion).

---

1 Glossing abbreviations: DEM — demonstrative, F — feminine, M — masculine, NOM — nominative, PAR — partitive case, PL — plural, PST — past, REFL — reflexive, SG — singular
2 I believe my use of hybrid agreement is a slight departure from its use by Bruening (2020), who variably uses the terms agreement with hybrid nouns and hybrid agreement. What I ultimately propose is that the pattern in Estonian/Finnish is not about nouns, so the term agreement with hybrid nouns is unsuitable for Estonian/Finnish. I use hybrid agreement here to mean that there are two values for the same feature in what is otherwise a single agreement domain.
3 Though Landau (2016) does discuss Finnish, it is only a cursory exploration, and as I discuss in Norris 2018a, there are non-trivial issues with the analysis of Finnish that Landau sketches.

---
The hybrid agreement pattern in BCS involves gender. Nouns like vladik ‘bishop’ can sometimes trigger either feminine or masculine agreement, as indicated on the verb posvadali/posvadale ‘argue’ in (3).

What is special about vladik (and a small number of other nouns) is that it bears the morphological form of a feminine noun, but its canonical referent is socioculturally male. In BCS, such a combination can result in hybrid agreement with some degree of optionality. However, the optionality disappears once an agree-er matches the semantic (i.e., sociocultural) gender. When this happens, every subsequent agree-er must follow suit. This is why the feminine plural verb form is ungrammatical in (4): the masculine demonstrative ovi precludes further feminine (i.e., grammatical) agreement.

Bruening (2020) proposes an analysis of BCS hybrid agreement without DP or any other nominal functional structure. In doing so, he argues against Salzmann (2018, 2020), who claims that BCS hybrid agreement cannot be generated without reference to nominal functional structure, and therefore, that BCS hybrid agreement is an argument in support of the DP hypothesis. In proposing an NP-only analysis, Bruening (2020) aims to show that hybrid agreement (as exemplified by BCS) is not an argument in favor of the DP hypothesis. It is this specific claim that I challenge here, arguing that some hybrid agreement—namely, that found in Estonian/Finnish NNCs—does, in fact, require functional structure; a fortiori, an NP-only theory of nominal structure is too restrictive.

Bruening’s analysis of hybrid agreement does not work for Estonian/Finnish because their hybrid agreement differs from that of BCS in two key ways. The first difference is that in Estonian/Finnish hybrid agreement, there is no optionality/variation with respect to its appearance. Syntactic position (as determined by linear order) determines the number value of the modifier, and if there is any deviation from this, I suspect it is either negligible or significantly smaller than the amount of variation in BCS hybrid agreement. The second difference is that Estonian/Finnish hybrid agreement is fully generalized, meaning that all nouns participate in the pattern (so long as they can be counted with cardinal numerals). In BCS, hybrid agreement emerges only with a small subset of nouns in the language—only in the feminine gender, and only if the entities denoted by the noun are (or can be) socioculturally male.

In section 3, I argue that Estonian/Finnish hybrid agreement cannot be treated as lexical; its source is structural. As a result, analyzing it without functional structure requires complex and/or idiosyncratic assumptions about features that are tantamount to admitting functional structure anyway (see also Salzmann (2020)). In section 4, I use what space is left to sketch what a structural analysis would look like, but the primary focus is to carefully show Estonian/Finnish hybrid agreement is unique (thus far) in the hybrid agreement literature and as such, some analyses of BCS agreement will not work for Estonian/Finnish, including Bruening’s.

2. Preserving endocentricity in NP-internal hybrid agreement

The key challenge for an NP-only analysis of hybrid agreement is preserving endocentricity, the tenet that the projections of a head must contain the same information as the head (i.e., they must not gain or lose information). For hybrid agreement in an NP-only analysis, the entire nominal phrase (from top to bottom) has the same head, but agreeing modifiers in the NP do not all reflect the same feature values. It is not clear how these modifiers can have different values without changing the features of the NP, which would violate endocentricity.

I am hedging here, because this study is based on descriptions in Estonian grammars and my fieldwork, but I am not aware of any corpus or experimental study of agreement in Estonian/Finnish numeral-noun constructions. However, the Estonian grammars cited here are quite thorough, and they do not mention any exceptions to the agreement patterns discussed herein.
Bruening’s (2020:9) analysis of BCS hybrid agreement comprises two proposals, represented in (5)-(6) below.

(5) 
\[ N \]
\[ Φ \]
\[ [\_ GENDER] \]
\[ Gr \]
\[ [F GENDER] \]
\[ Sem \]
\[ [M GENDER] \]

(6) Principle of Semantic Preference (hereafter **SemPref**):
When a probe copies a Sem value for any feature, this value is registered in the matched category’s Φ-node.5

There are two key pieces of the proposal in (5). First, grammatical (Gr) and semantic (Sem) gender features are on the same level of structure, meaning they are equidistant from any higher element in need of a gender feature. Second, they share a mother node, Φ, which has its own gender feature. Φ’s gender feature begins unvalued, but its value can be fixed during the course of a derivation. If Φ has a fixed gender value, that value is closer to any subsequent agree-ers than either the Gr or Sem values due to their hierarchical structure. The circumstances in which Φ’s gender feature is valued are partially dictated by the principle in (6). When a probe agrees with the Sem value of one of these nouns in BCS, the Φ-node is automatically fixed to match the Sem value. This codifies the observed pattern whereby it is not possible to switch back to grammatical agreement once the switch to semantic agreement is made.

With these two proposals in (5)-(6), the patterns seen in BCS (and, I conjecture, every similar form of agreement discussed by Landau (2016) apart from Finnish) fall out without complication. For reasons of space, I must refer the reader to Bruening (2020) for a full demonstration of the analysis. Importantly, Bruening (2020) contends that valuing the Φ-node somewhere in the middle of the NP does not violate endocentricity, because the NP is not gaining or losing information that the N did not have. Rather, it is a restructuring of that information. This raises questions about the concept of information with respect to endocentricity, but again space limits prevent further investigation; I only note this to underscore that it is the core challenge to an NP-only framework that is posed by hybrid agreement patterns.6

3. Estonian/Finnish agreement in an NP-only theory

Following Bruening’s (2020) proposals, Estonian/Finnish nominals like those in (1)-(2) have the structure in (7). A-hi and A-lo represent the two possible adjective positions. I have added corresponding number values for clarity:

(7) 
\[ NP \]
\[ Dem \]
\[ plural \]
\[ A-hi \]
\[ plural \]
\[ Num(eral) \]
\[ singular \]
\[ A-lo \]
\[ singular \]
\[ N \]

5 Bruening (2020) does not give a name to this proposal and does not set it off from the text as I have done. He says the following in text: “[T]here is a principle that says that as soon as something agrees with Sem, the features of Sem are copied onto the Φ node.”

* I must note here as well that Bruening (2020) does consider a variant whereby the new structured representation is, in fact, shared throughout the NP. This version would rely on derivational timing to explain why lower modifiers are still able to agree with the grammatical value, but Bruening (2020) does not formalize it in detail.
In order to preserve endocentricity given that this entire phrase is a single NP, the representation of number in a Bruening-style account of Estonian/Finnish must be maximally similar to BCS. Number values must be lexically hybrid, i.e., nouns would be marked for both [Gr:SG] and [Sem:PL] number values, as in (8).

(8) Lexical representation for hybrid number following Bruening (2020):

```
   N
    Φ
       [ _ NUMBER]
          Gr
          [SG NUMBER]  Sem
          [PL NUMBER]
```

When A-hi, the second-merged adjective, looks for a number value, it must find the semantic value [Sem:PL]. Due to SemPref, this then fixes the Φ node’s number value at [Pl], as shown in (9).

(9) When agree-er (here, A-hi) matches Sem value, Φ also matches Sem value due to SemPref.

```
   NP
    Φ: [ Pl NUMBER]
       Gr:[Sg] Sem:[Pl]
          Dem
          'these'
            [ Pl NUMBER]
          NP
            Φ: [ Pl NUMBER]
               Gr:[Sg] Sem:[Pl]
                  A-hi
                  'beautiful'
                    [ Pl NUMBER]
                 NP
                   Φ: [ _ NUMBER]
                      Gr:[Sg] Sem:[Pl]
                          Num
                          'five'
                            [ SG NUMBER]
                         N
                            Φ: [ _ NUMBER]
                               Gr:[Sg] Sem:[Pl]
```

Because the number features are structured (in a similar enough way to syntactic structure), the Φ node’s number value is closer to any future agreement targets than the individual [Gr:SG] and [Sem:PL] features. Thus, when the demonstrative merges, the closest value is [Pl], and so the demonstrative must be plural. This analysis also predicts that the predicate must be plural when agreeing with an NP like the one in (9). Though not shown here for space reasons, this is borne out. As with BCS, no (completely) new features are added to N/P; only their structural relationship is changed. In this way, endocentricity is preserved in the same way as in Bruening’s (2020) analysis of BCS.

There are technical and conceptual problems for this account of hybrid number agreement in Estonian/Finnish. Unlike BCS, there is no variation/choice in modifier feature values. Unlike BCS, this pattern of hybrid agreement is not a pattern restricted to a particular morphological class of nouns. Unlike BCS, this pattern is observed in only one syntactic context—the NNC—and putting hybrid features on nouns misses the connection to numerals. I discuss these in turn.

Importantly, the facts are clearest when there is a modifier manifesting plural morphology, e.g., a high adjective or demonstrative. So far as I know, in both Estonian and Finnish, the verb is obligatorily plural when agreeing with such an NP. However, the facts are complex when the NP has no higher modifiers. Rather than being exclusively plural like demonstratives, verbal predicates can be singular in such instances. Nelson & Toivonen (2000) claim that singular agreement is required in Finnish, but they provide only one minimal pair. In Estonian, plural is certainly possible; for discussion and examples, see Norris (2021).
3.1. No optionality: No choice in agreement in NNCs in Estonian/Finnish

I note first that, unlike BCS hybrid agreement, there is no choice in the number values within NNCs in Estonian/Finnish. Everything up to and including the numeral must be singular (for the Bruening-style analysis, the grammatical value). Everything after the numeral must be plural (for the Bruening-style analysis, the semantic value). This is different from the BCS pattern whereby the demonstrative could match either the grammatical value or the semantic value (at least, for some speakers).

Note that this cannot be about A-hi obligatorily agreeing “with the numeral” in such a way that the numeral can be morphologically singular but trigger plural agreement (e.g., due to their connection to non-atomic sets), because the feature values of the NP node dominating the numeral should be closer to the demonstrative. So, all else being equal, we expect the modifier merged after the numeral to be able to match either the grammatical singular value or the semantic plural value, but it cannot; it must be plural.8

While this empirical difference is important to note, I believe there is a solution that is reasonable within the confines of Bruening’s approach. Bruening (2020) stipulates a number of principles throughout the paper, including what I call SemPref. Thus, we could stipulate that the semantic value must be chosen after numeral is merged. We must also stipulate that the semantic value cannot be chosen before the numeral is merged. As Bruening observes, there is nothing inherently wrong with such stipulations, and he contends that his stipulations are no worse than those required by the DP analysis. Ultimately, perhaps adding these additional stipulations surrounding the numeral will still be better than those required of the DP analysis. However, this problem is an initial signal that the NP-only analysis might not be on the right track, and in fact, there is another issue that poses a stronger challenge to an NP-only analysis of this pattern.

3.2. Hybrid Number in Estonian/Finnish is not lexical

A key ingredient of Bruening’s (2020) NP-only analysis of hybrid agreement in BCS is representing both the grammatical and semantic gender features on the noun itself. This differs from most structural analyses where semantic gender features are introduced on heads higher in the structure. From a mechanical perspective, this choice is critical, as representing both features on the N0 enables Bruening to preserve endocentricity. However, it also has some intuitive appeal, because the words that participate in hybrid agreement in BCS form an exceptional class. Specifically, the only nouns exhibiting hybrid agreement in BCS are those which (i) denote entities that are (or can be) socioculturally male and (ii) belong to declension class II, which is otherwise largely feminine nouns (Puškar 2018: 277-278). Thus, since participation in hybrid agreement is in some ways a property of these nouns insofar as it is connected to both their lexical form and lexical meaning, it is reasonable to associate it with their lexical entries.

Turning back to Estonian/Finnish, I contend that encoding hybrid agreement in NNCs via the noun’s feature structure is untenable in Estonian/Finnish for two reasons. First, the pattern in Estonian/Finnish is not tied to an exceptional class of nouns as in BCS. If the noun can be used in an NNC—i.e., if it is a count noun or a mass noun that can be coerced into a count noun—then the noun participates in hybrid agreement.9 Second, unlike in BCS, hybrid agreement in Estonian/Finnish is construction-dependent. What I mean by this is it surfaces only in NNCs. Outside of NNCs, demonstratives must match their associated nouns in number (see (10)) and predicate number must match the grammatical number of the associated subject (see (11)).10

8 It is not clear to me that we could use anti-locality as an escape here given the syntactic assumptions required for an NP-only analysis. For example, the modifier that is sister to N0 must be able to agree with N0.
9 One additional caveat I must add is that Estonian and Finnish have some pluralia tantum (and other lexical-ish plurals) which do not show hybrid agreement in number (Norris 2018a). Instead, the nominal phrases containing them are plural through and through. This is not an issue for the analyses discussed here, as these nominal phrases simply do not show hybrid agreement.
10 In Erelt’s (2017:203-206) discussion of predicate number agreement, the only case of mismatch between grammatical form and predicate number is NNCs, where a plural predicate is possible (but not obligatory) in some instances with an NNC subject bearing no plural morphology. Even predicates agreeing with collective nouns like rahvas ‘people/folks’ obligatorily track the noun’s grammatical number (Erelt 1999: 36-37).
In (10a), the demonstrative is plural while the noun is singular, but this is not possible without the numeral (10b). In (11a), the verb can be either singular or plural when agreeing with an NNC that bears no morphological manifestation of plurality (with some restrictions). However, it is the numeral that makes this possible. A singular noun with no numeral can only control singular verb agreement.

If the analysis of hybrid agreement in NNCs is based on the analysis in BCS, then we must encode hybrid number on the nouns themselves. But if that were true, we would have a poor understanding of the different empirical patterns in BCS and Estonian/Finnish.

4. Hybrid agreement is structural

The defining properties of hybrid agreement in Estonian/Finnish suggest a structural analysis. Hybrid agreement is tied to a specific structure—the numeral-noun construction—and thus it should be something about that syntactic construction that yields the agreement patterns. However, when it comes to the particular implementation, I do not know of any examples from Estonian/Finnish that provide evidence for one proposed structural analysis over another.

The analysis I suggest is thus familiar. In the vein of existing analyses of hybrid agreement (Landau 2016, Matushansky 2013, Pesetsky 2013), I propose that the “semantic” plural feature is introduced on a functional head that is connected to the numeral-noun construction. The literature on the syntax of cardinal numerals is a pick-a-mix of analytical confections, and space considerations preclude a deep comparison of approaches here; one possible implementation is represented below.

In this structure, there is a number feature on a low head, which I represent as Num0 in (12). This head introduces the singular number feature in NNCs, and I also take it to be the location of the sole number feature in nominal phrases without numerals. There is a number feature on a higher head, which I represent

---

For some discussion of the restrictions surrounding verb agreement with NNCs, see Norris (2021).
with the label $F^0$. This head introduces the plural number feature that is represented on higher modifiers. And between these two heads is the numeral itself.

I proposed something very similar in Norris 2018a. The key differences are the identity of $F^0$ and $\text{Num}^0$, which I previously identified as $\#^0$ (the “counting plural” from Mathieu (2014)) and $\text{Div}^0$ (from Borer (2005)). I am not using new labels to signify a change in stance here— I am intending the more neutral labels to indicate that I do not view certain details of (12) to be crucial to the account. However, I do think that having the numeral sandwiched between two different number features is a necessary analytical choice. That being said, since I first proposed such an analysis in my dissertation (Norris 2014), I have had in my mind that a more elegant analysis is possible. However, I still see challenges that I find too big to simply set aside in pursuit of an approach with only one number feature.

I treat the remainder of this section as a brief FAQ, discussing representations for number that differ from the proposals in (12). My hope is that this can sharpen some of the key analytical points for any researchers who wish to develop or refute the analysis I proposed above.

4.1. Why does the plural feature merge after the numeral?

In (12), the numeral merges before the plural feature. A reasonable alternative one might consider is that the numeral merges after the plural feature (as proposed by Landau (2016)) or in some way, simultaneously. The main argument against such an approach, in my view, is that the numeral is not plural in an NNC; it is demonstrably singular. I say “demonstrably singular” because, in fact, cardinal numerals in Estonian/Finnish have singular and plural forms. The plural forms are used when counting, e.g., pluralia tantum nouns like _kahe-d kangastelje-d_ two-PL loom-PL ‘two looms’ (Norris 2018a). If the cardinal numeral merged after or with the [pt] in a standard NNC, it is not clear to me why the numeral would be singular.

Landau (2016) proposes that the numeral is singular in virtue of being inherently singular, but this misses the fact that cardinal numerals always match the number value of the nouns they count: singular when singular and plural when plural. Norris (2018a) proposes that this is just another instance of the concord system in Estonian; numerals match the number of the nouns they count because they show concord like most other nominal modifiers. Under all generative theories of concord I am aware of, if the [pt] feature merged before the numeral, we would expect the numeral to bear that feature. This is the challenge I see for positing a lower merge position for the plural feature as Landau (2016) does.

4.2. Why not get rid of the singular feature below the numeral?

Another option to consider is getting rid of the lower singular feature. Perhaps the head above the numeral is the only head that introduces a number feature in Estonian/Finnish nominals. In this view, the singular elements come to express this feature as a kind of morphological default, likely because the higher plural feature is somehow inaccessible. One reason to be skeptical is that this would also require an unusual location for cardinal numerals. I am aware of no other generative analyses of inflectional plural languages where the cardinal numeral merges before plurality is introduced. As Norris (2018b: 529-530) observes, the [sg] constituent in Estonian/Finnish NNCs is not particularly small, either. It can contain adjectives (as we have seen) and possessors.

A bigger challenge, in my view, is that hybrid number agreement only occurs in numeral-noun constructions. In the effort to simplify the number system, we end with an analysis where the presence of a cardinal numeral does not affect the structure in any other way. However, it’s not clear that this is correct. If the noun is singular in numeral-noun constructions due to the plural feature being inaccessible—if it has nothing to do with the numeral or associated syntax—then it should also be inaccessible in an ordinary noun phrase. As we saw in (10), this is not the case. In the absence of a numeral, all elements showing concord in a nominal phrase must match the number value of the noun. The presence of a numeral results

---

12 In fact, this may be what I proposed in Norris 2018a. I treated number as a privative feature, such that absence of [pl] meant that the noun would express what we call singular. The high location for plural is more plausible in that analysis because it builds on Borer’s (2005) theory of number as (in part) a tool for nominal individuation. However, I think that Norris (2018a) faces the same challenges I describe in this section as regards mechanisms of agreement.
in an apparent disruption of the agreement system. In the analysis I propose, this is because the syntax of numeral-noun constructions involves the introduction of a second number feature.

5. Conclusion

Ultimately, even when multiple analyses can be made to work, there is one important question that should be kept at the forefront: what is the nature of the generalization? In Estonian/Finnish NNCs, the generalization regarding different number values has to do with syntactic structure, not lexemes. Hybrid agreement in these languages is connected to numeral-noun constructions— the individual roots are immaterial. Given that hybrid agreement in Estonian/Finnish NNCs must be structural whereas hybrid agreement in BCS (and other gender-based hybrid agreement in the literature) can be analyzed in either way, this calls into question whether all forms of hybrid agreement should be analyzed structurally. It also returns hybrid agreement to the debates concerning the utility and necessity of nominal phrases headed by functional morphemes.

References


Norris, Mark. 2021. Estonian and Finnish show that hybrid agreement is structural (Supplemental materials, version 1). Archived at OSF.IO. DOI:10.17605/osf.io/9TYGZ.


