Evidence for DP in Estonian*

Mark Norris, UC-Santa Cruz

LSA 88, January 2–5, 2014

1 Introduction

Estonian is a so-called “articleless language,” meaning bare nouns have a relatively free syntactic distribution and a variety of possible semantic interpretations.1

- They can be interpreted as indefinite (1a) or definite (1b).
- They can be predicates in copular clauses (2).
- They can have generic interpretations (3).

(1) a. Õue-s on kass ja koer.
    yard-INE be.3 cat and dog
    ‘There is a cat and a dog outside.’

b. Koer aja-b kassi taga.
    dog drive-3SG cat.PAR back
    ‘The dog is chasing the cat.’

(2) Nee-d naise-d on advokaadi-d.
    those-PL woman-PL be.3 lawyer-PL
    ‘Those women are lawyers.’

(3) Jääkaru / Jääkaru-d on havimisohus.
    polar.bear / polar.bear-PL be.3 endangered
    ‘The polar bear is endangered.’ / ‘Polar bears are endangered.’

There are two kinds of analyses one might propose for the kinds of nominals that Estonian has.

- null D⁰: Estonian has a null determiner, i.e., the difference between nominals in Estonian and nominals in a language like English is lexical/phonological (4).

    – the widely assumed (Universal) DP Hypothesis (Abney 1987; Szabolcsi 1983, 1994, among many others)

---

*For helpful discussion of this work, I thank Sandra Chung, Amy Rose Deal, Jorge Hankamer, Boris Harizanov, Peter Jenks, Jim McCloskey, and Bern Samko. Thanks as well to my Estonian consultants for teaching me about their language: Katrin Jänese, Mervi Kalms, Leelo Kask, Miina Norvik, and Kärt Vahtramäe. Any remaining errors lie with me. See p. 13 for gloss abbreviations.

1Another way to say this would be that the language has nothing like the more familiar (in)definite articles from Indo-European languages.
• no $D^0$: Estonian nominals have no DP layer, i.e., the difference between nominals in Estonian and nominals in a language like English is syntactic (5).

\[
\begin{align*}
(4) & & \text{DP} & & (5) & & \text{NP} \\
 & & \text{D} & & \text{NP} & & \text{jääkaru} \\
 & & \emptyset & & \text{jääkaru}
\end{align*}
\]

Recently, the adoption of the DP Hypothesis for articleless languages has come under increased scrutiny.

• In particular, Bošković (2005, 2008, 2009) and Despić (2013) have argued that Serbo-Croatian nominals lack a DP layer.

• They argue that this leads to a wide array of syntactic consequences.

➤ Further, they suggest that this may be a universal property of articleless languages.

In this talk, I will argue for a null-$D^0$ analysis of Estonian by way of two related claims:

1. Estonian does not show any of the properties of articleless languages outlined by the work on Serbo-Croatian (section 2).
   – Furthermore, the claim that Estonian lacks DP would require an unmotivated analysis of possessors in Estonian.

2. Though Estonian lacks clear correlates of Indo-European articles, there are elements which, I argue, occupy the $D^0$ position (section 3).
   – In other words, the set of elements that can occupy $D^0$ is not identical to the the set of elements we call ‘articles’.

I will discuss the bigger picture and conclude in section 4.

2 Prospects for a no-$D^0$ analysis of Estonian

The no-$D^0$ analysis of Serbo-Croatian proposed by Bošković (2005), *et seq.* and Despić (2013) comes with syntactic consequences, which the authors are quite explicit about.

• Their hypothesis that all articleless languages lack DP makes clear predictions about the behavior of nominals in articleless languages.

• It has led to investigations of nominals in various other articleless languages (Gillon and Armoskaite, 2012, To Appear; Pereltsvaig, 2007, 2013)

In this section, I explore those predictions in the domain of possessors.

• I will show that an NP analysis does not provide much leverage in understanding Estonian nominal structure.
2.1 Estonian nominals can have (at least) two genitive modifiers

Bošković (2008) reveals a number of properties that many articleless languages share; it is claimed that these differences come from an NP/DP distinction.

- A summary of the implications in Bošković 2008 is given in Table 1, though for time reasons, I will only discuss one: double genitives.\(^2\)

<table>
<thead>
<tr>
<th>CRITERION</th>
<th>ESTONIAN</th>
<th>CONCLUSION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left branch extraction → NP</td>
<td>no Left branch extraction</td>
<td>—</td>
</tr>
<tr>
<td>Adjunct extraction → NP</td>
<td>no Adjunct Extraction</td>
<td>—</td>
</tr>
<tr>
<td>Double genitives → DP</td>
<td>Yes</td>
<td>DP</td>
</tr>
<tr>
<td>Majority for ‘most’ → DP</td>
<td>Yes/unclear</td>
<td>DP/—</td>
</tr>
<tr>
<td>Superiority in multiple wh-fronting → DP</td>
<td>N/A (no MWF)</td>
<td>—</td>
</tr>
<tr>
<td>Negative Raising → DP</td>
<td>N/A (no NPIs that I know of)</td>
<td>—</td>
</tr>
<tr>
<td>Clitic Doubling → DP</td>
<td>N/A (no clitic doubling)</td>
<td>—</td>
</tr>
<tr>
<td>Locality in IHRCs → NP</td>
<td>N/A (no IHRCs)</td>
<td>—</td>
</tr>
<tr>
<td>Polysynthesis → NP</td>
<td>N/A (not polysynthetic)</td>
<td>—</td>
</tr>
</tbody>
</table>

Table 1: Languages with and without articles (Bošković, 2008)

As noted by Willim (2000), some languages with articles allow two nominal genitive arguments, but articleless languages disallow this (see also Bošković 2008:n7).

(6) a. Hannibals *Eroberung Roms*  
Hannibal.GEN conquest Rome.GEN  
‘Hannibal’s conquest of Rome’ (German)

b. *podbicie Rzymu Hannibala  
conquest Rome.GEN Hannibal.GEN  
Intended: Hannibal’s conquest of Rome’ (Polish)

This follows straightforwardly from the NP analysis on the assumption that the relationship between case-assigning heads and case-bearing nominals is one-to-one.

(7)  
\[\text{NP} \quad \text{NP} \quad \underline{\text{N'}} \quad \underline{\text{N'}} \quad \underline{\text{XP}}\]

- Bošković (2008) thus suggests that only languages with articles (i.e., with DP) permit two non-lexical genitives.

\(^2\)Most (all?) of these are phrased as one-way implications, as I have shown in the table using “→ NP/DP.” Thus, the fact that Estonian does not exhibit a particular property does not necessarily tell us anything about its nominals. Any analysis developed for these phenomena based on the NP/DP parameter must accordingly be structured such that the implications are preserved.
In Estonian bare nominals can have two non-lexical genitive modifiers (cf. Erelt 2009).³

(8) lapse pidev panni-de löö-mine
    child.GEN continuous pan-PL.GEN hit-NMLZ
    ‘the child’s constant banging of pans’

(9) emis-te päevane proteiini tarbi-mine
    sow-PL.GEN diurnal protein.GEN consume-NMLZ
    ‘the sows’ diurnal consumption of protein’

• This means that, with respect to this property, Estonian must be a DP language.

• Under a DP analysis of Estonian, we could say that the extra genitive is made possible by the presence of D⁰ (with one genitive presumably made available by N⁰).

(10) DP
    D
    → DP
        F
    → NP
        AP
    → N

2.2 Estonian possessors are not adjectival

A consequence of the NP analyses proposed by Bošković (2005); Corver (1992); Despić (2013) is that demonstratives and possessors are syntactically analyzed as being very close to adjectives.

• For Bošković (2005), following Corver (1992), they actually are A⁰ heads.

• Otherwise, they are assumed to be NP adjuncts (11) or multiple NP specifiers (12).

(11) [NP Dem [NP Poss [NP Adj [NP N ]]]]
(12) [NP Dem [N′ Poss [N′ Adj [N′ N ]]]]

Bošković argues that this is actually a strength of the proposal, as possessors are ‘adjectival’ in Serbo-Croatian (for at least the following reasons):

1. Possessors in Serbo-Croatian show concord just like adjectives.

2. Possessors in Serbo-Croatian are very small— they cannot be modified, even by adjectives.

3. Possessors in Serbo-Croatian can serve as predicates.

³Erelt (2009) says this is impossible, but he provides no ungrammatical examples, and one of the sources he cites (Koptjevskaja-Tamm 1993) provides a grammatical example with two genitives.
If Estonian lacked DP, then demonstratives and possessors would have to be analyzed in the same way as in Serbo-Croatian.

Possessors in Estonian do not display any of these properties.

- They are one of the only elements that does not show concord in case or number (with elements of the possessed nominal) (13).
- They can be modified— in fact, they are full nominal extended projections (14).
- Furthermore, they cannot be predicates (15).

(13) Kõigi-l Kärdi(*l) rikas-te-l sõpra-de-l on auto-d.
    all.PL-ADE Kärt.GEN(*ADE) rich-PL-ADE friend-PL-ADE be car-PL.NOM
    ‘All of Kärt’s rich friends have cars.’

(14) . . . sest ta kaitse-b laia-s mõtte-s [ kõigi töötaja-te ]
    because s/he protect-3SG broad-INE thought-INE all.PL.GEN worker-PL.GEN
    huve. interest.PL.PAR
    ‘. . . because s/he is advocating, in a broad sense, for the interests of all workers.’
    (PARLIAMENT)

(15) a. See auto on Kärdi *(oma).
    DEM.NOM car.NOM be Kärt.GEN own
    ‘That car is Kärt’s.’

b. See auto on minu *(oma).
    DEM.NOM car.NOM be 1SG.GEN own
    ‘That car is mine.’

I propose that possessors in Estonian are normal nominal extended projections.

- As sketched before, they are introduced by a functional projection above NP, though some nouns can also license a possessor in their own specifier.
- I suggest they cannot be predicates on their own, because their genitive case comes as a result of their being a part of a more articulated nominal spine.

2.3 Against an NP analysis of Estonian

It may be that an NP analysis is illuminating for the properties of Serbo-Croatian, but I contend that it is not so for Estonian.

- According to Bošković’s (2008) tests, Estonian does not behave like an NP language.
- Furthermore, an NP analysis would require an unmotivated analysis of possessors in Estonian.
3 Overt material in D⁰ in Estonian

The traditional category of determiners was a broad class, including demonstratives, quantifiers, cardinal numerals, the word *which*, articles, and pronouns (among others).

- Even before the work of Abney (1987); Szabolcsi (1983, 1994), it was noted that these elements did not all occupy the same position (e.g., Jackendoff 1977; Perlmutter 1970).

Since then, research on the internal syntax of nominals has resulted in an expansion of functional categories and a more nuanced view of “Determiners.”

1. **Quantifiers** are often argued to occupy a separate projection above DP: Q⁰ (e.g., Matthewson 2001).

2. **Demonstratives** are argued to occupy a specifier position: Spec,DP, something lower (Deal, 2010), or generated lower and then moved to Spec,DP (Alexiadou et al., 2007; Giusti, 1997).

3. **Cardinal numerals** occupy a lower position (e.g., Spec,NumP) (Longobardi, 2001; Watanabe, 2006; Deal, Under Review)

This leaves us with articles, pronouns, and wh-determiners like *which* as elements that may occupy D⁰.

- Though Estonian lacks articles, it does have elements which have been argued/assumed to occupy D⁰ in other languages.

3.1 Estonian indefinite pronouns

The **INDEFINITE PRONOUN CONSTRUCTION** is the name sometimes given to nominals like those in (16) and (17):

- The typical examples involve the indefinite pronouns⁴ *miski* ‘something’ or *keegi* ‘someone’ modified by an adjective.⁵

- As with their English translations, the **adjective must follow** the pronominal.

(16) a. midagi huvitava-t PAR something.PAR interesting-PAR ‘something interesting’
   b. * huvitava-t midagi interesting-PAR something.PAR

(17) a. keegi uus somebody new
   b. * uus keegi new somebody

Note that in Estonian (as in English), **adjectives typically precede** the nouns they modify.

---

⁴This is the English translation of the term given to them by Erlet et al. (2000).
⁵As with their English correlates, *miski* and *keegi* can appear in isolation as well.
3.1.1 Towards an analysis of the indefinite pronoun construction

The classic (and to my mind, the simplest) analysis of these facts is that the indefinite pronoun does not occupy the same (surface) position as a normal noun.

- Movement (from N⁰) to a higher functional projection: Abney (1987); Kishimoto (2000)

In either case, the surface position of the indefinite pronoun is higher than N⁰.

The fact that the indefinite pronouns can co-occur with an overt noun suggests that indefinite pronouns cannot be generated in N⁰ (contra (19b)).

Thus, I propose there is a higher functional position in Estonian, and it can be occupied by indefinite pronouns.

3.1.2 Internal structure of Estonian indefinite pronouns

The indefinite pronouns miski and keegi are morphologically related to the wh-pronouns mis and kes.

- They are formed by adding the suffix -gi (with orthographic variant -ki) to the wh-pronoun.⁶

---

⁶The one exception is the nominative form, where kes+gi becomes keegi, not *keski.
• This suffix still exists in standard Estonian as a focus marker, but the -gi of indefinite pronouns contributes _indeterminateness_.

A unified analysis: _kes_ and _mis_ are generated in Num^0_ or φ^0_ (Déchaine and Wiltschko, 2002); “fossilized” indefinite -gi is D^0_.

• _kes_ and _mis_ always move to D^0_.

• If -gi is present, the result is an indefinite pronoun. If -gi is absent, we get a normal wh-pronoun.

(22) a. indefinite pronoun:

\[
\begin{array}{c}
\text{DP} \\
\phi P \\
\phi \\
\text{mis} \\
\end{array}
\begin{array}{c}
D \\
< \phi > \\
\text{NP} \\
\text{N} \\
\end{array}
\]

b. wh-pronoun:

\[
\begin{array}{c}
\text{DP} \\
\phi P \\
\phi \\
\text{mis} \\
\end{array}
\begin{array}{c}
D \\
< \phi > \\
\text{NP} \\
\text{N} \\
\end{array}
\]

This analysis posits the existence of two functional projections above NP.

• Since we have evidence that _mis(ki)_ can co-occur with a noun, it seems these pronouns are not generated in N^0_.

• Syntactically separating -gi from the indefinite pronouns requires an additional position, but it allows for a unified analysis of wh-pronouns and indefinite pronouns.

This analysis extends straightforwardly to the wh-determiner _milline_, which corresponds to English _which_.

(23) a. **Millise-d** **panga-d** **hakka-vad laene** **and-ma?**

which-PL.NOM bank-PL.NOM start-3PL loan.PL.PAR give-MA

‘Which banks will be giving out loans?’ (PARLIAMENT)

b. **Millise-s** **seriaali-s** **ta mängi-b?**

which-INE series-INE he play-3SG

‘Which series is he in?’ (BALANCED)

As with indefinite pronouns, _milline_ precedes the head noun.

• In the spirit of Abney (1987); Postal (1966/1969), I propose that _milline_, like _mis_ and _kes_, is located in D^0_.

• It is when these words select NP complements that they appear to look like wh-determiners.

---

7See Nevis 1984 for a comparison on phonological grounds.

8See Progovac 1998 for a similar analysis of personal pronouns in Serbo-Croatian.

9Like English, Estonian also allows the word _mis_ ‘what’ as a wh-determiner, though interestingly, this use of _mis_ does not inflect. Erelt et al. (2000) suggest it may be a truncation of _missugune_ ‘what kind’, in which the _mis_ morpheme is invariant.
3.2 Estonian demonstratives are not in D⁰

Research that addresses Finnic nominal syntax typically assumes that it is not unreasonable to locate demonstratives in the D⁰ position (e.g., Brattico 2010).

- It has in fact been claimed that there are definite and indefinite articles “developing” in Estonian (Hiietam and Börjars, 2003; Pajusalu, 1997, 2000).

(24) Ööse-l oli tuul. (See) tuul oli vinge.
    night-ADE be.PST.3SG wind.NOM DEM wind.NOM be.PST.3SG cold
    ‘There was wind at night. The wind was piercing.’ (Nemvalts, 1996)

In (24), a demonstrative can optionally be used when referring back to a previously established referent.¹⁰

- But this use does not mean that demonstratives are necessarily D⁰ heads.
- There is reason to think that demonstratives behave like phrasal elements in Estonian.

3.2.1 DP-internal phrasal movement in Estonian

The normal order of elements in Estonian is Q > Dem > Poss > Adj > N.

(25) kõik nee-d Kärdi punase-d auto-d
    all DEM-PL Kärt.GEN red-PL car-PL
    ‘all these red cars of Kärt’s’

However, in certain circumstances, demonstratives or possessors can come before strong quantifiers.

(26) Demonstratives before Q⁰:
    a. Tea-des nee-id kõiki keerukus-i,
        know-DES DEM-PL.NOM.PL complication-PL.NOM
        ‘Knowing all these complications, . . . ’ (PARLIAMENT)
    b. nee-d kõik ettevõtte-d
        DEM-PL.NOM all.NOM company-PL.NOM
        ‘all those companies’ (BALANCED)

(27) Possessors before Q⁰:
    a. Kärdi kõik poja-d käi-vad kooli-s.
        Kärt.GEN all.NOM son-PL.NOM go-3PL school-INE
        ‘All of Kärt’s sons go to school.’ (LK, Volunteered)
    b. selline akt, mille iga paragrahvi kohta on palju eriarvamusi.
        the.kind act whose.GEN each.GEN paragraph.GEN about be many dissent.PL.NOM
        ‘the kind of act for which there are many differing opinions about every paragraph.’
        (PARLIAMENT)

¹⁰Cf. Dayal (2004); Löbner (1985), who argue that demonstratives “used as definite articles” in articleless languages are still not exactly (English-like) definite articles.
Furthermore, for wh-possessors, this raising is obligatory.

(28) a. * Kõik kelle poja-d käi-vad kooli-s?
   all.NOM who.GEN son-PL.NOM go-3PL school-INE
   Intended: ‘All of whose children go to school?’

   b. Kelle kõik poja-d käi-vad kooli-s?
      who.GEN all.NOM son-PL.NOM go-3PL school-INE

   (29) a. * Kõik mille jala-d oli-d sinise-ks värvi-tud?
      all.NOM what.GEN leg-PL.NOM be.PST-3PL blue-TRL paint-PASS.PST.PCPL
      Intended: ‘All of what thing’s legs were painted blue?’
   
     b. Mille kõik jala-d oli-d sinise-ks värvi-tud?
        what.GEN all.NOM leg-PL.NOM be.PST-3PL blue-TRL paint-PASS.PST.PCPL

It seems to me that the simplest analysis is that the fronting of demonstratives and possessors is the same (see Alexiadou et al. 2007; Horrocks and Stavrou 1987).

• The fact that this is the position of wh-pronouns suggests to me that the position to the left of kõik and iga is an A′ position (i.e., the ‘edge’ of DP).
  
  – For non-wh-elements, I assume this is a movement is focus-based (see Horrocks and Stavrou 1987).

• Either kõik and iga are D0 heads and this is Spec,DP, or they are Q0 heads (à la Matthewson 2001) and this is Spec,QP.

(30) D/QP

An analysis that treated article-like demonstratives as exponents of D0 would have no clear path to treating these processes as the same.

4 Conclusions

I have argued here that Estonian, a language without articles, still has DP.

• Estonian nominals allow two non-lexical genitive arguments (one licensed by N0, one licensed by D0).
  
  – This is a DP behavior according to Bošković’s (2008) diagnostics.
• Furthermore, the syntactic assumptions Bošković (2005), et seq. must make for “possessors” do not extend so easily to Estonian.

• There is an edge position in Estonian DPs that is a target for DP-internal wh-movement as well the fronting of demonstratives.

• Estonian indefinite pronouns and wh-determiners— both a part of the larger class of pronouns— appear to occupy a position distinct from N0. I proposed this position was D0.

The null-D0 analysis proposed here shares something in common with the no-D0 research program outlined by Bošković (2005, 2008) and developed in subsequent work.

• Narrow(er) question: What are the properties of the D0 head?

• Broad(er) question: How much structure is present in the nominal extended projection?

Focusing on the narrow question, the no-D0 hypothesis suggests at least the following about the nature of D0:

• D0 is the position of (only?) articles.11

• As outlined by Bošković (2008), et seq., D0 has a number of universal syntactic properties (e.g., it makes Left Branch Extraction impossible; it licenses a second possessor)

In contrast, the null-D0 hypothesis that I adopt says the following:

• D0 is the position of a number of different descriptive categories (including, but not limited to, articles).

• Different D0 heads may have different syntactic properties (e.g., some D0 heads may license another possessor, but others may not)

If the conclusions I have reached here are reasonable, the hypothesis Bošković (2008) calls the stronger hypothesis — that languages without articles uniformly lack DP— needs to be revised.

• Estonian does not exhibit any NP language behavior; it exhibits some DP behavior.

• A version of the weaker hypothesis— that some languages without articles lack DP— may still obtain.

Another way forward is to assume that variation is not located in the presence or absence of functional material, but in the properties of that functional material.

11It seems worth pointing out that it is not universally accepted that so-called indefinite or definite articles actually occupy D0. For example, Perlmutter (1970) proposes that the English indefinite article is actually lower than D0, and the definite article has been argued to occupy Spec,Dp (Lyons, 1999). Perhaps this can be disentangled with a formal definition of an ‘article’ — I will not speculate on this issue here.
• In such a view, the differing properties of null-article languages with respect to, e.g., Bošković’s diagnostics, would come down to the particular properties of the language’s D₀ heads.

• In subsequent work, it seems that at least Bošković (2009); Despić (2013) do not endorse the view that there is never functional material above NP in Serbo-Croatian.
  – Perhaps the claim could be that languages without articles have less functional structure than languages with articles.
  – The evidence suggests Estonian has a “normal” amount of functional structure.

A fine-grained syntactic architecture is the most productive if we can latch broader semantic and syntactic generalizations to it.

• Balancing between too much or not enough structure will continue to be a difficult task.
  ➤ This work at its core can be viewed as an argument for the simplest version of Abney’s (1987) proposal: at least one functional head is universally available above NP.

**Abbreviations**

Naturally-occurring examples come from the following sources:

• BALANCED: A balanced literary corpus containing equal parts journalism, modern fiction, and scientific articles (15 million words) available here (in Estonian):
  http://www.keeleveeb.ee/dict/corpus/balanced/

• EKSS: *Eesti keele seletav sõnaraamat* ‘Estonian descriptive dictionary’, available here (in Estonian):
  http://www.eki.ee/dict/ekss/index.cgi

• PARLIAMENT: a corpus of stenographic records of parliamentary discussions (13 million words) available here (in Estonian):
  http://www.keeleveeb.ee/dict/corpus/riigikogu/

Gloss abbreviations: 3 third person, DEM demonstrative, DES *des*-gerund, GEN genitive, INE inessive, MA *ma*-infinitive, NMLZ nominalization, NOM nominative, PAR partitive, PASS “passive”, PL plural, PST past, PST.PCPL past participle, TRL translative, SG singular

**References**


Deal, Amy Rose. Under Review. Apportionment and the mass-count distinction in Nez Perce. ms, UCSC.

Mark Norris
University of California, Santa Cruz
1156 High Street
Department of Linguistics
Santa Cruz, CA 95060
mnorris@ucsc.edu
http://people.ucsc.edu/~mnorris/