Modeling declensions without declension features. The case of Russian

Pavel Caha
Masarykova univerzita

Abstract

This paper presents an analysis of the Russian declension in Nanosyntax (Starke 2009, 2018). The analysis has two theoretically important aspects. First, it makes no reference to language-particular declension features. This allows one to maintain the idea that morphosyntactic features are drawn from a set provided by the UG, i.e., language invariant. The analysis also does not use contextual rules. In order to correctly pair the right ending with a particular root, the analysis only relies on specifying each marker for the features it spells out. The correct pairing of roots and affixes falls out from such a specification and the Nanosyntax model of spellout.

Keywords: Nanosyntax; declension; Russian; morphology; syntax

1 The arbitrary nature of declensions

Let me start by introducing two Russian nouns, namely 'snowstorm' and 'week.' Their roots are given in (1).

(1) Russian
   a. *metel'-* 'snowstorm'
   b. *nedel'-* 'week'

What is interesting about the two nouns is that "both of these are feminine, and both have a stem ending in a soft consonant, but they decline differently" (Corbett 1982). Their differing declensions are illustrated in (2) and (3), where I list the nominative and the instrumental singular respectively.

(2) The nominative
   a. metel'-**Ø** (snowstorm)
   b. nedel'-**a** (week)

(3) The instrumental
   a. metel'-**ju** (snowstorm)
   b. nedel'-**ej** (week)

The issue is that there is no obvious difference between the two nouns (be it their phonology, morphology or semantics) that would allow us to explain why the root *nedel'-* takes one ending in INS, while *metel'-* takes a different one. More specifically, to the extent that there are semantic or other differences between the two nouns (e.g., 'week' denotes a time interval with a fixed duration, while 'snowstorm' does not), it is not clear how such a difference could be generalized

to reliably distinguish all relevant cases. Therefore, the conclusion that many linguists draw is that the difference in the endings is arbitrary, and has to be somehow stated on top of the information that the roots have a certain phonology, particular formal properties (gender) and refer to a particular concept.

The way this is usually encoded is by assigning roots into declensions with arbitrary labels. I shall use the labels I-IV, though the labels differ in different accounts. In the numbering I adopt (Corbett 1982), the noun *nedel'-* belongs to Declension II, while the noun *metel-'* belongs to Declension III. The singular paradigms of the two declensions are depicted side by side in (4), using the roots *gub-* 'lip' and *tetraď-* 'notebook' as the representatives of each declension.

(4) The singular of declensions II and III (Timberlake 2004)

	lip II (FEM)	notebook III (FEM)
NOM	gub-a	tetraď-Ø
ACC	gub-u	tetraď-Ø
GEN	gub-y	tetraď-i
LOC	gub-e	tetraď-i
DAT	gub-e	tetraď-i
INS	gub-oj	tetraď-ju

The point of the table is to show that the difference in NOM/INS is not an isolated fact; it is a pervasive property of the paradigm that the two classes of nouns combine with different endings throughout the declension.[1]

As already mentioned, Russian has in total four declensions. The remaining two are given below in (5). These show a difference in the nominative and in the accusative, but the oblique cases are the same, as the shading indicates. This is interesting in that this seems to suggest that declension class differences may be present in some cases, but obliterated in others, and any account of declension classes should take this into account.

(5) Declensions I and IV (Timberlake 2004)

[1]One thing that is worth pointing out is the issue of phonological alternations. In Russian, Declension II shows a regular alternation between the so-called hard stems and soft stems. They differ in the instrumental and in the genitive.

In the instrumental, soft stems have *-ej* and hard stems exhibit *-oj*. In other words, the INS ending *-ej* in *nedel'-ej* (recall (3b)) is a phonologically conditioned allomorph of *-oj* in *gub-oj*, see (4). I shall treat this as a regular phonological process (orthogonal to the issue of declension).

In the genitive, we find an alternation between (hard-stem) *gub-y* 'lip, GEN' and (soft-stem) *nedel'-i* 'week, GEN'. This could lead one to suspect that in the genitive in Table (4), there is actually syncretism between the two declensions, because the difference between *-y* and *-i* could be attributed to a regular phonological process.

However, this analysis turns out to be incorrect in this particular case for an independent reason, which is that the genitive ending of Declension II interacts differently with stress placement than the Declension III ending (Melvold 1989:21). In particular, the ending in Declension II is accented, while the one in Declension III is non-accented. I will therefore treat these two markers as independent endings.

	factory I (MASC)	place IV (NEUT)
NOM	zavod-Ø	mest-o
ACC	zavod-Ø	mest-o
GEN	zavod-a	mest-a
LOC	zavod-e	mest-e
DAT	zavod-u	mest-u
INS	zavod-om	mest-om

Syncretism among declension classes is also characteristic of the plural, which I shall discuss in Section 5 (after the analysis of the singular is in place).

1.1 Declension features

Let me now turn to the question of how differences in declension class are encoded theoretically. To have a concrete example to work with, let me come back to the instrumental. The specific question to be addressed is what forces the grammar to generate the words in (6), rather than the hypothetical words in (7), where the endings are swapped.

(6) The correct instrumental
   a. metel'-**ju** (snowstorm)
   b. nedel'-**ej** (week)

(7) The wrong instrumental[2]
   a. *metel'-**ej** (snowstorm)
   b. *nedel'-**ju** (week)

The standard approach to this issue is to assign an arbitrary index II and III to the roots, see (8).

(8) Roots carry declension features
   a. nedel' ⇔ [FEM, II]
   b. metel' ⇔ [FEM, III]

The endings are then specified for the case they pronounce (the instrumental), and they contain (in addition) a contextual restriction. The contextual restriction says that the relevant endings can only be used in the context of a Class II noun (*-ej* in (9a)) or a Class III noun (*-ju* in (9b)).

(9) Endings are specified for the context of insertion
   a. **ej** ⇔ INS / _ [II]
   b. **ju** ⇔ INS / _ [III]

The (arbitrary) features on the roots and the contextual restriction of the endings are the two tools that current theories use to generate the correct pairings of roots and affixes in (6), and, at the same time, rule out the incorrect pairings in (7) (see, e.g., Halle 1997, Halle & Vaux 1998).

An analysis along these lines has been applied to Russian by Müller (2004) (cf. Alexiadou & Müller 2008). What Müller proposes in addition is that declension classes on nouns are not primitive entities, but decompose into two equipolent features, namely [+/-α] and [+/-β], see (10).

[2]The form *nedel'-ju* would be acceptable as an accusative, but not as an instrumental.

(10) Müller (2004)

a. I: $[+\alpha, -\beta]$
b. II: $[-\alpha, +\beta]$
c. III: $[-\alpha, -\beta]$
d. IV: $[+\alpha, +\beta]$

The reason why Müller proposes to decompose the Class features I-IV is that this allows him to capture syncretism in between classes; recall, for instance, table (5) showing that declensions I and IV have the same endings in most cases. The decomposition in (10) is intended to capture this by proposing that the two declensions share the feature $[+\alpha]$. An ending specified for this feature as its contextual restriction will naturally appear in both classes that have this feature.

The important point for now is that underlying the proposal is the very same logic as described at the beginning: tag the roots with arbitrary class features and make the endings sensitive to those features. Müller makes this crystal clear when he writes that "[i]nflection class features are arbitrary and irreducible by definition; this is reflected in the labels. Still, it is worth emphasizing that the features $[+/-\alpha]$, $[+/-\beta]$ are no more arbitrary than standardly adopted features like [class I], [class II]."[3]

1.2 Arbitrary features vs. UG

Arbitrary class features are theoretically interesting entities, because they have implications for the architecture of grammar. Taking first a narrowly morphological perspective, a traditional architectural question in morphology is which words should be decomposed into component parts, which should not be, and how we can tell. For instance, should the past tense of *go* – namely *went* – be decomposed into an irregular root *wen-* and a non-productive past tense *-t*? Should the plural *children* be decomposed into an irregular root *child(r)-* and an irregular plural *-(r)en*?

Some approaches argue for full decomposition in such cases (see, e.g., Stockall & Marantz 2006). Others propose that morphological decomposition is justified only when we can productively combine the pieces together. Blevins (2016:69-73) mentions specifically Russian declension as a case where it is impossible to combine roots/stems productively with the correct ending. According to Blevins, the need to introduce arbitrary declension features is nothing but a blatant admission of this fact. According to him, such diacritic features cannot be treated like regular properties of lexical items analogous to 'plural'. Rather, they are

[3] An anonymous reviewer points out that an alternative way of looking at α and β is as stand-in features for a more substantive analysis. Under this interpretation, α and β are contentful features, we just do not know as yet which ones these are. A proposal along these lines is presented in Privizentseva (2020), who argues that the identity of the feature [-α] is [+FEM], because inanimate nouns of Declensions II and III are overwhelmingly feminine.

However, for such an analysis to go all the way, one would have to find a similar interpretation for the feature [β], which is the crucial one to explain away. This feature is the one that distinguishes Declensions II and III, i.e., our two nouns *nedel'-* 'week' and *metel'-* 'snowstorm' in (2). The point made by both Corbett and Müller is that it is very unlikely that we will ever be able to link this difference to some grammatical feature. The feature [+/-β] thus remains an arbitrary feature also in Privizentseva's (2020) work, as seems unavoidable.

"assembly instructions" in disguise, masquerading as properties of lexical items.

An even broader issue can be linked to the ontological status of such features. For example, Chomsky (2001:1,4) hypothesizes that the faculty of language "FL appears to be a species property, close to uniform across a broad range. It has a genetically-determined initial state S_0, which determines the possible states it can assume. [...] S_0 determines the set {F} of properties ("features") available for languages."

Clearly, arbitrary language-specific declension features like [I, II] or [+/-β] are not among those. So the need to postulate such features significantly weakens our theory of grammar: by postulating declension features, we implicitly claim that grammars do not only rely on the set of features determined by UG, but also on language-specific features. This in turn weakens the degree to which the "genetically-determined initial state S_0" actually restricts the shape of surface grammars.

The idea that UG relies on a language-invariant feature set has been recently advocated also in Cinque (2013:50-51). Specifically, Cinque stresses "the observation that of all the concepts and distinctions that populate our system of thought only a fragment receives a grammatical encoding in the languages of the world, arguably the same in all languages. [...] Most cognitive concepts and distinctions do not find any such encoding. [...] Verbal projections in clauses grammatically encode (through affixes, particles, auxiliaries, etc.) distinctions relating to the external and internal temporal constituency of events (tense and aspect) and the speaker's attitude toward the truth of the proposition (mood), but they are never found to grammatically encode such human cognitive universals as "shame", "mourning", "sexual taboos", etc." Such facts make declension features even more remarkable: why is it that grammars do not allow for features related to cognitive universals, while allowing for a feature like [+/-β]?

The positions entertained by Chomsky and Cinque are, of course, controversial, and some of that controversy relates precisely to language-particular categories like the Russian declension class. For example, Haspelmath (2007:119) says: "descriptive linguists still have no choice but to adopt the Boasian approach of positing special language-particular categories for each language. Theorists often resist it, but the crosslinguistic evidence is not converging on a smallish set of possibly innate categories." Arbitrary declension classes I, II, etc. are clearly an instance of language particular categories that need to be posited in our descriptions. They are therefore a relevant piece in the puzzle that we must consider in order to be able to address the larger question: what is the relationship between language particular classes and the invariant set of features? Is the existence of declensions compatible with a universal set of features at all?

In this article, I put forth one specific proposal as to how we can eliminate declension features while still accounting for the existence of declension classes. If this attempt turns out to be successful, it represents a small (but meaningful) step on the path towards a restrictive (or even contentful) theory of UG.

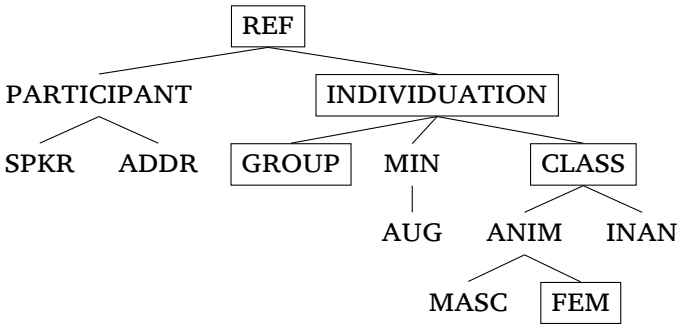
2 An alternative to declension features

The model I shall explore here has two major components. The first component is the morphosyntactic structure underlying nouns. What features are present? How are they arranged? I tackle this in Section 2.1. The second component concerns spellout. How do the underlying feature structures map onto specific exponents? This issue will be picked up in 2.2.

2.1 Features and hierarchies

The features I adopt in my analysis come from the cross-linguistic study of pronominal systems by Harley & Ritter (2002). The paper studied the formal distinctions expressed by pronouns in 110 languages. The study finds that the distinctions between various persons, numbers and genders are expressible in terms of the feature hierarchy depicted in (11).

(11)



In my proposal, I am going to rely on a subset of these features to model the Russian declension. The specific features I shall use are placed in rectangles. The main reason for only using features identified in an independent cross-linguistic study is to make sure that I avoid using features that are Russian specific. This is important in order to reach the desideratum, which is showing that it is possible to only rely on universal features and still account for (language-specific) declension classes.

The second reason for using the features in (11) is that nouns often mirror the grammatical categories of 3rd person pronouns (and *vice versa*). For instance, looking at Russian 3rd person pronouns, we see that they show very much the same distinctions as nouns: they fall into three genders (like nouns), they have a singular and plural (like nouns) and they exhibit the same set of case distinctions. Under the approach I adopt here, nouns and pronouns thus have the same grammatical features and differ mainly in that nouns have an encyclopedic content (concept) associated to them, while pronouns lack it.

The features used by Harley and Ritter are privative. Their presence indicates the presence of a particular grammatical meaning, the absence of the feature indicates the absence of that meaning.

The topmost node of the tree is labelled REF. This feature encodes the fact that pronouns are referential expressions: they always include this feature. I will assume that this feature is present also in nouns, and it is this feature what distinguishes nouns from other lexical categories. This is reminiscent of Baker (2003), who proposes that nouns are distinguished by bearing a referential index.

The features below REF are organized into two branches. On the left branch, there are person/number features. When present, they distinguish various types of first and second persons from each other. Their absence is characteristic of 3rd person pronouns. Since I am not interested here in 1st/2nd person expressions, these features are irrelevant for my analysis and I will not use them.

Moving on to the right branch of the tree (11), we find here information related to number and gender. Let me begin by number. The INDIVIDUATION node (IND) is a default number node. On its own, it yields singular interpretation. When further number features are present (these are below IND on the left branch), they influence the number interpretation. For example, when the feature GROUP is present, this yields plural interpretation. The features AUGmented and MINimal are used for special numbers (dual, paucal). These two are not relevant for my analysis and I will not use them.

The right-hand branch under the INDIVIDUATION node encodes gender. Russian has three genders (neuter, feminine, masculine) and it has an animate/inanimate distinction for masculine nouns. I will set animate nouns aside in my analysis and focus only on inanimate nouns.

It turns out that in order to account for inanimate declension classes in Russian, it is enough to distinguish only two genders, where feminine nouns form one class, and the neuter and masculine nouns belong to a second class. This is not to say that the Russian grammar does not distinguish between masculine nouns and neuter nouns in terms of features, but it is to say that the declension endings are oblivious to this distinction (as we shall see in the course of the discussion). In order to use as few features as possible, I will only use two of the gender features proposed by Harley and Ritter. Specifically, nouns of all genders will be given the feature CLASS, while feminine nouns add the feature FEM.[4]

In (12) and (13), I summarize the relevant features. In (12), I give my assumptions about number, in (13), I list the minimal amount of gender features that are necessary for my analysis.

(12) Number
- a. [IND] = singular
- b. [IND + GROUP] = plural

(13) Gender
- a. [CLASS] = masculine and neuter
- b. [CLASS + FEM] = feminine

In addition, I will need to rely on a number of case features. For the time being, I will use a single case feature K as a stand in for the nominative case, see (14a). I shall introduce additional case features in section 3.2.

(14) Case
- a. [K] = nominative
- b. [K + ???] = other cases, to be determined

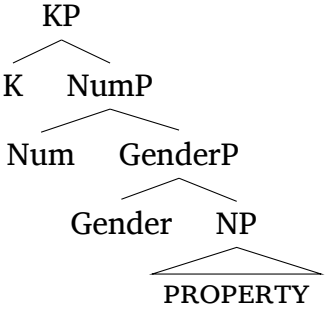
Following Harley & Ritter (2002), I shall be further assuming that the features laid out in (12) to (14) are structured. As for the specifics of such a structure, I

[4]The analysis is compatible with Russian having three genders. I return to this in Section 6.

am going to diverge from their proposal in several ways. Before I highlight the differences, let me say the following: the main reason why I am going to use a different structure is that this alternative structure will allow me to explain how declension classes work, while at the same time using an independently proposed theory of spellout. It is, however, possible that declension classes can also be understood without altering the structure. As to whether that is possible or not remains to be seen: the point I want to make is that there is a way of structuring these features such that an account of declension emerges without the need to introduce declension features.

Another general point is that while the structures I propose are motivated by the desire to provide a coherent proposal for declension classes, the structures are not *ad hoc*. They have both an internal logic and a relationship to the existing literature on the topic. The essence of my proposal is depicted in (15). At the bottom of this structure, we have the NP node that traditionally hosts the noun root. The NP node denotes a property that the referent of the noun has. Above the noun, we have the projections of gender, number and case respectively.

(15)

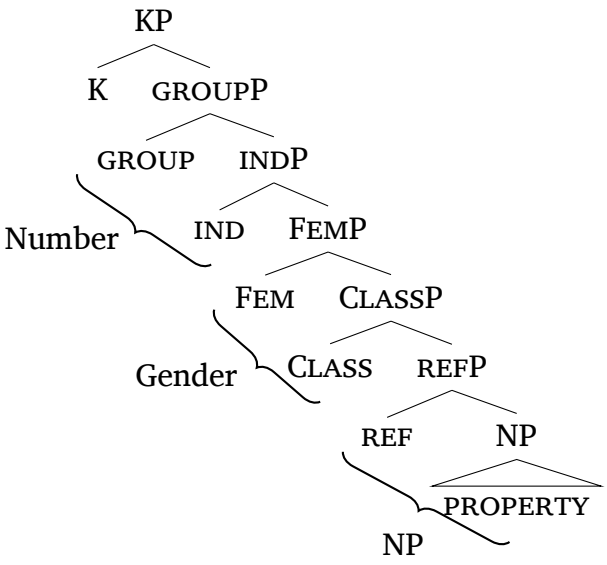


The lower part of the structure – with number and gender above the NP – can be traced back at least to work by Picallo (1991), and a version of it is also used more recently in Kramer (2015:43). Although the proposals differ in important details, both Picallo and Kramer converge on the idea that gender corresponds to a head that is syntactically independent of the noun root. Both approaches also consider gender to be a separate head from number, and both approaches place the number head higher than gender.

Similarly, whenever K is proposed to exist as an independent head, there is agreement that this head is high up in the noun phrase, definitely above number and quite likely also above D (which itself is higher than Num, see, e.g., Bittner & Hale 1996, Bayer et al. 2001, Kloudová 2020).

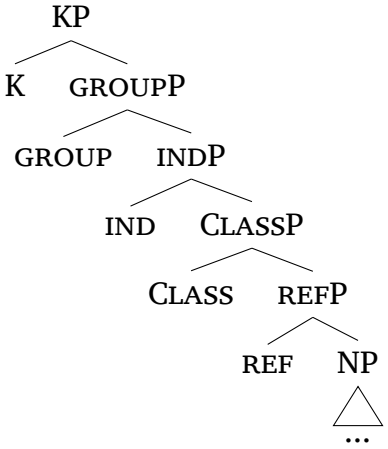
Now a specific property of my proposal is going to be that rather than locating gender/number features *inside* the heads in (15), I am going to rely on an approach where each such feature is an independent head. This leads me to adopt the tree in (16) as a 'split-XP' version of (15). The features that we see in (16) are exactly the same features as used by Harley and Ritter, and they are abstractly grouped into sets that represent the traditional division of the NP into its 'lexical' part (at the bottom), and then gender, number and case on top.

(16) The full structure (FEM.PL)

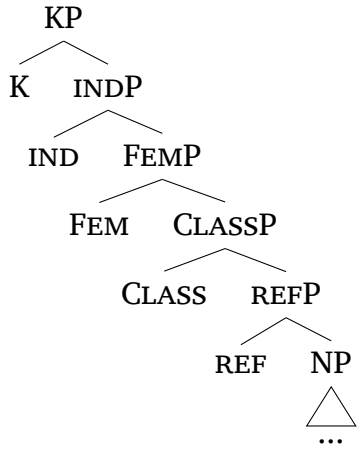


As in the original proposal by Harley & Ritter (2002), the features are privative. This means that they can be present or absent. The tree in (16) contains the maximal number of features, and corresponds therefore to the feminine plural. However, some of the features may be missing. For example, if the FEM head is missing, we get a non-feminine noun, i.e., either a masculine noun or a neuter plural, see (17). If the GROUP feature is missing, we get a feminine singular structure, as in (18).

(17) Masculine/neuter plural



(18) Feminine singular



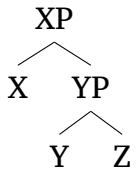
2.2 Phrasal spellout

In addition to the structures explored above, I will rely on phrasal spellout as the second ingredient that allows for the elimination of declension features. The specific approach to phrasal spellout I adopt has been developed within the Nanosyntax framework (Starke 2009, 2018, cf. Baunaz & Lander 2018, Taraldsen 2019, Caha 2020). This section describes my assumptions about spellout in detail. Since I motivate the spellout principles independently of the empirical topic of declension, the section is a digression on the path. Therefore, readers who are familiar with Nanosyntax should proceed directly to Section 3.

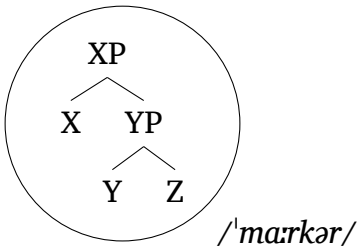
I start by introducing the basic concepts. Phrasal spellout refers to a situation

where a full syntactic phrase containing multiple terminals – such as the one in (19) – is pronounced by a single lexical item. This can be depicted as in (20). The circle indicates that the relevant phrase (containing the features X, Y and Z) has been pronounced by a single marker inserted at the top node. The phonology of the marker is placed below the circle.

(19)



(20)



Let me now describe in more detail how phrasal spellout works. To have a concrete case to work with, I shall look at the three singular-plural pairs given in (21). Accounting for these three pairs will allow me to showcase all the tools we shall need for the analysis of the Russian declension.

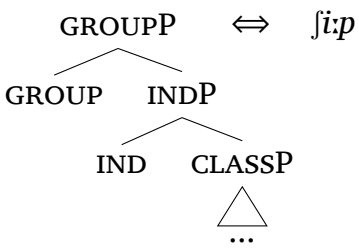
(21) English plural

a. sheep — sheep (syncretism)
b. mouse — mice (suppletion)
c. cat — cat-s (affixation)

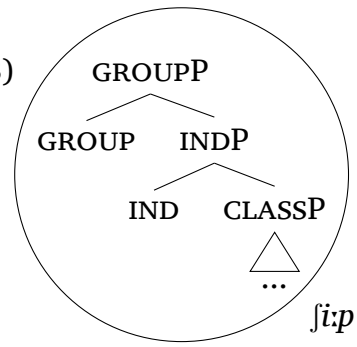
Each pair features a different type of morphological relation between the singular and plural. In (21a), we find syncretism. In (21b), we have suppletion. Finally, in (21c), the plural has an extra affix.

Let me now show how the three different cases are treated in Nanosyntax, beginning with (21a) (*sheep*). The starting point of my analysis is the observation that this noun has no ending in the plural. The lexical entry for this item is therefore construed in a way that it pronounces all the features contained inside the plural. What are these features? According to the proposal (16), the plural structure has two number features, IND and GROUP. I am further assuming that the noun *sheep* has a CLASSP below the number features, but no FEM. The lexical entry therefore looks as in (22).

(22)



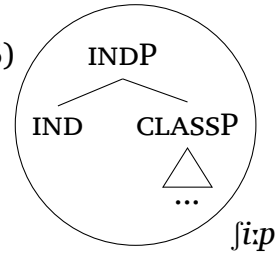
(23)

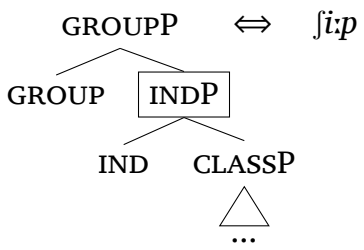


As a result, when syntax builds the relevant constituent corresponding to the plural, as in (23), this constituent can be pronounced using the lexical entry of *sheep*. The specific principle which determines when a lexical item matches a syntactic structure is called the Superset Principle, see (24).

(24) *Superset Principle* (Starke 2009)
A lexically stored tree L matches a syntactic node S iff L contains the syntactic tree dominated by S as a subtree.

The principle says that matching obtains when the tree built by syntax is fully contained inside the tree stored in the lexical item. This is obviously the case in (23). However, matching also obtains when syntax builds only the tree associated to the singular, such as the one in (25).

(25)


(26)


The tree in (25) is also contained in the lexically stored tree for *sheep*. The entry is repeated for convenience in (26), and the relevant constituent is marked by a rectangle in the lexical entry. As a result, the singular structure is also spelled out as *sheep* and we get the same spellout for singular as we did for plural. This is the intended consequence of matching based on the Superset Principle.

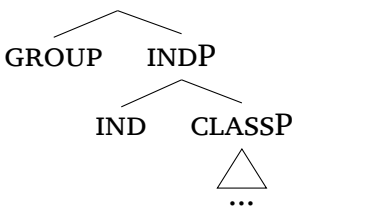
In this paper, I shall be using the so-called lexicalization tables to encode the result of spellout. A lexicalization table for the singular and for the plural of *sheep* looks as in (27).

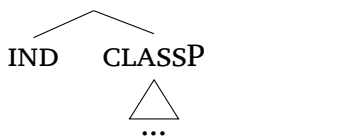
(27)

	CLASSP	IND	GROUP
'sheep' SG	ʃi:p		
'sheep' PL	ʃi:p		

In the header of the table, the grammatical features are given. The rows depict how they are spelled out. The first row of the table depicts the spellout of the singular. The singular has only two components: CLASSP and IND. The shading indicates that both of these are spelled out by *sheep*. The second row depicts the plural, which has three components. All three of them are once again pronounced by *sheep*, as indicated by the shading.

Let me now turn to the pair *mouse* – *mice*. What we need for such cases is to have a pair of lexical entries as in (28) and (29).

(28)


(29)


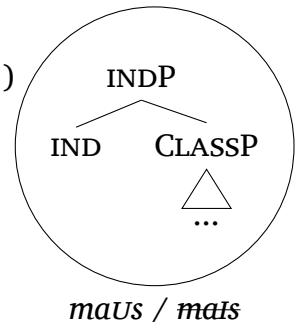
When syntax builds the tree corresponding to the singular, both lexical items match, because they both contain INDP. When two lexical items match, they com-

pete for insertion. The winner of the competition is determined by the Elsewhere condition (Kiparsky 1973). I give the Elsewhere Condition in (30).

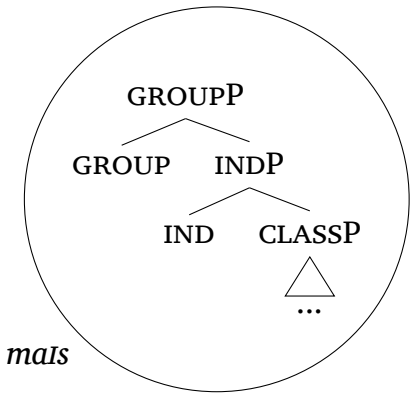
(30) *The Elsewhere Condition*:
When two entries can spell out a given node, the more specific entry wins. The more specific entry is the one which has fewer features.

In the case at hand, this means that *mouse* wins over *mice* in the singular. This is depicted in (31) under the circle, where *mice* (as a losing candidate) is placed under a strike-through. The spellout by *mouse* wins, and the singular therefore comes out correctly as *mouse*.

(31)



(32)



In the plural, see (32), only *mice* is a candidate. *Mouse* does not contain GROUP inside its specification, and it therefore cannot spell out the relevant phrase.The lexicalization table for the pair *mouse* – *mice* therefore looks as in (33). The difference in the shading gradient highlights the fact that the singular and plural are spelled out by different lexical entries.

(33)

	CLASSP	IND	GROUP
'mouse' SG	maʊs		
'mouse' PL	maɪs		

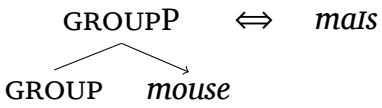
When I come back to the Russian declension, I shall be mainly using lexicalization tables like the one in (33). However, I want to make it clear that these tables arise as a result of a spellout algorithm that operates over trees (rather than stretches of cells).

Before we move on, let me also clarify that even though the choice between *mouse* and *mice* may be seen as the result of competition regulated by Elsewhere, insertion of roots is generally not subject to such competition. For instance, the root *mouse* does not compete with the root *sheep*. (Such a competition would wrongly lead to *mouse* blocking the singular use of *sheep*.) The spirit of this observation is compatible with The Elsewhere Condition: the condition says that when we have two appropriate candidates, we chose a better match. However, the selection of the root is generally guided by 'what we want to talk about:' if we want to talk about *sheep*, *mouse* is not an appropriate candidate, and therefore it does not even enter into the competition.

There are two ways how to allow for competition between *mouse* and *mice*, while blocking it for *mouse* and *sheep*. One option is to restrict competition to

items that are linked to the same concept. As a result, *mouse* and *mice* compete, because they are linked to the same concept, *mouse* and *sheep* do not. (Functional items, which are not linked to any concept, compete against each other, too.)

An alternative way of allowing for competition between two roots is to use the so-called pointers (see, e.g., Caha et al. 2019). Pointer is a device that explicitly tags one lexical item (*mice*) as a suppletive version of another item (*mouse*), see (34). The entry (34) almost literally says that *mice* is a plural version of *mouse*.

(34) 

Technically, this entry applies at GROUPP that contains the feature GROUP as one daughter, and the other daughter is spelled out by *mouse*. This interpretation of pointers presupposes the idea that spellout applies cyclically and proceeds bottom up. I state this in (35).

(35) *Cyclic phrasal spellout*
Spell out must successfully apply to the output of every Merge F operation. After successful spellout, the derivation may terminate, or proceed to another round of Merge F.

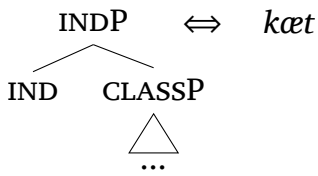
Cyclic spellout entails that as soon as a particular phrase is constructed by syntax (such as the singular structure in (31)), spellout must apply to the top node and find a matching entry for that node. We know that the spellout procedure does indeed find a matching item for (31), namely *mouse*. This, however, does not lead to an immediate pronunciation. If we want to add more features, the system does not send the item to PF as yet, but rather proceeds to further merging. For instance, if we want to derive the plural, the singular structure (31) is subject to further Merge. The GROUP feature is added, and spellout must apply again to the top node, as in (32). As the diagram shows, spellout is again successful in (32), because the lexical item *mice* provides a match.

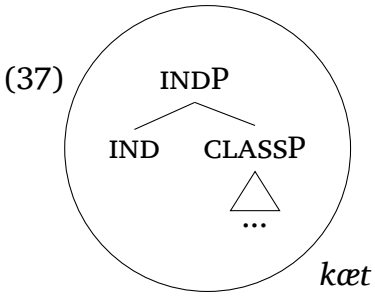
Matching is achieved regardless of whether we use the lexical item (28) or (34). In the latter case, *mice* applies because the plural feature is added to a constituent that has been spelled out by *mouse* at the previous cycle. While root suppletion remains outside of the domain of the current paper, the idea of pointers will resurface in the analysis of Russian plurals in Section 5. I shall be also relying on the Elsewhere Condition throughout the discussion.

It is important to note at this point that if we find a match for a particular node, this match (*mice*) replaces any items matching lower nodes, specifically the original spellout *mouse* obtained at the INDP mode. This is called Cyclic Override.

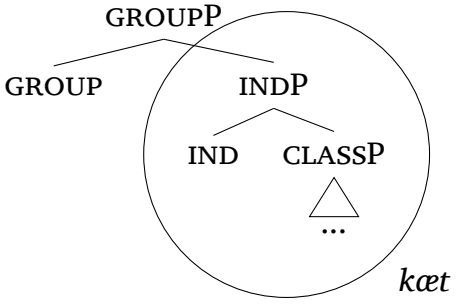
If there are no more features to be added, the derivation terminates and the spellout *mice* is sent to PF (else the derivation continues by adding more features).

Let me now turn to the singular-plural pair *cat* – *cat-s*. The main idea behind its analysis is that the root *cat* cannot spell out the full plural structure and that is why a suffix shows up in the plural. We can encode this by saying that the root is only specified as INDP, see (36). This lexical entry allows the root to spell out the relevant constituent on its own, see (37).

(36) 

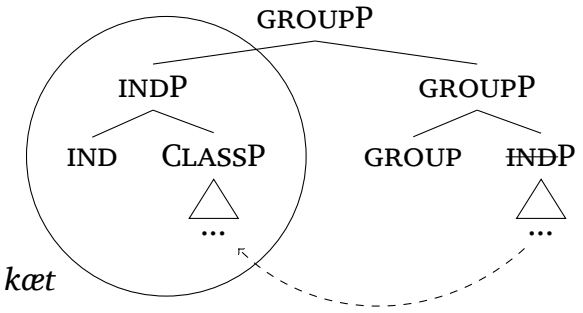
(37) 

Now recall that the spellout procedure is set up in a way that if no more features are added, the INDP (corresponding to the singular) would be pronounced as *cat*. If plural reading is intended, the GROUP feature must be added. The stage of the derivation where GROUPP is formed by merging GROUP to (37) is shown in (38). The root *cat*, as given in (36), does not contain this structure, and therefore, it cannot spell out the GROUPP in (38). The problem here is that the lexical entry for *cat* does not contain the feature GROUP and its projection GROUPP. These two nodes thus remain without spellout in (38). The plural marker *-s* will thus have to be introduced to spell them out. Its entry is as given in (39).

(38) 

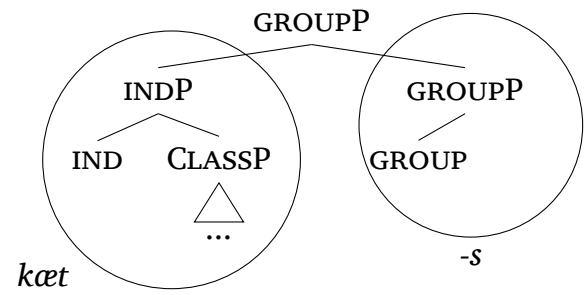
(39) 

However, even though the plural marker is specified for exactly those structural components of (38) that the root *cat* cannot spell out, the lexical item also cannot apply to the full phrasal node GROUPP in (38). This is because the entry for *-s* does not contain this node in its entirety (it contains only one of its daughters, namely the feature GROUP). Therefore, in order to achieve successful spellout at GROUPP, INDP has to evacuate, as in (40).

(40) 

After this movement (which renders the extracted INDP irrelevant for matching), the GROUPP can be spelled out by *-s* without any problems, see (41):

(41)



The final question we must address is why INDP moves. In Nanosyntax, this movement is triggered by the so-called spellout algorithm (see Starke 2018, c.f. De Clercq & Vanden Wyngaerd 2017, Caha et al. 2019, Vanden Wyngaerd et al. 2020). The algorithm is given below.

(42) Spellout Algorithm (based on Starke 2018)

a. Merge F and spell out.
b. If (a) fails, move the Spec of the complement and spell out.
c. If (b) fails, move the complement of F and spell out.

The algorithm at its core implements the idea of cyclic spellout: the initial clause says that we always Merge a feature F and we try to spell out the phrase we have created, see (42a). (42b,c) are new, and they tell us what happens when spellout fails, describing two types of rescue movements: namely Spec movement (42b) and complement movement (42c).

Consider how the algorithm applies to (38). Its first step, namely (42a), tells us that we should try to find a match for this phrase. However, recall that there is no matching item for the full phrase. As a consequence, the option (42b) is considered. The clause says that we should try to move the Spec of the complement. This will be relevant later, but in the case at hand in (38), the complement has no Spec. Therefore, this step is irrelevant and leads to no movement. As a consequence, we proceed to (42c). This clause says that we should try to move the whole complement of the newly added head, which is precisely the tree we saw in the tree (41). In this tree, the spell out of GROUPP succeeds and the derivation either terminates (if there are no more features), or it proceeds to the next cycle of Merge.

The lexicalization table for the pair *cat* – *cat-s* is given in (43). The new thing is that the spellout of the plural form has two morphemes, where each morpheme is responsible for pronouncing a part of the features.

(43)

	CLASSP	IND	GROUP
'cat' SG	kæt		
'cat' PL	kæt		*s*

To summarize: the model I assume (Nanosyntax) allows that morphemes – including root morphemes – spell out phrasal constituents. Different morphemes may be lexically specified for different numbers of features. For example, some roots (like *sheep* or *mice*) spell out the plural feature GROUP. As a result, they need no affix in the plural. Other roots (*cat*) are specified only as INDP and can-

not spell out the GROUP feature. These nouns need a special suffix in the plural (*-s*), which spells out GROUP. The division of labor between roots and suffixes is governed by the spellout algorithm, which implements the idea of cyclic spellout and spellout-driven movement. The result of the lexicalization procedure can be depicted as a lexicalization table where individual morphemes are placed under the features they spell out.

3 Declension as a function of root size

Recall now from (2) that one of the basic contrasts we are trying to derive is the one given in (44) and (45). What we see in (44) is that a Declension III noun 'snowstorm' has no audible ending in NOM.SG, while the Declension II noun 'week' has an *-a*. What we see in (45) is that in the instrumental, each noun simply has a different ending.

(44) The nominative
- a. metel'-**Ø** (snowstorm)
- b. nedel'-**a** (week)

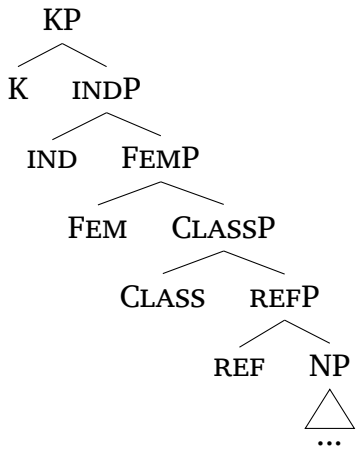
(45) The instrumental
- a. metel'-**ju** (snowstorm)
- b. nedel'-**ej** (week)

I will now show how we can derive this contrast without any reference to declension features. For clarity, the discussion is split across two separate sections. In section 3.1, I show how we can account for the difference in the nominative. In section 3.2, I show how we can extend this to account also for the instrumental.

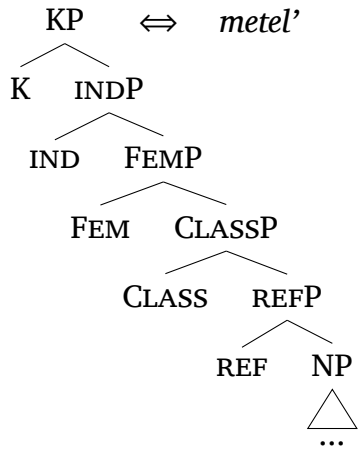
3.1 The nominative

Let me start my analysis of the example in (44) by drawing the structure of the nominative singular feminine, see (46). This structure is relevant for both nouns in (44), since both nouns are in the nominative singular, and both are feminine.

(46)



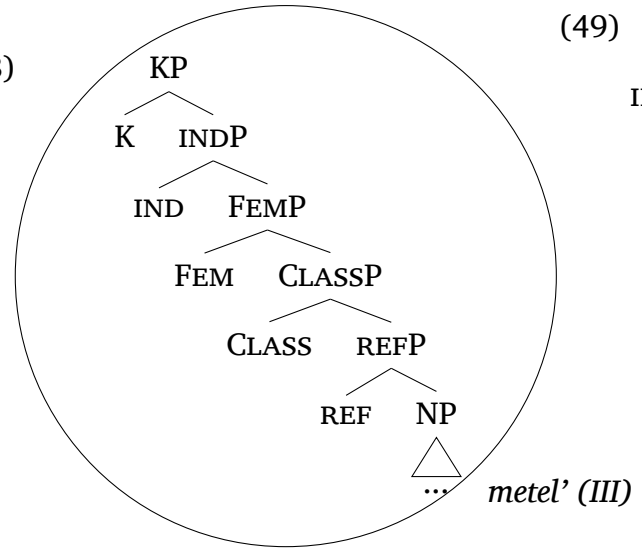
(47)



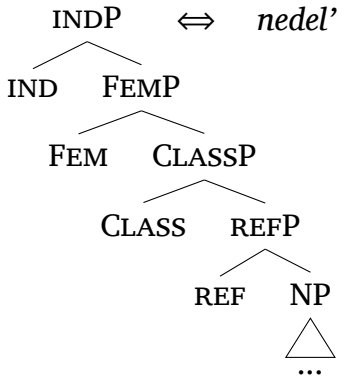
We know that the Declension III noun 'snowstorm' has no ending when spelling out a structure like (46). Therefore, I propose that what defines Declension III nouns is the fact that they are able to spell out such a structure in its entirety. This means that they are stored in the lexicon in the form shown in (47).

If that is so, Declension III nouns will be able to spell out all the features of the nominative without the need of any suffix, as in (48).

(48)



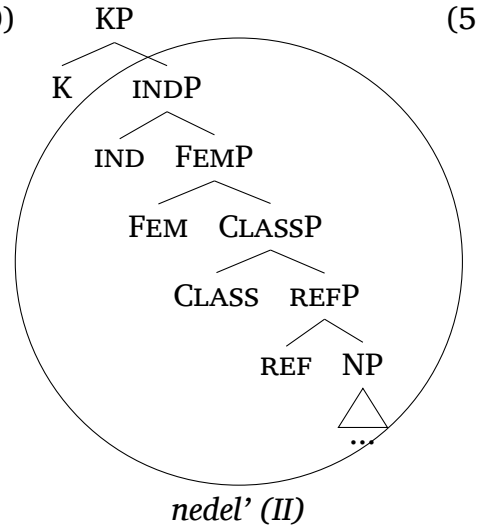
(49)



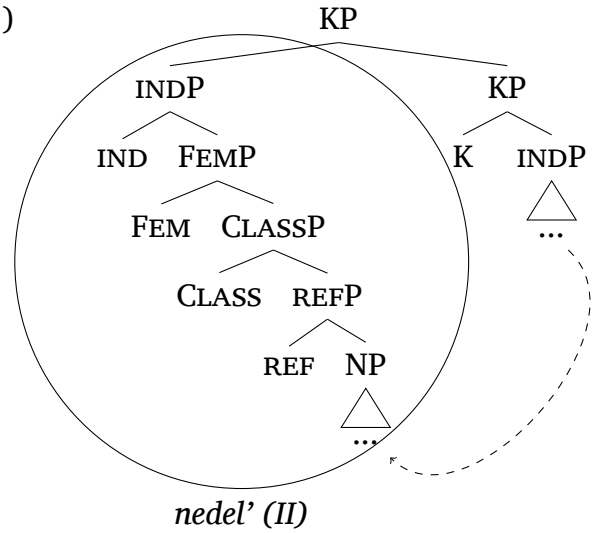
On the other hand, we know that the Declension II noun 'week' needs a suffix in the nominative. We can capture this by saying that the noun is stored in such a way that it does not spell out all the features of the nominative. We do not know exactly how many features the noun spells out, but let me assume for the start that it is lexically specified as spelling out the INDP projection, as in (49). (It could spell out also a lower projection. I will come back to this later and suggest that this is in fact the case.)

The crucial point is that with such an entry, Declension II nouns cannot spell out all the features. Specifically, they leave the K head and its projection without spellout, see (50).

(50)

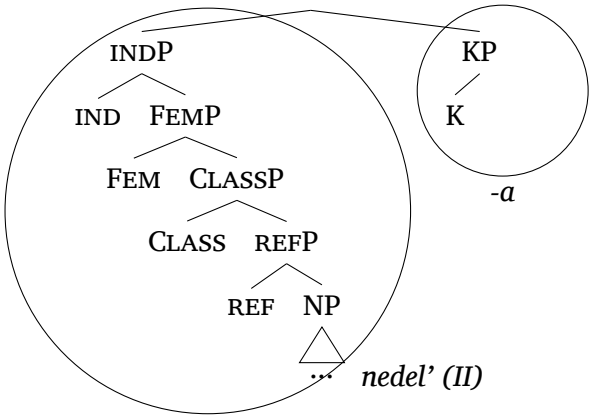


(51)



We saw during our discussion of the English plural that when a root cannot spell out a particular phrasal node, it moves to the left, as in (51). This movement makes it possible for the remnant KP to be spelled out by a suffix. The lexicalization is shown in (52), and the lexical entry of the suffix is as given in (53).

(52)



(53) KP ⇔ -*a*

This way, we capture the difference in the nominative case between Declension II and III without using declension-class features. The fact that the nominative of one class has a suffix where the other class has none is captured by associating each root with a structure of a different size. The rest follows from a general theory of spellout introduced in Section 2.2.

3.2 The instrumental

Let me now turn to the question as to whether (and how) we can extend the treatment of the nominative to also encompass the difference in the instrumental, see the basic dataset repeated again in (54) and (55).

(54) The nominative
- a. metel'-**Ø** (snowstorm)
- b. nedel'-**a** (week)

(55) The instrumental
- a. metel'-**ju**
- b. nedel'-**ej**

In order to be able to do so, I now have to say something about the representation of case in the grammar. My starting point is an observation about case syncretism in Russian (due to Chvany 1982, McCreight & Chvany 1991). The observation is that if we order the Russian cases in a particular way, then syncretism in case is restricted to contiguous stretches of paradigms. The particular ordering proposed by Chvany (1982) is NOM — ACC — GEN — LOC — DAT — INS. When cases are ordered in this way, as in the table (56), then all case syncretisms occupy contiguous regions within a paradigm.

(56) *Syncretism in Russian (McCreight & Chvany 1991)*

	window	teacher	both	book	hundred
	SG.	PL.	MASC.	SG.	
NOM	OKN-O	učitel-ja	dv-a	knig-a	st-o
ACC	OKN-O	UČITEL-EJ	dv-a	knig-u	st-o
GEN	okn-a	UČITEL-EJ	DV-UX	knig-y	ST-A
LOC	okn-e	učitel-jax	DV-UX	KNIG-E	ST-A
DAT	okn-u	učitel-am	dv-um	KNIG-E	ST-A
INS	okn-om	učitel-ami	dv-umja	knig-oj	ST-A

The columns of the table illustrate various types of syncretism. In the first column, we see the syncretism of NOM – ACC. In order for this syncretism to be

adjacent, NOM and ACC must not be separated by any other case. Similarly, the second column illustrates the fact that there are paradigms exhibitting an ACC–GEN syncretism. This requires that ACC – GEN must be next to each other. The same logic can be traced throughout the remaining columns with the result that the order must be as given in the table.

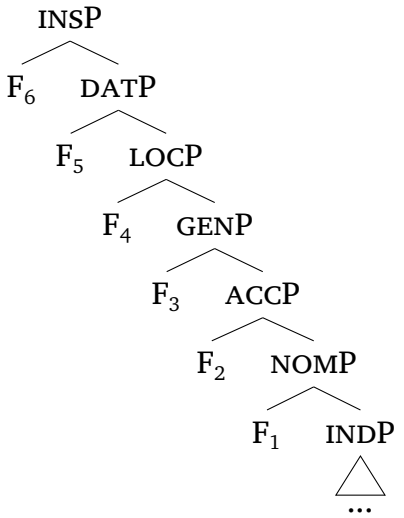
It has been argued in the literature (Caha 2009, Bobaljik 2012, De Clercq 2013) that *ABA constraints point to a type of feature decomposition that can be called 'cumulative.' The idea is that the number of features monotonically grows as we move from the nominative to the accusative and on. This is depicted in Table (57) (see Caha 2009, 2013, Starke 2017, McFadden 2018, Smith et al. 2019, Zompì 2019).

(57) Cumulative feature decomposition

CASE	FEATURES
NOM	F1
ACC	F1, F2
GEN	F1, F2, F3
LOC	F1, F2, F3, F4
DAT	F1, F2, F3, F4, F5
INS	F1, F2, F3, F4, F5, F6

If this type of decomposition is adopted (and regardless of the precise content of the case features), a *ABA pattern cannot be derived (though see Caha 2017, Bobaljik & Sauerland 2018, Vanden Wyngaerd et al. 2020 for other ways of representing the *ABA). In this article (following Caha 2009), I understand each of the case features in (57) as a separate syntactic head. Once this assumption is adopted, the structure of case in Russian looks as in (58).

(58)

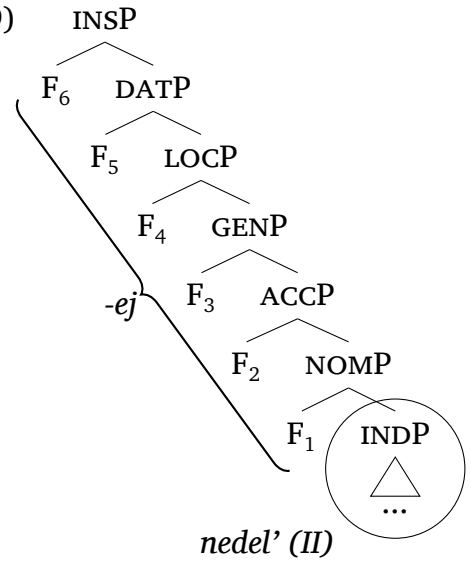


The tree encodes the proposal that when the feature F_1 is added on top of the extended NP (here represented by the singular number, INDP), we get the nominative case. When, in addition, we merge the feature F_2 on top of the nominative, the result is the accusative case, and so on.

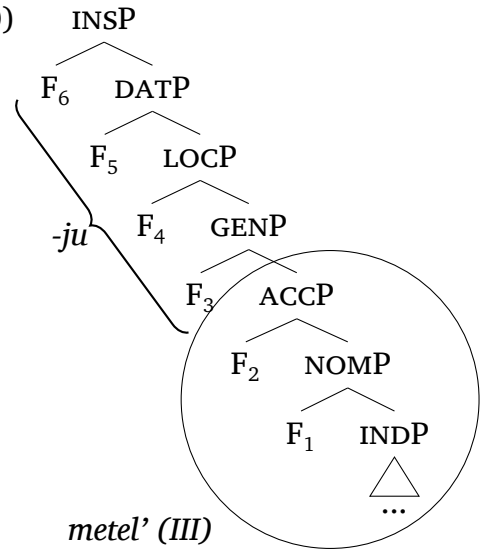
This proposal allows us first of all to capture all case syncretisms by a simple application of the Superset Principle (Caha 2009). Second, it also allows us to

extend our analysis of the declension differences from the nominative case to all the other cases, including the instrumental. To see how this works, consider the trees in (59) and (60).

(59)



(60)



What we see here are two trees, each of them containing features characteristic for a singular feminine noun in the instrumental case. At the bottom of each tree, we see a different root: we find the Declension II *nedel'*- on the left, and the Declension III *metel'*- on the right. We know from before that the Declension II *nedel'*- spells out a projection that is smaller than the projection of the nominative, because (recall) the noun requires an overt ending in the nominative. For now, the hypothesis is that *nedel'*- spells out INDP.

This contrasts with the situation in Declension III. Here we know from before that the root *metel'*- 'snowstorm' spells out minimally NOMP. In (60), it is depicted as spelling out ACCP. This is thus an update on the previous analysis, which takes into account the fact that *metel'*- has a zero ending also in the accusative, see (61), repeated from (4).[5]

(61) The singular of declensions II and III (Timberlake 2004)

	lip II (FEM)	notebook III (FEM)
NOM	gub-a	tetraď-Ø
ACC	gub-u	tetraď-Ø
GEN	gub-y	tetraď-i
LOC	gub-e	tetraď-i
DAT	gub-e	tetraď-i
INS	gub-oj	tetraď-ju

Now, the most important point is that due to the fact that the root *metel'*- and *nedel'*- spell out trees of different sizes, they leave different number of case features/projections left for the spellout by the ending. It can be seen from the

[5]Due to the Superset Principle, the fact that *metel'*- in (60) spells out ACCP is compatible with it spelling out also NOMP, as proposed in Section 3.1. Therefore, this update does not change anything about our previous conclusions.

diagram in (60) that the ending *-ju*, used with Declension III roots, only needs to spell out features F_3 to F_6. This is crucially a different set than the one spelled out by the Declension II ending *-ej/-oj*, which needs to spell out all the case features from the nominative F_1 up to the instrumental F_6. In other words, the difference in root size between *nedel'-* and *metel'-* allows us to capture not only the difference in the nominative, but also in the instrumental (and, by extension, in all the cases in between).

As the analysis now stands, the relationship between the two classes can be depicted by the lexicalization table in (62). What we see in the upper part of the table is a Declension III noun. I am using the root *tetrad'-* 'notebook' seen in (61). This root (just like all Declension III roots) is specified for all the features up to (and including) F_2. Due to the way the spellout algorithm works, this root always spells out all the features specified in its entry, i.e., up to F_2 (this is because spelling out without movement is always preferred). It is only when the root cannot spell out all the features that a suffix is introduced. The list of suffixes that occur after a Declension III root are given in the table: we see the GEN/LOC/DAT *-i* and the instrumental *-ju*. These endings are lexically specified for spelling out features from F_3 and up to F_5 (*-i*) and from F_3 to F_6 (*-ju*). Note that due to the Superset Principle, the ending *-i* (specified as F_3-F_5) spells out constituents of different sizes (namely the locative F_3-F_4 and the genitive F_3), leading to syncretism. Note that the instrumental *-ju* can also spell out DAT/LOC/GEN, but loses to *-i* in competition.

(62)

	xNP	REF	CLASS	FEM	IND	F1	F2	F3	F4	F5	F6
NOM	tetrad'										
ACC	tetrad'										
GEN	tetrad'							i			
LOC	tetrad'								i		
DAT	tetrad'									i	
INS	tetrad'										ju
NOM	gub					a					
ACC	gub						u				
GEN	gub							y			
LOC	gub								e		
DAT	gub									e	
INS	gub										oj

In the bottom part of the table, we see a Declension II root *gub-* 'lip.' This root does not spell out any case features. Therefore, the endings that combine with a Declension II root must actually spell out more features than Declension III endings (they need to spell out F_1 and F_2).

As a result of how the lexical entries are set up, the spellout algorithm does not allow us to generate hypothetical forms where a Declension III ending appears on a Declension II root. The reason is that this would leave the features F_1 and F_2 without spellout. Similarly, we cannot place a Declension II ending on a Declension III root, since that would lead to a double spellout of the very same features. Both of these options are ruled out by the Spellout algorithm (42).

The general conclusion is that if this kind of analysis turns out to be viable

in the context of the other two declensions, it shows that an arbitrary lexical difference among roots in terms of 'size' (the number of features associated to the root in its entry) is a tool that is powerful enough to model declension classes. In addition, the analysis has no need for contextual specifications: the correct combination of the roots and the endings falls out from the lexical specifications themselves.

Note again that while we eliminate language-particular declension *features*, we are not eliminating language-particular declension *classes*. Declension classes exist and correspond to a difference in root size. This theoretical possibility relativizes the relationship between the existence of surface 'language-particular categories' and a universal set of features. Specifically, once this proposal is adopted, the existence of language-particular classes is compatible with a universal set of features.

4 The pairing of declensions and genders

In the reminder of this paper, I provide an account of all four declension classes of Russian following the analytical guidelines adopted above. This will require me to change slightly the account of Declension II and III and introduce a new derivational option used in Nanosyntax, namely Backtracking. The analysis of the plural (to be presented in Section 5) will require the use of pointers.

4.1 The basic facts

So far, I have focussed on the difference between Declension II and III. The reason is that in these two declensions, it is possible to find minimal pairs of roots that have the same grammatical features and the same type of phonology, but belong to different declensions. Such pairs illustrate the fact that declension classes are (at least in the relevant cases) arbitrary.

However, there is another aspect of declension class membership in Russian, which is that while it is arbitrary to some extent, it is not fully arbitrary. If we focus on inanimate nouns, there is a rather neat implication relation between declensions and genders (Corbett 1982). In (63), I summarize Corbett's observations (see also Privizentseva 2020).

(63) Declensions and gender, inanimates (Corbett 1982:216)
- a. I → masculine
- b. II → feminine
- c. III → feminine[6]
- d. IV → neuter

The implicational relations in (63) hold for inanimate morphologically simplex (i.e., underived) nouns. Some derived nouns lead to minor complications that I come back to in section 6. However, in what follows, I will extend the account

[6]There are two apparent exceptions, namely the masculine noun *put'* 'journey' and a couple of neuters like *vremja* 'time'. However, the membership of these items in Declension III is spurious (they only share one single ending with Declension III). More on this below.

of the preceding section in a way that it can capture both the arbitrary nature of declension classes, as well as the implicational relations in (63).

4.2 Declension II → FEM

Let me start by the observation that so far, the implicational relations noted in (63) do not follow from our system at all. In order to see this, consider the table (64).

(64)

	xNP	REF	CLASS	FEM	IND	F1	F2	F3	F4	F5	F6
NOM			gub			a					
ACC			gub				u				
GEN			gub					y			
LOC			gub						e		
DAT			gub							e	
INS			gub								oj
	xNP	REF	CLASS		IND	F1	F2	F3	F4	F5	F6
NOM			meuter			a					
ACC			meuter				u				
GEN			meuter					y			
LOC			meuter						e		
DAT			meuter							e	
INS			meuter								oj

What I show in the upper part of the table is the current analysis of Declension II, where the root spells out all the features up to IND, and the ending spells out the case features. In the lower part of the table, I show a hypothetical masculine/neuter noun. The hypothetical noun is labelled as *meuter* (a blend of **m**asculine/n**euter**). Recall that masculine/neuter nouns lack the feature FEM; therefore, the FEM feature is lacking in the row of features to be lexicalized by such a noun, as depicted in the middle row of the table.

The issue with the hypothetical *meuter* noun is that if this noun was also specified for the size INDP in the lexicon as indicated in the table (64), it would be expected to combine with exactly the same endings as found with Declension II nouns. Therefore, the conclusion is that it is so far mysterious as to why there is no such noun as the hypothetical *meuter* in (64).

Can we change something about our analysis of Declension II in a way that we could rule out the hypothetical scenario in (64)? It turns out that this is possible if we adjust the size of the root of Declension II nouns. I have made it clear in Section 3.1 that the only thing we know about Declension II roots is that they do not spell out all the features of the nominative. However, it is not *a priori* clear where exactly the boundary between the root end the ending actually is. For example, there is so far nothing that would make us prefer the analysis in (64) to the one depicted in (65), where the declension endings spell out not only case, but also number. If anything, the latter analysis would be preferred on the grounds that Russian endings are portmanteau markers for case and number.

(65)

	xNP	REF	CLASS	FEM	IND	F1	F2	F3	F4	F5	F6
NOM	gub				a						
ACC	gub					u					
GEN	gub						y				
LOC	gub							e			
DAT	gub								e		
INS	gub									oj	

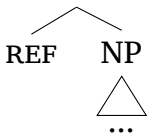
So, the question is where exactly the boundary between the root and the ending is. It turns out that if we push Declension II roots down to the level of REFP and let the ending spell out the remaining features, we end up with a system that actually derives the generalization that Declension II endings only appear on feminine nouns. The table (66) depicts this option. The Declension II root *gub-* 'lip' spells out REFP, the remaining features are spelled out by the ending. For instance, the ending *-a* spells out the features CLASS, FEM, IND and F_1.

(66)

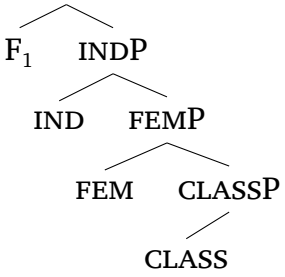
	xNP	REF	CLASS	FEM	IND	F1	F2	F3	F4	F5	F6
NOM	gub			a							
ACC	gub				u						
GEN	gub					y					
LOC	gub						e				
DAT	gub							e			
INS	gub								oj		

It can be demonstrated that such endings are unusable when the feature FEM is lacking in syntax, as is the case (by definition) with masculine and neuter nouns. In order to see why, consider first the lexical entry of the root *gub-* 'lip' and of the ending *-a*, see (67) and (68) respectively.

(67) REFP ⇔ *gub-*



(68) NOMP ⇔ *-a*

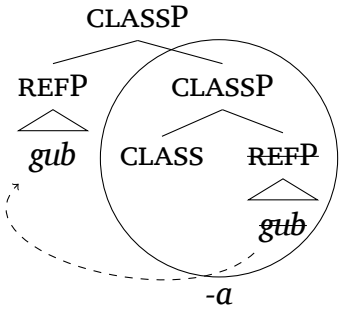


Let us now see how the root and the ending combine in an actual derivation guided by the spellout algorithm. Recall that in the course of the derivation, we are cyclically merging features and spelling out the result. The preferred option is to spell out without movement, and we do so as long as the root contains the relevant constituent. In the case of *gub-*, this works all the way to REFP, see the entry in (67).

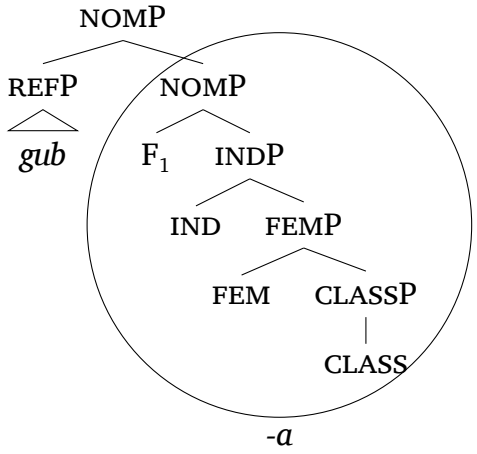
However, once we merge the CLASS feature, the root has to move to the left, and the ending *-a* spells out CLASSP, see (69). When more features (including FEM) are added, the root moves cyclically upwards, so that the nominative looks

as in (70). This is nothing but a structural rendering of the top row in the lexicalization table (66).[7]

(69)

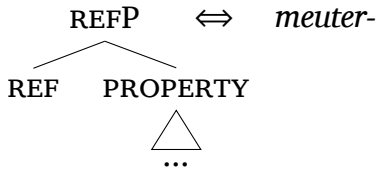


(70)

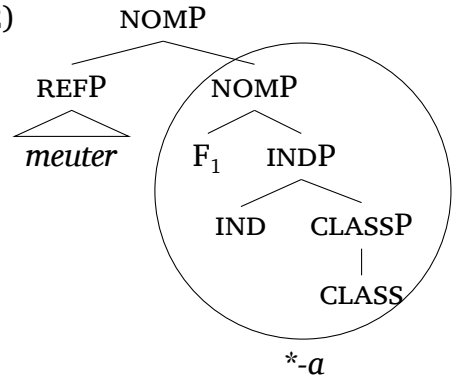


Importantly, the derivation (70) is unavailable for masculine/neuter nouns. To see that, let me assume that we have, once again, the hypothetical noun *meuter* with the lexical entry as in (71). This entry is of the same size as the corresponding feminine noun *gub-* 'lip.' If we now try to run the same derivation with this root – but crucially omitting FEM – we end up with the structure (72). This structure is just like (70), but it lacks FEM.

(71)



(72)



The interesting thing about (72) is that this structure cannot be lexicalized by the Declension II ending *-a*. The reason is that the constituent we are trying to lexicalize in (72) (namely the NOMP) is not a sub-constituent of the lexical entry for *-a*. Therefore, the ending *-a* cannot be used as a case marker of *meuter* nouns, which is what we were trying to derive.[8]

To summarize: under the new analysis, the roots of Declension II only spell out REFP. Consequently, the endings of Declension II must spell out the rest of the features, crucially including the FEM feature. As a consequence, such endings are

[7]It is an important question to ask as to why the feature FEM must be added with roots like *gub-* 'lip.' If FEM were not added, the root would not be feminine. In section 6, I speculate that this has to do with the idea that roots are 'phrasal idioms' (Marantz 1996), and they only become associated to a concept when they combine with FEM.

[8]It is relevant to note that this result presupposes strict constituent matching; cf. Vanden Wyngaerd (2018) for an alternative view. For the analysis of plurals in 5, it is also going to be relevant that *-a* can spell out CLASSP without spelling out FEM. In other words, *-a* can appear with *meuters* as long as it only spells out CLASS.

unable to lexicalize a constituent that has case features on top, but it is missing the feature FEM. This observation amounts to deriving the implicational relationship between using Declension II endings – and the result being a feminine noun.

4.3 Declension III → FEM

Let me now turn to Declension III. The analysis we have so far is given in (73). This analysis was motivated by the fact that the root needs no ending in NOM/ACC.

(73)

	xNP	REF	CLASS	FEM	IND	F1	F2	F3	F4	F5	F6
NOM	tetrad’										
ACC	tetrad’										
GEN	tetrad’							i			
LOC	tetrad’								i		
DAT	tetrad’									i	
INS	tetrad’										ju

However, recall from (63) that Declension III nouns are always feminine (with the exception of the noun, *put’* ‘journey,’ and also a small group of neuters of a special type, e.g., *vremja* ‘time’).[9] The problem is that the particular analysis in (73) does not capture this. To see that, consider the fact that the Declension III endings only spell out case features. The question is: why can’t the same endings spell out case features also on *meuter* nouns? One possibility would be that *meuter* nouns never spell out a constituent of the size ACCP, i.e., including all the features up to and including F_2. If that was so, the Declension III endings would be unusable with *meuter*s, because their use with a root smaller than ACCP would leave some case features without spellout.

However, it is not plausible that there are no nouns like that. In fact, there are good candidates among the masculine nouns for spelling out a constituent of the relevant size. To see that, consider the table (74), repeated from (5).

(74) Declensions I and IV (Timberlake 2004)

	factory I (MASC)	place IV (NEUT)
NOM	zavod-Ø	mest-o
ACC	zavod-Ø	mest-o
GEN	zavod-a	mest-a
LOC	zavod-e	mest-e
DAT	zavod-u	mest-u
INS	zavod-om	mest-om

What we see here is that the masculine nouns of Declension I have no overt ending

[9]While traditionally assigned to Declension III, it is not clear to me that these nouns should be classified as such. The reason is that such nouns only share one ending with Declension III (the GEN/DAT/LOC *-i*), while they inflect differently in all other cases. For example, the neuters take Declension IV endings in the whole of plural and also in the instrumental singular. All things considered, the membership of these nouns in Declension III is not clear. As a result, I am taking the generalization about the link between declension III and FEM at face value.

in the nominative and accusative. Therefore, it is not unreasonable to think that they are able to spell out all the features going up to F_2, though they presumably lack the FEM feature compared to the nouns belonging to Declension III. But if that is so, then nothing else said, we would expect the masculines of Declension I to combine with the endings of Declension III (contrary to fact).

Therefore, in order to make sure that no *meuter* noun combines with Declension III endings, we have to revise our hypothesis about Declension III endings. This is what I turn to now.

The safest way to prevent Declension III endings from combining with *meuter* nouns would be to let the endings spell out the FEM feature. Recall that if the FEM feature is located inside the lexical entry of an ending, the ending cannot spell out case features without simultaneously spelling out the FEM feature. Therefore, such endings can spell out case only if the feature FEM is present in the structure; and this is what Corbett's generalization boils down to in the current system.

Table (75) shows one way we can set up a system like this. The idea is that the feature FEM is the lowest feature of the endings *-i/-ju* belonging to Declension III. Compared to the table (73), the endings therefore also spell out the IND feature, and the case features F_1 and F_2. This has the desired effect that endings such as *-i* or *-ju* will not combine with masculine nouns.

(75)

	xNP	REF	CLASS	FEM	IND	F1	F2	F3	F4	F5	F6
NOM		tetrad'			Ø						
ACC		tetrad'				Ø					
GEN		tetrad'					i				
LOC		tetrad'						i			
DAT		tetrad'							i		
INS		tetrad'								ju	
NOM		gub		a							
ACC		gub			u						
GEN		gub				y					
LOC		gub					e				
DAT		gub						e			
INS		gub							oj		

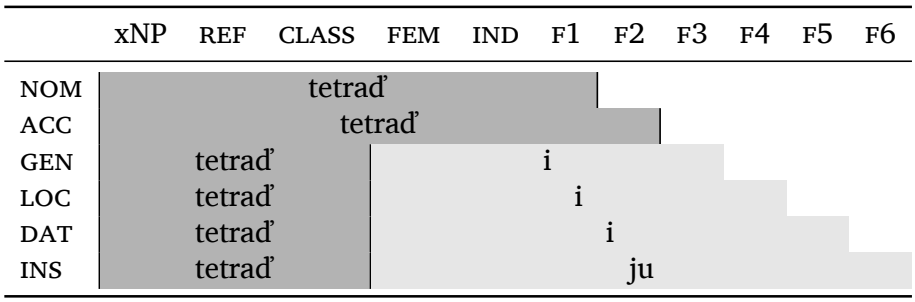
This proposal leads us further to revise the size of Declension III roots: in order to fit under FEM (which is the lowest feature of *-i/-ju*), these nouns are now proposed to be of the size CLASSP. This in turn leads us to the introduction of two silent endings in the nominative and the accusative. For now, I place these in the table, but I will have more to say about this later.

In the lower part of the table, I repeat the analysis of Declension II as proposed above in Section 4.2. The purpose of this is to show that there is still a difference in size between the two roots: the Declension III root *tetraď-* spells out CLASSP, while a Declension II root *gub-* 'lip' spells out only REFP. This is enough to ensure that the two roots combine with different endings.

Let me now come back to the zero ending in the NOM/ACC of Declension III. While this is not necessary for the analysis of the feminine declension, the currently used Nanosyntax theory actually has the tools that allow us to eliminate the zero marker. The updated analysis is depicted in (76). The table features the

idea that the endings of Declension III spell out the FEM feature, while at the same time also maintaining the idea that the root of Declension III nouns spells out all the features in NOM/ACC.

(76)



Combining these two ideas within a single table leads to a phenomenon that we can refer to as root shrinking: a root that spells out quite a lot of features in NOM/ACC 'shrinks' down to the level of CLASSP. Note first that 'shrinking' is consistent with the Superset Principle: there is no issue for our theory of matching in (76). However, while we have an idea why the root *may* shrink, it is not clear as yet why it *does* shrink.

The reason for this is because of the endings at our disposal: all the endings that can spell out F_3 contain the FEM feature. Recall that for such endings to actually spell out F_3, they must also spell out FEM. Since no feature can be spelled out twice, the root has to shrink. This is the only way to spell out F_3.

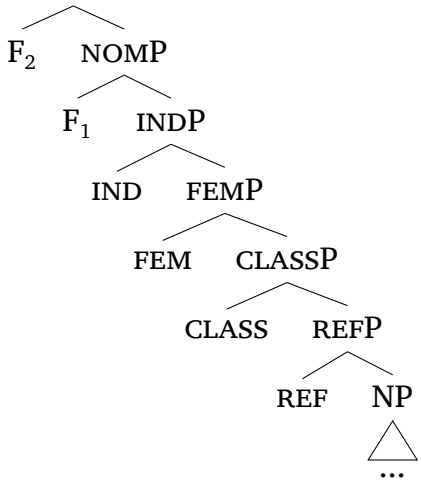
Root shrinking is going to be relevant in the analysis of the non-feminine declension, and I will therefore make a small detour here in order to explain how root shrinking is technically implemented. In the recent Nanosyntax literature, this mechanism is referred to as Backtracking (see Starke 2018, Vanden Wyngaerd et al. 2020). Backtracking is a derivational option that is activated when the spellout algorithm (repeated in (77)) fails to provide a spellout for a particular phrase. When that happens, backtracking brings the derivation one step back and tries some different derivational step at this earlier stage.

(77) Spellout Algorithm (based on Starke 2018)
   a. Merge F and spell out.
   b. If (a) fails, move the Spec of the complement and spell out.
   c. If (b) fails, move the complement of F and spell out.

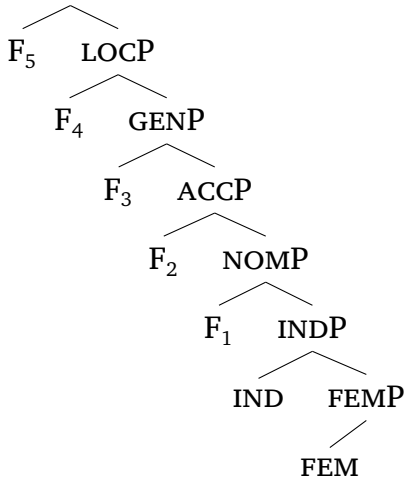
(78) Backtracking (Starke 2018)
   When spellout fails, go back to the previous cycle, and try the next option for that cycle.

To see how this works, suppose that we have the lexical entries in (79) (the root) and (80) (the ending). These entries reflect the specifications as given in (76).

(79) ACCP ⇔ *tetrad'-*

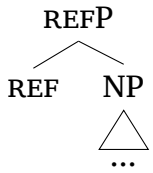


(80) DATP ⇔ *-i*

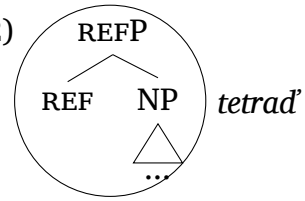


The point is to show how these endings lead to the lexicalization result as depicted in the table, using the algorithm in (77) and (78). Suppose, then, that we are building the genitive singular case of the noun *tetrad'-* (which is *tetrad'-i*). The derivation starts by assembling the REFP, see (81). This structure can be lexicalized without any movement (82).

(81)

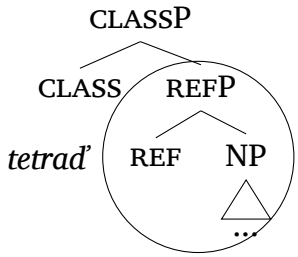


(82)

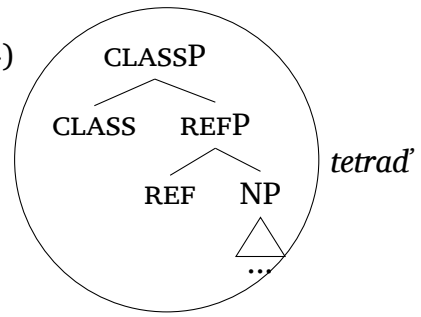


The derivation then continues by adding the CLASS feature, see (83). Spellout applies without any movement (84).

(83)

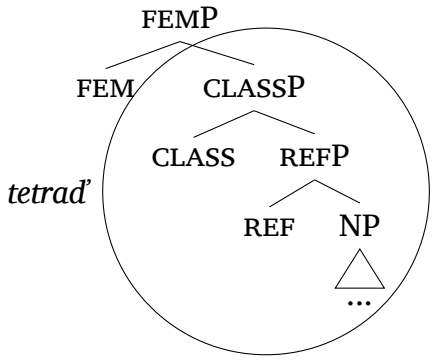


(84)

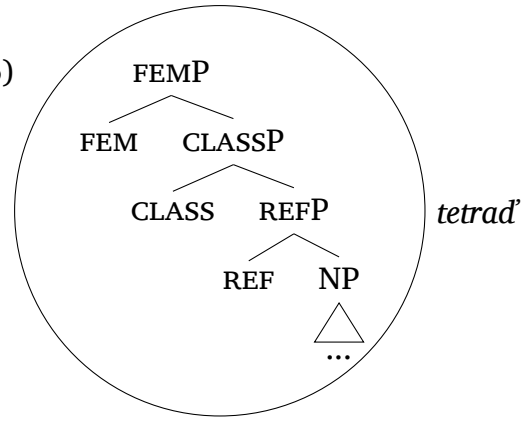


Let me show one more step of the same procedure: the feature FEM is added to (84), yielding (85). Spellout applies, and we get (86).

(85)

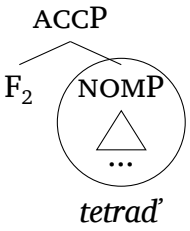


(86)

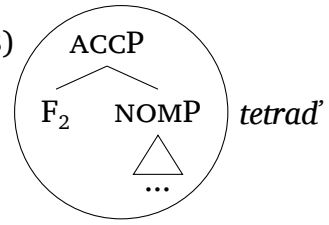


Following a sequence of analogous steps, the derivation ultimately reaches the stage where the accusative case feature is added, see (87), and the whole constituent is spelled out simply by inserting the Declension III root *tetrad'* (88).

(87)

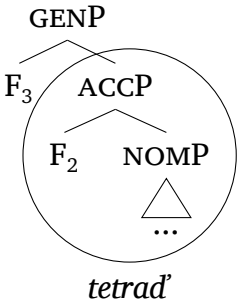


(88)

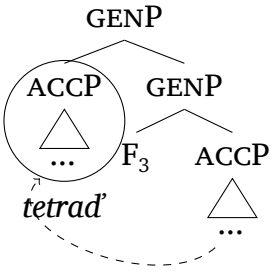


When the genitive feature F_3 is added, we get (89). Spellout without movement is impossible, since the lexical entry for *tetrad'* does not contain F_3. Therefore, spellout movements are activated following the spellout algorithm. This leads to (90).

(89)

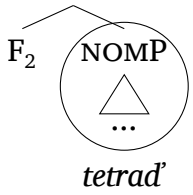


(90)

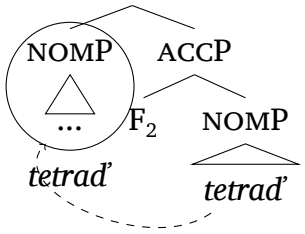


However, the structure in (90) does not lead to a successful spellout: there is no item that matches the remnant GENP (containing just F_3) in (90). Therefore the spellout algorithm fails to produce any output whatsoever. This is when Backtracking is activated. Backtracking says we should go to the previous cycle (which is ACCP). The relevant stage to which we backtrack is given in (91), repeated from (87). Once back at this stage, we should try to spell out ACCP in some other way than before (this is what Backtracking is about). On the first try, we spelled out ACCP without any movement, but that failed later on; so now we must try alternative options of the spellout algorithm. Trying something else means trying first Spec movement, as required by the spellout algorithm. However, there is no Spec, and so we have to try complement movement, as in (92).

(91)



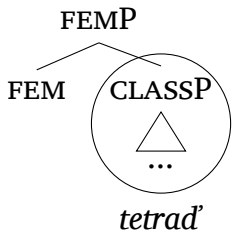
(92)



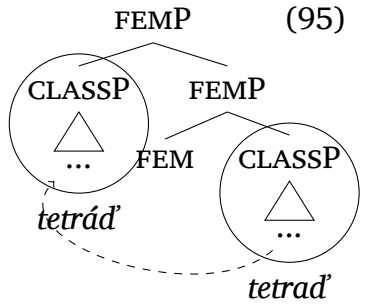
However, this does not lead to a success either, since there is no lexical item matching the remnant ACCP in (92). Since (92) does not lead to a successful derivation, we will need more backtracking. Note, however, that as a result of backtracking, the root is slowly giving up on the features it had swallowed earlier, and it is releasing them one by one in an attempt to have them spelled out by a suffix, thereby opening some alternative lexicalization path.

Since (92) failed to find a matching item for ACCP, backtracking is triggered again, then again, until we succeed and open a new derivational option. This happens only once we backtrack all the way back to FEMP. This stage is depicted in (93), repeated from (85). Here we have first spelled out FEMP without any movement, recall (86). However, this ultimately lead nowhere at the genitive level, so we are now trying a different option. Since Spec movement is impossible, complement movement is tried, yielding (94).

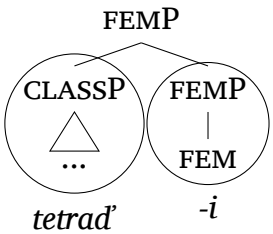
(93)



(94)



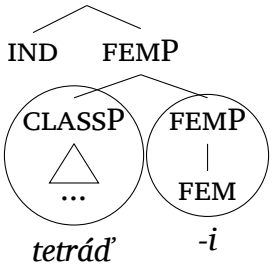
(95)



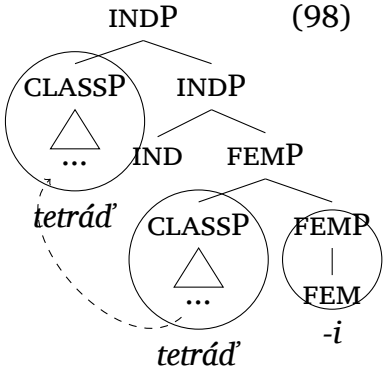
This option is successful, because it yields a structure where the remnant FEMP constituent is matched by the lexical entry for *-i*. This is shown in (95). Notice that by now, the root has given up on quite a lot of features and it has shrunk to the size of CLASSP. This is still a larger size compared to a Declension II root, but nevertheless significantly smaller than ACCP, which is the size for which the root is lexically specified. So this is how backtracking leads to shrinking, as depicted in Table (76).

Once we have opened a new derivational path, the derivation now proceeds as usual. We first add IND on top of (95), yielding (96). This structure cannot be spelled out without movement, so we must move the Spec of the complement, yielding (97).

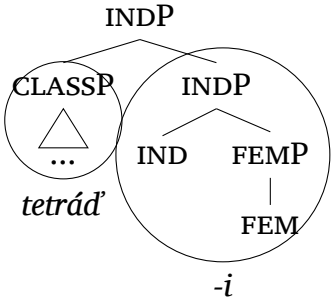
(96)



(97)



(98)



This structure leads to a successful lexicalization of the remnant INDP by *-i* (98). The derivation then proceeds in an analogous fashion, namely by adding case features on top of the structure (98). These case features always trigger the Spec-movement option of the algorithm in a fashion analogous to (97). The *-i* always spells out the relevant remnant constituent and ultimately lexicalizes all the relevant case features F_1 to F_3. This is what is depicted in the original table (76).

4.4 Declensions I and IV

Backtracking is also important in the analysis of Declensions I and IV. The phenomenon which backtracking will be needed for is inter-class syncretism as observed in the table (99). What we see here is that the two declensions differ in NOM/ACC, but they have the same endings in the oblique cases.

(99)

	factory I (MASC)	place IV (NEUT)
NOM	zavod-Ø	mest-o
ACC	zavod-Ø	mest-o
GEN	zavod-a	mest-a
LOC	zavod-e	mest-e
DAT	zavod-u	mest-u
INS	zavod-om	mest-om

Recall that this phenomenon has motivated Müller (2004) to propose that declension features should be decomposed. The point I demonstrate in this section is that this effect can also be modeled using Backtracking.

Let me start first with an analysis of Declension I. The simplest analytical option (which will turn out to be wrong) is sketched in the table (100).[10]

[10]Note that this table differs from the tables used for Declension II and III in that there is no FEM feature to be lexicalized.

(100)

	xNP	REF	CLASS	IND	F1	F2	F3	F4	F5	F6
NOM	zavod									
ACC	zavod									
GEN	zavod						a			
LOC	zavod						e			
DAT	zavod						u			
INS	zavod						om			

The starting point of the analysis depicted in the table is the observation that the root *zavod-* has no ending in NOM/ACC. This is reflected in the table by letting the root spell out all the features of these forms. In the oblique cases, the size of the root is kept constant in (100), and the case endings simply spell out the leftover case features.

However, this analysis has the same problem as the original proposal for Declension III: the endings depicted in (100) are pure case markers, and we would expect them to appear on feminine nouns. More specifically, the problem is that we would expect these endings to combine with the feminines of Declension III like *metel'-*. According to the latest (and final) proposal in (76), this noun spells out all the features up to F_2. So if the case endings were as shown in the table (100), we would predict that nouns like *metel'-* will combine with the endings in (100), contrary to fact.

This state of affairs forces us to reconsider the lexical entries of the endings in (100). What we must achieve is that these endings cannot combine with Declension III nouns. Once again, the strategy is to say that there are many ways in which the features of the oblique cases can be divided between the root and the suffix, and we must look for a way to do this division in a way that leads to the correct distribution.

Given this desideratum, it will not help (for instance) to say that the endings of Declension I spell out all case features and in addition also the IND node, but not the CLASS node. This would still allow them to combine with (feminine) Declension III roots.

The strategy that works is to propose that in addition to case and IND, the endings spell out also CLASS. Once we encode their specification like this, these endings will be inapplicable in any structure that has the feature FEM in between CLASS and IND. The reason is that a specification CLASS + IND + CASE will never match a constituent containing CLASS + FEM + IND + CASE. As a result, such endings will never appear with feminine nouns.

This type of analysis is depicted in Table (101). Note that I am assuming that the root of Declension I has to shrink under the case endings in the oblique cases. When these roots shrink, they spell out the REFP, and the ending spells out the remaining features.

(101)

	xNP	REF	CLASS	IND	F1	F2	F3	F4	F5	F6
NOM	zavod									
ACC	zavod									
GEN	zavod				a					
LOC	zavod				e					
DAT	zavod				u					
INS	zavod				om					

In principle, we could also analyze Declension I as in (102). This analysis does not rely on backtracking, but it has a zero ending in NOM/ACC.

(102)

	xNP	REF	CLASS	IND	F1	F2	F3	F4	F5	F6
NOM	zavod			Ø						
ACC	zavod			Ø						
GEN	zavod			a						
LOC	zavod			e						
DAT	zavod			u						
INS	zavod			om						

The two analyses ((101) and (102)) appear to be on a par (both capture the data), but only one of them is compatible with the fact that the endings of the oblique cases appear also in Declension IV. The most straightforward analysis of this final declension is depicted in (103). Since this declension has the very same endings in GEN, LOC, DAT and INS as Declension I, the lower part of the table (103) is identical to (101) and (102). The only difference is in the nominative and accusative case.

(103)

	xNP	REF	CLASS	IND	F1	F2	F3	F4	F5	F6
NOM	mest			o						
ACC	mest			o						
GEN	mest			a						
LOC	mest			e						
DAT	mest			u						
INS	mest			om						

Specifically, Declension IV has the ending *-o* here. The simplest analysis would say that the size of the root is invariant throughout the paradigm, and that the ending *-o* therefore spells out the features CLASS + IND + F_1 + F_2. This analysis is, however, incompatible with the analysis in (102), which has exactly the same specification for the (spurious) -Ø marker. Therefore, the two analyses clash, because we have the same specification for two different markers. However, it is not possible to have the same specification for them, because they have a different distribution. Hence, juxtaposing the analyses in (102) and (103) leads to a problem.

On the other hand, the analysis in (101) is directly compatible with the analysis in (103). The reason why there is no problem with (101) is because there

is no zero ending; its effect is captured by having the root grow and shrink as allowed by Backtracking.

Note at this point that backtracking has an important role to play in syncretism across different declensions. This is because Backtracking leads to the consequence that a root with a large tree in the lexicon shrinks down to a different size. For example, Declension I root like *zavod* 'factory' shrinks from spelling out ACCP in the accusative to spelling out just REFP in the genitive, recall (101). Once the (formerly large) root shrinks, it attains the same size that the roots of Declension IV have; these are associated to a tree of the size REFP in the lexicon. Once the two different roots attain the same size in a derivation, they combine with the same endings as a function of this. Backtracking (root shrinking) thus allows us to model cross-class syncretism, which arises when a large root shrinks down to the size of a different root class.

5 Plural

With the analysis of the singular in place, let me now turn to the analysis of the plural. I show a subset of the relevant paradigms in (104). The most obvious fact about the plural declension is that the endings in the LOC, DAT, INS are the same for all nouns.

(104) Declensions I-IV, plural (Timberlake 2004)

	notebook III (FEM)	factory I (MASC)	lip II (FEM)	place IV (NEUT)
NOM	tetraď-i	zavod-y	gub-y	mest-a
ACC	tetraď-i	zavod-y	gub-y	mest-a
GEN	tetraď-ej	zavod-ov	gub-Ø	mest-Ø
LOC	tetraď-ax	zavod-ax	gub-ax	mest-ax
DAT	tetraď-am	zavod-am	gub-am	mest-am
INS	tetraď-ami	zavod-ami	gub-ami	mest-ami

The order of the declensions (III-I-II-IV) reflects the commonalities that can be observed in the remaining cases (NOM, ACC, GEN). In particular, Declension II and IV have the same form of the GEN.PL, and therefore, they are placed next to each other.

In the NOM.PL, we have the ending *-y* in Declension I and II, and these two declensions are therefore placed next to each other as well.

However, the close affinity between declension I and II (to the exclusion of Declension III) is only apparent. To see that, recall from the discussion in footnote 1 that *-y/-i* are phonologically conditioned allomorphs, where *-i* is found after 'soft-stem' nouns. This means that if we take a soft-stem noun of Declension I (like *kon'* 'horse'), we shall see an *-i* in the NOM.PL (*kon'-i*). And similarly, when we look at a soft-stem noun of Declension II (e.g., *nedel'-a* 'week'), the NOM.PL will again have an *-i*, i.e., *nedel'-i*.

Therefore, the shading in the NOM/ACC.PL in (104) should really also encompasses Declension III, because the endings *-i* and *-y* are phonologically conditioned allomorphs. The table (105) brings this out.

(105) Declensions I-IV, plural (Timberlake 2004)

	notebook III (FEM)	knife I (MASC)	weel II (FEM)	place IV (NEUT)
NOM	tetraď-i	nož-i	nedel'-i	mest-a
ACC	tetraď-i	nož-i	nedel'-i	mest-a
GEN	tetraď-ej	nož-ej	nedel'-Ø	mest-Ø
LOC	tetraď-ax	nož-ax	nedel'-ax	mest-ax
DAT	tetraď-am	nož-am	nedel'-am	mest-am
INS	tetraď-ami	nož-ami	nedel'-ami	mest-ami

The table (105) also brings out a new fact, which is that in GEN.PL, Declensions I and III pattern the same, because the noun *nož-* 'knife' has a phonologically conditioned allomorph *-ej* in GEN.PL. This observation leads us to the final generalization about the plural declension, which is that once phonologically-conditioned allomorphy is controlled for, the endings in Declension I are fully identical to Declension III.[11]

I summarize the main observations in (106).

(106) a. Full inter-class syncretism in LOC, DAT, INS
b. Declension II and IV have the same ending in GEN (as well as in LOC, DAT, INS)
c. Declensions I and III have the same endings throughout the plural

These observations give rise to a bifurcation at the empirical level between declensions I + III (106c) vs. II + IV (106b). This fact is reflected in Müller's (2004) analysis by assigning the (arbitrary) feature [-β] to declensions I + III. Complementarily, II and IV are assigned the feature [+β], recall (10) (see also Privizentseva 2020).

In this section, the point I want to make is that Müller's feature [+/-β] perfectly corresponds to a difference in root size as established on the basis of the singular. In particular, declensions I and III, assigned [-β] by Müller, are declension classes that under the current analysis contain 'large' roots. (These are the classes with a zero ending in NOM.SG, namely *zavod-Ø* 'factory' and *tetraď-Ø* 'notebook'.) Complementarily, the feature [+β] is characteristic for classes with 'small' roots.

In sum, the conclusion is that root size (as identified on the basis of the nominative singular: Ø vs. not) is the main organizing factor behind inter-class syncretism in the plural.

Before I present a detailed analysis, let me explicitly say that I shall consider singular and plural markers as unrelated. For instance, I treat the GEN/LOC/DAT.SG *-i* found in Declension III as unrelated to the NOM/ACC.PL of the same declension. (This approach is also adopted in Müller 2004.)

[11] There are also nouns in Declension I that have the ending *-ev* in GEN.PL, which is yet another allomorph of *-ov* (Timberlake 2004:135). I analyze *-ev* as an allomorph that is derived by a regular phonological process from *-ov*. On the other hand, I analyze *-ej* as an independently stored allomorph (with exactly the same syntactic structure as *-ov*) that is selected on the basis of phonological considerations, as in Bye & Svenonius 2012.

5.1 Declension I and III

This section addresses the generalization that Declension I and III inflect the same in the plural. The reason for this, I propose, is that these declensions are based on 'large' roots. Table (107) shows my proposal for the two declensions. I shall explain the content below.

(107)

	xNP	REF	CLASS	IND	GROUP	F1	F2	F3	F4	F5	F6
NOM			nož			i_{pl} (y)					
ACC			nož			i_{pl} (y)					
GEN			nož			ej (ov)					
LOC		nož				ax$_{\rightarrow a}$					
DAT		nož				am$_{\rightarrow a}$					
INS		nož				ami$_{\rightarrow a}$					

	xNP	REF	CLASS	FEM	IND	GROUP	F1	F2	F3	F4	F5	F6
NOM			tetrad'				i_{pl} (y)					
ACC			tetrad'				i_{pl} (y)					
GEN			tetrad'				ej (ov)					
LOC		tetrad'					ax$_{\rightarrow a}$					
DAT		tetrad'					am$_{\rightarrow a}$					
INS		tetrad'					ami$_{\rightarrow a}$					

Let me start from the structural cases (NOM, ACC, GEN). What we see here is that the the roots spell out features all the way up to IND. However, they cannot spell out the feature GROUP (it is not a part of their lexical entry). This feature must therefore be spelled out by the plural case endings. The ending we see here is the NOM/ACC.PL *-i* with an allomorph (*-y*) and in the genitive we have the GEN.PL *-ej* (*-ov*).

In the oblique cases, we see endings that are coded in a special format because they contain a pointer. I am going to explain what pointers are shortly; for now, the main idea here is that when the locative case feature F_4 is added, the ending *-ax* must be used (none of the singular endings matches). Since the lowest feature of *-ax* is CLASS, the root has to backtrack below CLASS. The analysis where *-ax* spells out CLASS is forced by the fact that *-ax* is also found with the small roots of Declension II and IV. The latter roots are lexically stored in the REFP size, and the oblique endings must therefore start no higher than CLASS.

In the table, I am including all the phonologically conditioned allomorphs for each marker. Such markers are assumed to have identical morphosyntactic specifications. One option is that they are related by a productive phonological process. Another option is that each of them is independently stored, but linked to the exact same structure as its allomorph. In the latter case, the choice between them is made at PF, depending on the final consonant of the root. I remain agnostic on this here.

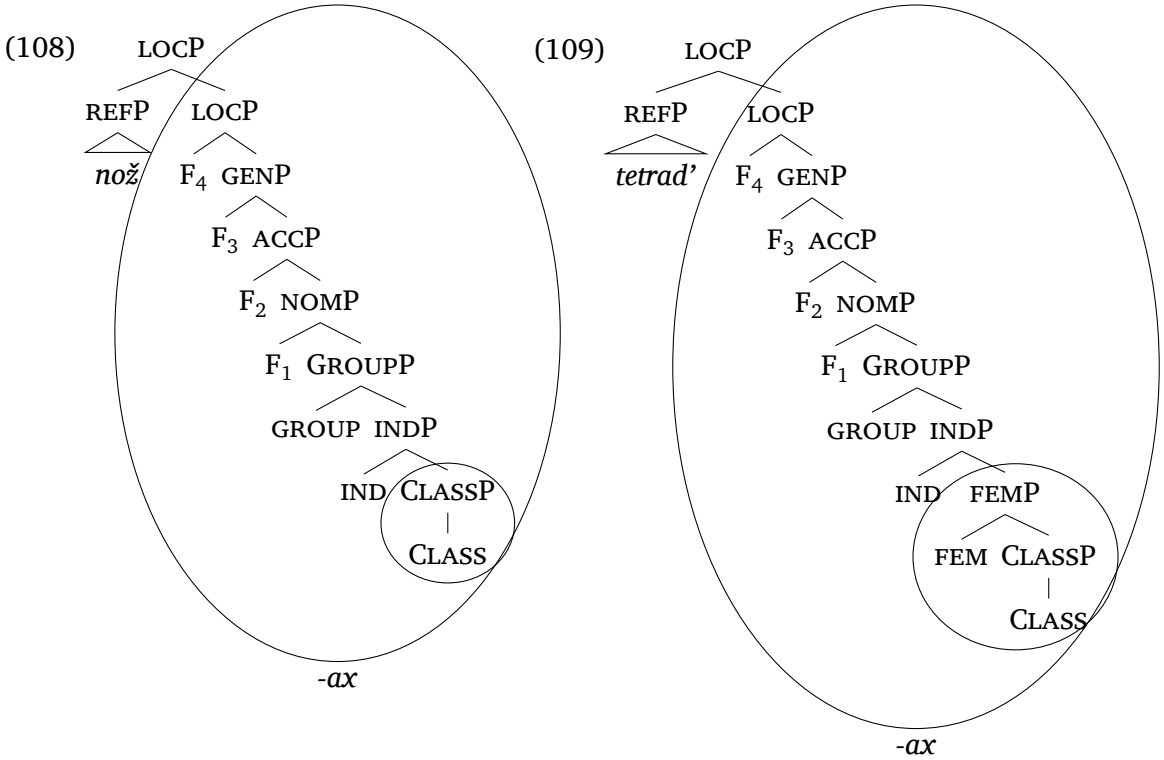
The subscript *pl* on *-*i_{pl} is there to remind us of the general strategy that plural markers are considered distinct from singular markers, i.e., the marker *-*i_{pl} is different from the *-i* found in GEN, LOC, DAT of the singular Declension III.

I shall now state two beneficial consequences of this proposal. The first consequence is that if things work as depicted in (107), then the structural case endings

only spell out number and case features. Gender features are spelled out by the relevant roots. This has the effect that these endings are not bound to a particular gender, which is correct (*tetrad'* is feminine, *nož* is masculine, yet they both combine with *-i* and *-ej*).

The second good consequence is that all the endings spell out the feature GROUP. This has the effect that they cannot appear in the singular (where GROUP is absent). As a result, we do not need to worry that these endings will make our analysis of the singular collapse (and vice versa).

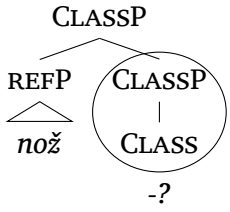
Turning now to the non-structural cases (GEN, LOC, INS), there is one non-trivial aspect of the proposal as depicted in Table (107). To see it, consider the locative plural *-ax*. As the table indicates, *-ax* spells out all the features from CLASS up to F_4. With masculine nouns like *nož-* 'knife,' this corresponds to the LOCP constituent circled in (108). With feminine nouns, this amounts to the LOCP constituent circled in (109).



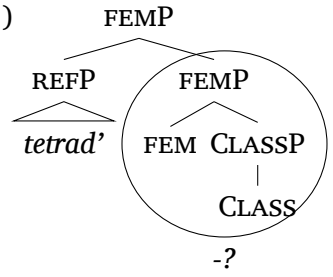
The difficult point in the analysis is that these constituents differ in the gender features that are found at the bottom of the trees. Specifically, the constituents below the IND feature are not identical, as the inner circle brings out. This makes it challenging to come up with the right entry for *-ax*. If we specify it as in (109), it will not match in (108), since (108) is not a sub-constituent of (109). The same is true if we specify *-ax* as in (108): such a tree is also not a sub-constituent of (109).

The solution to this conundrum emerges when we ask ourselves the question how the gender structures are pronounced when there are no number features on top of them. In other words, the question I am asking now is how the structures in (110) and (111) are pronounced. These structures correspond to the inner circles in (108) and (109).

(110)



(111)

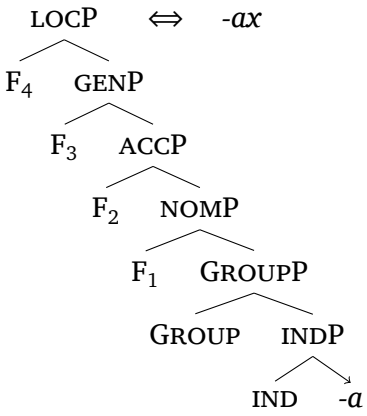


For the tree (111), the clear winner is the singular marker *-a*, which appears in the NOM.SG of Declension II. Out of all the markers that can spell out a constituent with CLASS at the bottom and FEM right above, *-a* has the smallest number of unused features (see (66) for all the markers that match (111)). Therefore, the constituent (111) is spelled out as *-a*.[12]

Let us now turn to the spellout of the ClassP in (110). Note that *-a* is clearly a candidate here, since the relevant ClassP in (110) is a subset of the FemP that is spelled out by *-a* in (111). When *-a* spells out ClassP in (110), it has three unused features (FEM, IND and F_1). Other candidates for the spellout of ClassP have either the same number or more. The best candidate out of the singular markers used in Declensions II and IV is the NOM/ACC.SG *-o*, recall (103). This marker also has three unused features (IND, F_1 and and F_2). Therefore, there seems to be a draw here, but clearly, *-a* is a candidate at least as good as any other marker. In fact, the unused features of *-o* are more marked than those of *-a*, and therefore, *-a* is the most likely winner here.[13]

Let us now come back to the entry of *-ax*. Originally, we had problems formulating its lexical entry in a way that it matches both structures in (108) and (109). Now that we know that both of the inner circles are (or at least can be) spelled out as *-a*, we can formulate the entry as in (112), using a pointer.

(112)



The lowest part of the entry reads: if one of the daughters is spelled out by *-a*, and if the other daughter corresponds to the feature IND, then *-ax* can be inserted. Therefore, this entry matches the IndP in both (108) and (109). Following the same logic, this entry matches also all higher nodes in (108) and (109), and it

[12]This marker is never used with Declension III in the singular, since the root *tetrad'*- 'notebook' never backtracks so deep in the singular, recall (76). However, in the plural, it has to (there is no other LOC.PL marker than *-ax*).

[13]'More marked' refers to structural height. The extra feature F_2 in the lexical entry of *-o* is higher up in the feature hierarchy than any other feature of *-a*.

can then ultimately spell out both (108) and (109).

I am assuming that case markers in DAT and INS also have a pointer to *-a*. These entries are depicted in (107) by using a subscript (i.e., $ax_{\rightarrow a}$). The part of the structure that is spelled out by the item pointed to (i.e., *-a*) is in a darker shade. However, it is worth stressing that *-ax* is a single non-decomposable marker like *mice*.

The entries with pointers are the only entries that can spell out oblique case features in the plural, and they therefore force all the roots to shrink below their lowest feature. As we shall see, this leads to the effect that all declensions are unified in the oblique cases.

Before we move on, let me highlight the most important consequence of this analysis, which is that it explains why declensions I and III pattern together. This is because they have a big root, which allows each of them to spell out the INDP constituent in the plural. The effect of both roots spelling out the same constituent is that they take the same suffixes (including when they backtrack even lower). This result supports the idea that declension class is the effect of root size.

5.2 Declension II and IV

Let me now turn to declensions II and IV. Their analysis is given in Table (113).

(113)

	xNP	REF	CLASS	IND	GROUP	F1	F2	F3	F4	F5	F6
NOM	mest			a_{pl}							
ACC	mest			a_{pl}							
GEN	mest				$Ø_{\rightarrow a}$						
LOC	mest				$ax_{\rightarrow a}$						
DAT	mest				$am_{\rightarrow a}$						
INS	mest				$ami_{\rightarrow a}$						

	xNP	REF	CLASS	FEM	IND	GROUP	F1	F2	F3
NOM	žen				y_{pl} (i)				
ACC	žen				y_{pl} (i)				
GEN	žen					$Ø_{\rightarrow a}$			
LOC	žen					$ax_{\rightarrow a}$			
DAT	žen					$am_{\rightarrow a}$			
INS	žen					$ami_{\rightarrow a}$			

Let me start from the Declension IV noun *mest-o* 'place.' As is clear from the table, I am assuming that the NOM/ACC.PL *-a* (found in Declension IV) is a portmanteau realization of all the relevant heads, starting on top of the root (which spells out REFP). The marker *-a* thus spells out CLASS, IND, GROUP and then case features up to F_2. This marker will be found only with Declension IV nouns. It cannot be used with feminine nouns, because it lacks the FEM feature (and it has no pointer). In principle, it could be used in Declension I (see the masculine *nož* 'knife' in Table (107)). However, since the roots of Declension I are 'large,' using *-a* with them would require backtracking. Therefore, this spellout option is dispreferred in Declension I, and *-y/-i* is used instead.

Moving on to GEN, I am assuming that there is a lexical entry here introducing

a phonological zero (e.g., an empty CV skeleton), and that this entry has a pointer to *-a*. As a result, we find it also in Declension II, since it matches both structures.

Since this entry has the same abstract structure as the entry for *-ax*, it could in principle be used also in the genitive of Declensions I and III. However, that would again require backtracking, and this option is therefore dispreferred in those declensions. In declension II and IV, however, no backtracking is needed, since their root is small to begin with. As a result, the zero marker is used with 'small-root' declensions II and IV, while *-ej* is used with 'big-root' declensions. This way, root size replaces the diacritic [+/-β] used in Müller's (2004) work.

The entries in the LOC, DAT, INS are already familiar to us, and I have explained in detail how they spell out the relevant structure.

In the lower part of the table, we see the feminine Declension II (illustrated by *žen-a* 'woman'). Starting with the NOM/ACC, I am assuming that the *-i/-y* that we see here is a portmanteau analogous to the Declension IV *-a*, spelling out CLASS, FEM, IND, GROUP and case. This is thus a different specification than the one I have postulated for the *-i/-y* that we find in NOM/ACC.PL in Declension I and III. This amounts to accidental homophony, which would be ideally been removed. However, I have to leave this for future research.

Starting with the genitive, we find the same markers as in declension IV. This is as expected, since the oblique case markers match the relevant structure.

This concludes my analysis of the Russian declension. The main point was to show that we can capture the distribution of case markers in all four declension classes of Russian without making use of language-specific declension features. What I have demonstrated is that this is not only possible, but we can in fact do so in a way that provides an explanation for both vertical and horizontal syncretism among the markers in individual declensions.

As always, there remain a number of open issues.

6 Open ends

The first issue, noted in footnote 7, is that so far, we have no straightforward way of enforcing that feminine roots actually combine with the FEM head in the course of the derivation, and conversely, that non-feminine roots do not combine with this head. To take the most obvious example, consider roots belonging to Declension II (like *žen-a*). As the theory stands, these roots are associated to the structure of exactly the same size as Declension IV roots (e.g., *mest-o*); both are associated to REFP. Now given that these two roots have an entry of the exact same size, the question is why one of the roots requires the presence of FEM in the derivation, while the other does not.[14]

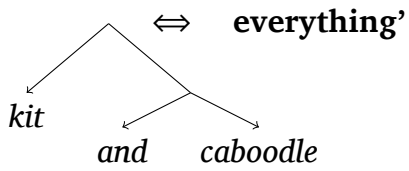
The answer I suggested in footnote 7 was to treat this as an instance of idiomaticity. The gist of the idea is that feminine roots can only be interpreted in the context of the feature FEM. If the feature FEM is not merged, the root cannot be used. This idea is inspired by Harley (2014). Harley notes that certain roots can only be interpreted in specific context (inside an idiom); as a case in point, she mentions the phrase *kit and caboodle*. Here the root *caboodle* cannot be used

[14] It is possible that these roots differ in their internal composition, but the question still stands: how does the internal composition make the presence/absence of the FEM head obligatory?

outside of this idiom. In a similar sense, I would like to explore the option that feminine roots can only be used in the context of the feature FEM.

In Nanosyntax, idioms contain pointers. The idiom *kit and caboodle* would look as in (114), where a non-compositional meaning is associated to a phrase containing these three items.

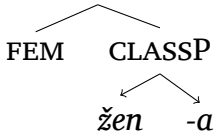
(114)



In a similar sense, the entry for *mice* in (28) is a 'phonological' idiom, making sure that *mouse* combined with PL is pronounced non-compositionally as *mice*.

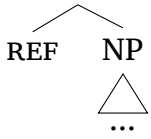
On analogy with these cases, I am proposing that *žen-* (a *caboodle*-style root with no interpretation outside of the idiom) is interpreted as **woman'** when combined with FEM, see (115). The tree in (115) depicts the stage of the derivation where the feature FEM has been added on top of the CLASSP constituent, which we know spells out as *žen-a*.

(115)



⇔ **woman'**

(116)



⇔ *žen-*

This type of an approach presupposes that phonological and conceptual information may be associated to constituents of different sizes. In particular, while I am assuming that the concept **woman'** is associated to FEMP, I am assuming that the phonology of *žen-* is associated to a different constituent, namely REFP, see (116).

The suggestion (115) and (116) is tentative, yet such a bifurcation between phonological and conceptual information is likely to exist independently. Recently, it has been explored in an ongoing work by Omer Preminger. He illustrates this on the example of the verb *went off*, where the phonology *wɛnt* is assigned to [PAST + GO], which is a different constituent from [GO OFF], to which the conceptual information is assigned, see (117).

(117) The spellout of *went off*

wɛnt		*vf*
PAST	GO	OFF
before utterance	explode	

Another open issue relates to the difference between nouns of masculine and neuter gender. Specifically, I have set up my analysis in a way that it does not need to distinguish between masculine and neuter nouns. Still, this property of nouns is clearly accessed by agreement, and so the question arises whether this fact is compatible with my analysis at all. While I shall not provide a specific proposal, the general solution must be that this distinction must be made somewhere within the low part of the NP (specifically REFP) that is always spelled out

by a root, and which is therefore unaccessible to the endings.

This claim (that the endings do not distinguish masculine and neuter gender) may appear problematic in view of the fact that there is an implicational relationship between Declension I/IV and masculine/neuter gender, respectively, recall (118).

(118) Declensions and gender, inanimates (Corbett 1982:216)

a. I → masculine
b. II → feminine
c. III → feminine
d. IV → neuter

At this point, it becomes relevant that the generalizations (118) are intended to characterize underived nouns, even though as far as I can tell, they hold also for a large majority of derived nouns. Interestingly, among the few types of derived nouns where the generalizations fail to hold, we find cases where masculine nouns end in *-o*. This is the case with augmentatives such as*dom-iš'-e* 'house, augmentative,' which is derived from *dom* 'house' (Timberlake 2004:146, Privizentseva 2020). The augmentative noun keeps the masculine gender of the base *dom* 'house,' which is a Declension I noun. However, the ending *-e* is a phonologically conditioned allomorph of the Declension IV NOM/ACC.SG *-o*.[15]

Therefore, the fact that the theory allows masculine nouns ending in *-o* is in fact a good thing. The apparently exceptional augmentatives can be relatively easily fitted into our current theory by assuming that the augmentative *iš'* spells out a head that comes below the CLASS node, thereby confining Class I roots to the spellout of REFP. As a result, the remaining features must be spelled out by the suffix *-e* (a phonologically conditioned allomorph of *-o*). This analysis is depicted in Table (119).

(119)

	xNP	REF	AUG	CLASS	IND	F1
NOM	dom		iš'	e (o)		

If the analysis above is correct, it supports the decision to locate the distinction between neuter and masculine nouns inside the constituent called REFP.

A different set of issues arises for animate nouns. The first thing to know about animates is that generally, they combine with the same endings as inanimates. Given this general tendency, I am assuming that the animacy feature must be located low down in the structure as well, more specifically in the part of the hierarchy that is always spelled out by the root (REFP).

That said, Russian animate nouns exhibit two points of divergence from inanimates in their declension. The first difference is that with animates of Declension I, the form of the direct object is identical to the genitive case (rather than to the nominative, as with inanimates).

[15]Similarly, in the plural, the augmentative has the Declension IV shape *dom-iš'-a* in NOM/ACC, and *dom-iš'-Ø* in the GEN.PL, exhibiting all the typical shapes for a Declension IV noun, despite its masculine gender. A puzzling fact is that in the NOM.PL (and only there), the augmentative may also take on the Declension I ending *-i*.

(120)

	factory I (MASC)	student I (MASC, ANIM)
NOM	zavod-Ø	student-Ø
ACC	zavod-Ø	student-a
GEN	zavod-a	student-a

In the current setting, I do not treat this as a case of allomorphy triggered by animacy, but rather as a genuine difference in the marking of the object along the lines of Starke (2017). What Starke proposes is that differentially marked objects carry a special case that he calls BIG ACCUSATIVE, which happens to be identical to the genitive in Russian. Simplifying, then, the idea is that the special shape of the accusative in (120) is not due to animacy-conditioned allomorphy, the animate direct objects are simply marked by a different case than inanimates. If that is so, we still do not have any grounds to locate the animacy feature anywhere else but within the low structure spelled out by the root.

The second area where animates diverge from inanimates concerns the pairing between declensions and genders. Specifically, there are masculine animate nouns that belong in Declension II. (This contrasts with Declension II inanimates, which are exclusively feminine.) What I suspect may be the case here is that despite appearances, these masculine animate nouns actually do contain the feature FEM in their structure, and the endings reflect this. The reason why we get masculine agreement on modifiers with such nouns would then have to be attributed to a special property of animate nouns, namely that they allow for agreement 'according to sense.'

This idea further requires that the feature FEM is not interpreted literally as 'female,' but more abstractly (which must anyway be the case for inanimates). What I tentatively propose is that the more general interpretation of the feature FEM could be related to (small) size. It has been observed in work by Jurafsky (1996) that a recurrent pattern across languages is one where morphology marking female-denoting nouns is the same as (small) size morphology (diminutives). Therefore, the idea is that the feature FEM has the interpretation of 'small size.' The essence of the idea would be that when such a feature applies to an animate noun, it yields a female-denoting noun, while with inanimate nouns, it is interpreted as a size morpheme. This more abstract interpretation of FEM may in fact be reflected in some special properties of Declension II masculine nouns. Specifically, Timberlake (2004:131) mentions two classes of such cases. The first class contains what Timberlake calls 'diminutive names' like *Saš-a*, where the supposed 'diminutive-style' interpretation of FEM seems in fact relevant. The other class contains mildly derogatory nouns like *neposed-a* 'fidgety person' or *neved-a* 'ignoramus,' where again the more abstract meaning of FEM may be relevant. Needless to say, all these questions remain open for future research.

7 Conclusions

This article started from the observation that standard accounts of declension classes often rely on postulating arbitrary declension-class features. Such features

are problematic for a restrictive theory of UG. Consequently, the goal of this article was to formulate an analysis of Russian declension that removes from the grammar any reference to these language-particular declension features.

In the account I have proposed, I rely on a universal set of features, arranged in a hierarchy. Declensions arise because different roots spell out different number of features. As a result of spelling out a different number of features, each root leaves a different residue, where by residue I mean those features that are present in the structure but not spelled out by the root. The residue is spelled out by the ending. When two roots leave a different residue, they combine with different endings.

The bulk of the article then discussed an algorithmic implementation of this idea in a Nanosyntax model of spellout. We have seen that it is possible to formulate an analysis that captures virtually all syncretisms (both vertical and horizontal) and does not rely either on declension features, or contextual specifications of lexical entries.

In sum, since the residue determines the ending, and the root determines the residue, the root controls which ending it combines with. However, the selection is not encoded by contextual restrictions and/or by declension features; all of this is an automatic consequence of encoding in the lexicon the set of features spelled out by the root. The rest follows from the general theory of spellout.

Acknowledgements

The work on this article has been supported by a grant from the Czech Research Foundation (GAČR), grant number GC21-12611J. I also want to thank three anonymous reviewers of this paper for their helpful comments and suggestions, as well as the audience at Sinfonija 13 in Budapest.

References

ALEXIADOU, A. & MÜLLER, G. 2008. Class features as probes. *Inflectional Identity*, eds. A. Bachrach & A. Nevins, 101–155. Oxford: Oxford University Press.

BAKER, M. 2003. *Lexical categories: Verbs, nouns and adjectives*. Cambridge University Press.

BAUNAZ, L. & LANDER, E. 2018. Nanosyntax: the basics. *Exploring Nanosyntax*, eds. L. Baunaz, K. De Clercq, L. Haegeman & E. Lander, 3–56. Oxford: Oxford University Press.

BAYER, J., BADER, M. & MENG, M. 2001. Morphological underspecification meets oblique case: Syntactic and processing effects in German. *Lingua* 111, 465–514.

BITTNER, M. & HALE, K. 1996. The structural determination of case and agreement. *Linguistic Inquiry* 27(1), 1–68.

BLEVINS, J.P. 2016. *Word and paradigm morphology*. Oxford: Oxford University Press.

BOBALJIK, J. 2012. *Universals in Comparative Morphology: Suppletion, superlatives, and the structure of words*. Cambridge, Mass: MIT Press.

BOBALJIK, J. & SAUERLAND, U. 2018. *ABA and the combinatorics of morphological features. *Glossa* 3(1), 15.1–34.

BYE, P. & SVENONIUS, P. 2012. Non-concatenative morphology as epiphenomenon. *The Morphology and Phonology of Exponence*, ed. J. Trommer, 427–495. Oxford: Oxford University Press.

CAHA, P. 2009. *The Nanosyntax of Case*. Ph.D. thesis, CASTL, University of Tromsø.

CAHA, P. 2013. Explaining the structure of case paradigms through the mechanisms of Nanosyntax. *Natural Language and Linguistic Theory* 31, 1015–1066.

CAHA, P. 2017. How (not) to derive a *ABA: The case of Blansitt's generalisation. *Glossa: a journal of general linguistics* 2, 84.

CAHA, P. 2020. Nanosyntax: some key features. *The Cambridge Handbook of Distributed Morphology*, eds. A. Alexiadou, R. Kramer, A. Marantz & I.O. Massuet. Cambridge: Cambridge University Press.

CAHA, P., DE CLERCQ, K. & VANDEN WYNGAERD, G. 2019. The fine structure of the comparative. *Studia Linguistica* 73(2).

CHOMSKY, N. 2001. Beyond explanatory adequacy. *MITWPL* 20.

CHVANY, C.V. 1982. Hierarchies in the russian case system: For N-A-G-P-D-I, against N-G-D-A-I-P. *Russian Language Journal / Русский язык* 36(125), 133–147.

CINQUE, G. 2013. Cognition, universal grammar, and typological generalizations. *Lingua* 130, 50–65.

CORBETT, G.G. 1982. Gender in Russian: An account of gender specification and its relationship to declension. *Russian linguistics* 6, 197–232.

DE CLERCQ, K. 2013. *A unified syntax of negation*. Ph.D. thesis, Ghent University.

DE CLERCQ, K. & VANDEN WYNGAERD, G. 2017. *ABA revisited: evidence from Czech and Latin degree morphology. *Glossa* 2(1), 69: 1–32.

HALLE, M. 1997. Impoverishment and fission. *Papers at the Interface*, eds. B. Bruening, Y. Kang & M. McGinnis, 425 – 449. No. 30 in MITWPL, Cambridge, MA.: MIT.

HALLE, M. & VAUX, B. 1998. Theoretical aspects of Indo-European nominal morphology: The nominal declensions of Latin and Armenian. *Mir Curad: Studies in Honor of Clavert Watkins*, eds. J. Jasanoff, H.C. Melchert & L. Olivier, 223–240. Innsbrucker Beiträge zur Sprachwissenschaft, Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.

HARLEY, H. 2014. On the identity of roots. *Theoretical Linguistics* 40, 225–276.

HARLEY, H. & RITTER, E. 2002. Person and number in pronouns: A feature-geometric analysis. *Language* 78(3), 482–526.

HASPELMATH, M. 2007. Pre-established categories don't exist: Consequences for language description and typology. *Linguistic typology* 11(1), 119–132.

JURAFSKY, D. 1996. Universal tendencies in the semantics of the diminutive. *Language* 533–578.

KIPARSKY, P. 1973. 'Elsewhere' in phonology. *A Festschrift for Morris Halle*, eds. P. Kiparsky & S. Anderson. New York: Holt, Rinehart and Winston.

KLOUDOVÁ, V. 2020. *The Interaction of Functional Morphemes inside the Nominal Phrase*. München: LINCOM GmbH.

KRAMER, R.T. 2015. *The morphosyntax of gender*. Oxford: Oxford University Press.

MARANTZ, A. 1996. Cat as a phrasal idiom: consequences of late insertion in Distributed Morphology. Ms., MIT.

McCreight, K. & Chvany, C.V. 1991. Geometric representation of paradigms in a modular theory of grammar. *Paradigms: The Economy of Inflection*, ed. F. Plank, 91–112. Berlin, New York: Mouton de Gruyter.
McFadden, T. 2018. *ABA in stem-allomorphy and the emptiness of the nominative. *Glossa: a journal of general linguistics* 3, 8.
Melvold, J.L. 1989. *STRUCTURE AND STRESS IN THE PHONOLOGY OF RUSSIAN*. Ph.D. thesis, CUNY, MIT.
Müller, G. 2004. A distributed morphology approach to syncretism in russian noun inflection. *Proceedings of FASL*, vol. 12, 353–373.
Picallo, M.C. 1991. Nominals and nominalizations in catalan. *Probus* 3(3), 279–316.
Privizentseva, M. 2020. Declension tracks gender: Insights from mixed agreement in Russian. Handout of a talk at NELS 51.
Smith, P.W., Moskal, B., Xu, T., Kang, J. & Bobaljik, J.D. 2019. Case and number suppletion in pronouns. *Natural Language & Linguistic Theory* 37(3), 1029–1101.
Starke, M. 2009. Nanosyntax. A short primer to a new approach to language. *Nordlyd 36: Special issue on Nanosyntax*, eds. P. Svenonius, G. Ramchand, M. Starke & T. Taraldsen, 1–6. Tromsø: University of Tromsø.
Starke, M. 2017. Resolving (DAT = ACC) ≠ GEN. *Glossa: a journal of general linguistics* 2(1), 104.1–8.
Starke, M. 2018. Complex left branches, spellout, and prefixes. *Exploring Nanosyntax*, eds. L. Baunaz, K. De Clercq, L. Haegeman & E. Lander, 239–249. Oxford: Oxford University Press.
Stockall, L. & Marantz, A. 2006. A single route, full decomposition model of morphological complexity: MEG evidence. *The mental lexicon* 1(1), 85–123.
Taraldsen, K.T. 2019. An introduction to Nanosyntax. *Linguistics Vanguard* 5(1).
Timberlake, A. 2004. *A reference grammar of Russian*. Cambridge: Cambridge University Press.
Vanden Wyngaerd, G. 2018. The feature structure of pronouns: a probe into multidimensional paradigms. *Exploring Nanosyntax*, eds. L. Baunaz, K.D. Clercq, L. Haegeman & E. Lander, 277–304. Oxford: Oxford University Press.
Vanden Wyngaerd, G., Starke, M., De Clercq, K. & Caha, P. 2020. How to be positive. *Glossa* 5(1).
Zompì, S. 2019. Ergative is not inherent: Evidence from* aba in suppletion and syncretism. *Glossa: a journal of general linguistics* 4(1), 73.

Appendix

In this appendix, I summarize my final analysis of the patterns. Table (121) contains the feminine singular declensions II and III. One thing to be kept in mind is that the *-i* in DAT/LOC/INS of Declension III is not the same marker as *-y* in Declension II (they differ in accentual properties).

(121) Feminine declensions, singular

	xNP	REF	CLASS	FEM	IND	F1	F2	F3	F4	F5	F6
NOM	tetrad’										
ACC	tetrad’										
GEN	tetrad’					i					
LOC	tetrad’						i				
DAT	tetrad’							i			
INS	tetrad’								ju		

	xNP	REF	CLASS	FEM	IND	F1	F2	F3	F4	F5	F6
NOM	gub					a					
ACC	gub						u				
GEN	gub							y			
LOC	gub								e		
DAT	gub									e	
INS	gub										oj

In (122), I summarize my analysis of the meuter singular declensions. It is worth pointing out that the *-a* in GEN.SG of the meuter declensions is different from the NOM.SG *-a* in Declension II. Also the locative *-e* of the meuter declensions is different from the *-e* in Declension II. However, this is not a shortcoming of the analysis, since these markers have different accentual properties, and they should therefore not be unified.

(122) Meuter declensions, singular

	xNP	REF	CLASS	IND	F1	F2	F3	F4	F5	F6
NOM	zavod									
ACC	zavod									
GEN	zavod					a				
LOC	zavod						e			
DAT	zavod							u		
INS	zavod								om	

	xNP	REF	CLASS	IND	F1	F2	F3	F4	F5	F6
NOM	mest									
ACC	mest									
GEN	mest					a				
LOC	mest						e			
DAT	mest							u		
INS	mest								om	

In (123), I summarize my analysis of the feminine plural declensions. These endings are considered different from the singular endings. There is one instance of accidental homophony, which is the NOM/ACC *-i/-y*. This ending can be seen in both declensions, but each of them has a different entry. In Section **??**, I showed one way how they could be unified using Hagen Blix' idea of structured lexical entries. I leave this now as something for future research.

(123) Feinine declensions, plural

	xNP	REF	CLASS	FEM	IND	GROUP	F1	F2	F3	F4	F5	F6
NOM	žen				y							
ACC	žen				y							
GEN	žen						$Ø_{\rightarrow a}$					
LOC	žen						$ax_{\rightarrow a}$					
DAT	žen						$am_{\rightarrow a}$					
INS	žen						$ami_{\rightarrow a}$					

	xNP	REF	CLASS	FEM	IND	GROUP	F1	F2	F3	F4	F5	F6
NOM			tetrad'				i					
ACC			tetrad'				i					
GEN			tetrad'				ej					
LOC	tetrad'						$ax_{\rightarrow a}$					
DAT	tetrad'						$am_{\rightarrow a}$					
INS	tetrad'						$ami_{\rightarrow a}$					

In (124), we see the analysis of the meuter declensions in the plural. The oblique endings are identical to those in (123). There is no accidental homophony introduced in (124).

(124) Meuter declensions, plural

	xNP	REF	CLASS	IND	GROUP	F1	F2	F3	F4	F5	F6
NOM			nož		i_{pl}						
ACC			nož		i_{pl}						
GEN			nož		ej						
LOC	nož				$ax_{\rightarrow a}$						
DAT	nož				$am_{\rightarrow a}$						
INS	nož				$ami_{\rightarrow a}$						

	xNP	REF	CLASS	IND	GROUP	F1	F2	F3	F4	F5	F6
NOM	mest			a_{pl}							
ACC	mest			a_{pl}							
GEN	mest				$Ø_{\rightarrow a}$						
LOC	mest				$ax_{\rightarrow a}$						
DAT	mest				$am_{\rightarrow a}$						
INS	mest				$ami_{\rightarrow a}$						