**Syncretism and markedness paradoxes in the Russian demonstrative declension**

Pavel Caha & Lucie Taraldsen Medová (Masaryk University)

## 1. Introduction

This chapter investigates the singular declension of Russian demonstratives and nouns. A complete overview of the paradigms is provided in Table 1 (Timberlake 2004). The table lists four major nominal paradigms (classes $I_A$, $I_B$, II and III), along with the demonstratives.[1]

| | **this** | **wife (II)** | **notebook (III)** | **this** | **place ($I_B$)** | **this** | **factory ($I_A$)** |
|---|---|---|---|---|---|---|---|
| | **FEMININE SINGULAR** | | | **NEUTER SINGULAR** | | **MASCULINE SINGULAR** | |
| **NOM** | et-**a** | žen-**a** | tetrad' | et-**o** | mest-**o** | etot-**ø** | zavod-**ø** |
| **ACC** | et-**u** | žen-**u** | tetrad' | et-**o** | mest-**o** | etot-**ø** | zavod-**ø** |
| **GEN** | et-oj | žen-y | tetrad-i | et-**ogo** | mest-**a** | et-**ogo** | zavod-**a** |
| **LOC** | et-oj | žen-e | tetrad-i | et-**om** | mest-**e** | et-**om** | zavod-**e** |
| **DAT** | et-oj | žen-e | tetrad-i | et-**omu** | mest-**u** | et-**omu** | zavod-**u** |
| **INS** | et-**oj(u)** | žen-**oj(u)** | tetrad-ju | et-**im** | mest-**om** | et-**im** | zavod-**om** |

TABLE 1: Russian demonstrative and nominal paradigms

The table has three vertical parts, one for each gender (feminine, neuter and masculine).[2] Declensions II and III are almost exclusively occupied by feminines, Declension $I_B$ is neuter, and $I_A$ only contains masculines. In the table, different types of shading indicate regions where the inflection is the same for different paradigms.

These paradigms pose an interesting set of challenges for standard theories of paradigm

[1] The instrumental of the feminine demonstrative and the noun 'wife' is *-oj*, with an archaic form *-oju*. We indicate this by placing the *u* in parenthesis.
[2] Russian has also special forms for masculine animates which we do not discuss here for reasons of space.

structure. To see this, consider the two paradigms in Table 2. The table shows that the neuter demonstrative ‘that’ is different from the masculine in the nominative and accusative, while they are the same in the oblique cases (shaded). The lack of a gender distinction in the oblique cases can be called syncretism or neutralisation of an opposition between the masculine and neuter forms of the demonstrative.

| | **that, NEUT** | **that, MASC** |
|---|---|---|
| **NOM** | et-o | etot |
| **ACC** | et-o | etot |
| **GEN** | et-**ogo** | et-**ogo** |
| **LOC** | et-**om** | et-**om** |
| **DAT** | et-**omu** | et-**omu** |
| **INS** | et-**im** | et-**im** |

TABLE 2: The masculine/neuter syncretism

Table 2 illustrates a common property of paradigms, which is that differentiation is found in “unmarked” environments (NOM/ACC), while neutralisation is found in “marked” environments (obliques). Croft's (2003:97) typology textbook expresses this as follows: “If the marked value has a certain number of formal distinctions in an inflectional paradigm, then the unmarked value will have at least as many formal distinctions in the same paradigm.”

The tendency for neutralisation to happen in marked environmentrs has also recieved attention in Distributed Morphology (DM, Halle and Marantz 1993). In DM, a systematic neutralisation of a morphosyntactic distinction across a range of environments (e.g., in all oblique cases) is accounted for by Impoverishment Rules (see Bonet 1991, Bobaljik 2002, Harley 2008, Arregi and Nevins 2012).

To see how Impoverishment works, suppose, for instance, that masculine is marked compared to the neuter, assuming a feature decomposition like the one in (1a,b).[3]

(1) a. neuter = [CLASS]

b. masculine = [CLASS, MASC]

c. *Impoverishment rule*: MASC → Ø \ _OBL

d. [CLASS, MASC, OBL] → [CLASS, OBL]

In (1c), we postulate an Impoverishment rule deleting MASC in OBL cases. This rule makes masculine obliques indistinguishable from neuter obliques, see (1d). As a result, insertion inevitably choses the same exponent for the two genders (whichever exponent that happens to be, depending on case).

The effects of the Impoverishment Rule (1c) can be also observed in the declension of nouns, which are subject to the same syncretism, see the striped shading in Table 1.

Given the tendency to neutralise inflectional distinctions in marked environments, it comes as a surprise that when we compare demonstratives and nouns, we observe the exact opposite pattern: they are syncretic in the unmarked, structural cases, but differentiated throughout the marked, oblique cases (with the exception of the instrumental singular of the feminine gender). This is illustrated in Table 3, where, again, the identical forms are shaded. The neutralisation is systematic in NOM/ACC across all genders, regardless of the specific marker (*-a*, *-u*, *-o*, *-Ø*). For DM, such systematicity is a hallmark of Impoverishment.

[3] The idea that masculine is marked with respect to the neuter is supported by impersonal sentences, where (in the absence of a referential subject) the verb bears the default neuter form.

| | this | wife (II) | this | place ($I_B$) | this | factory ($I_A$) |
|---|---|---|---|---|---|---|
| | **FEMININE SINGULAR** | | **NEUTER SINGULAR** | | **MASCULINE SINGULAR** | |
| **NOM** | et-a | žen-a | et-**o** | mest-**o** | etot-**ø** | zavod-**ø** |
| **ACC** | et-u | žen-u | et-**o** | mest-**o** | etot-**ø** | zavod-**ø** |
| **GEN** | et-oj | žen-y | et-ogo | mest-a | et-ogo | zavod-a |
| **LOC** | et-oj | žen-e | et-om | mest-e | et-om | zavod-e |
| **DAT** | et-oj | žen-e | et-omu | mest-u | et-omu | zavod-u |
| **INS** | et-oj(u) | žen-oj(u) | et-im | mest-om | et-im | zavod-om |

TABLE 3: Russian demonstrative and nominal paradigms

However, if Impoverishment only applies in marked cases, the facts lead to a paradox. Namely, from the perspective of gender neutralisation, oblique cases appear to be marked (recall Table 2). However, Table 3 suggests the opposite: from the perspective of the demonstrative/noun distinction, the structural cases are marked. Since both statements cannot be true at the same time, a contradiction arises.

This paper shows that the complex and apparently unusual patterns of morpheme distributions as attested in Russian can be insightfully modelled using the cross-linguistically invariant lexicalisation procedure proposed within Nanosyntax. As is customary in the framework, we refrain from using any dedicated post-syntactic operation for modelling syncretism. The patterns of marker distribution emerge as a function of the lexical entries interacting with the language-invariant lexicalisation procedure.

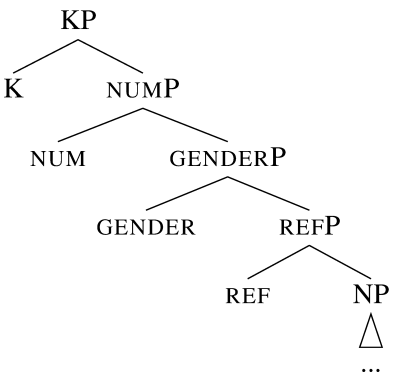
The chapter is organised as follows. In Section 2, we introduce the relevant morphosyntactic features and their hierarchy. In Section 3, we discuss masculine and neuter demonstrative paradigms, showing how backtracking leads to neutralisation of inflectional differences in the marked cases. In Section 4, we focus on the feminine declensions, and show

how the subextracting algorithm of Starke (2022) allows us to capture the fact that demonstratives inflect like nouns in NOM/ACC, but not in the oblique cases. Section 5 extends this account to neuter and masculine nouns. Section 6 concludes.

## 2 The morphosyntactic features

This section introduces the morphosyntactic features relevant for our analysis, i.e., features related to gender, number and case. In the literature, it is commonly assumed that gender is located lower than number (NUM), which is in turn located lower than case (K), as in (2); see Lamontagne and Travis (1987), Picallo (1991).

(2) [tree 2]

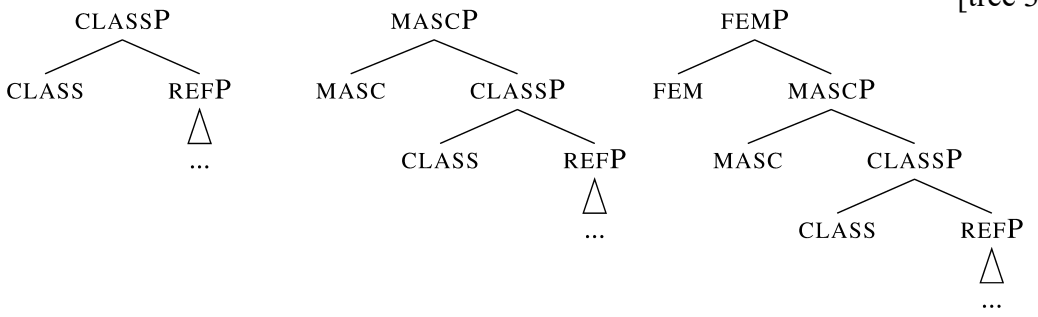



At the bottom of the tree, we include the feature REF, which is a feature that introduces the referential index, common to all refering expressions (see Harley and Ritter 2002).

With the basic hierarchy of projections established, let us look more closely at the individual categories. As is customary in Nanosyntax (Starke 2009), we decompose the values of each of the categories (gender, number and case) into privative features, where each feature is a head of its own (see also Cinque and Rizzi 2010).

Starting with gender, we adopt the view that neuter is the unmarked value, characterised by a single gender feature that we call CLASS, see (3a).

(3) a. NEUTER b. MASCULINE c. FEMININE [tree 3]




Support for the idea comes from the fact that in sentences with expletive (non-referential) subjects, the verb takes the neuter gender ending as a default, see (4).

(4) Gremel-o

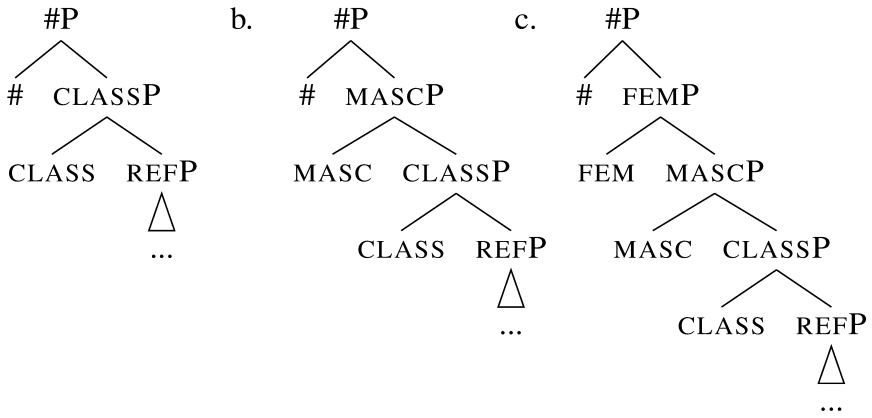
thunder-NEUT.SG

'It thundered.'

Moving on to the masculine, we propose that it is characterised by the presence of an extra feature MASC compared to the neuter, see (3b). We analyse the feminine gender as the most marked one, with the feature FEM added on top of the masculine as in (3c). The marked status of the feminine with respect to the mascusline is supported by the traditional observation (going back at least to Jakobson's 1932 work) that in most cases, nouns of masculine gender actually carry no presuppositions about the actual gender of the referent, while the feminine gender presupposes a feminine referent.

Turning now to number, we shall only discuss the singular forms, and we therefore only work with a single number head #. The singular head may come on top of various genders: (5a-c) shows, respectively, the neuter, masculine and feminine singular.

(5) a. b. c. [tree 5]




Finally, moving on to case, we use the so-called cumulative case decomposition proposed in Caha (2009). The cumulative decomposition has been motivated by the fact that case syncretism in Russian is restricted by a linear-adjacency constraint (Chvany 1982). The constraint says that in the sequence NOM-ACC-GEN-LOC-DAT-INS, only adjacent cases can be syncretic.

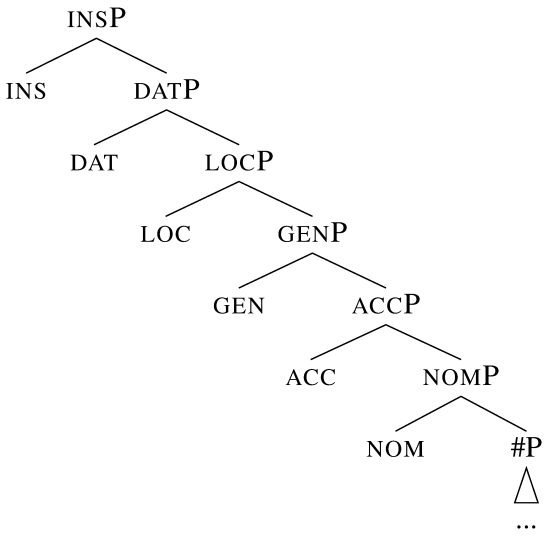
We illustrate this in Table 4. The ordering of cases observed by Chvany is represented in the top-down ordering of the cases in the table. The "staircase-like" shading highlights the syncretisms of, respectively, NOM-ACC, ACC-GEN, GEN-LOC, LOC-DAT, and DAT-INS. No other order than the one proposed by Chvany can accommodate all these syncretisms in a way that the syncretic forms are adjacent. No syncretism violates this order.

| | **place** | **horse** | **beautiful pl** | **this** | **hundred** |
|---|---|---|---|---|---|
| **NOM** | mest-o | kon' | krasn-ye | žen-a | st-o |
| **ACC** | mest-o | kon-ja | krasn-ye | žen-u | st-o |
| **GEN** | mest-**a** | kon-ja | krasn-yx | žen-y | st-a |
| **LOC** | mest-**e** | kon-e | krasn-yx | žen-e | st-a |
| **DAT** | mest-**u** | kon-ju | krasn-ym | žen-e | st-a |
| **INS** | mest-**om** | kon'-om | krasn-ymi | žen-oj(u) | st-a |

TABLE 4: Syncretism in Russian

To capture this, Caha (2009) argued that we need to adopt the so-called cumulative decomposition, where the number of features characterising individual cases monotonically grows as we go down in the paradigm from top to bottom. The decomposition is shown in (6).

(6) [tree 4]

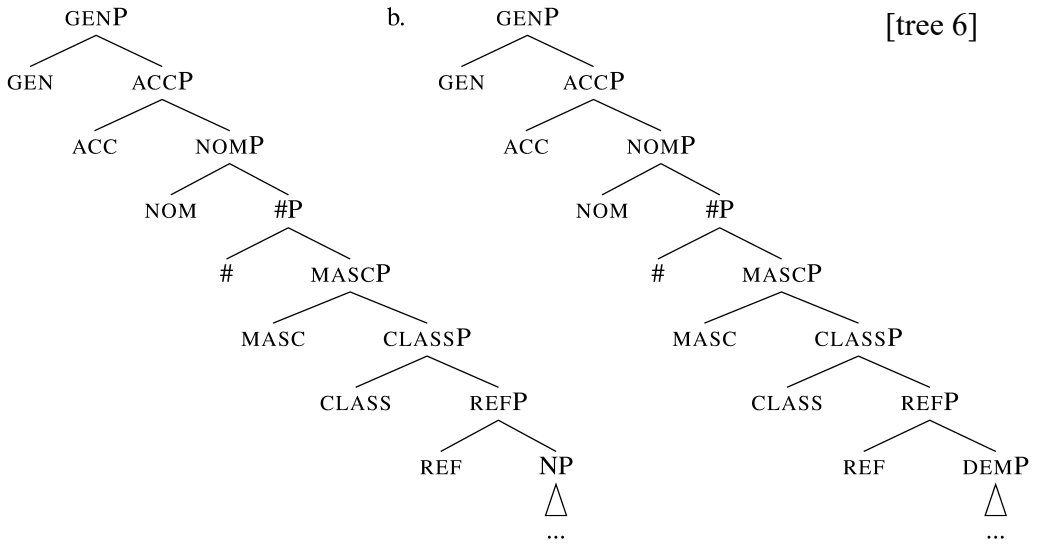



The tree conveys the proposal that the nominative singular corresponds to a #P with a single case feature on top of it, see the NOMP in (6). The accusative singular corresponds to the ACCP, i.e., to a #P with two case features (NOM and ACC). The genitive corresponds to GENP and so on.

Summarising, we assume that gender features reside low in the functional sequence, right above the referential feature REF. Above gender, we propose the singular feature #, and then a differing number of case features on top, depending on the case. Before we turn to the issue how these features are realised morphologically, two more assumptions must fall in place.

The first assumption is that nouns and demonstratives have the same structure of phi features (see Caha 2023 and references there). In this approach, nouns and demonstratives differ only in the type of projection found at the bottom of the tree, everything else is the same. For example, the genitive singular case of a masculine noun would have the structure as in (7a). The genitive singular of a masculine demonstrative would have the structure as in (7b), where the only difference is that (7a) has an NP at the bottom, while (7b) has a DEMP.

(7) a. b. [tree 6]




We assume that the features are base-generated on the modifier and they introduce presuppositions on the nature of the modified noun (Spathas and Sudo 2020). Since these presuppositions must be maximised, the features on the modifier end up being the same as those on the modified noun.

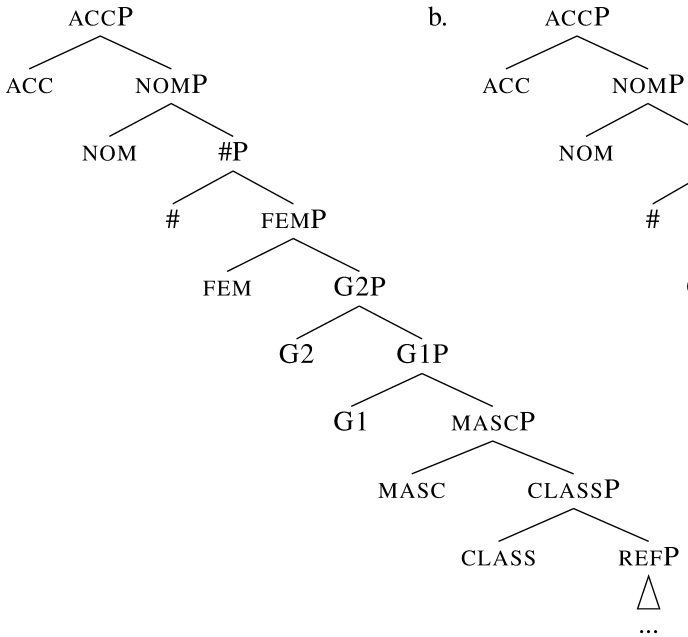
One reason for assuming this parallel is that phi features can be realised the same on nouns as on demonstratives (a point made in Matushansky and Halle 2006:382). We illustrate this by placing side by side the declension of a feminine surname ‘Panova’ and the feminine singular demonstrative. The identity of marking is neatly captured under the proposal in (7): same features give the same endings.

| | **Panova, FEM.SG** | **that, FEM.SG** |
|---|---|---|
| **NOM** | Panov-a | et-a |
| **ACC** | Panov-u | et-u |
| **GEN** | Panov-oj | et-oj |
| **LOC** | Panov-oj | et-oj |
| **DAT** | Panov-oj | et-oj |
| **INS** | Panov-oj(u) | et-oj(u) |

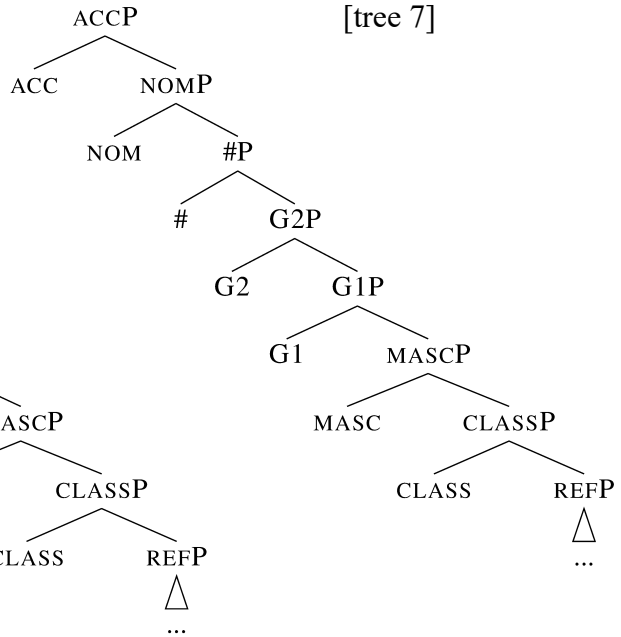
TABLE 5: Nouns can have the same endings as demonstratives

The final assumption arises due to certain technical aspects of the analysis to be presented. Specifically, we need to introduce two additional features, G1 and G2, possessed by all nouns, which are located in between the MASC and FEM features, see (8a).

(8) a.




b.




[tree 7]

(8a) depicts a projection containing all the gender features, representing the accusative of a feminine singular noun/demonstrative. (8b) shows the accusative singular of a masculine noun/demonstrative, which lacks FEM, but has G1 and G2. Neuter nouns (not shown) would

lack MASC in addition to lacking FEM, but they too would have G1 and G2.

The names of the features (G1 and G2) reflect their position inside the gender sequence of the nouns (i.e., G1/2 stands for GENDER 1/2). The existence of such features is not problematic in the sense that their presence does not lead to any contradictions, but we admit that so far, these features lack an independent motivation. We leave this issue for future research.

## 3. Masculine and neuter demonstratives

This section provides our account of the masculine and neuter demonstratives. We repeat the relevant facts in Table 6. Compared to Table 2, we split the oblique endings into a "stem marker" *-o* and an additional suffix. We do not segment the instrumental, and analyse *-im* as a portmanteau.[4]

| | **that, neut** | **that, masc** |
|---|---|---|
| **NOM** | et-o | etot-ø |
| **ACC** | et-o | etot-ø |
| **GEN** | et-**o-go** | et-**o-go** |
| **LOC** | et-**o-m** | et-**o-m** |
| **DAT** | et-**o-mu** | et-**o-mu** |
| **INS** | et-**im** | et-**im** |

TABLE 6

Following Halle and Matushansky (2006), we assume that the demonstrative has the underlying form *ett-*, realised as *et-* before vowels (due to degemination) and as *etot* (with an epenthetic vowel) word finally. Vowel~zero alternations of this kind are independently attested in Russian;

[4] We could also assume that the instrumental is *-o-im*, with *-o* deleted due to a V-V sequence; this would be compatible with our analysis.

for instance, the noun *tur**o**k* ‘Turk’ has the vowel *o* between *r* and *k* word finally, see (9a). However, the vowel is absent when the root is followed by another vowel, see (9b). Following Halle and Matushansky, we assume the same alternation for the paradigm in Table 6.[5]

(9) a. tur**o**k (= et**o**t) b. turk-a (= ett-o → et-o)
Turk-NOM Turk-GEN

Our account starts from the fact that the masculine demonstrative has no ending in the nominative or accusative. This suggests that the demonstrative is capable of lexicalising all the features of these forms, as depicted in the top part of the lexicalisation table in (10).

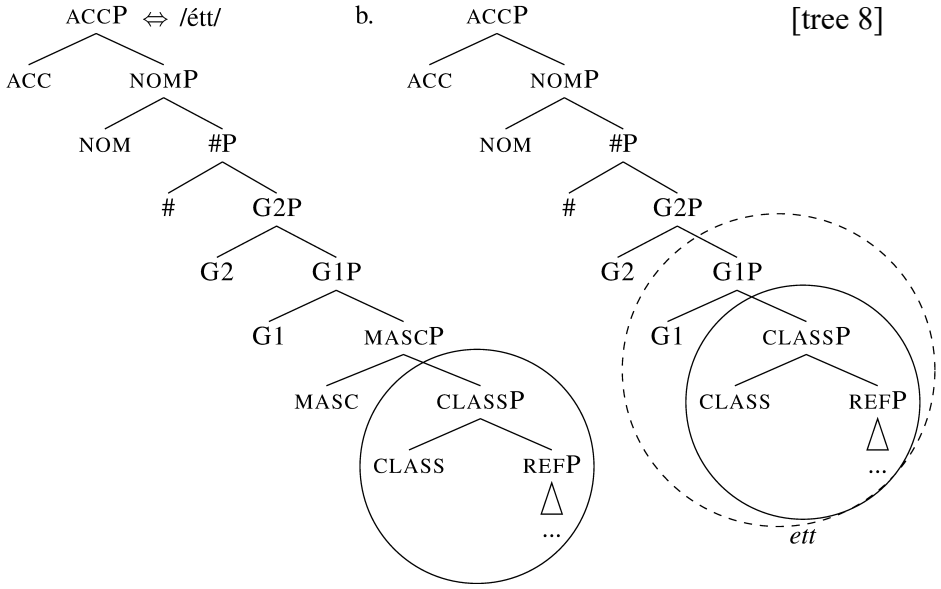
(10)

| | DemP | REF | CLASS | MASC | G1 | G2 | # | NOM | ACC |
|---|---|---|---|---|---|---|---|---|---|
| NOM | etot | | | | | | | | |
| ACC | etot | | | | | | | | |
| | DemP | REF | CLASS | MASC | G1 | G2 | # | NOM | ACC |
| NOM | | ett | | | | o | | | |
| ACC | | ett | | | | | o | | |

The simplest way to accomplish such a lexicalisation is through the lexical entry in (11a).

[5] Unlike Halle and Matushansky (2006), we do not assume an account in terms of yers, but in terms of empty nuclei (Scheer 2004).

(11) a. ACCP ⇔ /ett/ b. [tree 8]




Such an entry allows *ett-* to lexicalise all the features of the masculine ACC because the lexical entry is identical to its morphosyntactic structure. The lexicalisation of the nominative is possible due to the Superset Principle, see (12). This principle allows matching whenever the lexical entry contains a constituent identical to the morphosyntactic structure. As can be seen in (11a), the entry does indeed contain the masculine nominative.

(12) Superset Principle (Starke 2009): A lexically stored tree matches a syntactic node iff the lexically stored tree contains the syntactic node.

Let us now turn to the analysis of the neuter, as shown on the bottom of Table (10). Recall first that neuters lack MASC, which is indicated in Table (10) by the black coloring under MASC. The absence of a feature is also referred to as a gap.

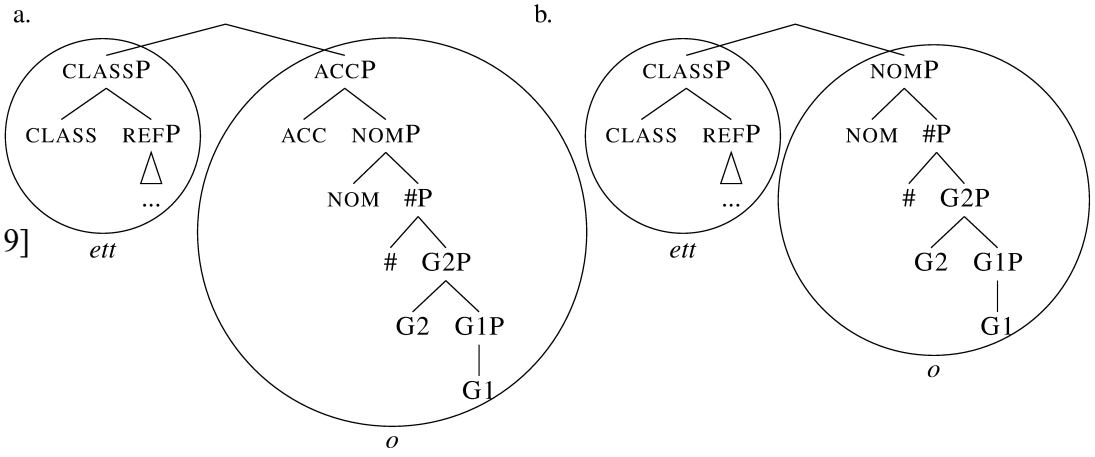
Because of the gap, *ett-* cannot lexicalise all the features of the NOM/ACC neuter. To see why, compare the entry of *ett-* in (11a) and the structure of the neuter in (11b). Because the structure in (11b) lacks MASC, the lexical entry only matches the CLASSP in (11b). The identity between the CLASSP and the entry is highlighted by the full circle. However, the lexical entry

does not match G1P (marked by the dashed circle), because it does not contain a constituent identical to it (the entry has MASC inside G1P).

What happens to the features that the root cannot lexicalise, i.e., G1 and higher? The idea depicted in Table (10) is that they are lexicalised by *-o*. For this to happen, the root of the demonstrative must move to the top of the structure (11b), see (13). This has the effect that all the remaining features form a constituent, This constituent is lexicalised by the ending. The entry of *-o* corresponds to the accusative tree circled in (13a). The nominative tree circled in (13b) correspond to a subconstituent, which means that the Superset Principle is satisfied.

(13) a. b.

[tree 9]




Let us now comment on how the movement of the root is triggered. In Nanosyntax, this happens due to the lexicalisation algorithm (Starke 2018). The general logic of the Nanosyntax lexicalisation algorithm is that whenever a new feature F is merged into the derivation, the FP produced must be externalisable, i.e., there must be a lexical entry in the lexicon that matches it. If this is not the case, the structure is rejected at the interface, and syntax must perform evacuation movements. The evacuation movements keep changing the phrase marker until lexicalisation suceeds (else the structure is "ineffable").

Nanosyntax assumes that the syntax and the lexicon are separate systems. Therfore, syntax performs the movements blindly, i.e., without knowing whether a particular movement

makes the phrase marker externalisable or not. However, the movements are not random either; they happen in a particular order, following a syntax-internal logic that is valid across all derivations and specified in the lexicalisation algorithm.

In Starke (2018) and much related work, the lexicalisation algorithm first tries to evacuate the specifier of the complement, followed by the movement of the whole complement. Here we explore an updated version of this algorithm that also includes subextractions from within specifiers (Starke 2022, Wiland 2019). To distinguish the two proposals, we refer to the new version as the subextracting algorithm.

At a general level, the subextracting algorithm preserves all the previous results because it allows all the movement options of the previous algorithm. It also preserves the ordering of the operations (Spec movement before complement movement). However, it opens additional options (subextraction from within specifiers), which seem to be required for our analysis.

We give a formalisation of the subextracting algorithm in (14). The sequence of steps that this algorithm triggers is the same as proposed in Starke (2022). However, the wording is ours.

(14) a. Merge F and lexicalise FP

b. If fail, evacuate the closest labelled non-remnant constituent and lexicalise FP

c. If fail, evacuate the immediately dominating node and lexicalise FP (recursive)

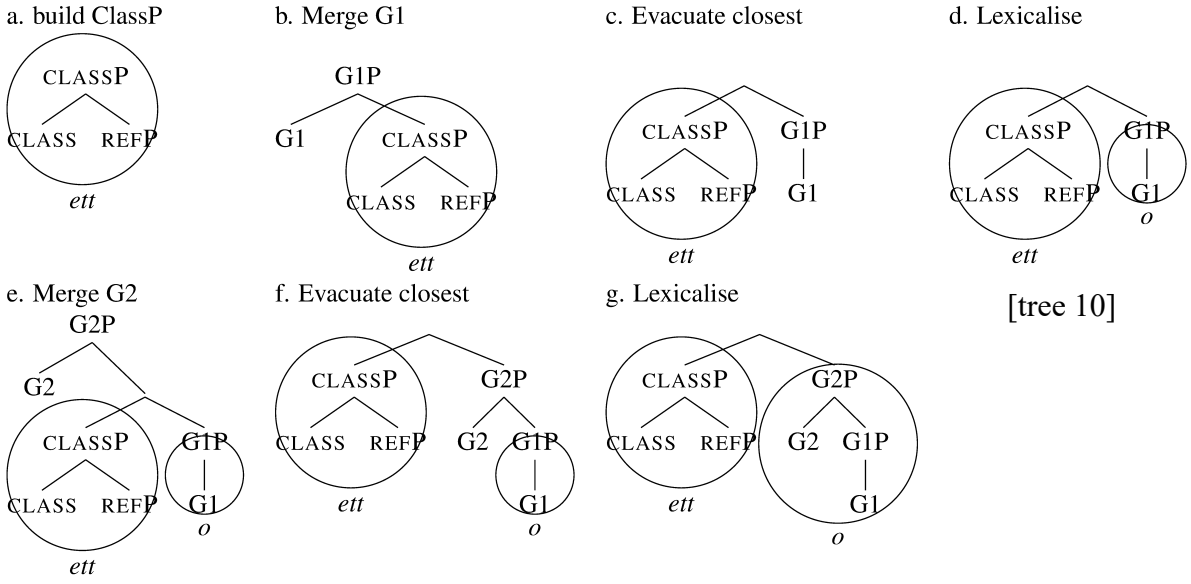
d. If fail, go to the previous cycle and try the next option for that cycle

The phrasing in (14) is based on the idea that if a phrase marker can be lexicalised, syntax prefers not to move anything (14a); but when the phrase marker cannot be lexicalised, syntax performs evacuation movements (14b,c). Syntax first tries to move the closest available object (14b); cf. Move Closest (Chomsky 1995). The definition of an available object in (14b) encodes

some restrictions, e.g., movement of remnant constituents is not allowed (see Cinque 2005). When movement does not suceed, pied-piping larger and larger phrases is attempted (14c).

Let us illustrate the working of the algorithm on the derivation of the NOM/ACC forms of the neuter demonstrative. First, the DEMP is constructed. Following (14a), it must be lexicalised. The DEMP is contained inside the lexical entry of the demonstrative *ett-*, hence, DEMP is lexicalised as *ett-*. After this, REF is merged, and REFP lexicalised as *ett-* again. When CLASS is merged, CLASSP is again lexicalised by *ett-*, see (15a). However, when G1 is merged, as in (15b), the resulting G1P cannot be lexicalised, recall the discussion surrounding (12b).

(15)




When lexicalisation of G1P fails in (15b), the step (14b) of the lexicalisation algorithm (14) is activated. (14b) says that the closest labelled non-remnant constituent must be evacuated from within G1P, where 'closest' is calculated as the distance from the root node G1P. In (15b), the relevant "closest labelled non-remnant constituent" is CLASSP, which is therefore evacuated out of G1P, yielding (15c). In (15c), as customary in Nanosyntax, the trace in the original position of CLASSP is eliminated, and the new root node has no label (spellout movements do not label the phrase marker). Note that the distinction between labelled and unlabelled nodes then feeds

into the definitions in (14).[6]

After the movement, the G1P can be lexicalised by *-o*, see (15d). Once matching succeeds, a new feature (G2) is introduced in the derivation, see (15e). G2P has no match, so the step (14b) is activated. Looking at (15e), the closest labelled non-remnant constituent is not the complement, since the complement has no label. The relevant constituent is not G1P either, since G1P is a remnant constituent (containing a unary-branching node). Therefore, the "closest non-remnant labelled constituent" in (15e) is CLASSP. CLASSP is therefore evacuated, yielding (15f), where there is again no trace of CLASSP inside G2P, and the root node again lacks a label. G2P is now matched by *-o* (see (15g)).

The derivation beyond G2P then continues in similar steps. After G2P in (15g) is lexicalised, we merge # on top of it. Lexicalisation fails, so we evacuate CLASSP out of #P as the closest non-remnant labelled constituent, and spell out #P. Then we add the nominative feature, move CLASSP again and obtain the structure (13b), which is also lexicalised as *ett-o*. The accusative is then realised as in (13a).

Summarising, neuter demonstratives differ from masculine ones in NOM/ACC (*et-o* vs. *etot*). This is explained by proposing the entry (11a): in the masculine structure, the entry lexicalises all the features, and there is no ending. In the neuter, due to the gap, the entry cannot lexicalise number and case, and an ending is needed. Due to the lexicalisation algorithm, the demonstrative cyclically moves to the top of the tree and the correct constituency is achieved.

Let us now turn to the genitive, where MASC and NEUT paradigms unify. We begin by presenting our account in terms of the lexicalistaion table (16). The key to the analysis is the proposal that the oblique endings (e.g., *-o-go, -o-m*, etc.) start at G1. In both genders, the demonstrative *ett-* is thus able to spell out the complement of G1 (CLASSP and MASCP respectively, both contained in its entry in (11a)). This is why the two genders are syncretic.

[6] If the root node were labelled as G1P, the definitions in the algorithm (14) would need to use the segment/category distinction of Kayne (1994).

(16)

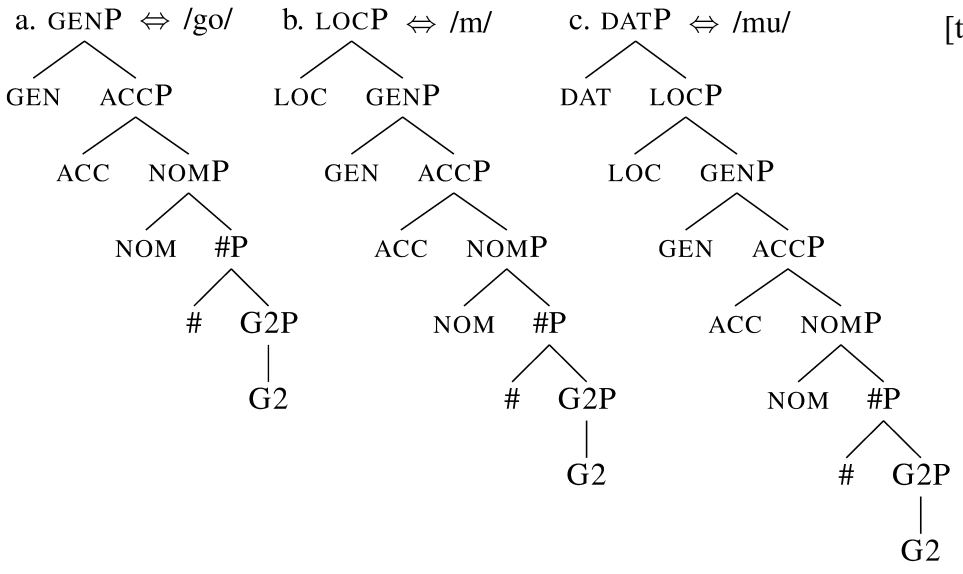
| | DemP | REF | CLASS | MASC | G1 | G2 | # | NOM | ACC |
|---|---|---|---|---|---|---|---|---|---|
| NOM | | ett | | | | o | | | |
| ACC | | ett | | | | | o | | |
| GEN | | ett | | | o | | | go | |
| LOC | | ett | | | o | | | m | |
| DAT | | ett | | | o | | | mu | |
| INS | | ett | | | | | | | im |

[lex table 2]

| | DemP | REF | CLASS | MASC | G1 | G2 | # | NOM | ACC | GEN | LOC | DAT | INS |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NOM | | | | etot | | | | | | | | | |
| ACC | | | | | etot | | | | | | | | |
| GEN | | | ett | | | o | | | go | | | | |
| LOC | | | ett | | | o | | | m | | | | |
| DAT | | | ett | | | o | | | mu | | | | |
| INS | | | ett | | | | | | | im | | | |

Let us now show the derivations in detail, starting from the neuter paradigm. In the genitive, the root is followed by two additional markers, *-o* and *-go*. We assume that *-o* is the same marker as the NOM/ACC *-o*, but realising only a part of its full potential, namely G1P. We analyse the markers *-go, -m, -mu* as dedicated singular case markers (i.e., they lexicalise the singular # along with the case features). We also let all the suffixes lexicalise G2, see (17).

(17)



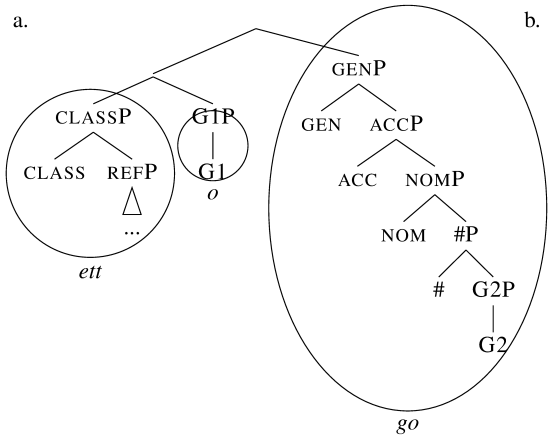

[tree 11]

This specification has two consequences. The first consequence is that these suffixes are restricted to the masculine and neuter gender. This is because these two genders have G2 immediately below #, just like the lexical entries in (17). These suffixes cannot be used in the

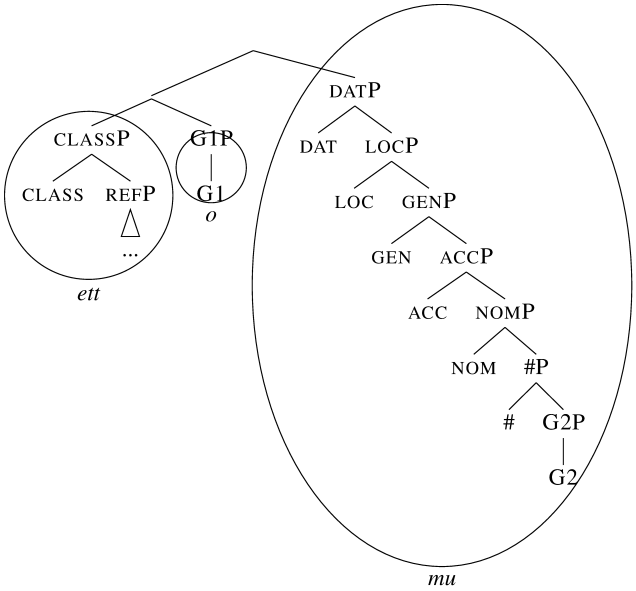
feminine because feminines have the FEM feature intervening between G2 and #, recall (8a). This prevents any match between the entries in (17) and any structure that includes FEM.

The second consequence is the following. Due to the fact that the entries in (17) are (so far) the only affixes capable of lexicalising the oblique cases, using them is the only chance of lexiclising the genitive, locative and dative. Now in order to use these suffixes, the derivation must create the right type of constituent structure for these affixes to match the morphosyntactic structure. The only way this can happen is for the complement of G2 (the lowest feature of the suffixes in (17)) to evacuate all the way to the the top of the tree. The relevant structures are given in (18), where the complement of G2, i.e., G1P, has (cyclically) moved to the left. As we know from (15d), G1P is realised as *ett-o*, so we get the correct forms *ett-o-go* (for the genitive), see (18a), and *ett-o-mu* for the dative, see (18b). The locative is not shown for reasons of space.

(18) a.



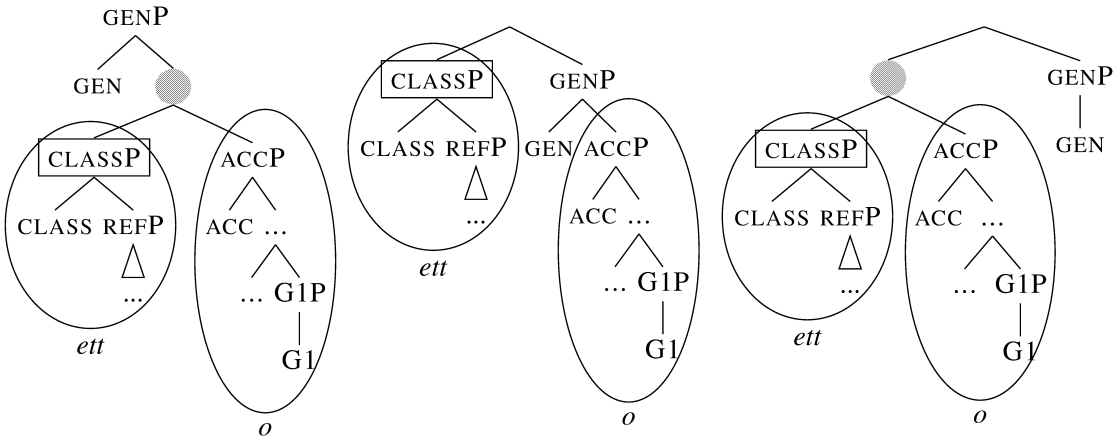

b.




**[tree 12]**

Let us now explain how exactly the derivation gets from the accusative (depicted in (14a)) to the genitive in (18a). The first step is that the genitive feature is added to the accusative, yielding (20a).

(20) a. merge GEN b. Evacuate closest c. Pied-pipe [tree 13]



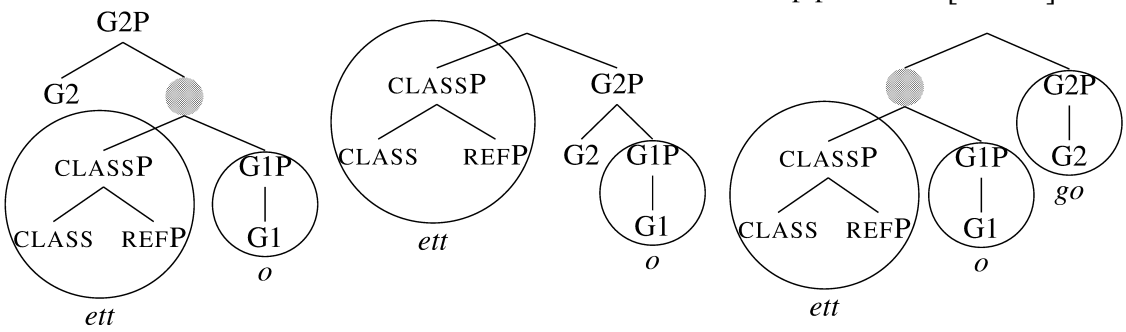

There is no lexical item matching this structure, which means that evacuation movements are triggered, following the algorithm (14). The closest labelled non-remnant constituent in (20a) is CLASSP, see the rectangle in (20a). After CLASSP evacuates from GENP, we get (20b). There is still no item matching the GENP in (20b), since *-go* in (17a) cannot spell out G1, which is at the bottom of the GENP constituent in (20b).

When matching fails in (20b), the CLASSP is restored to its original position, so we agin have the tree in (20a). The next step of the algorithm is to pied-pipe the immediately larger constituent, i.e., the unlabelled node marked by the gray circle in (20a). When we evacuate this node, we have the structure in (20c), but again, GENP cannot be lexicalised. The lexical item (17a) does not contain the GENP in (20c) as a sub-constituent.

When matching fails in (20c), we again restore the moved constituent to the original position in (20a). At this point, there is no more pied-piping that could be tried, and we therefore have to backtrack (14d). Backtracking has the following logic: up to now, we have followed a series of steps and ended up merging the GEN feature to a particular structure, corresponding to the accusative. However, all evacuation steps failed when GEN was merged. Therefore, the only chance is to go back in the derivation, change the structure at some prior step, rebuild the structure back to ACC, and then finally merge GEN to a different configuration, hoping that this time, matching is going to succeed.

Backtracking goes back in the derivation step by step, trying to find a new realisation for a structure that had already been lexicalised. If no new option for lexicalisation exists at a cycle preceding the one where lexicalisation stalled, we must backtrack one more step, and so on. The relevant moment for us is when the derivation backtracks all the way down to G2P, see (20a).

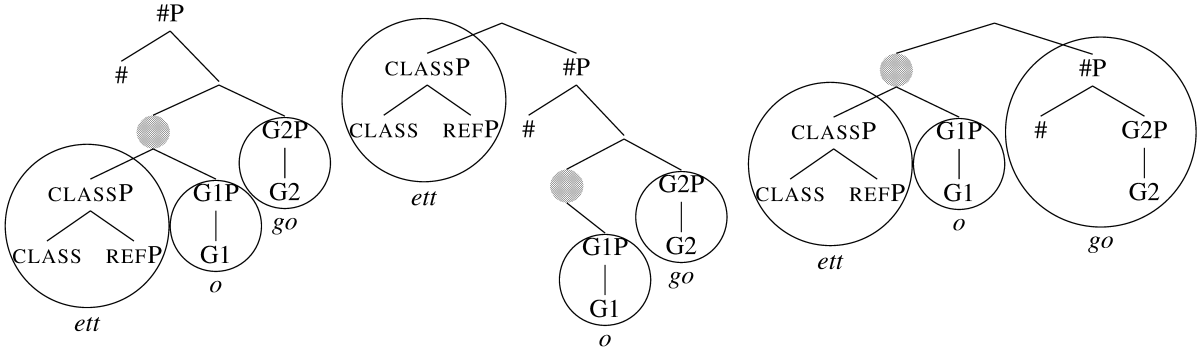
(20) a. Backtrack b. Evacuate closest c. Pied-pipe [tree 14]




When we first merged G2, we evacuated the closest non-remnant labelled constituent, i.e., CLASSP. The structure we got is in (20b), repeated from (15f). Originally, G2P was lexicalised as *-o*, recall (15g). However, continuig in that line of derivation led nowhere at GENP, so what we need to do now is to skip the step (20b), restore the CLASSP in the original position, and we are back at (20a). Following the step of the algorithm (14c), we now pied-pipe the constituent marked by the gray circle in (20a), yielding (20c). At this point, the G2P can be spelled out by *-go*, because *-go* contains G2P (with a single daughter) as a sub-constituent, see (17a).

The derivation now continues by merging #, yielding (21a). There is no lexical item matching this constituent, so we must evacuate the closest non-remnant constituent, i.e., the CLASSP. The result of the evacuation is in (21b).

(21) a. Backtrack b. Evacuate closest c. Pied-pipe [tree 15]

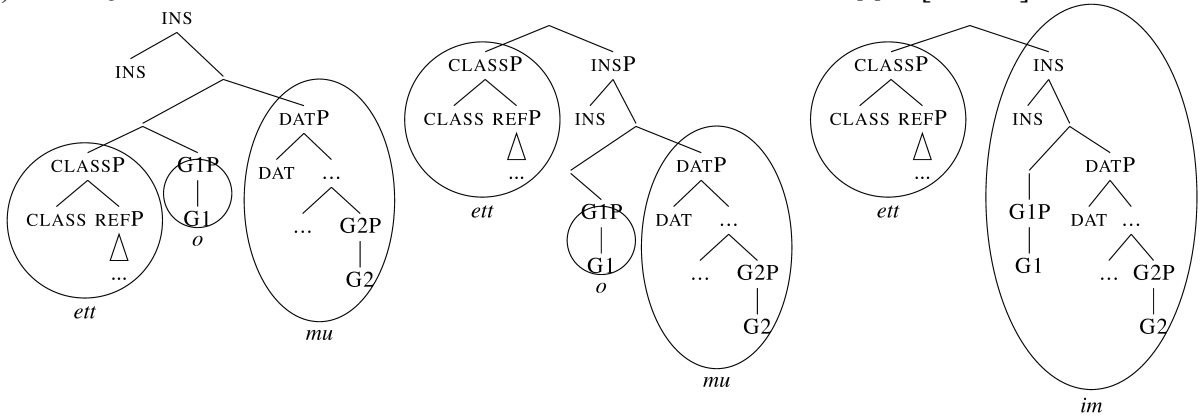



This movement step represent a movement type traditionally called subextraction (we move the Spec of a Spec). One of the goals of this paper is to explore this new type of derivations, which are not a part of the algorithm proposed in Starke (2018) (see also Wiland 2019, Cortiula 2023). The subextraction step will become important in several derivations. However, in (21b), subextraction fails, as there is no lexical item matching the #P in (21b). The CLASSP is therefore placed back, and the immediately dominating constituent (marked by the gray circle in (21a)) is evacuated. After evacuation, the structure is as in (21c), where *-go* spells out #P.

The derivation now proceeds along similar lines, adding NOM, ACC, where *ett-o* (corresponding to the constituent marked by the gray circle in (21c)) cyclically moves up. The suffix *-go* lexicalises larger and larger constituents, eventually leading to the genitive structure in (18a). The derivation of the locative and the dative proceeds similarly, ultimately yielding the structure (18b) for the dative.

When the instrumental feature is merged on top of the dative (as in (22a)), there is no matching lexical item. Therefore, we must evacuate the closest non-remnant labelled constituent, which is CLASSP.

(22) a. Merge INS b. Evacuate closest c. Pied-pipe [tree 16]




The result of the subextraction is shown in (22b). Our analysis of the instrumental, as depicted in the lexiclaisation table (23) (repeated from (16)), says that the instrumentsal *-im* is a portmanteau that lexicalises the INSP constituent as given in (22b), see (22c). This is the first point of the analysis where subextraction is required.

(23)

| | DemP | REF | CLASS | MASC | G1 | G2 | # | NOM | ACC |
|---|---|---|---|---|---|---|---|---|---|
| NOM | | ett | | | | o | | | |
| ACC | | ett | | | | o | | | |
| GEN | | ett | | | o | | go | | |
| LOC | | ett | | | o | | m | | |
| DAT | | ett | | | o | | mu | | |
| INS | | ett | | | | | im | | |

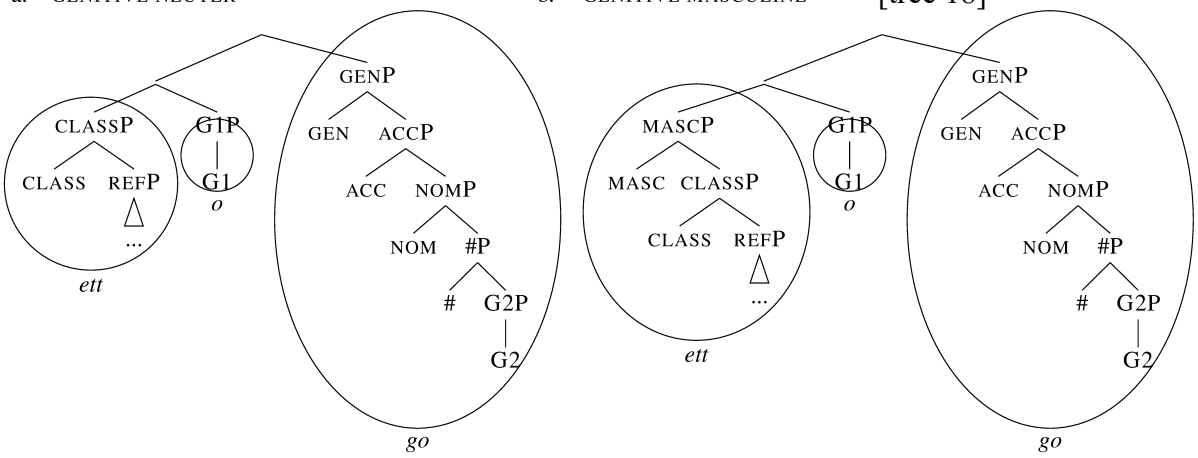
| | DemP | REF | CLASS | MASC | G1 | G2 | # | NOM | ACC | GEN | LOC | DAT | INS |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NOM | | | | etot | | | | | | | | | |
| ACC | | | | etot | | | | | | | | | |
| GEN | | | ett | | o | | go | | | | | | |
| LOC | | | ett | | o | | m | | | | | | |
| DAT | | | ett | | o | | mu | | | | | | |
| INS | | | ett | | | | im | | | | | | |

To see the reasons for this proposal, consider first the fact that due to subextraction, the instrumental ending *-im* spells out G1 in a stranded position above DAT, rather than in the base position (see (22c)). If subextraction did not exist, we would need to propose an entry of *-im* with G1 below G2. However that would effectively prevent the bimorphemic combination *-o-go* in the genitive. The reason for that is that if *-im* had a regular entry and did not rely on

subextraction, it would match the structure in (20b), and the genitive would be realised as *ett-im*, the worng outcome. It is therefore important that *-im* spells out G1 in the high position, above DAT, something which is only possible with subextractions.

Let us now turn to the discussion of the masculine oblique cases. Our idea is that the syncretism arises due to the ambiguity of the root. We show the proposed structures for the neuter and masculine genitive side by side in (24a,b) respectively. We can se that the MASC feature in (24b) is realised by the root. The root is also capable of realising a structure without it, see (24a). As a result, we get syncretism: the presence/absence of MASC is never reflected on the surface.

(24) a. GENITIVE NEUTER b. GENITIVE MASCULINE [tree 18]

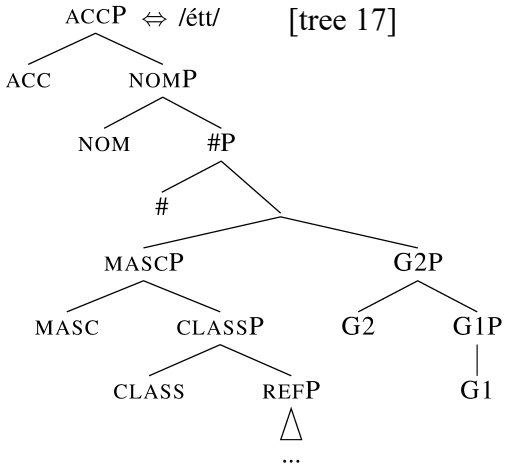



Note also that by placing the ambiguity inside the root, we avoid the standard argument for Impoverishment, which is that if the ambiguity was located insight the endings, each ending would have to be ambiguous in the same way. This is not the case here, it's just the root that is ambiguous.

However, one issue that needs to be addressed is the following: we can see in Table (23) that the root of the demonstrative is capable of lexicalising G1 in the masculine NOM/ACC cases. The question therefore is why *-o* is needed in (24b) to spell out G1: why can't the masculine

demonstrative lexicalise the whole G1P?

Let us first make it clear that if the entry of the root was as in (11a), as originally proposed, we would indeed predict that G1P is lexicalised by the root, deriving ungrammatical forms such as **et(o)t-go*. Drawing on the ideas explored in Blix (2022), we therefore propose that the lexical entry of the demonstrative is as in (25).

(25) 

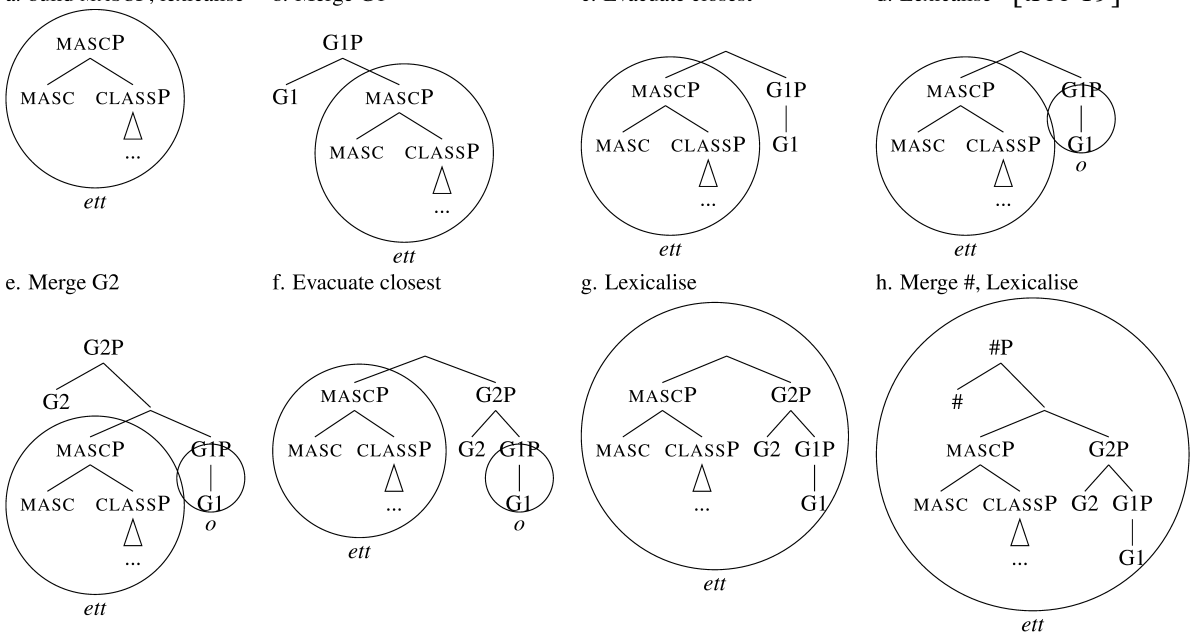
[tree 17]

The make-up of the entry rests on the idea that lexical items pair well-formed syntactic structures with phonological representations. Since well-formed syntactic structures may contain movement (the movement of MASCP across G2 in (25)), lexical items may link such structures to phonology.

Let us now discuss what effects such entries have. The first remark is that this entry changes nothing on the derivation of the neuter paradigm, where the root consistently spells out CLASSP. This is still possible, and the update therefore changes nothing of substance.

Let us now turn to the masculine derivation, see (26). We start following it at the point when MASCP is formed and lexicalised as *ett*, see (26a). Lexicalisation is possible since MASCP is contained in (25). When G1 is merged, yielding (26b), there is no match, since the entry of the demonstrative in (25) does not contain the G1P constituent as given in (26b). Therefore, MASCP is evacuated (see (26c)) and G1P is lexicalised by *-o*, see (26d).

(26) a. build MASCP, lexicalise b. Merge G1 c. Evacuate closest d. Lexicalise [tree 19]




e. Merge G2 f. Evacuate closest g. Lexicalise h. Merge #, Lexicalise




When G2 is merged, see (26e), there is still no match. Therefore, MASCP evacuates, yielding (26f). Since the whole constituent in (26f) is contained in the entry (25), we can lexicalise it as *ett*, see (26g). When # is merged on top of G2, see (26h), matching suceeds again. Adding case features on top of (26h) will lead to successful lexicalisation by the root as well, and so the nominative and accusative of the masculine correspond to the bare root.

When the genitive feature is added, backtracking is triggered. Since the only endings capable of spelling out the oblique features (GEN and up) have their lowest feature at G2, G1P will need to evacuate from below G2, exactly as in the derivation of the neuter, recall (20c). Since G1P in the masculine is lexicalised as *ett-o*, see (26d), we correctly derive the syncretic forms for neuter and masculine oblique cases as given in (24).

## 3. Deriving the feminine paradigms

In this section, we turn to the feminine paradigm, repeated in an updated form in Table 7. The table has the demonstrative in the middle, flanked by the paradigms of declensions II and III. Recall that our main goal is to explain why the demonstrative inflects the same as the noun

'wife' in NOM, ACC and INS, but not in the other cases.

| | wife (II) | this | notebook (III) |
|---|---|---|---|
| | | **FEMININE SINGULAR** | |
| **NOM** | žen-a | et-a | tetrad' |
| **ACC** | žen-u | et-u | tetrad' |
| **GEN** | žen-y | et-o-**j** | tetrad-**i** |
| **LOC** | žen-e | et-o-**j** | tetrad-**i** |
| **DAT** | žen-e | et-o-**j** | tetrad-**i** |
| **INS** | žen-o-j(u) | et-o-**j(u)** | tetrad-**ju** |

TABLE 7: The feminine paradigms

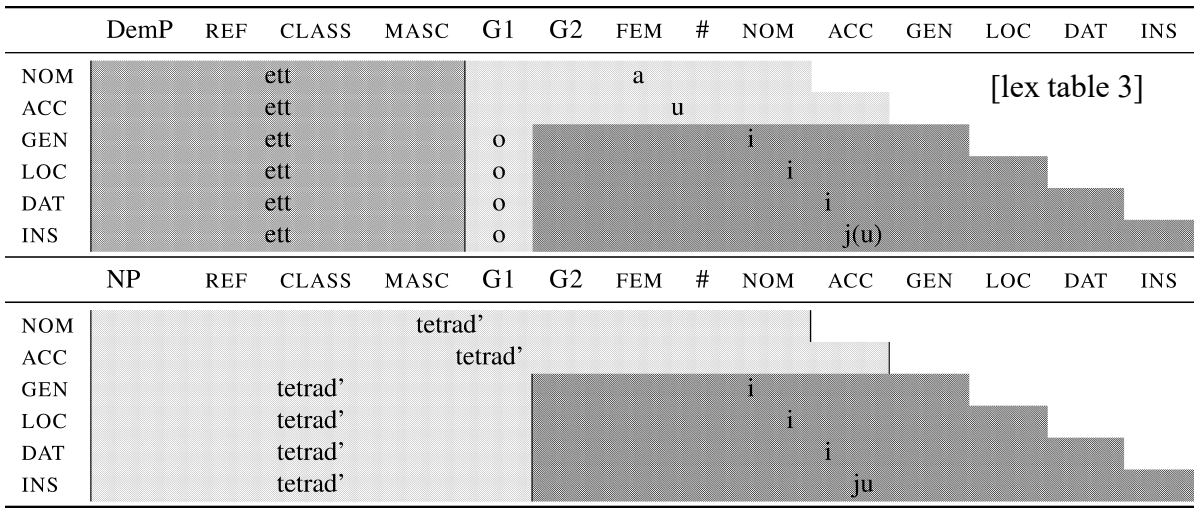
In the demonstrative paradigm, we again segment out the initial *-o* of the demonstrative ending as a special morpheme. Once the demonstrative is segmented, it can be observed that the morphemes that follow the marker *-o* can be understood as identical to the third declension endings. The identity is indicated by boldface, assuming that the ending *-i* in GEN, LOC and DAT is realised as *-j* after vowels.

At this point, the archaic ending *-ju* becomes significant because it helps us see the parallel between the endings of the third declension and the endings of the demonstrative. We assume that also the non-archaic endings are identical to the ending of the third declension, i.e., they are underlyingly *-ju*, but the final *u* is lost due to a phonological process (apocope). Apocope only applies when *-ju* is preceded by a vowel (as in the demonstrative paradigm and also in the paradigm 'wife'), but not when it is preceded by a consosnant (as in Declension III).[7]

[7] In more technical terms, we analyse the *u* of *-ju* as a floating vowel, which is only attached to the skeleton if it must govern a preceding empty nucleus. In the strict CVCV model of Scheer (2004), an empty nucleus is found at the end of each morpheme that ends in a consonant. Thus, after consonants, we find *-ju* (where *u* attaches to

Let us now turn to the morphosyntactic analysis of the feminine demonstrative. Our goal is to capture the fact that in the oblique cases, the demonstrative combines with a set of endings that are found on third declension nouns. We also want to capture the fact that the third-declension endings are preceded by *-o* on the demonstrative, which should ideally be the same *-o* as in the masculine/neuter paradigm. All of these facts can be captured by an analysis that we depict in Table (27).

(27)



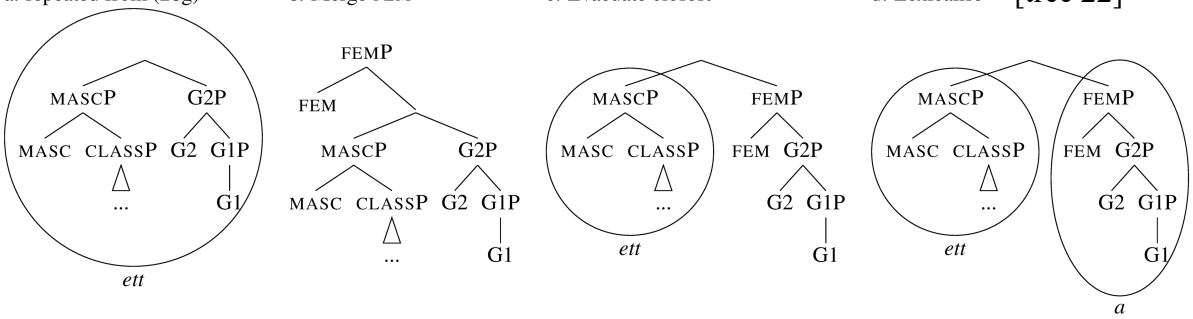

Let us now explain how this analysis works. Recall first from (3) that the MASC feature if found also in the feminine gender; MASC is therefore included in Table (27). Given this, the derivation of the feminine demonstrative proceeds initially the same as the derivation of the masculine demonstrative in (26). A consequence of this, which is worth noting, is that the G1P of the feminine demonstrative is realised the same as the masculine one, namely as as *ett-o*, recall (26d). This fact is relevant for the analysis of the feminine obliques, where, as we propose, the feminine endings *-i* and *-j(u)* attach to the G1P base *ett-o*, see Table (27).

Let us now see how the derivation proceeds from FEM on. Recall first that G2P is realised as (28a), repeated for convenience from (26g). When we merge FEM, a difference between the maculine and the feminine appears. Namely, the structure can no longer be spelled out by the

govern the preceding empty nucleus), and *-j* elsewhere. In this analysis, there is no expectation that all final vowels in Russian are elided: only floating vowels are.

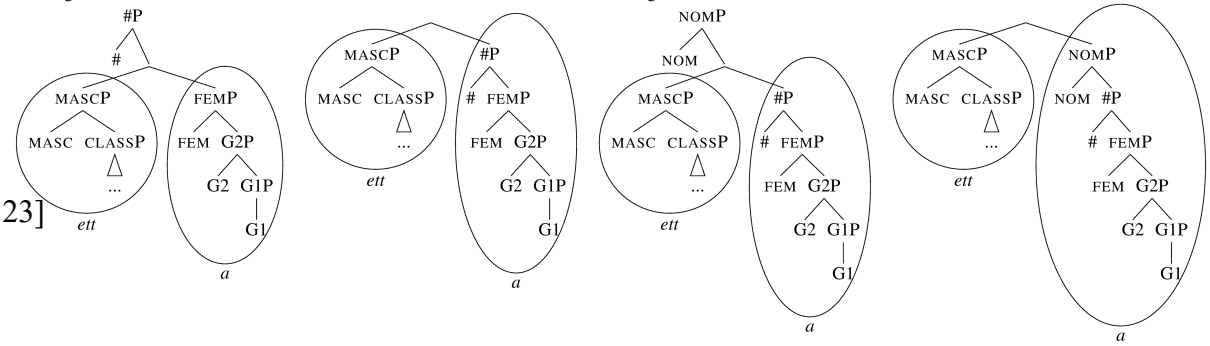
root, and an evacuation movement is triggered, see (28c). We hypothesise that the constituent on the right branch is lexicalised as *-a*, see (28d).

(28) a. repeated from (28g) b. Merge FEM c. Evacuate closest d. Lexicalise [tree 22]




The derivation then continues by merging #, see (29a), which forces another evacuation movement, see (29b). When the NOM feature is added (29c), we agin get an evacuation movement and lexicalise as in (29d). The result is as per the lexicalisation table (27).

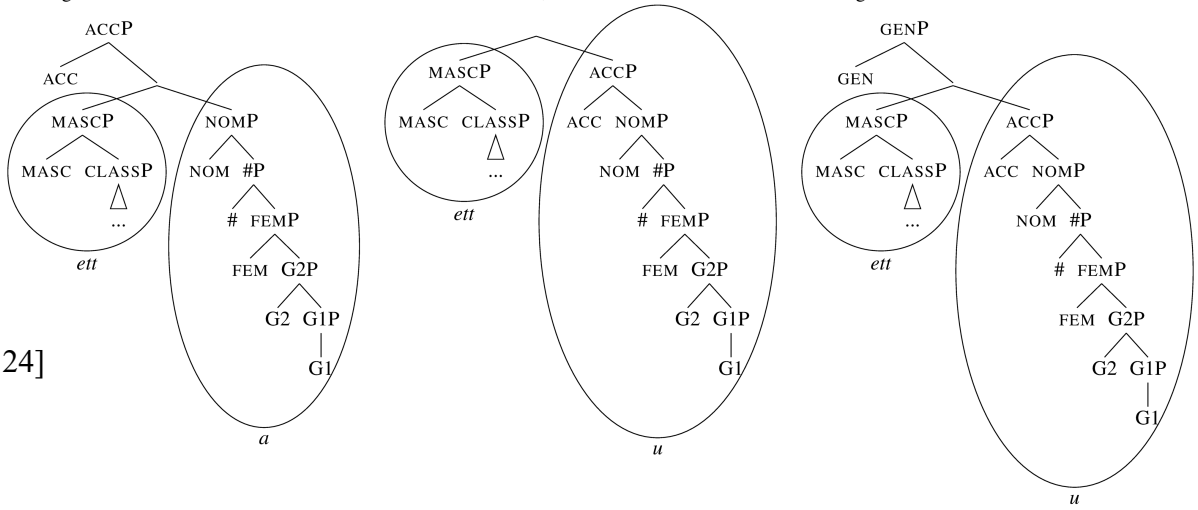
(29) a. Merge # b. Evacuate closest, Lexicalise c. Merge NOM d. Evacuate closest, Lexicalise




[tree 23]

When the ACC feature is added as in (30a), MASCP evacuates again as in (30b) and we get the form *ett-u*.

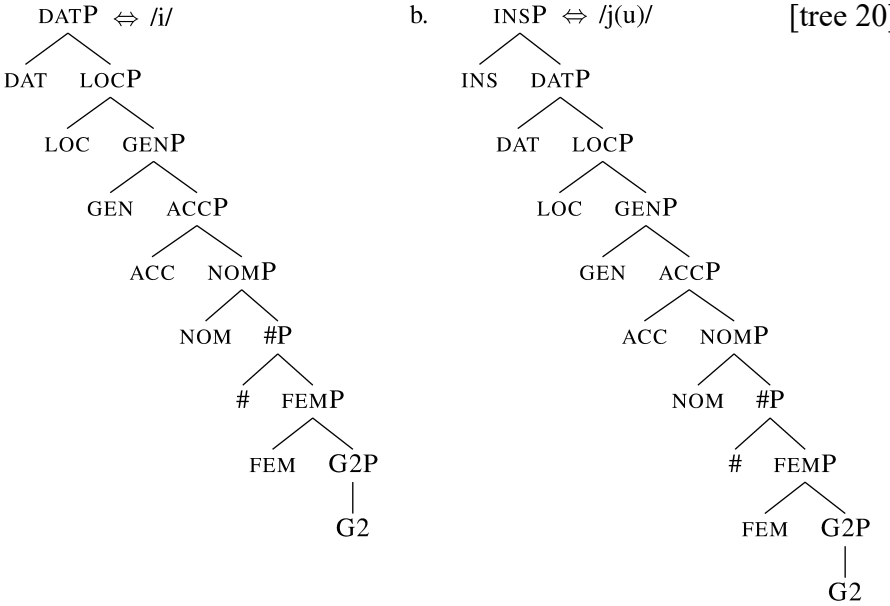
(30) a. Merge ACC b. Evacuate closest, Lexicalise c. Merge GEN




[tree 24]

The question is what happens when the GEN feature is added, as in (30c). We know that masculine/neuter oblique endings *-go, -mu*, etc. are unusable in the feminine gender (recall the discussion surrounding (17)), which is a good thing, because they don't play any role in these derivations. Instead, we must postulate a new set of feminine endings, namely *-i* and *-ju*. We propose that these start at G2, and reach all the way up to the oblique features. We give the new endings in (31).

(31) a. DATP ⇔ /i/ b. INSP ⇔ /j(u)/ [tree 20]

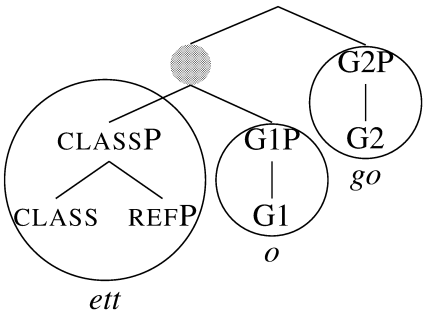



Due to the fact that their lowest feature is G2, the oblique endings trigger backtracking down to the G2 cycle, which is analogous to the masculine/neuter demonstratives, recall (20a). In the

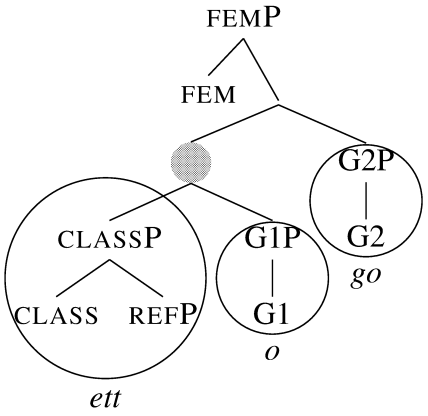
derivation of the feminine, the original realisation of G2P looks as in (28a), which is a structure derived by moving the MASCP across G2 by Evacuate Closest. However, since this lexicalisation ultimately led to a crash at GEN, we now have to resort to pied-piping, moving the whole complement of G2, as in (32a). The G2P here is spelled out as *-go*, despite the fact that the feminine *-i* in (31a) is also a match, but *-go* has the smallest number of unused features, see (17a). When FEM is merged, we get (32b). Since (32b) cannot be lexicalised, Evacuate closest applies, yielding (32c). This structure, however, cannot be lexicalised either. Therefore, we go back to (32b), and pied-pipe the whole Spec, yielding (32d). Here the FEMP is lexicalise is *-i* (realised as *-j* for phonological reasons).

(32) [tree 25]

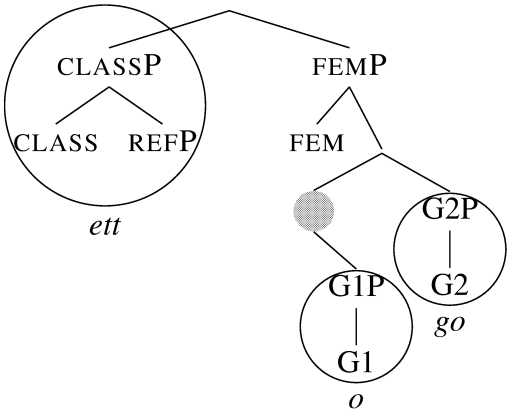
a. Stage (22c)



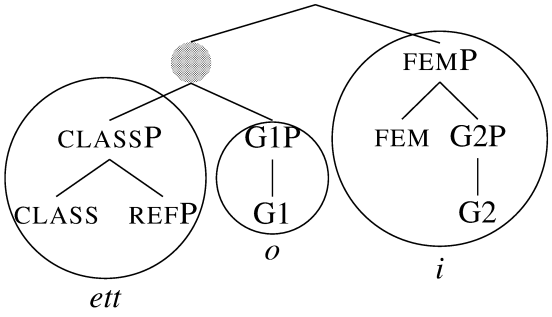

b. Merge FEM




c. Evacuate closest

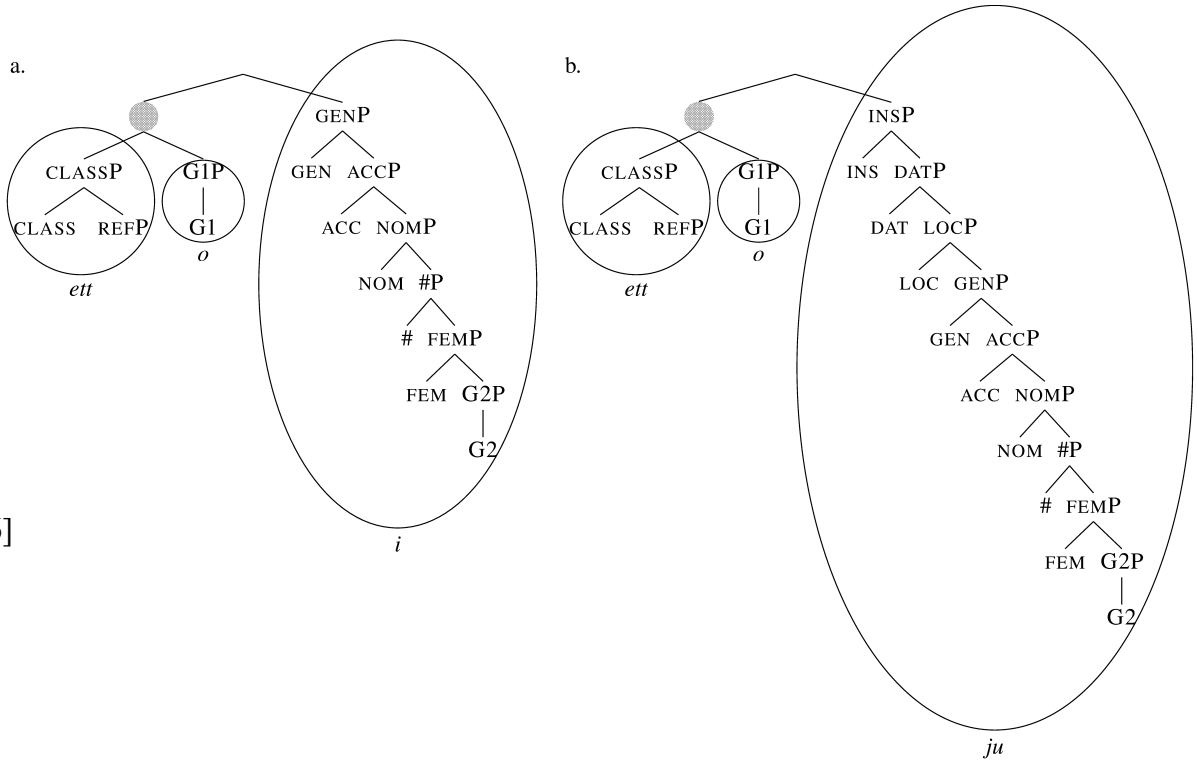



d. Pied-pipe, lexicalise




The derivation now continues by merging # to (32d), pied-piping the whole Spec, etc., until all the features from G2 up to GEN are realised by the oblique feminine singular ending *-i*, as in (33a). (33b) shows the instrumental.

(33)




[tree 26]

The lexicalisation table in (34) (reapeated from (27)) summarises the derivations.

(34)

| | DemP | REF | CLASS | MASC | G1 | G2 | FEM | # | NOM | ACC | GEN | LOC | DAT | INS |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NOM | ett | | | | | a | | | | | | | | |
| ACC | ett | | | | | u | | | | | | | | |
| GEN | ett | | | | o | | | | i | | | | | |
| LOC | ett | | | | o | | | | | i | | | | |
| DAT | ett | | | | o | | | | | | i | | | |
| INS | ett | | | | o | | | | | | | j(u) | | |

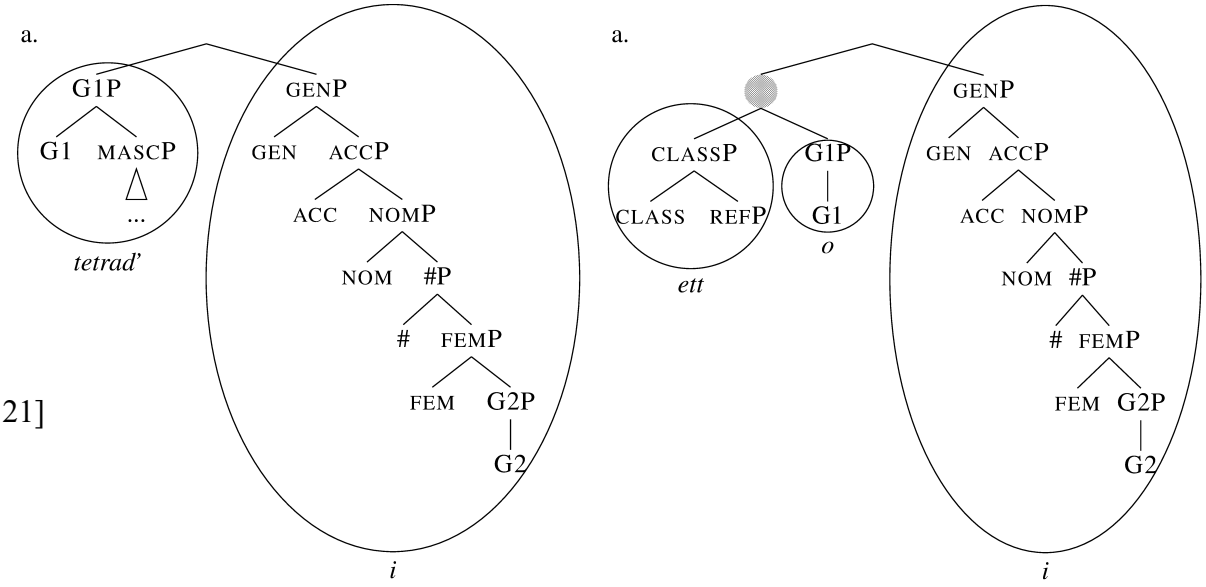
[lex table 3]

| | NP | REF | CLASS | MASC | G1 | G2 | FEM | # | NOM | ACC | GEN | LOC | DAT | INS |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NOM | tetrad’ | | | | | | | | | | | | | |
| ACC | tetrad’ | | | | | | | | | | | | | |
| GEN | tetrad’ | | | | | | | | i | | | | | |
| LOC | tetrad’ | | | | | | | | | i | | | | |
| DAT | tetrad’ | | | | | | | | | | i | | | |
| INS | tetrad’ | | | | | | | | | | | ju | | |

Let us now turn to the analysis of the noun *tetrad*’ ‘notebook.’ The only new entry we need is that of the root. The root (we propose) is capable of lexicalising all the features up to the accusative, and, therefore, no endings are needed in NOM/ACC. Since the noun root cannot lexicalise the oblique case features from the genitive up, it needs an ending. Since all the oblique endings start at G2, backtracking is triggered down to G1P, as in the case of the demonstrative.

The only difference between the noun and the demonstrative rests in the fact that the noun lexicalises G1P as a single constituent, see (35a). The demonstrative given in (35b) spells out G1P in two pieces (recall (33a)).

(35)




[tree 21]

With the analysis of the demonstrative in place, let us now turn to the paradigm ‘wife', repeated for convenience in Table 8 along with the other feminine paradigms.

| | **wife (II)** | **this** | **notebook (III)** |
|---|---|---|---|
| | **FEMININE SINGULAR** | | |
| **NOM** | žen-a | et-a | tetraď |
| **ACC** | žen-u | et-u | tetraď |
| **GEN** | žen-y | et-o-**j** | tetrad-**i** |
| **LOC** | žen-e | et-o-**j** | tetrad-**i** |
| **DAT** | žen-e | et-o-**j** | tetrad-**i** |
| **INS** | žen-o-j(u) | et-o-**j(u)** | tetrad-**ju** |

TABLE 8: The feminine paradigms

The main challenge (recall) is to capture the fact that the paradigm ‘wife’ is syncretic with the demonstrative in NOM/ACC, then the paradigms diverge and finally unify again in the INS.

Our analysis is depicted in (36). As we shall see, it heavily relies on subextractions, namely in GEN, LOC and DAT.

(36)

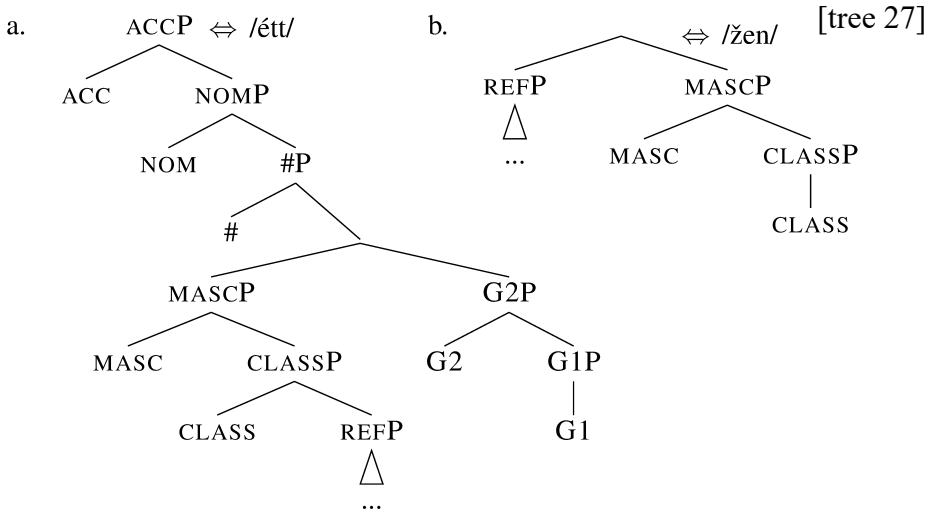
| | DemP | REF | CLASS | MASC | G1 | G2 | FEM | # | NOM | ACC | GEN | LOC | DAT | INS |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NOM | | | ett | | | | a | | | | | | | |
| ACC | | | ett | | | | | u | | | | | | |
| GEN | | | ett | | o | | | | i | | | | | |
| LOC | | | ett | | o | | | | | i | | | | |
| DAT | | | ett | | o | | | | | | i | | | |
| INS | | | ett | | o | | | | | | | j(u) | | |

[lex table 4]

| | NP | REF | CLASS | MASC | G1 | G2 | FEM | # | NOM | ACC | GEN | LOC | DAT | INS |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NOM | | | žen | | | | a | | | | | | | |
| ACC | | | žen | | | | | u | | | | | | |
| GEN | | žen | | | | | | | y | | | | | |
| LOC | | žen | | | | | | | | e | | | | |
| DAT | | žen | | | | | | | | | e | | | |
| INS | | | žen | | o | | | | | | | j(u) | | |

In (37), we provide the lexical entries for the roots in Table (36). Both of these entries contain a complex left branch that has been displaced from the base position (Blix 2022). An important difference between them rests in the shape of the MASCP constituent: both roots are able to realise this constituent, but the shape of the constituent is different. Namely, the root ‘wife’ moves its REFP out of MASCP, while the demonstrative has its REFP below CLASS.

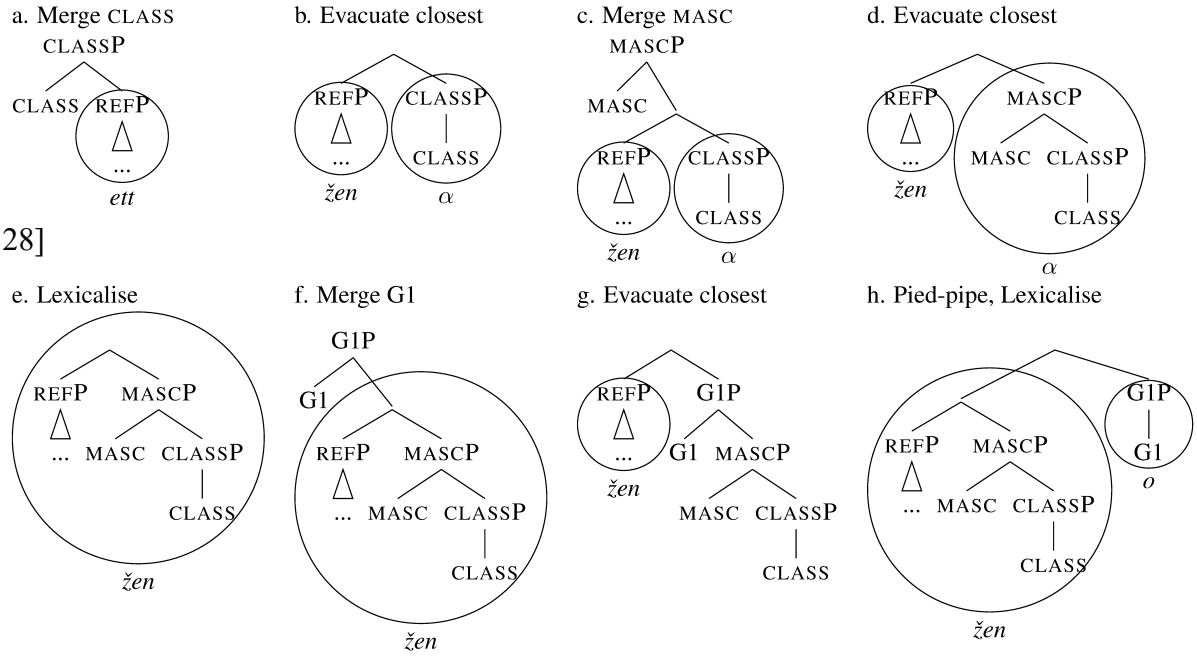
(37)




[tree 27]

Let us now discuss how the paradigm ‘wife’ is generated, highlighting interesting points of divergence with the demonstrative. Let us start describing the derivation at the point when CLASS is merged on top of REFP, see (38a). Since the lexical entry of ‘wife’ in (37b) does not contain this constituent, an evacuation movement is triggered, see (38b). We so far have no lexical item to lexicalise CLASSP in (38b), so we introduce a new lexical item α, which can lexicalise CLASSP. When MASC is meged in (38b), evacuation is triggered again, yielding (38d). For reasons that will become clear, we assume that α matches MASCP.

(38) [tree 28]

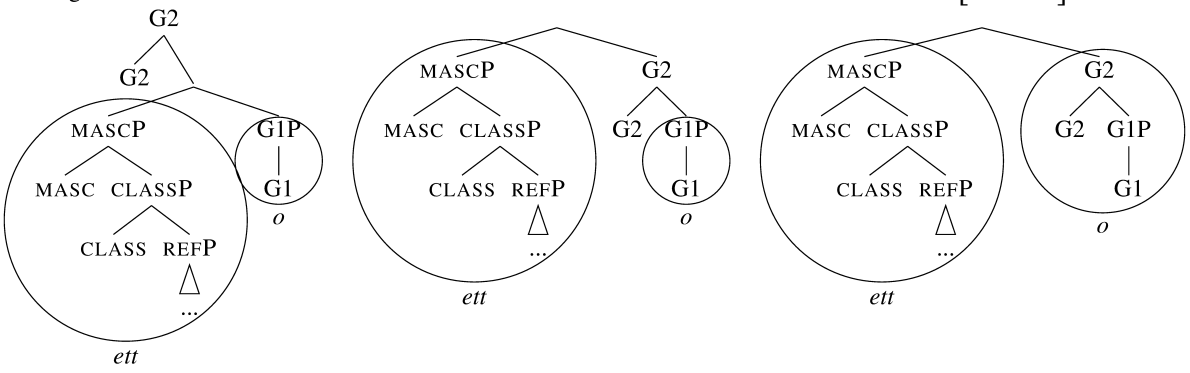



However, since the noun ‘wife’ matches also the whole constituent (38d), this lexicalisation is preferred, see (38e). When G1 is merged (see (38f)), Evacuate Closest applies, producing (38g), where REFP has moved out of G1P. However, we assume that α is no longer a match, so G1P in (38g) cannot be lexicalised. Therefor, the REFP is placed back, and we move the whole complement of G1 out of G1P, yielding (38h). Here G1P is realised as *-o*.

It is worth noting that the G1P of a feminine demonstrative is also realised as a root followed by an *-o*, i.e., *ett-o*, recall (26d). However, despite being realised the same, the

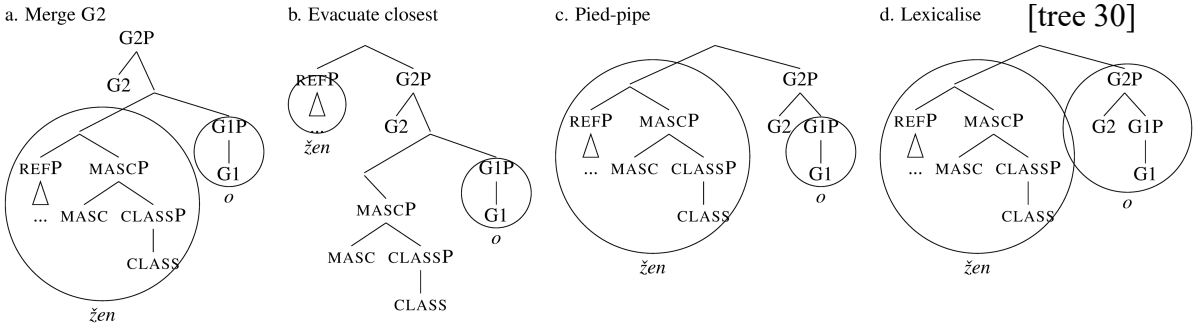
structures differ slightly in one crucial respect. To see this, consider what happens when we merged G2 on top of the demonstrative, as in (39a). The structure finds no match, so Evacuate Closest applies. The closest labelled non-remnant constituent is MASCP, and so Evacuate Closest yields (39b). The structure is successfully spelled out, see (39c).

(39) a. Merge G2 b. Evacuate closest c. Lexicalise [tree 29]




The derivation of the feminine noun works differently. Namely, when G2 is merged, Evacuate Closest leads to the movement of REFP, yielding (40b). Note that the step in (40b) has no parallel in the demonstrative paradigm in (39). There is, however, no match for G2P in (40b), so the movement is undone. The Pied-piping step of the algorithm then moves the whole Spec, yielding (40c), which leads to a successful matching, see (40d).

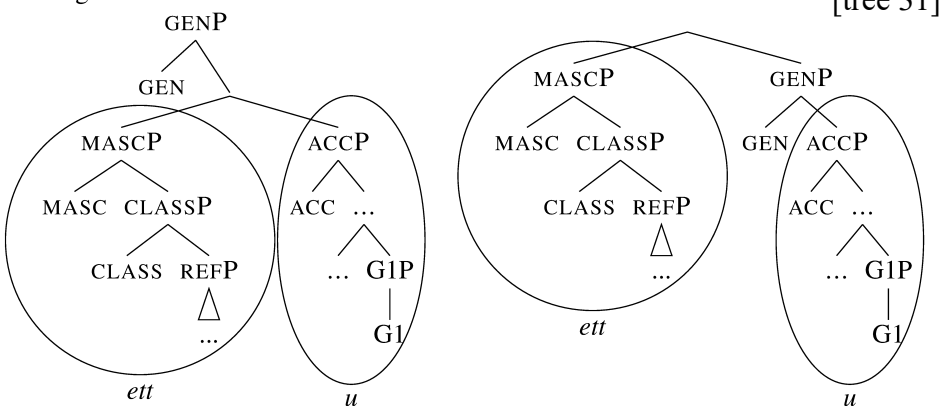
(40) a. Merge G2 b. Evacuate closest c. Pied-pipe d. Lexicalise [tree 30]




The net result is that the noun in (40) ends up with the same ending as the demonstrative (39),

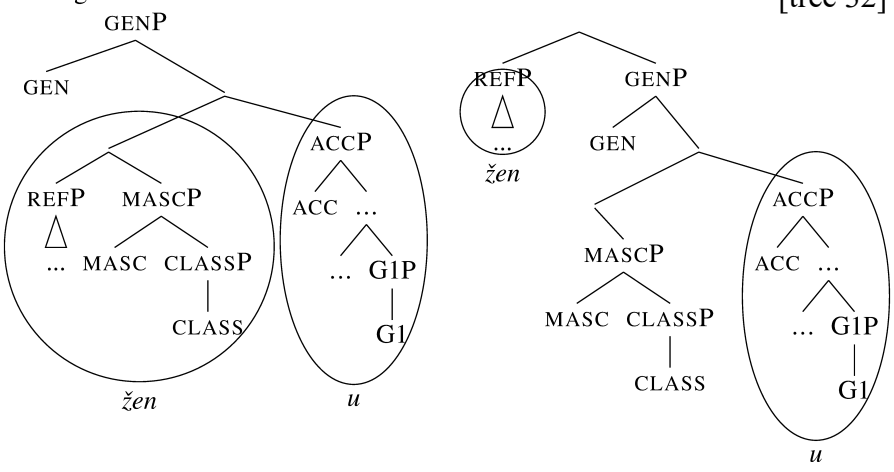
but the noun always goes through an extra step of subextraction, see (40b). As long as this extra intermediate step fails to lexicalise, the noun will have the same inflection as the demonstrative. This is so all the way to the accusative, but things work out differently in the genitive. When the genitive feature is merged on top of the demonstrative in (41a), lexicalisation fails even after evacuation movements (like the one in (41b)), and backtracking applies (recall (31)).

(41) a. Merge GEN b. Evacuate closest [tree 31]




However, as we already know, the noun has a slightly different sequence of evacuation steps. So when the GEN feature is merged on the noun (as in (42a)), the Evacuate Closest step of the algorithm subextracts REFP from within the Spec, see (42b).

(42) a. Merge GEN b. Evacuate closest [tree 32]



Given the fact that the structure (42b) is unique to nouns, it is tempting to propose that the GENP in (42b) is realised by the genitive ending *-y*, as shown in the lexicalisation table (43), repeated from (36). We can see that as a result of subextraction, *-y* is marked in the table as realising CLASS and MASC, which are the features that have been stranded in (42b) due to subextraction.

(43)

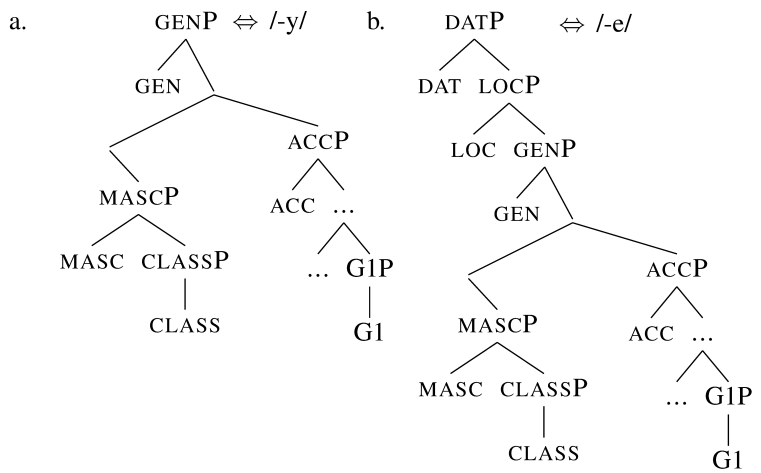
| | DemP | REF | CLASS | MASC | G1 | G2 | FEM | # | NOM | ACC | GEN | LOC | DAT | INS |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NOM | | | ett | | | | a | | | | | | | |
| ACC | | | ett | | | | u | | | | | | | |
| GEN | | | ett | | o | | | | i | | | | | |
| LOC | | | ett | | o | | | | | i | | | | |
| DAT | | | ett | | o | | | | | | i | | | |
| INS | | | ett | | o | | | | | | | j(u) | | |

[lex table 4]

| | NP | REF | CLASS | MASC | G1 | G2 | FEM | # | NOM | ACC | GEN | LOC | DAT | INS |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NOM | | | žen | | | | a | | | | | | | |
| ACC | | | žen | | | | u | | | | | | | |
| GEN | žen | | | | | | | | y | | | | | |
| LOC | žen | | | | | | | | | e | | | | |
| DAT | žen | | | | | | | | | | e | | | |
| INS | | | žen | | o | | | | | | | j(u) | | |

As the table makes it clear, we assume that also the LOC/DAT *-e*, which is specific to the paradigm ‘wife', lexicalises CLASS and MASC. The lexical shape of these endings copies the derivation provided in (42b), with CLASS and MASC stranded below GEN. The constituency requirement on lexicalisation has the consequence that these markers can only lexicalise the oblique case features if they also lexicalise the stranded CLASS and MASC features, see (44).

(44)



Since the demonstrative fails to strand these features (recall (41)), these endings cannot be used in the demonstrative declension, and we get the effect, coded in the lexicalisation table, that the noun and the demonstrative inflect intially alike, but diverge in GEN, LOC, DAT.

Finally, when the derivation hits the instrumental, the only available ending is *-ju* in (31b). This means that the nominal derivation must backtrack down to G2P, and move the complement of G2 cyclically up, as in the demonstrative derivation, recall (33b). Since G1P is realised as *žen-o-* (recall (39a)), the form we correctly derive is *žen-o-j(u)*.

## 5. Neuter and masculine nominal paradigms

The final task that lies ahead is capturing the paradigms 'factory' (*zavod*) and 'place' (*mest-o*), see Table 9.

| | **this** | **place ($I_B$)** | **this** | **factory ($I_A$)** |
|---|---|---|---|---|
| | **NEUTER SINGULAR** | | **MASCULINE SINGULAR** | |
| **NOM** | et-o | mest-o | etot-ø | zavod-ø |
| **ACC** | et-o | mest-o | etot-ø | zavod-ø |
| **GEN** | et-ogo | mest-**a** | et-ogo | zavod-**a** |
| **LOC** | et-om | mest-**e** | et-om | zavod-**e** |
| **DAT** | et-omu | mest-**u** | et-omu | zavod-**u** |
| **INS** | et-im | mest-**om** | et-im | zavod-**om** |

TABLE 9

Recall that what we have to capture is that the nouns are the same as the demonstratives in NOM/ACC, but different in the oblique cases. The analysis is going to be analogous to the feminine paradigms: the nouns will be associated to a lexical entry with a moved phrase. This property then interacts with subextractions in a way that it unlocks a set of endings for the nouns

that are not available for the demonstratives. Due to space limitations, we cannot guide the reader fully through the derivations, but we nevertheless want to give here a taste of the analysis including one potentially tricky issue that is going to be resolved via pointers.

We start with the neuter paradigms. We give the lexicalisation table in (45). Recall that neuters lack the projections of both MASC and FEM, which we leave out of the table.

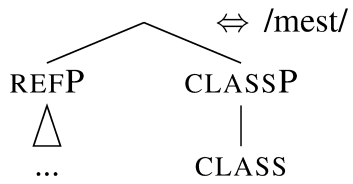
(45)

| | DemP | REF | CLASS | G1 | G2 | # | NOM | ACC | GEN | LOC | DAT | INS |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NOM | | ett | | | | o | | | | | | |
| ACC | | ett | | | | | o | | | | | |
| GEN | | ett | | o | | | | go | | | | |
| LOC | | ett | | o | | | | | m | | | |
| DAT | | ett | | o | | | | | | mu | | |
| INS | | ett | | | | | | | | im | | |

| | NP | REF | CLASS | G1 | G2 | # | NOM | ACC | GEN | LOC | DAT | INS |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NOM | | mest | | | | o | | | | | | |
| ACC | | mest | | | | | o | | | | | |
| GEN | mest | | | | | | a | | | | | |
| LOC | mest | | | | | | | e | | | | |
| DAT | mest | | | | | | | | u | | | |
| INS | mest | | | | | | | | | om | | |

We see that the nominal root -- just like the demonstrative root -- realises CLASSP in NOM/ACC. However, the demonstrative roots and the nominal roots differ in that the nominal root has its REFP in the Spec of CLASS, see (46).

(46)

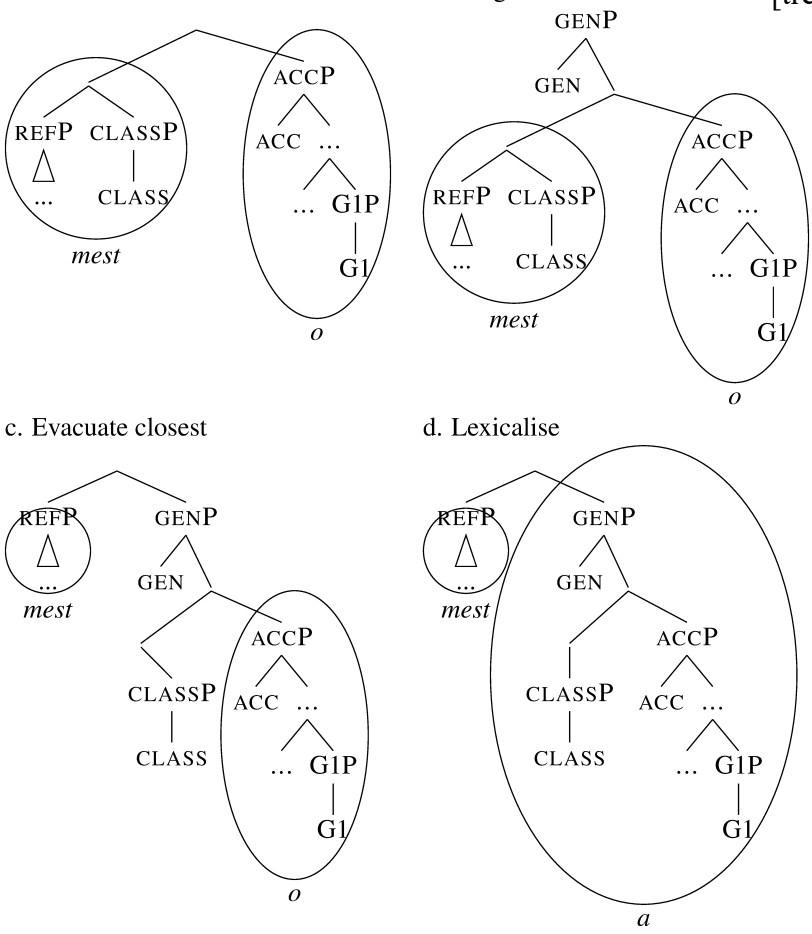



[tree 34]

We follow the derivation of the neuter noun starting with the accusative, which looks as in (47a). When the GEN feature is merged on top of it, see (47b), evacuate closest subextracts the REFP, yielding (47c). We assume that the genitive *-a* spells out precisely this constituent, see

(47d).

(47) a. The structure of ACC b. Merge GEN [tree 35]




We further assume that when the locative feature is added, REFP moves one notch up and the LOCP (with CLASSP stranded in between GEN and ACC) is realised as *-e*. Analogous derivation is then assumed for all the oblique endings. As a result, all the oblique endings are only available for nouns, but not for demonstratives, which lack the subextraction option due to their lexical shape.

Before we move on to the discussion of masculines, please note that the CLASSP stranded in the Spec of ACCP in (47c) can be, in principle, realised by the morpheme α, utilised in (38c,d). This fact will become important in the discussion of the masculines.

For masculines, we propose the lexicalisation table in (48).

(48)

| | DemP | REF | CLASS | MASC | G1 | G2 | # | NOM | ACC | GEN | LOC | DAT | INS |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NOM | etot | | | | | | | | | | | | |
| ACC | etot | | | | | | | | | | | | |
| GEN | ett | | | | o | | | go | | | | | |
| LOC | ett | | | | o | | | m | | | | | |
| DAT | ett | | | | o | | | mu | | | | | |
| INS | ett | | | | im | | | | | | | | |

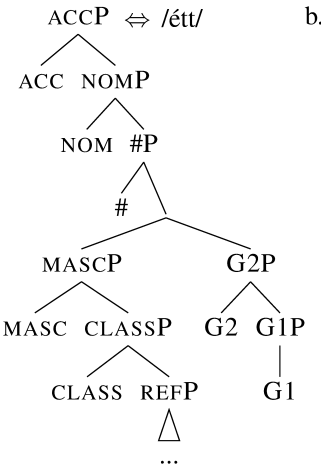
[lex table 5]

| | DemP | REF | CLASS | MASC | G1 | G2 | # | NOM | ACC | GEN | LOC | DAT | INS |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NOM | zavod | | | | | | | | | | | | |
| ACC | zavod | | | | | | | | | | | | |
| GEN | zavod | | | | a | | | | | | | | |
| LOC | zavod | | | | e | | | | | | | | |
| DAT | zavod | | | | u | | | | | | | | |
| INS | zavod | | | | om | | | | | | | | |

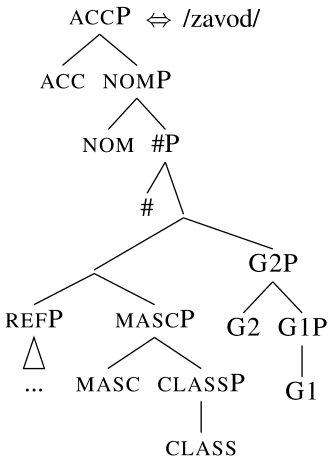
We give the entries for the demonstrative and the noun side by side in (49). We can see that the only difference is the position of REFP relative to CLASS and MASC inside the left branch.

(49)

a.




b.

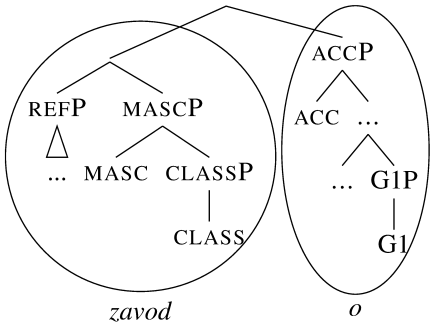



[tree 36]

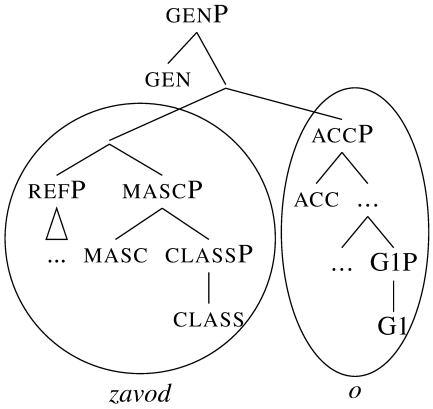
Both of these entries will be able to lexicalise the NOM and ACC cases without any suffixes. The difference between them shows up in the genitive. In order to lexicalise the genitive, both derivations backtrack down to the #P cycle, which originally required no movement, see (26h). However, since that led to a dead end, evacuation movements now start moving the MASCP left branch up the tree, from below #, leading to an intermediate representation such as the one in (50a) for the noun (the demonstrative would have a different structure of the Spec). When the GEN feature is added, we get (50b). There is no match, so evacuate closest performs a

subextraction of REFP, see (50c). As a result, the noun has the oblique ending *-a*, see (50d).

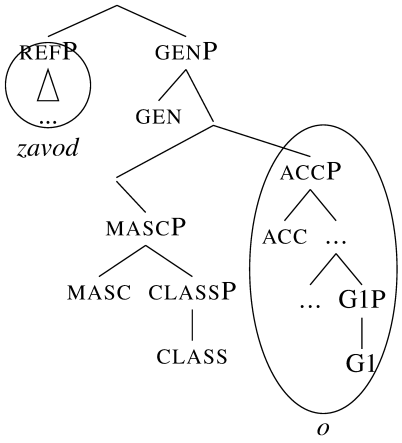
(50) a. The structure of ACC



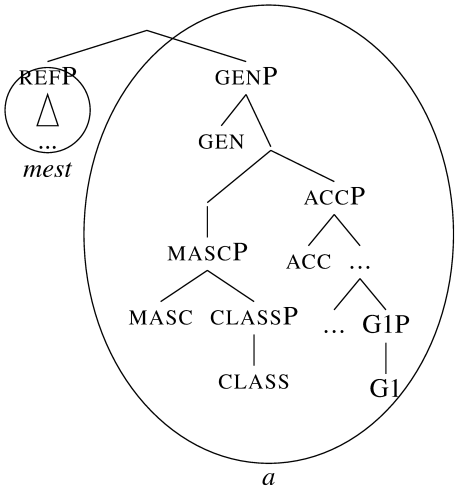

b. Merge GEN




c. Evacuate closest




d. Lexicalise

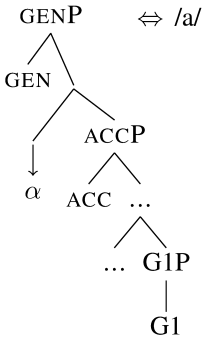



Importantly, since this subextraction is unavailable for the demonstrative, *-a* will only be available for nouns. As a result, the masculine demonstrative will need to undergo further backtracking, which has been described in (20)-(22) for the neuter.

At this point, we need to remove the final hurdle. When we compare the trees in (47d) and (50d), we see that the left branches in the trees lexicalised by *-a* differ in the presence/absence of the feature MASC. This creates an issue for matching. If we specified *-a* as realising the structure circled in (50d), we would not expect it to match in (47d), because the structures are not in a containment relation.

We can, however, notice that in both cases, the left branch can be realised by the lexical

item α, which we had to introduce for independent reasons in (38). We can now use this fact to propose a lexical entry for *-a*, which has a pointer to α, as in (51). This lexical entry makes sure that *-a* appears both with masculines and neuters, which is what the facts require.

(51) 

[tree 38]

At this point, we have accounted for why the derivation of masculine nouns and denonstratives diverge in the oblique cases, which was one of the initial puzzles. We believe that the readers will be able to complete the derivations of the locative, dative and instrumental, which only differ in the number of case features in an ending like the one in (51).

**5. Conclusions**

This paper proposes an analysis of the singular demonstrative declension of Russian in comparison to nouns. These paradigms offer a curious mix of some well-known as well as unusual properties. The well-known aspect of the paradigms is that neuter and masculines demonstratived are syncretic in all oblique cases. Traditionally, this is understood as a classical markedness effect, where "marked" cases give rise to the neutralisation of a gender opposition. However, when we compare nouns and demonstratives, we find the complete opposite pattern: they are identical in structural cases, but different in oblique cases.

The goal of this paper was to show that these complex patterns can be modelled within a particular version of Nanosyntax based on the subextracting lexicalisation algorithm (Starke

2022, Wiland 2019). This paper argued that this model gives us sufficient power to generate the puzzling patterns, and correctly derive some additional observations, including the identity between the stem marker *-o* used in all gender, and the neuter NOM/ACC case marker *-o*, and the identity between the oblique feminine endings on the demonstratives and Declension III nouns.

**References**


Arregi, K., & Nevins, A. (2012). *Morphotactics: Basque auxiliaries and the structure of spellout.* Dordrecht: Springer.

Blix, Hagen. (2022). Interface Legibility and Nominal Classification: A Nanosyntactic Account of Kipsigis Singulatives. *Glossa: a journal of general linguistics*, *7*(1).

Bobaljik, J. D. (2002). Syncretism without paradigms: Remarks on Williams 1981, 1994. *Yearbook of morphology 2001*, 53-85.

Bonet, Eulàlia. (1991). *Morphology after syntax. Pronominal clitics in Romance.* PhD thesis, MIT.

Caha, Pavel. (2009). *The nanosyntax of case*. PhD thesis, Tromso.

Caha, Pavel. (2023). The layered structure of concord. In *The many facets of agreement*, edited by Pavel Caha, Pavel Rudnev, Svetlana Toldova, 3-36. München: LINCOM.

Chomsky, Noam. (1995). *The minimalist program*. MIT press.

Chvany, Cathrine. V. (1982). Hierarchies in the Russian case system: For NAGPDI, against NGDAIP. *Russian language journal/Русский язык*, *36*(125), 133-147.

Cinque, Guglielmo. (2005). Deriving Greenberg's Universal 20 and its exceptions. *Linguistic inquiry*, *36*(3), 315-332.

Cinque Guglielmo and Rizzi Luigi. (2010). The cartography of syntactic structures. In *The Oxford Handbook of Linguistic Analysis*, ed. B Heine, H Narrog, pp. 51–65. New York: Oxford Univ. Press

Cortiula, Maria. (2023). *The Nanosyntax of Friulian Verbs*. PhD thesis, Masaryk University, Brno.

Croft, W. (2002). *Typology and universals*. Cambridge University Press.

Halle, Morris and Marantz, Alec. (1993). Distributed morphology and the pieces of inflection. In Kenneth Hale & Samuel J. Keyser (eds.) *The view from Building 20*, 111-176. Cambridge, MA: MIT Press.

Halle, Morris and Matushansky, Ora. (2006). The morphophonology of Russian adjectival inflection. *Linguistic inquiry 37*: 351-404.

Harley, Heidi. (2008). When is a syncretism more than a syncretism? Impoverishment, metasyncretism, and underspecification. In Daniel Harbour, David Adger & Susan Béjar (eds.) *Phi theory: Phi-features across modules and interfaces*, 251-294. Oxford: Oxford University Press.

Harley, Heidi and Ritter, Elisabeth. (2002). Person and number in pronouns: A feature-geometric analysis. *Language*, 482-526.

Jakobson, Roman. (1932). Zur Struktur des russischen Verbums. In *Charisteria Gvilelmo Mathesio qvinqvagenario a discipulis et Circuli Lingvistici Pragensis sodalibus oblata*, Praha.

Kayne, Richard S. (1994). *The antisymmetry of syntax*. Cambridge, MA: MIT press.

Lamontagne, Greg and Travis, Lisa. (1987) The syntax of adjacency. In: *Proceedings of the West Coast conference on formal linguistics*. Vol. 6. Stanford, CA: Stanford Linguistics Association.

Picallo, S. (1991). Nominals and nominalization in Catalan. *Probus* 3, 279–316.

Scheer, Tobias. (2004). *A lateral theory of phonology: What is CVCV and why should it be?* Mouton de Gruyter.

Spathas, Gregory & Sudo, Yasutada. (2020). Gender on animal nouns in Greek. *Catalan journal*

*of linguistics* 19, 25-48.

Starke, Michal. (2018). Complex left branches, spellout, and prefixes. In *Exploring nanosyntax*, edited by Lena Baunaz, Karen De Clercq, Liliane Haegeman, Eric Lander, 239– 249. Oxford: Oxford University Press.

Starke, Michal. 2022. Nanoseminar. Online lecture series at Masaryk university, Brno.

Timberlake, Alan. (2004). *A reference grammar of Russian*. Cambridge: Cambridge University Press.

Wiland, Bartosz. (2019). *The spell-out algorithm and lexicalization patterns: Slavic verbs and complementizers*. Berlin: Language Science Press.