Interpreting Syntactic Structures (at PF)

by

Hagen Blix

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Department of Linguistics

New York University

May, 2023

Alec Marantz

Acknowledgments

When I entered the graduate program at NYU, I (rather naively) thought the difficulties that would come with doing a PhD would primarily be those inherent to research. I certainly have spent some sleepless nights over the distribution of this phrase and that affix. The difficulties I had failed to expect, were those arising a) from the academic structure itself, and b) from making one's own research comprehensible to others. These acknowledgments are for the friends and colleagues who helped with all these things.

Let me do the genre's customary thing, and begin with my committee. Alec, who served as my primary advisor throughout my years at NYU, has in many ways been the deepest influence on my research. It took a surprisingly long time for us to figure out what we agreed and disagreed on. Learning to both understand and express these issues has given a depth to the analyses in this dissertation, that they would surely have lacked without him. Thank you, Alec, it was fun. Since before I began graduate school, Pavel has been extremely generous with his time and ideas. Much of what is in here would have been impossible without him, and seeing him take up some of my ideas in return has been among the great pleasures of publishing linguistic papers. I've told Jonathan about various ideas on both sides of the Atlantic, and he has always engaged them with a combination of encouragement, curiosity, and a critical eye that I can only hope to emulate in any future endeavors. It does probably not need to be said, but few linguists are as fun and creative as Richie. In any meeting, there would unexpected connections and data that Richie could pull out of his metaphorical hat – it would take me another decade to think through half of

them, but I hope I managed to do justice to at least a small fraction of the things that have come up in our meetings. Last, but most certainly not least, thanks to Gary, who lent me his sharp eye for which parts of an analysis are explanatory, and who made his vast knowledge of syntax available to me whenever I asked him to meet, or to read a draft of mine.

Beyond the committee, the ideas in here owe their debt to plenty of other people who wanted to listen and discuss things. Michal Starke, Yining Nie, Maria Kouneli, Naomi Lee, Guy Tabachnick, Sarah Phillips, Stanislao Zompì, Ömer Demorik, Paloma Jeretič, Kate Mooney, Lucie Janků, Nina Haslinger, and Aaron Doliana (in no particular order) have all contributed thoughtful questions, criticisms, and/or relevant data to the issues discussed here. Thanks to Michal and Pavel for inviting me to two lovely weeks in Brno, in addition to the many other meetings and talks. Thanks to Yining for the years of sharing an office, of bento boxes, and for reading my earliest drafts. Thanks to Maria for many discussions of linguistic facts, theories, and the field at large, and the state of the world outside of that. Thanks to Naomi for subjecting all the theories to experiments, and for plenty of evenings of linguistics, politics, and literature. Thanks to Guy for being more skeptical of papers than me, and for discussions about a better world. Thanks to Sarah for all the projects that we have started and left unfinished. Thanks to Stan and Ömer for each hosting me multiple times in Boston, and for being friends that talk about morphosyntax too much. Thanks to Paloma for her unique way of combining thinking about small things and the big picture. Thanks to Kate for teaching me stuff about phonology that I actually wanted to learn. Thanks to Lucie for technical discussions where, for once, I got to be the one who got to be confused by phrasal spellout. Thanks to Nina who always sees the importance of any small detail. Thanks to Aaron for many years of friendship, linguistics, and investigations of funky corners of German.

The administrative staff at NYU have always done an amazing job at supporting us wherever they could. I don't think anything could get done in the department without Teresa Leung. Seena Berg, Teresa Colaizzo, and Star Jimenez, who were all always amazing. Thanks to Aura Holguin

who, in addition to her role as an administrator, also introduced me to youtube videos of keyboard sounds. Gillian Gallagher taught me unexpected things about anti-immigrant workplace harassment in her role as director of graduate studies.

It's hard to imagine anyone could make it through a PhD program without a supportive cohort – I certainly couldn't have. Thanks to Maxime Tulling, Maddie Gilbert, Suhail Matar, and Sandy Abu El Adas, variously for liquorice, for picketing, and for rigorously enforced birthdays. Thanks for listening to me about LaTeX and zotero, for remembering there is a world outside of research,for starting a fiction book club cut short by a pandemic, and (I suppose) most importantly: for commiserating.

Thanks to the many people who shared an office with me through the years – in addition to some of the people mentioned above, Soo-Hwan Lee, Alicia Parish, Chiara Repetti-Ludlow, and Dunja Veselinović have had to deal with mad ramblings about the syntax of very, very small things. I would also like to thank Daniel Duncan, Masha Esipova, Itamar Kastner, Mary Robinson, Ildiko Szabo, Daniel Szeredi for continuing occasional happy hours, and for their efforts in making the department a welcoming space before I had even arrived. Thanks to Omar Agha, Alex Warstadt, and Chris Klehm, for the joy of an unfinished python implementation of a Collins & Stabler style minimalism. Thanks to Kaustubh Ghoshal and Zhuoye Zhao, for taking care of my cats (thanks also to Pour Over & Simit!). Thanks to Alicia Chatten, Dylan Bumford, Jeff Parrott, Isaac Bleaman, José Álvarez Retamales, Kimberley Baxter, Philip Shushurin, Sheng-Fu Wang, Tanya Bondarenko, Sigwan Thivierge, Sonia Kasyanenko, Nicole Holliday, Nigel Flower, Vishal Sunil Arvindam.

Thanks to other faculty at NYU – thanks to Ailís Cournane for discussions about modality and the history of particular commodities, thanks to Anna Szabolcsi for making me think about semantics, thanks to Chris Collins for introducing me to Khoekhoe and discussing the syntax of German, thanks to Hadas Kotek for teaching me about questions and questions about about academia, thanks to Stephanie Harves, for many discussions about spans and pointers, for putting

a human face on things, for drinks, and for that poop-shaped stress-ball that expanded unexpectedly when left in the office during the pandemic. Thanks to Maria Gouskova for fun during morphbeer (whatever happened to that squib on knitting verbs?).

Thanks to my oldest friends, the ones that I've known quite a bit longer than I've known about linguistics, Kilian, Jens, and Daniel. Thanks to my family who supported me across an ocean, and who couldn't have expected that they would not see me for a whole global pandemic. Thanks to Claudius and Cornelius, Gerda, Reinhard, Hans-Jürgen.

In all likelihood, nobody has engaged more with the stuff in these pages than Adina. Thanks for all the linguisticing, but more importantly – thanks for always wanting more from life, and for wanting it with me.

Abstract

This dissertation is concerned with the nature of the basic building blocks of syntax, and their relation to the relevant pieces of phonology ("morphemes", or "exponents") – that is, the nature of *interpretation* at the PF interface. The investigation proceeds through a series of case studies, exploring and advancing the following three hypotheses. (1) The basic building blocks of syntax are features (as opposed to objects that *have* features). (2) The process that associates syntactic objects with phonology is can be characterized as a function, in the mathematical sense, which maps syntactic objects to phonological objects. The domain of this function ranges over syntactically derived objects (trees), rather than basic building blocks (heads). (3) the PF interface function contains sub-functions; each sub-function ranges over a tree and its sub-trees as its domain, and a particular phonology as its range. Sub-functions are selected in a target-maximizing, "best match" manner – that is, we try to externalize trees that are as big as possible, using the most specific function (that is, minimizing the number of many-to-one mappings).

Contents

List of Tables

1 Introduction

This dissertation is concerned with the nature of the basic building blocks of syntax, and their relation to the relevant pieces of phonology ("morphemes", or "exponents") – that is, the nature of "interpretation" at the PF interface. It investigates this relation through a series of case studies.

The starting point of the research conducted here, is the death of lexicalism. That is to say, I take it for granted that there is no dedicated mental machinery for the construction of words that would in turn feed into the syntactic machinery – instead all structure building, above and below the word level, is entirely syntactic (see in particular Marantz 1997 for a succinct argument, and Bruening 2018 for a recent summary of arguments against the lexicalist hypothesis). Since words are not the input to syntax, i.e., not its basic building blocks, they can no longer plausibly provide the immediate link between syntactic substance and phonological substance. This raises questions such as the following: What are the basic building blocks of syntax, if not words? What syntactic objects are associated with pieces of phonology? What is the nature of that association?[1]

Consider briefly, how lexicalist and post-lexicalist models – in particular, Distributed Morphology (Halle & Marantz, 1993) – have addressed the relation between syntactic building blocks and phonology. For a lexicalist model, the answers were clear and without much complication: Words are the basic building blocks of syntax. By definition, a word has phonology associated with it. It also must have a set of syntactic features that make the object legible to syntax, and

[1] That is, of course, not to say, that nobody has asked/answered these questions, much of this dissertation is essentially a dialogue with standard (or at least common) assumptions of Distributed Morphology (Halle & Marantz, 1993).

determine the word's syntactic behavior (say, the features that identify a word as a noun, make it animate, etc).

In what follows below, I will explore and advance the hypothesis that the basic building blocks of syntax are privative features, and exponents spell out syntactic phrases. That is, there is no distinction between a feature and a head, and the syntactic objects that get paired with pieces of phonology are generally larger than individual features.

Or, to put it another way, I will entertain the suspicion that two alternative hypotheses about the syntactic building blocks that are commonly endorsed by those who killed lexicalism (practitioners of Distributed Morphology, DM) may themselves be vestiges of lexicalism with unclear standing. First, that the syntactic formants *are* heads that *have* features (as opposed to features simply being the syntactic primitives that get assembled into larger structures by the usual syntactic operation, Merge). Second, that it is these building blocks of syntax that are a privileged syntactic object in that they are paired with phonology (as opposed to the spellout algorithm determining *which* syntactic objects get associated with phonology).[2] Note that both of these hypotheses about the syntactic building blocks follow by necessity from a lexicalist model: At the heart of lexicalism is the notion of the lexicon as the machinery that constructs the basic building blocks of syntax, words. And if words are the syntactic building blocks, the input to syntactic computation (say, as part of the enumeration), then the building blocks do by definition have phonology associated with them, and they necessarily do have features that make them legible to the syntax, i.e., features that determine their behavior as particular syntactic objects (such as the features that identify a word as a noun, make it animate, etc). If lexicalism had been right, it could not have been otherwise.

Lexicalism, however, was wrong. The word is not the basic building block of syntax. DM gave us the insight that the list of syntactic building blocks (DM's formants) is distinct from the list

[2]There is, of course, a class of theories that rejects internally complex heads and instead endorses the *one feature one head* hypothesis, without abandoning the notion that the basic building blocks are the object paired with phonology, e.g., Collins and Kayne (2020), Koopman (2017a), a variety of cartographic approaches, etc.

of syntax-phonology pairings (DM's vocabulary items). The hypotheses that the building blocks have features, but are distinct from them, and that the building blocks are the privileged syntactic object that gets paired with phonology (the object that is the target of vocabulary insertion) no longer follow from general principles, once lexicalism is jettisoned.

In the series of case studies that follows, I argue that we should jettison these hypotheses along with the lexicalism that generated them. For both conceptual and empirical reasons, we should instead strive to take Distributed Morphology to its radical conclusion: The syntax operates entirely on individual features that are combined by Merge, i.e., the linguistic system exhibits *hierarchical structure all the way down*, with the corrolary that externalization targets derived syntactic objects (trees), and not syntactic primitives (features).

While all chapters of this dissertation are ulimately concerned with the nature of syntactic structure building and its externalization, Chapter 2 (published as *Spans in South Caucasian Agreement: Revisiting the Pieces of Inflection* in Natural Language & Linguistic Theory 39(1), see Blix 2021b) and Chapter 3 could be more narrowly characterized as being focused the nature of *positions of exponence*. They are concerned with data like the ones in (1), or Table 1.1.

(1) a. ***gv**-xedav-da* 3SG>1PL

1PL.OBJ-see-IMPF.SBJ:3

'He saw **us**.'

b. ***g**-xedav-da-**t*** 3SG>2PL

2.OBJ-see-IMPF.SBJ:3-PL

'He saw **you$_{pl}$**.' GEORGIAN

Aronson (1990: 171)

About such data, they raise questions such as these: How come that first person plural objects in Georgian are marked in a single locus (the prefix), while second person plural objects are marked in two positions, a prefix for the person, and a suffix for the plurality? Or, in case of the Arabic data: How come that gender in the third person is marked prefixally, but gender in the second person is marked suffixally?

In classic DM, the syntactic primitive (a head) is the locus of vocabulary insertion – that is, by

Subject	3rd Person	2nd Person
ms	***y****-aktub-a*	*t-aktub-****a***
fs	***t****-aktub-a*	*t-aktub-****ii***

Table 1.1: Some Arabic Subjunctive Forms

hypothesis, a syntactic primitive correlates with a position of exponence. Or to put it another way: The relation between features and phonological objects is thought to be *templatic*, albeit in a more abstract way than in traditional approaches – a position of exponence is generally characterized by a particular set of features (say, person features, or person and number features, etc, depending on the particular hypothesis regarding that particular head). The arguments that chapters 2 and 3 engage in, begin with the fact that the Georgian and Arabic data above militate against such a view: If a head is a bundle of features that serves as the locus of vocabulary insertion and thus provides us with a position of exponence, then the data in (1a) suggests that person and number form a single head, while the data in (1b) suggests the opposite – that they are two distinct heads. I propose that we solve this conundrum by abandoning the idea that a head is the locus of insertion. In the radically *atemplatic* theory I propose, all features are merged individually, and the association with phonology targets derived syntactic objects in a target-maximizing fashion – whatever the (locally) biggest structure is that can be externalized/spelled out, will be spelled out; the position of exponence is dynamically derived by the algorithm that matches syntactic structure with phonological exponents. No mystery arises – the Georgian first person plural object is spelled out by a single exponent, because one exists, while no single exponent exists for the second person, which thus requires two exponents, as would be expected under an approach where syntax-phonology pairing is not mediated by the notion of an internally complex head.

The chapters take this basic step (let's jettison the head-as-locus-of-phonological-association hypothesis), and show that a variety of properties can be capture systematically that would otherwise require a list of redundant stipulations, or even run into contradictions.

Chapters 4 (published as *Phrasal Spellout and Partial Overwrite: On an alternative to backtrack-*

ing in Glossa 6(1), see Blix 2021a) changes gears somewhat, in that it is not so much concerned with *what* the object is that gets associated with phonology, but rather with how it is computed locally, how we might derive affixation from set theoretical notions and the details of Merge, and what kind of vocabulary items are required to cover some relevant empirical space. I propose a particular mode of conjoining two vocabulary items, called a *Pointer* that allows for a significantly simpler spellout algorithm that do competing proposals.

It is worth commenting briefly on a change in perspective that distinguishes Chapters 2 and 3 on the one hand, from Chapters 4 and 5 on the other: In the first two papers, the working hypothesis is that the object associated with phonology is a *span*, i.e., a set of heads that stand in a head-complement relations. In contrast, the latter two treat syntactic *phrases* as the relevant object. The papers take up, and modify a hypothesis advanced by Starke (2018) which employ's Kayne's (1994) Linear Correspondence Axiom to derive suffixation from a spec-head relation. That is, these two papers are concerned, among other things, with the possibility that spellout constraints some syntactic operations (enforcing the building of PF-legible objects, i.e., objects with corresponding vocabulary items). Nothing in Chapters 2 and 3 is in principle incompatible with such a view, but it is worth highlighting here that not all movement operations that I employ in those chapters follow immediately from a phrasal-spellout perspective either, suggesting at the very least that spellout-driven movement is not the only operation that is required.

Finally, Chapter 5 was published as *Interface Legibility and Nominal Classification: A Nanosyntactic Account of Kipsigis Singulatives* in Glossa 7(1), see Blix (2022). It is concerned with the nature of (un)interpretability, as well as apparent extra-syntactic information in the syntax-morphology interface (such as declension classes). It takes as its starting point a particular notion of nominal classification as an uninterpretable feature that was proposed by Kouneli (2020), and argues that we can derive the (un)interpretability of this feature directly from the notion of phrasal spellout: Whenever a feature is in a configuration that does not correspond to any vocabulary item, it is uninterpretable – that is, (un)interpretability is explained in terms of the PF interpretation

function directly.

While they do form part of a (mostly) coherent whole, the individual chapters that follow are each essentially self-contained. Thus, the reader should feel free to engage with them individually. It was a pleasure to think through, and write these chapters, I hope they are fun to read as well.

2 South Caucasian Agreement

2.1 Introduction

Georgian verbal agreement has figured prominently in linguistic theory, and has been variously analyzed from a morphological perspective (e.g. Anderson, 1992; Foley, 2017; Halle & Marantz, 1993; Stump, 2001), or a syntactic one (e.g. Béjar, 2003; Béjar & Rezac, 2009; Lomashvili & Harley, 2011; McGinnis, 2008, 2013). The morphological approaches generally focus on deriving the correct distribution of all affixes that co-vary with Tense/Aspect and/or phi-features. In contrast, the syntactic ones are concerned with the fact that the agreement paradigm exhibits a person asymmetry: Local (i.e., first/second person) object agreement always has a dedicated exponent, as in (2), whereas third person object agreement does not (3). In (2), first/second person objects are marked with the prefixes *m-* and *g-*, respectively, regardless of the properties of the subject.[1]

(2) a. ***g**-xedav-di* 1>2

OBJ.2-see-IMPF.SBJ:LOCAL

'I saw **you**$_{sg}$.'

b. ***m**-xedav-di* 2>1

OBJ.1-see-IMPF.SBJ:LOCAL

'You$_{sg}$ saw **me**.'

[1] I use the notation X>Y for a transitive agreement context (such that X denotes the phi-features of the subject and Y denotes those of the object), as well as syntactic selection/sisterhood.

c. ***g***-*xedav-da* 3>2

OBJ.2-see-IMPF.SBJ:3

'S/he saw **you**$_{sg}$.'

d. ***m***-*xedav-da* 3>1

OBJ.1-see-IMPF.SBJ:3

'S/he saw **me**.' GEORGIAN

Aronson (1990: 169ff)

In contrast, the exponent we find in the same morphological position when the object is third person *does* depend on the subject person, i.e., we find *v*-/*Ø*- if the subject is local and the object is third person, as shown in (3).

(3) a. ***v***-*xedav-di* 1>3

1>3-kill-IMPF.SBJ:LOCAL

'**I** saw **him/her**.'

b. ***Ø***-*xedav-di* 2>3

2>3-kill-IMPF.SBJ:LOCAL

'**You$_{sg}$** saw **him/her**.' GEORGIAN

Aronson (1990: 169ff)

I propose that *v*-/*Ø*- are portmanteau morphemes for 1>3 and 2>3 contexts, respectively, i.e., that they spell out subject agreement and object agreement simultaneously. From this novel perspective, a language like Georgian always agrees with both subject and object, and the person asymmetry is a PF-effect of the available vocabulary items that interpret the abstract syntactic structure. I argue that previous approaches are mistaken in assuming that a morphological position of exponence is immediately reflective of a syntactic head/probe. Instead, a position of exponence is an effect of PF interpreting syntactic structure in a cyclical, bottom-up, *spanwise* manner, and that *v*- and *Ø*- can spell out a span that includes the agreement of a third person object, as well as subject agreement. Under such a perspective, all agreement affixes spell out a contiguous span in a fixed Tense-Agreement hierarchy given in (4), where Tense and AGR stand in for "regions" of maximally simple heads with no internal structure.

(4) Tense > AGR$_S$ > AGR$_O$

Such a portmanteau approach, under which *v-*/*Ø-* spell out AGR_S and AGR_O simultaneously, accounts for the person-asymmetry in terms of *person containment*, the notion that first/second person structurally contain third person (Béjar & Rezac, 2009; Harley & Ritter, 2002a). I implement this containment in terms of syntactic structure, and couple it with the Nanosyntactic assumption that vocabulary items are *overspecified* for the span of syntactic heads they can spell out, i.e., that they match spans that are sub-spans of the ones they lexicalize (Starke, 2009). Since *v-*/*Ø-* are portmanteau morphemes for 1>3 and 2>3 contexts, respectively, they are *under*specified with respect to a context with local objects, given the person containment hypothesis. That is to say, they cannot match a span containing both subject agreement and local object agreement – and therefore the first/second person markers *g-* and *m-* for objects surface.

Two lines of evidence in favor of such a view will be provided: *i)* Georgian data that pertains to prefix/suffix interactions, and *ii)* comparative data from a closely related South Caucasian language, the Pazar dialect of Laz which sheds further light on such interactions, in particular with respect to the expression of number agreement.

The Georgian-internal evidence comes from complex interactions between the prefixal and the suffixal marking. These interactions receive no adequate explanation under previous analyses, but they do receive a principled account if number agreement is represented by a null/PL contrast above person in the internal structure of the AGR region, as in (5).

(5) *The AGR region, partly decomposed*

(PL) > PERSON

Expanding on the local subject and object agreement data from (2), consider the 2-by-2 paradigm in (6), which provides all four possible number configurations for a 2>1 agreement context: With a first person object, the object's plurality is marked at the prefix *gv-*, and subject plurality is marked suffixally by *-t*, with both markings being independent.

(6) a. *m-xedav-di* 2SG>1SG

1SG.OBJ-see-IMPF.SBJ:LOCAL

'You$_{sg}$ saw me. '

b. ***gv**-xedav-di* 2SG>1PL

1PL.OBJ-see-IMPF.SBJ:LOCAL

'You$_{sg}$ saw **us**.'

c. *m-xedav-di-**t*** 2PL>1SG

1SG.OBJ-see-IMPF.SBJ:LOCAL-PL

'**You$_{pl}$** saw me. '

d. ***gv**-xedav-di-**t*** 2PL>1PL

1PL.OBJ-see-IMPF.SBJ:LOCAL-PL

'**You$_{pl}$** saw **us**.' GEORGIAN

Aronson (1990: 171)

In contrast, a second person object is unable to mark number prefixally (7b) – we find the same prefix *g-* with both singular and plural second person objects. In this context, *omnivorous number agreement* arises, i.e., the same suffix, here *-t*, occurs when the subject, or the object, or both arguments are plural.

(7) a. *g-xedav-di*

2.OBJ-see-IMPF.SBJ:LOCAL

'She/he saw you$_{sg}$.'

b. *g-xedav-di-**t***

2.OBJ-kill-IMPF.SBJ:LOCAL-PL

i. 'I saw **you$_{pl}$**.'

ii. '**We** saw you$_{sg}$.'

iii. '**We** saw **you$_{pl}$**.' GEORGIAN

Aronson (1990: 171)

That is to say, a *number-insensitive* object prefix correlates with omnivorous number agreement arising in the suffix position. Comparing Georgian to Laz further corroborates this correlation: Unlike Georgian, Laz does not have a number sensitive first person object prefix, and consequently, omnivorous number *also* occurs with first person objects, as shown in (8).

(8) a. *ce-m-ç-i*

PV-1.OBJ-beat-PST.SBJ:LOCAL

'You$_{sg}$ beat me.'

b. *ce-**m**-ç-i-**t***

PV-1.OBJ-beat-PST.SBJ:LOCAL-PL

i. 'You$_{sg}$ beat **us**.'

ii. '**You$_{pl}$** beat me.'

iii. '**You$_{pl}$** beat **us**.' LAZ

In an approach where spellout operates on spans of contiguous heads in a bottom-up fashion, this is easily captured. For first person agreement, Georgian has a vocabulary item *gv-* that spells out PL together with the object person representation; the first person object agreement in (6c,d) can be represented as in (9). For both first person singular and plural object agreement, the suffix *-t* corresponds only to a later/higher cycle of spellout, and therefore indexes only subject plurality.

(9) *Georgian First Person Objects*

a. Tense > $\overbrace{\text{PL}_\text{S} > \text{PERSON}_\text{S}\text{:2}}^{\textit{-t}}$ > $\overbrace{\text{PERSON}_\text{O}\text{:1}}^{\textit{m-}}$ 2PL>1SG

b. Tense > $\overbrace{\text{PL}_\text{S} > \text{PERSON}_\text{S}\text{:2}}^{\textit{-t}}$ > $\overbrace{\text{PL}_\text{O} > \text{PERSON}_\text{O}\text{:1}}^{\textit{gv-}}$ 2PL>1PL

In contrast, in the second person, no such portmanteau vocabulary item spanning object person and object number is available. Spellout targets the largest span that a vocabulary item is available for. Therefore, only the person structure of the object – but not its the number structure – is spelled out in the first cycle. Consequently, PL$_\text{O}$ is left for a later cycle of spellout.

(10) *Georgian Second Person Objects*

a. Tense > $\overbrace{\text{PL}_\text{S} > \text{PERSON}_\text{S}\text{:1} > \text{PL}_\text{O}}^{\textit{-t}}$ > $\overbrace{\text{PERSON}_\text{O}\text{:2}}^{\textit{g-}}$ 1PL>2PL

b. Tense > $\overbrace{\text{PL}_\text{S} > \text{PERSON}_\text{S}\text{:1}}^{\textit{-t}}$ > $\overbrace{\text{PERSON}_\text{O}\text{:2}}^{\textit{g-}}$ 1PL>2SG

c. Tense > $\overbrace{\text{PERSON}_\text{S}\text{:1} > \text{PL}_\text{O}}^{\textit{-t}}$ > $\overbrace{\text{PERSON}_\text{O}\text{:2}}^{\textit{g-}}$ 1SG>2PL

The omnivorosity of suffixal *-t* comes from its overspecification for both PL nodes, but the bottom-up nature of spellout blocks it from realizing that potential in (9). In contrast, the lower PL_{O} with second person objects (and with first person objects in Laz) is never spelled out in the first cycle, and consequently omnivorous number effects arise, as in (10). The fact that Laz first person objects pattern with Georgian second person objects reduces to Laz lacking a vocabulary item corresponding to *gv-*. In other words, variation reduces to interpretation at the interface (Chomsky 1995).

Note that the spanning account of omnivorous number in (10) suggests that *-t* spans not only both plural nodes, but also the subject person agreement (being overspecified for the structural difference between first/second person). Indeed, Laz provides crucial, corroborating evidence: While it is the local object (prefix) that is relevant for determining *whether* omnivorous number agreement occurs, it is the subject person that determines *what form* this omnivorous number expression takes. As indicated in Table 2.1, there are two different omnivorous number patterns in Laz (shaded), and the occurrence of one or the other is determined by the person of the subject. Note further, that – as with the prefixes – we find a person-conditioned asymmetry between these two omnivorous number patterns: Only with third person subjects does the omnivorous expression of number co-vary with Tense, as evident from the contrast *-es*/*-an* in Table 2.1b. If the subject is local, however, the omnivorous number expression is independent of Tense, i.e., *-t* appears across Tenses in Table 2.1a. Crucially, for both these person asymmetries, the data can be described as exhibiting fusional morphology with third person, but not local arguments: For third person objects, the spellout depends on the subject, but local object spellout does not, and the spellout of omnivorous number with third person subjects depends on Tense, but omnivorous number with local subjects does not.

Again, the same kind of explanation as for the object agreement spellout will be advanced for omnivorous number: Since first/second person subject agreement is structurally larger than third person agreement, and *-es*/*-an* are not overspecified for the structurally larger local subject

Table 2.1: Laz – Omnivorous Number

a. Local Subject

	Subject			
	PST		PRS	
Object	2Sg	2Pl	2Sg	2Pl
1Sg	m- -i	m- -i-t	m- -Ø	m- -Ø-t
1Pl	m- -i-t	m- -i-t	m- -Ø-t	m- -Ø-t

b. Third Person Subject

	Subject			
	PST		PRS	
Object	3Sg	3Pl	3Sg	3Pl
1Sg	m- -u	m- -es	m- -s	m- -an
1Pl	m- -es	m- -es	m- -an	m- -an

agreement, they are blocked from spanning the whole structure with Tense.

All in all, four interrelated phenomena are in need of explanation: Two person-asymmetries with respect to fusional morphology, the paradigmatic distribution of omnivorous number, and finally, the variation between Laz and Georgian. Assuming that post-syntactic spellout targets the largest spans that a vocabulary item is available for, in a bottom-up manner, all of these are effects of *interpreting a fixed hierarchy*.

I argue against the family of approaches that link the first person asymmetry to various complex forms of the operation *Agree* and the derivation of Person-Case Constraints (in particular, the *Cyclic Agree* approaches in Béjar 2003, Béjar and Rezac 2009, though the critique extends to *Multiple Agree* approaches such as Nevins 2011). In these approaches, the properties of one DP systematically determine whether the other one can be found by a syntactic probe. I show that these approaches extend neither to the prefix-suffix interactions discussed above, nor to the other person asymmetry. Since the former makes these approaches empirically inadequate, and the latter shows that their explanatory aspect does not generalize, I argue that the *Agree*-based family

of approaches is an incorrect perspective on the properties of Georgian agreement. Against the competing morphological approach of Distributed Morphology (Halle & Marantz, 1993), I will argue that – while deriving the paradigms correctly – it offers no account for any of the explananda. It instead treats them as arbitrary results of post-syntactic structure modification and contextual deletion that fails to capture the systematic nature of these systems. I argue that DM's larger theoretical insights into the structure of grammar, and post-syntactic interpretative morphology are correct, but that a theory in which it is the Vocabulary that drives the bundling of heads into exponents is more restrictive and provides clearer explanations. In DM parlance, what I propose here is that *Fusion* is the only mechanism that is needed to understand Georgian, and that Fusion itself reduces to *Matching*, i.e., it is driven by how the Vocabulary interprets abstract syntactic structure, not by dedicated rules distinct from the vocabulary items.

With respect to the larger picture, this paper is part of a research program that proposes that complex heads do not exist as pre-syntactic objects (cf. e.g., Bobaljik 2012). Instead, all composition of multiple features is the result of binary Merge, down to the individual feature level. This paper also abandons the notion of a template, as it is retained in the subset-based matching approach of DM where one head provides one position of exponence. Since the templatic stance requires extensive structure manipulation by means such as Fusion or Fission, I argue that we are better off abandoning it in favor of a theory where positions of exponence result from matching vocabulary items to syntactic contexts.

The paper is organized as follows: Section 2.2 introduces the precise technology for the interpretation of syntactic structure at the PF interface. Section 2.3 introduces the Laz/Georgian agreement paradigms. Section 2.4 lays out the concrete analysis. It shows how the span-based account derives the four aforementioned explananda. After accounting for the distributional properties of all agreement sensitive affixes, Section 2.5 derives their linear distribution in terms of phrasal movement, span-based pied-piping, and Antisymmetry (Caha, 2009; Kayne, 1994, 2017; Koopman, 2017a; Starke, 2009), providing evidence in favor of such an analysis from the linear

positions of various markers of morphologically complex Tenses. Section 2.6 offers a comparison with previous accounts. Finally, Section 2.7 concludes.

2.2 Theoretical Background: Nanosyntax

Nanosyntax, like DM, assumes that morphology is at its heart syntactic, and that the object of study is the mechanism of post-syntactic interpretation of abstract syntactic structure. In the process of PF-interpretation – a process likely to be cyclically intertwined with syntax proper (a notion I briefly touch upon in Section 2.5, but largely leave aside) – syntactic structure is therefore translated into (morpho-)phonological structure. To this end, both theories assume that there are items in the post-syntactic vocabulary that match certain syntactic structures which they can translate. However, the notion of PF-interpretation in DM is *templatic*, insofar as internally complex heads provide positions of exponence. In contrast, Nanosyntax is *atemplatic*. Instead of internally complex formatives/heads, every formative is conceived of as maximally simple, with no internal structure, with *Merge* being the only combinatory device available in natural languages (Chomsky, 1995). To derive the effects of bundling and positions of exponence, vocabulary items (usually termed *lexicalized tree structures* (LTS) in Nanosyntax) are hypothesized to interpret contiguous pieces of syntactic structure, i.e., positions of exponence are effects of interpretation. Syncretisms are derived as the result of overspecification, i.e., matching is subject to a *Superset Principle*, such that an item matches certain substructures. This section introduces a formalization of such a system that is based on Starke (2009), Caha (2009), Pantcheva (2011), Taraldsen (2018), as well as Pantcheva and Caha (2012).[2]

[2]See also the collection of papers in Baunaz et al. (2018) for more recent work in this framework, Williams (2003) for a precursor that introduced a related notion of spanning, as well as Mirror Theory (Adger et al., 2009; Brody, 2000) for related ideas.

2.2.1 PF Interpretation

The *Superset Principle* (SP) governs the matching of exponents and a syntactic structure, i.e., it relates any item in the post-syntactic vocabulary to the set of syntactic structures that it can potentially spell out. The targets of spellout are contiguous spans of heads, as per (11). Two heads form a span, if the phrase headed by the lower head is the complement of the higher head, and the notion applies transitively.

(11) *Span*

An n-tuple of heads $< X_n, \ldots, X_1 >$ is a span in a syntactic structure S, if and only if X_{n-1}P is the complement of X_n in S.

Adapted from Taraldsen (2018: 90)

A vocabulary item itself *lexicalizes* such a span of heads, and such a lexicalized span of length n, $< X_n, \ldots, X_1 >$, characterizes a matching set of spans, namely the set of its subspans, as per (12). Matching is constrained by an *Anchoring* requirement (12a) that demands that the bottommost element of the vocabulary item, X_1, and the syntactic span subject to spellout be identical, as well as a *Contiguity* requirement (12b), requiring the sequence of heads to be identical between the VI and the syntactic structure.

(12) *Superset Principle* (Matching)

A Vocabulary Item that lexicalizes a span $< X_n, \ldots, X_1 >$ matches any syntactic span $< Y_m, \ldots, Y_1 >$, s.t.:

(i) $X_1 = Y_1$, and *Anchoring*

(ii) for any Y_p, s.t. $p < m$: $(X_p = Y_p) \rightarrow (X_{p+1} = Y_{p+1})$ *Contiguity*

Consequently, every vocabulary item lexicalizes a span, and a lexicalized span characterizes a set of syntactic spans that it matches, as in (13). A span $< X_3, X_2, X_1 >$ thus characterizes the set of

matching contexts $\{< X_3, X_2, X_1 >, < X_2, X_1 >, < X_1 >\}$ (i.e., if we think of a span as an ordered list, it matches the set of its tails), but, say $< X_3, X_2 >$ would not be matched because it is not properly anchored, and $< X_3, X_1 >$ would fail the contiguity requirement.

(13) *Subspan*

Any vocabulary item $< X_n, \ldots, X_1 >$ characterizes a set of contiguous spans $\{< X_m, \ldots, X_1 > \mid m \leq n\}$ that it matches.

A primary goal of this notion is to derive classes of possible syncretisms, and argue that such syncretisms are not mere morphological curiosities, but are in fact informative with respect the internal syntactic structures whose PF interpretations we observe: If two forms are syncretic, i.e., if two different structures receive the same phonological interpretation, then a containment relationship should hold between them.[3]

For illustrative purposes, let us consider a toy grammar of the English copula. This example is provided to acquaint the readers with the relevant technology in a familiar context, and is not intended to provide an actual analysis of the English tense/agreement system. We begin with the first and third person singular, as in Table 2.2. In the past tense singular, first and third person are marked surface identically by *was*. Keeping in mind the Superset Principle, we might

Table 2.2: English Copula: Past Tense, Singular

	SG
3	was
1	was
2	were

suspect there to be a containment relation between the first and third person agreement. For our illustrative purposes, let us abstract away from most of the material in the structure, as well as plausible further segmentation, and focus on the agreement features. Let us assume for the

[3]That is, modulo notions such as zero affixes, zero heads/operators, etc.

moment that person in English follows a containment hierarchy as in (14), a point based on Harley and Ritter (2002a) and Béjar and Rezac (2009) that I will return to in more detail below:

(14) [2 [1 [3]]]

For the 1=3 syncretism of *was*, it now suffices to postulate a vocabulary entry such as (15), which lexicalizes first person agreement, which in turn properly contains the agreement structure of third person.[4]

(15) *was* ⇔ '[1 [3]]'

The span that *was* lexicalizes characterizes both the first and the third person, i.e., [1[3]] as well as [3], deriving the syncretism from the Superset Principle, as in (16):

(16) a. $\overbrace{[\ 3\ ]}^{was}$ *Superset Principle*

b. $\overbrace{[\ 1\ [\ 3\ ]]}^{was}$

Both the third person structure [3] and first person structure [1[3]] have [3] as their bottommost element, thus fulfilling the Anchoring Requirement, and both are contiguous subspans of [1[3]], and therefore both can be matched by the toy entry for *was* in (15).

While the Superset Principle determines a set of syntactic contexts that a VI matches, a second pricinple, *Cyclic Overwrite* (CO), determines, *which* syntactic spans actually receive spellout. The basic idea behind this notion is that the spellout algorithm targets the largest spans that it can find vocabulary items for. To accomplish this, spellout is implemented in a way that parallels Merge (17a), overwriting its own results at any step (17b), until no further overwriting is possible, due to the lack of an appropriate vocabulary item. When this happens, the previous cycle is *effective*, and a new cycle begins, anchored at the head that could not be matched in the previous cycle.

[4] As a notational convention, I will use the usual bracketing structure for spans throughout this paper, avoiding the n-tuple notation, < . . . >, despite the fact that a span is not usually/necessarily a constituent.

(17) *Cyclic Overwrite*

a. Spellout operates in a cyclic, bottom-up fashion, with each application of *Merge* being paralleled by a spellout operation on the resulting span.

b. Any spellout operation that finds a matching VI in the lexicon overwrites the previous cycle that spelled out syntactic structure contained in the current node.

c. A spellout operation is *effective*, if the next cycle of spellout fails to find a matching VI. A new cycle begins, anchored at the next node.

This derives what has been dubbed the 'biggest wins theorem', namely that the size of the syntactic structure that receives an interpretation by a morphological object, depends solely on the size of a language's matching items. The spellout mechanism continues to overwrite its own results for increasingly larger parts, until no 'bigger' VI can be found.

To continue with our toy grammar for the English copula, we can now model the *was*/*were* contrast in terms of vocabulary item size, i.e., by arguing that *were* lexicalizes the second person structure [2[1[3]]], as in (18a), resulting in the spellout in (18b).

(18) a. *were* ⇔ '[2 [1 [3]]]'

b. $\overbrace{[2\underbrace{[\ 1\ [\ 3\ ]]}_{\textit{\sout{was}}}]}^{\textit{were}}$ *Cyclic Overwrite*

Note, however, that within our toy grammar, we now run into a potential conflict that finds parallels in other modules: Both *was* and *were* are now matching candidates for the spellout of the first and third person singular in (16). To this end, the largely uncontroversial concept of an Elsewhere Principle (EP), famously ascribed to the ancient Sanskrit grammarian Pāṇini (Kiparsky, 1973), is adopted, which states that if at any given point more than one rule of grammar has its conditions for application matched, the more specific one applies. In the spanning/superset terms developed so far, more specific simply equals fewer heads, as in (19).

(19) *Elsewhere Principle*

If multiple VIs match a given syntactic span, the VI with the fewest "unused" heads (heads that are part of the VI but not the matched syntactic span) wins.

Since *were* lexicalizes a head [2] that is not present in the agreement structure of first and third person, it will now lose out to *was* in these cases, but will still win out in case of the second person, thanks to cyclic overwrite.

Finally, a less widely adopted device will be used here, namely the notion of a *Pointer*, as developed by Caha and Pantcheva (2012) and Pantcheva and Caha (2012).[5] argue, certain classes of L-shaped syncretisms are highly prevalent across multi-dimensional paradigms, while others appear to be absent.[6] In the discussion of Laz and Georgian, various empirical phenomena will be re-cast as such L-shaped syncretisms, and shown to be derivable under the Pointer approach. For illustration, consider one such L-shaped syncretism, once again from our English past tense copula example, as in Table 2.3. We can see that the plural forms of all three person configurations are syncretic with the second person singular, i.e., that the syncretism extends across the two paradigm "dimensions" of person and number. Continuing with our toy example, let us assume, that the singular/plural contrast is encoded simply by the presence/absence of a PL(ural)

[5]The original purpose of Pointers is related to certain types of idioms (Michal Starke, p.c.), but the idea to employ it for the present purposes goes back to Pantcheva and Caha (2012) The basic idea of this notion (as adopted here) is to let vocabulary items lexicalize more than one (contiguous) span; the basic empirical aim is to account for certain classes of cross-categorial syncretisms, and the absence of other such syncretisms. As Pantcheva and Caha (2012),Caha.Pantcheva.2012

[6]Since their proposal has not been published, I will give an extremely short recapitulation here: Blansitt's (1988) generalization states that if a dative can be used for locative purposes in a language, then it can also be used for allative purposes. Caha and Pantcheva (2012) propose, based on a crosslinguistic study that these four cases correspond to the structures in (i).

(i) Case Structures

a.	Genitive	[GEN]	b.	Dative	[DAT[GEN]]
c.	Locative	[GEN[P]]	d.	Allative	[DAT[GEN[P]]]

Under the Pointer approach, the LOC=DAT=ALL syncretism arises when the case affix lexicalizes [DAT[GEN → [P]]]. Crucially, Blansitt's generalization falls out from this, since matching both the dative structure and the locative structure is only possible if the VI also lexicalizes the allative structure without "unused" heads. Unattested syncretisms like LOC=GEN=DAT≠ALL, or LOC=DAT≠ALL, on the other hand, cannot be derived under this approach, i.e., in contrast to a subset-based approach to syncretisms that allows for free cross-classification of two inependent systems, the Pointer approach is restrictive in a predictive fashion.

Table 2.3: English Copula: Past Tense

	SG	PL
3	was	were
1	was	were
2	were	were

head above the person structure. To capture the fact that *were* can spell out the second person plural is now easy enough: We simply assume that the item *were* is to be reanalyzed as *were* ⇔ '[PL[2[1[3]]]]', of which the second person singular, [2[1[3]]], is still a matching sub-span. The issue arises, however, with the other two plural structures, since in the absence of [2], PL and the lower person structure would no longer be contiguously matched. To account for this, we will allow a weaker matching requirement, with the person and number regions of the structure to be matched individually ([1[3]] is a subspan of the lexicalized person span [2[1[3]]]), as we saw before, but with the additional caveat that the two matched spans, person and number, have to be contiguous with respect to each other (i.e., [PL[1[3]]] still shows contiguity between the matched number span, and the matched person span). That is to say, the span in the syntactic structure that is matched is still a contiguous span in the structure. The contiguity requirement for matching, however, is relaxed, in case such a pointer is present: The Superset Principle in its core form gives rise to the possibility of the cross-number syncretism for the second person, and the Pointer gives rise to the possibility of cross-person syncretisms; in case we have both types of syncretism simultaneously (i.e., here, in case there is no smaller VI for second person singular, but there is one for first/third person singular), the syncretism will cross person and number, in an L-shape. A formal implementation of this idea is given in (20a). Note in passing that I will assume here, as per (20b), that Pointers have a syntactic correlate: They can occur only in places where heads select for a region – e.g., in our current example, PL selects for a person phrase of any size (first/second/third person), whereas [1] always selects for [3].

(20) *Pointers*

a. A VI that lexicalizes two spans X $< X_n, \ldots, X_1 >$, Y $< Y_m, \ldots, Y_1 >$ by means of a pointer X→Y matches any syntactic span that is formed by contiguity between a subspan characterized by X and a subspan characterized by Y (including the empty ones): $\{< X_q, \ldots, X_1 >, < Y_r, \ldots, Y_1 >, < X_q, \ldots, X_1, Y_r, \ldots, Y_1 > \mid q \leq n \wedge r \leq m\}$

b. A Pointer may occur in a lexical item under a head X_1 only in case the syntactic head X_1 c-selects for a region, rather than a specific head.

We thus revise our toy *were* in (18a) as (21a), which now matches the plural agreement structure of all person configurations. As indicated in (21b-d), the PL head (i.e., the minimal span PL) is contiguous with the relevant subspans of [2[1[3]]], and *were* thus matches all these structures. Note that it will cyclically overwrite *was* in the plural cases, as well as the second person singular, but not interfere with our analysis for the singular, as it still loses out to *was*, due to the EP.[7]

(21) a. *were* ⇔ '[PL → [2 [1 [3]]]]'

b. $\overbrace{[\text{PL}[3]]}^{were}$ c. $\overbrace{[\text{PL}[1[3]]]}^{were}$ d. $\overbrace{[\text{PL}[2[1[3]]]]}^{were}$

In the remainder of the paper, I will apply these tools to the agreement paradigms of Laz and Georgian, showing that such approach allows for an interpretation in which a complex agreement system is derived by "cutting" an essentially one-dimensional structure (a span is a linear sequence of heads) into different pieces. Since the paradigms vary along five dimensions (object person, object number, subject person, subject number, tense), this is a highly restrictive interpretation of the system.

[7]Note that the Pointer approach is limited in the kinds of cross-categorial syncretisms it allows for: If there is a cross-person syncretism in the plural, that forms an L-shape with one of the singular forms, it has to be the largest one, since a hypothetical item /α/ ⇔ '[PL → [3]]' could not apply in any person structure that is larger than the third person. The current system is thus more restrictive than a subset-based approach that allows independent cross-selection of the two systems.

2.3 Laz & Georgian: The Data

This section first introduces two verbal agreement paradigms from Laz, and discusses the distribution of the individual affixes. In light of that description, I will lay out the properties of the Georgian paradigms as interesting deviations from some of the regularities we find in Laz. A set of descriptive generalizations about the exponence of person and number agreement will be introduced as a set of explananda. In section 2.4, I will then show that these generalizations can be derived under the restrictive notions provided above, and that the set of explananda can be understood in terms of contiguity: Person asymmetries arise when the smaller third person agreement can be spelled out with higher material, but the larger first and second person cannot; omnivorous number effects arise when both PL nodes are spelled out in a single effective cycle.

Before I turn to the distributional properties of the affixes, however, a general property of the two languages under discussion should be pointed out: Unlike third person subjects, third person plural objects do not trigger plural agreement in either language. While the spellout theory to be presented below could be adapted to derive these facts in terms of a set of syncretisms, I believe this would be not only stipulative but a mistake for the following reason: Both Laz and Georgian show additional restrictions on third person subjects triggering plural agreement. In Laz, inanimate third person subjects can never trigger plural agreement, and in Georgian they do so optionally. Silverstein Scale/Animacy Hierarchy effects of this kind are cross-linguistically common in number marking, and often define a cut-off point below which number remains unmarked (Corbett, 2000, Chap. 3, 4). The same scale has also been argued to show interaction with grammatical function in e.g., Aissen (1999, 2003), Keine (2010), Kiparsky (2008). That is to say, the inability to mark number appears to be correlated with being low on the Silverstein Hierarchy and/or being low on a scale related to grammatical function, and the Laz/Georgian case appears to be an instantiation of that general phenomenon. Insofar as the system advanced here does not capture these effects, I will treat the inability of third person objects to trigger plural agreement

as a fact in need of independent explanation, rather than a morphological phenomenon of the same type as the ones discussed here. Consequently all tables will be presented with a single row for third person object agreement, without distinguishing the third person object number.

It is also worth noting that this paper discusses transitive agreement as its primary focus. Intransitives agree with their sole argument as subjects, and are morphologically identical to transitives with third person objects. I will assume that this is a default, and that they thus fall under the same morphological analysis.

2.3.1 Pazar Laz

Laz is a South Caucasian language spoken mostly in the Black Sea region of Turkey. The data described in this section have been gathered in fieldwork with a speaker of the Pazar dialect of Laz, but identical facts have been described for the Arhavi dialect (Lacroix 2009, chap. 9; Lacroix 2011, p. 80).[8] Since third person objects do not trigger plural agreement, as discussed above, I present all data with a single row for third person objects.[9,10]

I first discuss the affixes that we find across both paradigms, i.e., the Tense-invariant ones. These fall into two descriptive categories, the prefixes (*m-*, emfg-, *v-*, and *Ø-*), and one suffix (*-t*).

The local object prefixes *m-* and *g-* are biuniquely related to a first and second person object, respectively: Whenever the object is first person, there is a prefix *m-*, and whenever there is a prefix *m-* the object is first person; the same relation holds for *g-* and second person objects. Nei-

[8] On the Pazar dialect in particular, see also Öztürk and Pöchtrager, 2011. Note that it contains a small number of errors with respect to the agreement accessibility of arguments, corrected in Demirok (2013).

[9] The paradigms lack the cells that would correspond to first and second person reflexive forms. There are two ways to express reflexivity in Laz, either with reflexive pronouns, which uniformly trigger third person agreement, or with valency changing verbal morphology that makes the verb intransitive. First and second person can therefore never simultaneously trigger subject agreement and object agreement within a single verb form, i.e. the lacking cells are not a part of the language.

[10] Note that the present stem is derived with a thematic suffix that is sensitive to argument structure, thematic roles, and lexical aspect (cf. Öztürk & Erguvanlı Taylan, 2017; Öztürk & Pöchtrager, 2011). The corresponding set of suffixes in Georgian have been dubbed *present/future stem formants* (cf. Aronson 1990, p. 40 and Harris 1982). Given that these suffixes are sensitive to argument structure, I assume that these spell out of Voice heads, contextually conditioned by linearly adjacent Tense in the spirit of Embick (2015) Kastner (2018), but due to their invariance with respect to agreement, I abstract away from them here.

Table 2.4: Laz Verbal Agreement

Past Tense

	Subject					
Object	1sg	1pl	2sg	2pl	3sg	3pl
1sg	—	—	m- -i	m- -i-t	m- -u	m- -es
1pl	—	—	m- -i-t	m- -i-t	m- -es	m- -es
2sg	g- -i	g- -i-t	—	—	g- -u	g- -es
2pl	g- -i-t	g- -i-t	—	—	g- -es	g- -es
3	v- -i	v- -i-t	-i	-i-t	-u	-es

Present Tense

	Subject					
Object	1sg	1pl	2sg	2pl	3sg	3pl
1sg	—	—	m- -Ø	m- -Ø-t	m- -s	m- -an
1pl	—	—	m- -Ø-t	m- -Ø-t	m- -an	m- -an
2sg	g- -Ø	g- -Ø-t	—	—	g- -s	g- -an
2pl	g- -Ø-t	g- -Ø-t	—	—	g- -an	g- -an
3	v- -Ø	v- -Ø-t	-Ø	-Ø-t	-s	-an

[a] The distribution of affixes in past and present Tense is identical, with the exception of a present Tense counterpart of *-i*. To highlight this symmetry, the paradigm is given with a zero suffix.

ther prefix co-varies with Tense or number, and both are sensitive only to the object's respective person features.

The third overt prefix, *v-* (phonologically conditioned variants: *p-*, *p'-*, *b-*) and a null counterpart *Ø-* show a related kind of distribution. They are also invariant with respect to Tense and number, and can also be characterized as having a biunique relation to their contexts: The prefix *v-* occurs in all and only those cases where the subject is first person with a *concurrent* third person object, and in the same 2>3 contexts we find *Ø-*. This asymmetry between the local object prefixes on the one hand, and the subject-sensitive third person object prefixes constitutes our *first explanandum*, i.e., the prefixal alternation. We can characterize this as follows: If the object is

local, it is sufficient to know the properties of the object to determine the prefix, but if the object is third person, determining what prefix is used requires taking into account properties of both the subject, and the object.

The final affix that does not vary across Tenses is *-t*. This suffix marks plurality whenever the subject is local. Note that it is insensitive to the source of plurality, occurring whenever at least one agreeing argument is plural.[11] Number marking is therefore *omnivorous* in the sense of Nevins (2011).

When the subject is local, Tense is uniformly marked by *-i* (PST) and *-Ø* (PRS) respectively. When the subject is third person, however, the expression of Tense interacts with number: In SG > SG contexts, *-u* (PST) and *-s* (PRS) are used, but if at least one argument is plural, we find *-es* (PRS) and *-an* (PST), respectively. As with *-t*, number marking is omnivorous. Note that this can descriptively be characterized as recurrent L-shaped syncretism, which will play crucially into my account of these data. In the form that omnivorous number takes, we thus find another person asymmetry, and thus a *second explanandum*: With local subjects, the expression of omnivorous number is independent of Tense (uniformly *-t*), but with third person subjects, its expression depends on Tense (*-es* vs. *-an*). Note that these explananda are parallel: In both cases the exponence of third person is dependent on additional features (the subject in the first explanandum, tense in the second one) when compared to the expression of local arguments.

[11] Conversely, if an argument does not agree for person, it is also excluded from agreeing for number, as discussed by Demirok (2013, p. 79): In ditransitive constructions, it is the indirect object that triggers object agreement, and the direct object can mark neither number nor person:

(i) *nana-k* *ma* *t'k'va* *m-ots'ir-u/-*es*
mother-ERG 1SG.DAT 2PL.NOM 1.OBJ-show-3SG.PST/-*3PL.PST

'Mother showed you$_{pl}$ to me' Demirok (2013)

2.3.2 Georgian

While Laz exhibits omnivorous number effect uniformly, the Georgian paradigm in Table 2.5 parallels this fully only in the 1>2 corner, as indicated by the shaded L-shaped syncretism. Tables 2.4 (Laz) and 2.5 (Georgian), show the two language's paradigms to be different in 5 out of 22 paradigm cells, indicated in bold.[12]

Table 2.5: Georgian Verbal Agreement, Present/Future Tense (based on Aronson (1990, p. 169ff))

	Subject					
Object	1sg	1pl	2sg	2pl	3sg	3pl
1sg	—	—	m- -Ø	m- -Ø-t	m- -s	m- -en
1pl	—	—	**gv- -Ø**	**gv- -Ø-t**	**gv- -s**	**gv- -en**
2sg	g- -Ø	g- -Ø-t	—	—	g- -s	g- -en
2pl	g- -Ø-t	g- -Ø-t	—	—	**g- -t**	g- -en
3	v- -Ø	v- -Ø-t	-Ø	-Ø-t	-s	-en

The most obvious difference between the two paradigms of Laz and Georgian is the presence of an additional prefix *gv-* that occurs whenever the object is first person plural. As discussed in Section 2.1, this is accompanied by a disappearance of omnivorous number effects, when compared to Laz – with Georgian first person objects, object plurality is marked at the prefix only, and it is exclusively the subject's plurality that gets marked in suffixal position. The plural agreement of each argument is exponed independently and distinctly. While there are differences with respect to the prefixal exponence of *number*, however, the prefixal exponence of *person* is identical to the one we find in Laz, i.e., we find first/second person object prefixes that are bi-uniquely related to their object context, but with third person objects, we find an *v-*/*Ø-* contrast that depends on the subject.

[12] It should be noted that both languages show a second type of transitive agreement paradigm called *Inversion*. In these paradigms a dative subject triggers the kind of prefixal agreement usually found with objects. These are much more divergent between the two languages (cf. Öztürk and Pöchtrager 2011, p. 60ff and Aronson 1990 sections 10.1, 12.1), and I follow various other authors in excluding these here.

The second difference concerns the 3SG>2PL cell. There, the plural suffix *-t* occurs, whose distribution is limited to contexts with local subjects in Laz (which has *-an* in this cell) – that is to say, we see another instance of Georgian "breaking" the symmetry of the omnivorous plural marking that we see in Laz. Note further that the suffix *-s* that occurs in all other cells with a third person singular subject is absent in this context.

Omnivorous number effects do not completely disappear in 3PL>2 contexts, however: A second person plural object triggers *-t* if the subject is third person singular. A third person plural subject triggers *-en* in 3PL>2SG, but if both arguments are plural, the object's plurality can no longer be marked with *-t*. Instead, we find a syncretism between the 3PL>2SG and 3PL>2PL cells. That is to say, in the total absence of omnivorous number effects we would expect the 3PL>2PL form in Table 2.5 to be **g- -en-t*, or **g- -t-en*, with each suffix marking one argument's plurality. Instead, we find what can be characterized as a *conditional omnivorous number effect* with third person subjects and second person objects: Only if both the subject and the object are plural, can *-en* mark number omnivorously. In contrast, no similar effect appears in the 1>2 corner – here, the suffix *-t* has the same omnivorous distribution as in Laz, i.e., it is "simply" occurring if one or both of the arguments are plural. We thus have a *third explanandum*: The paradigmatic distribution of omnivorous number effects within Georgian, and its comparative distribution, i.e., its absence in certain parts of the Georgian paradigm, and its presence in the same parts of the Laz counterparts.

The distributional facts about the spellout of third person subject agreement are not fully identical in the imperfect, in Table 2.6. While the conditional omnivorous number effect for 3PL>2 can be found in both paradigms, the imperfect counterpart of *-s*, *-da*, co-occurs with the plural marker *-t* (bolded). All other distributional facts, however, are parallel to the ones in Table 2.5. Viewed through the lense provided by Laz, Georgian can be characerized by a breakdown of plural symmetry: There is an additional, number sensitive first person *plural* object prefix *gv-*, as well as a conditional omnivorous number effect in 3>2 contexts. Otherwise, however, the

Table 2.6: Georgian Verbal Agreement, Imperfect (based on Aronson (1990, p. 171))

	Subject					
Object	1sg	1pl	2sg	2pl	3sg	3pl
1sg	—	—	m- -di	m- -di-t	m- -da	m- -dnen
1pl	—	—	gv- -di	gv- -di-t	gv- -da	gv- -dnen
2sg	g- -di	g- -di-t	—	—	g- -da	g- -dnen
2pl	g- -di-t	g- -di-t	—	—	**g- -da-t**	g- -dnen
3	v- -di	v- -di-t	-di	-di-t	-da	-dnen

two languages are remarkably similar. They show the same kind of asymmetry with respect to dedicated markers for local objects, but not third person objects, and they both show Tense dependency of *omnivorous* plural spellout only if the subject is third person. We thus arrive at a *fourth explanandum*: The variation between the two agreement systems. In (22), I provide a list of the four explananda.

(22) *Explananda*

a. Dedicated local object exponents, but subject-dependent third person object exponents (Prefixal Alternation)

b. Omnivorous plural exponents co-vary with Tense, if the subject is third person, but are independent of Tense with local subjects

c. Distribution of omnivorous number effects

d. Laz/Georgian variation

In the next section I will employ the machinery introduced above to derive these.

2.4 Deriving the Paradigms

I now turn to the paradigmatic distribution of the agreement affixes introduced above, abstracting away from linear order, to which I return to in Section 2.5. I first motivate the structure that I assume spellout to be operating on. I then offer an in-depth analysis of Laz, focusing on the explananda introduced above. I then move on to Georgian, showing that the differences between Laz and Georgian can be modeled simply in terms of slight variations in the Vocabulary.

2.4.1 The Structure

Given the adopted perspective on spanning, it is informative to consider a few structures that certain affixes appear to spell out, since being spelled out by a single vocabulary item requires contiguity. From the bi-unique relation that the first/second person object affixes *m-* and *g-* have to their contexts, we can safely conclude that the object's person structure is, by itself, a contiguous span. Through that lens, consider the Georgian differences between first and second person objects in Table 2.7.

Table 2.7: Georgian – First/Second Person Objects

	Subject	
Object	1Sg	1Pl
2Sg	g- -Ø	g- -Ø-t
2Pl	g- -Ø-t	g- -Ø-t

	Subject	
Object	2Sg	2Pl
1Sg	m- -Ø	m- -Ø-t
1Pl	gv- -Ø	gv- -Ø-t

From the distribution of *gv-*, bi-uniquely related to first person plural objects, we can now conclude that, perhaps unsurprisingly, the object's person and number agreement structures likewise form a contiguous span, as in (23), where "–" denotes contiguity.

(23) *Contiguity (1)*

$\text{PL}_\text{O} - \text{PERSON}_\text{O}$

The more curious case is obviously the omnivorous number effect exemplified by *-t* in the 1>2 sub-paradigm in Table 2.7. Recall that Laz shows exactly that omnivorous distribution of *-t* throughout all local subject sub-paradigms.

Table 2.8: Laz – Tense (in)dependent Omnivorous Number (SBJ = local)

a. Tense Independent (SBJ = local)

	Subject			
	PST		PRS	
Object	1Sg	1Pl	1Sg	1Pl
2Sg	g- -i	g- -i-t	g- -Ø	g- -Ø-t
2Pl	g- -i-t	g- -i-t	g- -Ø-t	g- -Ø-t

b. Tense Dependent (SBJ = 3)

	Subject			
	PST		PRS	
Object	3Sg	3Pl	3Sg	3Pl
2Sg	g- -u	g- -es	g- -s	g- -an
2Pl	g- -es	g- -es	g- -an	g- -an

As Table 2.8 shows, an omnivorous number effect is also found with *-es*/*-an*, where the subject is third person, i.e., the contrast *-t* vs *-es*/*-an* depends on subject person. I will interpret omnivorous number effects as the ability of a VI to span a structure that includes the subject's and/or the object's number agreement structure in addition to the subject's person features. The largest structure that *-t* can spell out is thus at least as big as (24).

(24) *Contiguity (2)*

$\text{PL}_\text{S} - \text{PERSON}_\text{S} - \text{PL}_\text{O}$

Building on that, the *-an*/*-es* contrast in Table 2.8 is informative as well: Since both affixes show the same omnivorous number effect as *t*, they too must lexicalize the contiguous span in (24). Since the contrast between *-an* and *-es* itself, however, depends on Tense, they must also lexicalize

this part of the structure, thus extending (24) to (25). Note that PL_{obj} is contiguous with $\text{PERSON}_{\text{obj}}$, as per (23), as well as $\text{PERSON}_{\text{sbj}}$, and that Tense cannot break the contiguity of (24). The restrictive notion of contiguous matching thus forces us to conclude that Tense is at the PL_{s} edge, and that it cannot be at the edge where we find PL_{obj} (since we know PERSON_{o} to be found there, from (23)).

(25) *Contiguity (3)*

Tense – PL_{s} – PERSON_{s} – PL_{o}

We can now combine (23) with (25) to we arrive at a complete characterization of the contiguity involved with all five aspects, Tense, Number of Subject, Person of Subject, Number of Object, and Person of object, as in (26).

(26) *Contiguity (Final)*

TENSE – PL_{s} – PERSON_{s} – PL_{o} – PERSON_{o}

Finally, assuming that subject agreement is higher than object agreement, we can translate (26) into a complete ordering of the individual regions that are involved in the agreement morphology of Laz and Georgian (27).

(27) *The Structure*

$\text{TENSE} > \text{PL}_{\text{s}} > \text{PERSON}_{\text{s}} > \text{PL}_{\text{o}} > \text{PERSON}_{\text{o}}$

Next, let us consider the internal structure of the regions Tense/Number/Person that are involved in the agreement system's morphology. I follow (cf. e.g., Béjar, 2003; Béjar & Rezac, 2009; Harley & Ritter, 2002a, 2002b) in assuming that local person contains third person, and therefore an asymmetry arises: A vocabulary item that can spell out the agreement structure of a local argument also matches the agreement structure of a third person argument (due to the superset principle), but the inverse is not true. I will argue in detail below, that a third person object is spelled out in a span with other material, but first/second person objects are "too large", i.e., they

break the relevant contiguity with the higher material. Concretely, I will assume that languages vary parametrically in which local person is more marked, and that the structural containment re-interpretation of the (partial) feature geometry of Harley and Ritter (2002a) that is given in Table 2.9 holds. I assume that Laz and Georgian are ADD selecting languages, but nothing crucial hinges on this.[13]

Table 2.9: Person Specifications

Person	AUTH(or) selecting	ADD(ressee) selecting
3rd	[REF]	[REF]
2nd	[PART[REF]]	[ADD[PART[REF]]]
1sg	[AUTH[PART[REF]]]	[PART[REF]]

For Number, I simply assume a minimal privative alternation: A PL(ural) head is optionally merged on top of the person structure, and the absence of such a node results in default singular interpretation. In this sense, I adopt a version of Nevins' (2011) argument that number contrasts are encoded via presence/absence of material, but person is always specified. Here, that translates into assuming that a minimal person structure, REF, is always present, but the singular simply corresponds to the absence of PL.

As for Tense, I will refrain from making claims about its internal structure in this section, given that I do not intend to undertake an analysis of the Tense system here. I limit myself to placeholders such as PRS and PST for what is likely to be be internally complex structure.

An example of a complete structure is given in (28), which illustrates the contrast between omnivorous number marking and independent number marking internal to Georgian: Whether omnivorous number effects arise does now depend on the way the structure is "cut up" by the vocabulary items in the course of spellout – if the two PL nodes are spelled out in two independent

[13]The system I propose below can be implemented in either variant. However, since an AUTH selecting implementation requires an additional zero affix, the ADD selecting implementation is slightly more elegant. The relevant notions for the analysis, however are that there is a containment relation between participants, and that participant agreement properly contains the structure of third person agreement, i.e., independent of this choice. However, since both 1=3≠2 syncretisms and 2=3≠1 ones exist crosslinguistically (Cysouw, 2003), this is one way for the present system to accommodate these facts.

cycles, the result is independent number marking, if they are spelled out in a single cycle, i.e., by the same VI, omnivorous number may arise.

(28) a. $\overbrace{[\text{IMPF}}^{\textit{-di}}\overbrace{[\text{PL}_\text{S}[\text{ADD}_\text{S}[\text{PART}_\text{S}[\text{REF}_\text{S}}^{\textit{-t}}\overbrace{[\text{PL}_\text{O}[\text{PART}_\text{O}[\text{REF}_\text{O}]]]]]]]}^{\textit{gv-}}$ Georgian: 2PL>1PL

b. $\overbrace{[\text{IMPF}}^{\textit{-di}}\overbrace{[\text{PL}_\text{S}[\text{PART}_\text{S}[\text{REF}_\text{S}[\text{PL}_\text{O}}^{\textit{-t}}\overbrace{[\text{ADD}_\text{O}[\text{PART}_\text{O}[\text{REF}_\text{O}]]]]]]}^{\textit{g-}}$ Georgian: 1PL>2PL

In what follows, I show that we can characterize all affixes as spanning a contiguous part of this type of structure.[14]

2.4.2 Laz

For ease of reference, I repeat the past Tense paradigm of Laz as Table 2.10 here; affixes that do not covary with Tense are italicized.

2.4.2.1 Laz – Prefixal Alternation (First Explanandum)

In this subsection I account for the prefix alternation. I argue that local object agreement is spelled out on its own, but that *v-* and *Ø-* are portmanteaus for 1>3, 2>3 contexts, respectively, thus spelling out third person objects with higher material, an option that is blocked by PART_O

[14]The structure raises obvious questions as to its nature and the way it is constructed. One might interpret this structure either as the result of successive cyclic head movement of a set of person/number agreement heads (along the lines of Preminger 2011), or possibly Multiple Agree (Hiraiwa, 2005) with the resulting structure reflecting relative heights. Under such a view, the target of spellout would be spans within a complex head, formed by the syntax, and a probe would be an instruction to the syntax to build such a structure under a matching requirement that pertains to a specific syntactic configuration, such as c-command. Alternatively, the agreement structure might in fact be part of the extended projection of the verb, possibly heads that provide phi values to initially unvalued pronominal elements, along the lines of Kratzer (2009), or Stegovec (2019), which might imply that agreement is somewhat reminiscent of Sportiche (2005, 2006) style determiners in the extended projection of the verb. The subpart of the analysis in the current section is agnostic about this question, as long as the relevant structural containment relations hold, i.e., as long as a specific theory of agreement is compatible with the syntax building the kind of structure given in (27) , either view is compatible with the results derived here. In Section 2.5, I provide some arguments in favor of the latter hypothesis, based on morpheme order; the dependency, however, is asymmetric: While the account of linear order will crucially rely on phrasal movement, and therefore on these heads being part of phrasal syntax, the account of their paradigmatic distribution is independent of such an interpretation. I will largely leave the larger questions that this raises for the nature of Agree or agreement untouched, for now, and hope that future research may shed more light on these questions.

Table 2.10: Laz Verbal Agreement, Past Tense (repeated)

	Subject					
Object	1sg	1pl	2sg	2pl	3sg	3pl
1sg	—	—	*m-* -i	*m-* -i-*t*	*m-* -u	*m-* -es
1pl	—	—	*m-* -i-*t*	*m-* -i-*t*	*m-* -es	*m-* -es
2sg	*g-* -i	*g-* -i-*t*	—	—	*g-* -u	*g-* -es
2pl	*g-* -i-*t*	*g-* -i-*t*	—	—	*g-* -es	*g-* -es
3	*v-* -i	*v-* -i-*t*	-i	-i-*t*	-u	-es

Tense-invariant affixes italicized.

breaking the relevant contiguity, deriving the asymmetry. The possibility of such an asymmetry between the subject (in)dependent spellout of local/third person arises because objects are low, and local persons are structurally larger than third person; the Vocabulary of Laz exploits this asymmetry, giving rise to the prefixal alternation.

First, let us consider the case of *m-* and *g-*, both of which have a bi-unique relation to their contexts, i.e., the occur in all and only those contexts where the object is first or second person, respectively. Let us assume that they lexicalize precisely the structure that is specific to these contexts, and that there are no relevant competitors that could overwrite them.

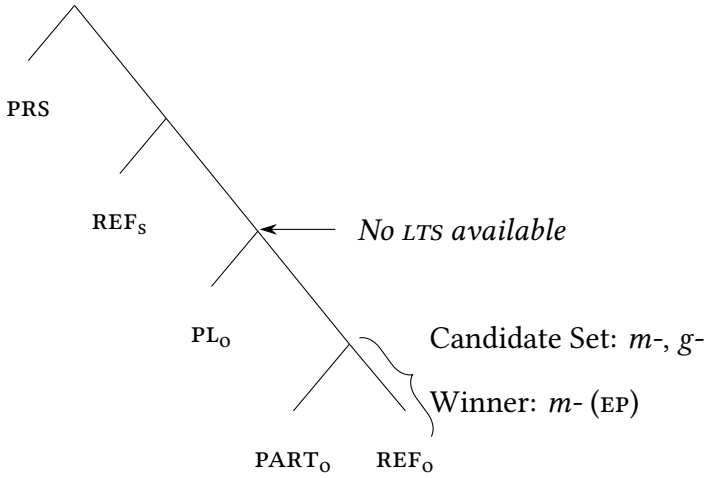
(29) a. *m-* ⇔ '[PART_O [REF_O]]'

b. *g-* ⇔ '[ADD_O [PART_O [REF_O]]]'

This trivially accounts for their distribution: Whenever we have first/second person object agreement, said agreement structure will be spelled out by *m-*/*g-*. This is exemplified in (30): Cyclic Overwrite tells us to spell out the largest structure that can be matched by a VI. The Vocabulary does not contain any item that could spell out the structure headed by PL_O, but it does have two items that match the one headed by PART_O, namely *m-* and *g-*. The elsewhere principle decides between the two candidates: Since *g-* contains an unused feature, ADD_O, but *m-* does not, the latter wins. The syntactic structure is marked as interpreted, and a second cycle of spellout can

begin, with PL_{O} as its bottommost element/anchor, as we will see in a moment. Note that this structure will provide a crucial point of comparison between Laz and Georgian, as I will argue that the latter does in fact have an affix that can lexicalize a first person plural object, *gv-*, thus starting the second cycle of spellout with a different bottom element.

(30) Laz: 3SG>1PL



Recall that *v-*, too, has a bi-unique relation to a specific context, and occurs in all and only 1>3 contexts. We may thus regard *v-* as a portmanteau morpheme, as encoded in (31a): It lexicalizes the person structure of both a third person object, and a first person subject. I follow previous analyses (Béjar, 2003; Béjar & Rezac, 2009; Halle & Marantz, 1993) in assuming that in addition to *v-*, there is a corresponding zero affix for second person subjects. Note that *Ø-*, unlike *v-*, contains a pointer, thus allowing it to be anchored at REF_{S}.[15]

[15]Phonologically null affixes are obviously motivated primarily on theory internal grounds. Georgian has an overt counterpart to *Ø-* in the copula's present Tense paradigm (Aronson, 1990, p. 66), thus providing independent evidence for such an affix.

(i) a. *v-ar* (1.SUBJ-COP) b. *x-ar* (2.SUBJ-COP)

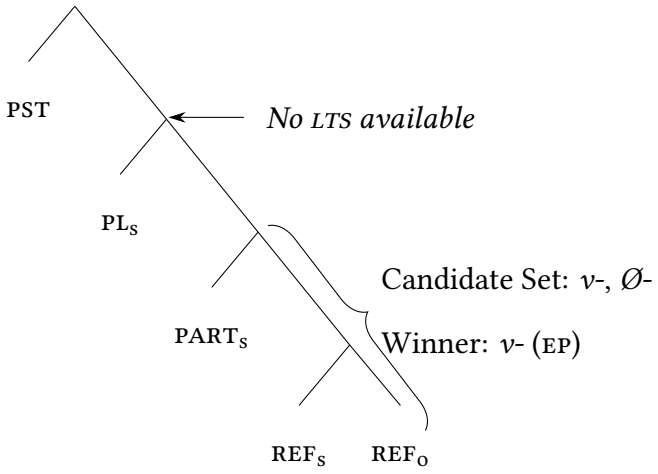
Note also that there is a rather curious prediction of this theory: This affix is able to spread into the 1SG>2SG cell, spelling out the subject features, although this could easily be avoided by postulating a second zero affix. I hope that, that these kinds of unexpected distributions will turn out to be useful in analyzing other complex agreement systems, but as it stands this is an unusual possibility predicted by the system.

(31) a. *v-* ⇔ '[PART_S [REF_S[REF_O]]]'

b. *Ø-* ⇔ '[ADD_S [PART_S [REF_S → [REF_O]]]]'

Consider a 1PL>3 context, as in (32), i.e the structure in which the portmanteau *v-* is found in the paradigm. The biggest structure that can be matched is the one headed by PART_S. Once more, two items lexicalize this structure, *Ø-* and *v-*, and the EP decides in favor of the smaller one of these, namely *v-*. Subject and object person are spelled out in tandem, since REF_S and REF_O are contiguous.

(32) Laz: 1PL>3



This derives the basic pattern of the prefix alternation: If the object is third person, material of the subject and the object are spelled out together; if the object is local, however, PART_O disrupts the relevant contiguity, and *m-*/*g-* instead spell out the object's person agreement as we saw above.

After seeing examples of the object-prefixes, as well as the portmanteau 1/2>3 ones, let us consider another simple case, 3SG>3, which shows only one affix, *-u* (*-s* in the present tense). Under the assumptions laid out in the previous section, the fact that there is a single affix, may be interpreted as *-u*/*-s* spelling out the whole Tense/Agreement structure, as in (33) – that is to say, these affixes, too, are potential portmanteaus.

(33) a. $\overbrace{[\textsc{pst}[\textsc{ref}_s[\textsc{ref}_o]]]}^{\textit{-u}}$ PST: 3SG>3 b. $\overbrace{[\textsc{prs}[\textsc{ref}_s[\textsc{ref}_o]]]}^{\textit{-s}}$ PRS: 3SG>3

However, *-u*/*-s* also occur with first/second person objects, in which case the resulting morphological spellout is bimorphemic. Given the VIs in (29), coupled with the fact that *-u*/*-s* occur in all 3SG>XSG contexts, we may further interpret this as *-u* spelling out either $[\textsc{pst}[\textsc{ref}_s[\textsc{ref}_o]]]$, as in (33) or $[\textsc{pst}[\textsc{ref}_s]]$, as in (34):

(34) a. $\overbrace{[\textsc{pst}[\textsc{ref}_s}^{\textit{-u}}\overbrace{[\textsc{add}[\textsc{part}[\textsc{ref}_o]]]]}^{\textit{g-}}]$ PST: 3SG>2SG

b. $\overbrace{[\textsc{pst}[\textsc{ref}_s}^{\textit{-u}}\overbrace{[\textsc{part}[\textsc{ref}_o]]]}^{\textit{m-}}]$ PST: 3SG>2SG

Under the current assumptions about spellout, the data leads us to the conclusion, that *-u*/*-s* lexicalize $\textsc{ref}_o$ under a pointer, and allowing them to match the two structures in (33) and (34), respectively:

(35) a. *-u* ⇔ '[PST [$\textsc{ref}_s$ → [$\textsc{ref}_o$]]]'

b. *-s* ⇔ '[PRS [$\textsc{ref}_s$ → [$\textsc{ref}_o$]]]'

Again, we see the same alternation: A local object is structurally big, and this triggers an additional cycle of spellout. The third person object, in contrast, can be spelled out together with subject material.

The account of the first explanandum – the fact that first/second person object agreement has dedicated exponents, while third person object exponence depends on the subject – is now complete, consisting of three parts. First, objects are low in the agreement structure, and thus subject to the first relevant cycle of spellout. Secondly, local objects correspond to a larger structure that properly contains the agreement structure of a third person object. Thirdly, given the structure I argued for, affixes may lexicalize $\textsc{ref}_o$ under a pointer, together with higher material, i.e., within a portmanteau. It is only when the object is local, that these specific portmanteau VIs

are blocked from effectively spelling out REF_{O} by cyclic overwrite – in the presence of an item that can spell out the larger object structure, they will be overwritten by these, i.e., neither *v-*, nor *-u*/*-s* are competitors for such a span. In the present system, the existence of portmanteau morphemes in the Vocabulary blocks a potential dedicated third person object marker from participating in the system: Even if a hypothetical VI existed that spelled out only REF_{O}, it would always be overwritten, and thus never surface. The same, however, does not apply to local objects, which block the relevant contiguity between REF_{O} and the subject agreement, and thus the VIs *m-* and *g-* always surface if the object is local. The explanatory load is thus divided: On the one hand, the syntactic structures that I argued for encode subject/object asymmetries in terms of height, and person asymmetries in terms of size, but it is the set of vocabulary items, i.e., the language specific PF-interpretation, that give rise to a system that exploits these asymmetries.

2.4.2.2 Laz – The Suffixes and Omnivorous Number

As we saw earlier, Laz exhibits omnivorous number effects throughout its agreement paradigm. This omnivorous expression of number is furthermore dependent on the subject person: If the subject is local, plural is spelled out by *-t*, regardless of Tense. If the subject is third person, another split comes about: In the past Tense, plural is spelled out by *-es*, and in the present Tense it is spelled out as *-an*. This subsection will provide a unified account of these facts.

Under the current approach, omnivorous number agreement can arise if a suffix is able to interpret a span that includes the subject's number and person, as well as the object's number – I now turn to the lexical entry of *-t* that will allow for such cycles of spellout. The lexical entry of *-t* must include local subject features. Given that we know it to occur with both first and second person features, i.e., given that *-t* can spell out both both the contexts in (36a) and (36b), we know that it can spell out PL_{S} even if it does not spell out ADD_{S} – that is to say, we see evidence for region-specific application of the superset principle. Therefore, the person structure must be under a pointer, for *-t* to be able to match both first and second person subject agreement.

Similarly, (36c) shows that the low PL_O must be under a pointer, as *-t* can spell out PL_S even when only the subject is plural, and *-t* is anchored at REF_S. I return to the spellout of PST by *-i* shortly, but indicate it here for completeness.

(36) a. $\overbrace{[\text{PST}}^{\text{-}i}\ \overbrace{[\text{PL}_\text{S}[\text{PART}_\text{S}[\text{REF}_\text{S}[\text{PL}_\text{O}}^{\text{-}t}\overbrace{[\text{ADD}_\text{S}[\text{PART}_\text{O}[\text{REF}_\text{O}]]]]]]]}^{g\text{-}}$ Laz: 1PL>2PL

b. $\overbrace{[\text{PST}}^{\text{-}i}\ \overbrace{[\text{PL}_\text{S}[\text{ADD}_\text{S}[\text{PART}_\text{S}[\text{REF}_\text{S}[\text{PL}_\text{O}}^{\text{-}t}\overbrace{[\text{PART}_\text{O}[\text{REF}_\text{O}]]]]]]]}^{m\text{-}}$ Laz: 2PL>1PL

c. $\overbrace{[\text{PST}}^{\text{-}i}\ \overbrace{[\text{PL}_\text{S}[\text{ADD}_\text{S}[\text{PART}_\text{S}[\text{REF}_\text{S}}^{\text{-}t}\overbrace{[\text{PART}_\text{O}[\text{REF}_\text{O}]]]]]]}^{m\text{-}}$ Laz: 2PL>1SG

The distributional facts about *-t* thus lead us to the structure in (37), allowing it to spell out a structure with either the subject, or the object, or both being plural.

(37) *-t* ⇔ '[PL_S ⟶ [ADD_S [PART_S [REF_S ⟶ [PL_O]]]]]'

Note that, crucially, the L-shaped syncretism only arises if we have a smaller competing affix, that occurs when neither argument is plural, and that *Ø-*, as defined above, fulfills this role, as exemplified in (38): In three out of four cases, i.e., the ones with a PL head, *-t* spells out the person structure. In (38a,b), *Ø-* cannot be anchored at PL_O, and thus does not compete. In (38c), *-t* overwrites *Ø-*, which cannot spell out PL_S, but in (38d), *Ø-* and *-t* are both candidates for the spellout of the subject person structure, and *Ø-*, the smaller affix, wins due to the elsewhere principle.

(38) *The L-shaped Syncretism with* -t

a. $\overbrace{[\text{PST}}^{\text{-}i}\ \overbrace{[\text{PL}_\text{S}[\text{ADD}_\text{S}[\text{PART}_\text{S}[\text{REF}_\text{S}[\text{PL}_\text{O}}^{\text{-}t}\ \overbrace{[\text{PART}_\text{O}[\text{REF}_\text{O}]]]]]]]}^{m\text{-}}$ Laz: 2PL>1PL

b. $\overbrace{[\text{PST}}^{\text{-}i}\ \overbrace{[\text{ADD}_\text{S}[\text{PART}_\text{S}[\text{REF}_\text{S}[\text{PL}_\text{O}}^{\text{-}t}\overbrace{[\text{PART}_\text{O}[\text{REF}_\text{O}]]]]]]}^{m\text{-}}$ Laz: 2SG>1PL

c. $\overbrace{[\text{PST}}^{\text{-}i}\ \overbrace{[\text{PL}_\text{S}[\text{ADD}_\text{S}[\text{PART}_\text{S}[\text{REF}_\text{S}}^{\text{-}t}\overbrace{[\text{PART}_\text{O}[\text{REF}_\text{O}]]]]]]}^{m\text{-}}$ Laz: 2PL>1SG

d. $\overbrace{[\textsc{pst}}^{\textit{-i}}\ \overbrace{[\textsc{add}_{\textsc{s}}[\textsc{part}_{\textsc{s}}[\textsc{ref}_{\textsc{s}}}^{\textit{Ø-}}\overbrace{[\textsc{part}_{\textsc{o}}[\textsc{ref}_{\textsc{o}}]]]]]}^{\textit{m-}}]$ Laz: 2SG>1SG

The Tense-dependent third person subject suffixes *-es* and *-an*, are omnivorous as well, and thus a parallel argument holds with respect to number: Given that we know the affixes to be able to occur with either or both of the arguments being plural, both $\textsc{pl}_{\textsc{s}}$ and $\textsc{pl}_{\textsc{o}}$ must have a pointer above them, in the former case under the respective Tense structures, given that we know these affixes to express Tense. Furthermore, paralleling the argument for *-u*/*-s*, *-es*/*-an* must be able to spell out $\textsc{ref}_{\textsc{o}}$, given a) that they are the only affix in the 3PL>3 context, and b) that we would otherwise expect *v-* to spread into these cells, since that is, so far, the smallest marker that is able to spell out $[\textsc{ref}_{\textsc{s}}[\textsc{ref}_{\textsc{o}}]]$. We therefore arrive at (39):[16]

(39) a. *-es* ⇔ '[PST → [$\textsc{pl}_{\textsc{s}}$ [$\textsc{ref}_{\textsc{s}}$ → [$\textsc{pl}_{\textsc{o}}$ → [$\textsc{ref}_{\textsc{o}}$]]]]]'

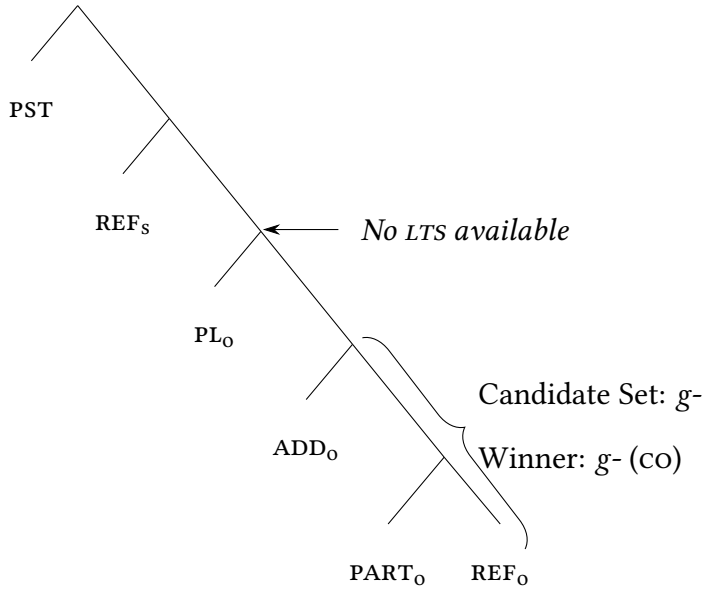
b. *-an* ⇔ '[PRS → [$\textsc{pl}_{\textsc{s}}$ [$\textsc{ref}_{\textsc{s}}$ → [$\textsc{pl}_{\textsc{o}}$ → [$\textsc{ref}_{\textsc{o}}$]]]]]'

[16]Note that these affixes embed a structure $[\textsc{pl}_{\textsc{o}} \rightarrow [\textsc{ref}_{\textsc{o}}]]$, despite my claim that a configuration $[\textsc{pl}_{\textsc{o}}\ [\textsc{ref}_{\textsc{o}}]]$ does not arise in transitive agreement. In the Laz inverse paradigms (discussed briefly in fn. 12), however, third person dative subjects do trigger plural object agreement. Whatever their precise syntax, their agreement morphology can therefore be captured in this system.

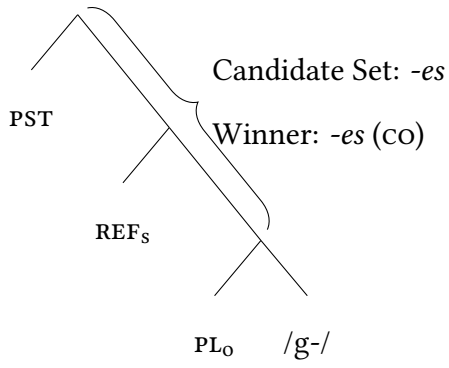
To see *-es* at play in spelling out a low plural feature, consider (40).

(40) Laz 3SG>2PL: *g- -es*

a. *First Effective Cycle*



b. *Second Effective Cycle*



The first effective cycle of spellout in (40a) operates as we have seen above, with *g-* spelling out the object's person structure. In (40b) we see that the second cycle of spellout becomes effective at the Tense node. There is no competing item that could lexicalize the same span: No other element that lexicalizes PST can also be anchored at PL$_O$, and we therefore successfully derive the

surface form *g- -es*.

As was the case for *-t* and *Ø-*, it is the competition with the smaller affixes, *-u*/*-s* that creates the L-shaped syncretism, and the remaining three cells from the 3>2 corner are given in (41). In the same fashion as above, *-u* only wins out when both affixes compete for the same structure, i.e., when both affixes can be anchored, and neither can overwrite the other. Only in this case does the Elsewhere Principle decide the competition, with *-u* becoming the effective spellout.

(40') *The L-shaped Syncretism with* -es (continued)

a. $\overbrace{[\textsc{pst}[\textsc{pl}_{\textsc{s}}[\textsc{ref}_{\textsc{s}}[\textsc{pl}_{\textsc{o}}}^{\textit{-es}}\overbrace{[\textsc{add}_{\textsc{o}}[\textsc{part}_{\textsc{o}}[\textsc{ref}_{\textsc{o}}]]]]]]]}^{\textit{g-}}$ Laz: 3PL>2PL

b. $\overbrace{[\textsc{pst}[\textsc{pl}_{\textsc{s}}[\textsc{ref}_{\textsc{s}}}^{\textit{-es}}\overbrace{[\textsc{add}_{\textsc{o}}[\textsc{part}_{\textsc{o}}[\textsc{ref}_{\textsc{o}}]]]]]}^{\textit{g-}}$ Laz: 3PL>2SG

c. $\overbrace{[\textsc{pst}[\textsc{ref}_{\textsc{s}}}^{\textit{-u}}\overbrace{[\textsc{add}_{\textsc{o}}[\textsc{part}_{\textsc{o}}[\textsc{ref}_{\textsc{o}}]]]]}^{\textit{g-}}$ Laz: 3SG>2SG

Recall that with local subjects, omnivorous number is not dependent on Tense, and that unlike the cases in (41) (which are bi-morphemic) with local subjects, we find a tri-morphemic agreement structure instead: Tense is spelled out in its own cycle, following the spellout of the phi-structure. The contrast in Tense is simply encoded by the affixes in (41):[17]

(41) a. *-i* ⇔ '[PST]'

b. *-Ø* ⇔ '[PRS]'

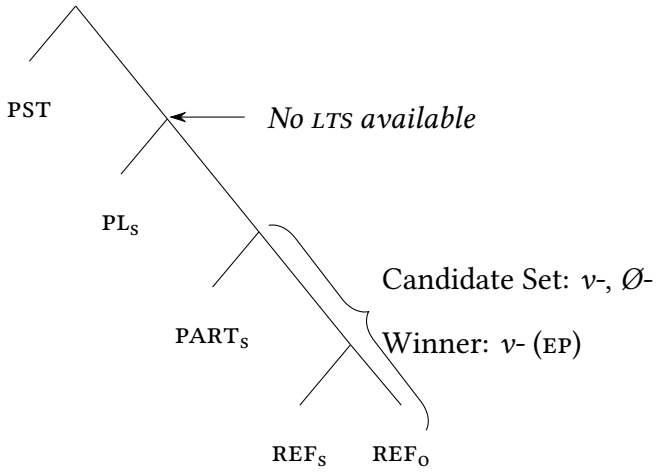
As a final illustration, now that all pieces are in place, consider the structure in (42), with a third person object and a first person plural subject. The structure headed by $\textsc{part}_{\textsc{s}}$ is the largest one that any item can spell out. As we saw in (32), we have two candidates, *Ø-* and *v-*, the winner of which is determined by the elsewhere principle. The second effective cycle of spellout targets

[17] An interesting alternative to the zero affix would be to argue that the past tense properly contains the present tense, and to specify the relevant affixes for the local subjects for PRS, but not PST, but the ones for third person for both. Under this perspective, *-i* would spell out only [PST], and we would not need another zero affix.

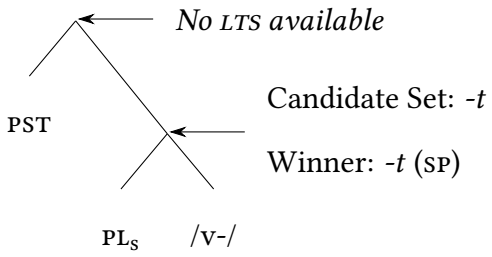
only PL_S. While *-es* does contain a span [PST[PL_S]], it is not a candidate due to the anchoring condition, and therefore no element is capable of spelling out a larger span. We therefore need three cycles of spellout, and since *-i* matches PST without any superfluous heads, the elsewhere principle decides in its favor, over *-es*, giving rise to the tri-morphemic result *v- -i-t*.

(42) Laz 1PL>3: *v- -i-t*

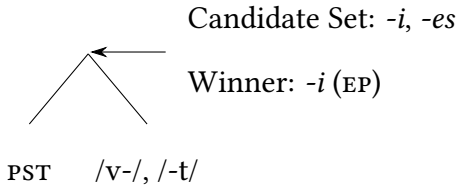
a. *First Effective Cycle*



b. *Second Effective Cycle*



c. *Third Effective Cycle*



This analysis gives an account of explanandum number two, i.e., the fact that omnivorous number

co-varies with Tense with third person subjects, but not with local subjects. With the latter, instead, Tense is spelled out by itself. Crucially, this is, once again, partly an effect of the lexical specifications of the VIs, but as was the case with the absence of dedicated third person markers, the structural claims restrict the space of possibilities. In fact, the hypothetical inverse of the current system, where omnivorous number co-varies with Tense only if the subject is local, but not if the subject is third person, is impossible to derive under the current assumptions.

To make this argument clear, consider the following: Firstly, we accounted for the fact that *-t* occurs with both first and second person, as well as omnivorously, by arguing that both subject person, and PL_O occur under a pointer. Since local persons contain the third person, it is thus compatible with all three subject persons, but since *-es*/*-an* spell out a larger span that includes Tense, this never comes to have effect if the subject is third person. The opposite split, however, would necessarily mean that the third person omnivorous number affixes were "tenseless", and the local subject ones were tensed. Consider three such hypothetical affixes, *t'*, *t"*, and *es'*.

(43) *Hypothetical Affixes*

a. *t'* ⇔ '[PST → [PL_S → [ADD_S [PART_S [REF_S → [PL_O]]]]]'

b. *t"* ⇔ '[PRS → [PL_S → [ADD_S [PART_S [REF_S → [PL_O]]]]]'

c. *es'* ⇔ '[PL_S [REF_S → [PL_O]]]'

Since the third person is properly contained in first/second person, the Tense-sensitive *t'*/*t"* would *always* overwrite *es'* at the Tense level, i.e. *es'* would never surface. No split would arise. The two third/local asymmetries are thus tied to the same structural properties, accounting for explananda one and two.

This, then, completes the account of the Laz agreement paradigms, with the affixes, "carving up" an agreement structure into pieces of varying sizes that are translated into morpho-phonological units, resulting in one to three positions of exponence. I showed that a complex, five-dimensional system can be interpreted as underlyingly one-dimensional, structured only by

the order of merge; a significant reduction in complexity. In doing so, I have accounted for two person-based splits, one pertaining to the presence of dedicated local object markers, and absence of dedicated third person object markers, and one pertaining to the co-variance of omnivorous number with tense. In the next subsection, I show that the analysis does not only carry over to Georgian, but that it offers an account of the remaining two explananda, namely the Laz/Georgian variation, and the distribution of omnivorous number effects within Georgian, and between the two languages. All that is needed is a slight variation in the Vocabulary of the two languages.

2.4.3 Georgian

Table 2.11 repeats the agreement paradigm, with affixes that do not co-vary with Tense italicized, and cells that differ from their Laz counterparts in bold.

Table 2.11: Georgian Verbal Agreement, Present/Future Tense (repeated)

	Subject					
Object	1sg	1pl	2sg	2pl	3sg	3pl
1sg	—	—	*m-* -Ø	*m-* -Ø-*t*	*m-* -s	*m-* -en
1pl	—	—	***gv-* -Ø**	***gv-* -Ø-*t***	***gv-* -s**	***gv-* -en**
2sg	g- -Ø	g- -Ø-*t*	—	—	*g-* -s	*g-* -en
2pl	*g-* -Ø-*t*	*g-* -Ø-*t*	—	—	***g-* -*t***	*g-* -en
3	*v-* -Ø	*v-* -Ø-*t*	-Ø	-Ø-*t*	-s	-en

Tense-invariant affixes italicized.

2.4.3.1 The Prefix Alternation & The Breakdown of Omnivorous Number

Recall that out of the five paradigm cells that show differences between the two languages, four concerned those with first person plural objects. Table 2.12 repeats a relevant part of the two language for an easy comparison. In Georgian, the subject/object symmetry in the spellout of number that we saw in Laz is broken down, when the object is first person: Subject and object

mark their plurality independently, and no omnivorous number effects arise. Crucially, *gv-* has the same one-to-one correspondence with a context that we saw with the other prefixes: It occurs in all contexts with first person plural objects, and only in these.

Table 2.12: First Person Objects

Laz			**Georgian**		
	Subject			Subject	
Object	2Sg	2Pl	Object	2Sg	2Pl
1Sg	m- -Ø	m- -Ø-t	1Sg	m- -Ø	m- -Ø-t
1Pl	m- -Ø-t	m- -Ø-t	1Pl	gv- -Ø	gv- -Ø-t

As before, I take this bi-unique relation to be indicative of the span *gv-* spells out, and thus arrive at the conclusion that Georgian *gv-* lexicalizes the structure of a first person plural object, as in (44a). Note that the prefixes in Georgian do otherwise work the same way, and are thus analyzed the same way as in Laz (44b-e).

(44) *Prefixes (Georgian)*

a. *gv-* ⇔ '[$_{\text{PL}_\text{O}}$ [$_{\text{PART}_\text{O}}$ [$_{\text{REF}_\text{O}}$]]]'

b. *m-* ⇔ '[$_{\text{PART}_\text{O}}$ [$_{\text{REF}_\text{O}}$]]'

c. *g-* ⇔ '[$_{\text{ADD}_\text{O}}$ [$_{\text{PART}_\text{O}}$ [$_{\text{REF}_\text{O}}$]]]'

d. *v-* ⇔ '[$_{\text{PART}_\text{S}}$ [$_{\text{REF}_\text{S}}$[$_{\text{REF}_\text{O}}$]]]'

e. *Ø-* ⇔ '[$_{\text{ADD}_\text{S}}$ [$_{\text{PART}_\text{S}}$ [$_{\text{REF}_\text{S}}$ →[$_{\text{REF}_\text{O}}$]]]]'

Consequently, we get the same alternation effect across all objects, with respect to *person*. With a first person object, however, *gv-* now breaks the omnivorous *number* effect – but not the person alternation. In Georgian, but not in Laz, a first person object's $_{\text{PL}_\text{O}}$ can be spelled out in the *first* cycle, as shown for the 2>1 cases in (45). Since $_{\text{PL}_\text{S}}$ is not interpreted in the same cycle as $_{\text{PL}_\text{O}}$, there is now no omnivorous number effect. When the object is first person plural (45a,b), *gv-*

overwrites *m-*. If the first person object is singular, however, as in (45c,d), *m-* wins, due to the elsewhere principle. Crucially, the subject plurality (45a,c) remains uninterpreted, regardless of the prefix, and thus, we get the potential for a four-way number contrast with first person objects in Georgian, but not in Laz. We have thus made a first step in accounting for the third and fourth explanandum, in accounting for the (variation in) distribution of omnivorous number effects.

(45) *Object-only Number Marking with* gv- *(First Cycle)*

a. $[_{\text{PRS}}[_{\text{PL}_\text{S}}[_{\text{ADD}_\text{S}}[_{\text{PART}_\text{S}}[_{\text{REF}_\text{S}}\overbrace{[_{\text{PL}_\text{O}}[_{\text{PART}_\text{O}}[_{\text{REF}_\text{O}}]]]}^{gv\text{-}}]]]]]]]$ Georgian: 2PL>1PL

b. $[_{\text{PRS}}[_{\text{ADD}_\text{S}}[_{\text{PART}_\text{S}}[_{\text{REF}_\text{S}}\overbrace{[_{\text{PL}_\text{O}}[_{\text{PART}_\text{O}}[_{\text{REF}_\text{O}}]]]}^{gv\text{-}}]]]]]]$ Georgian: 2SG>1PL

c. $[_{\text{PRS}}[_{\text{PL}_\text{S}}[_{\text{ADD}_\text{S}}[_{\text{PART}_\text{S}}[_{\text{REF}_\text{S}}\overbrace{[_{\text{PART}_\text{O}}[_{\text{REF}_\text{O}}]]]}^{m\text{-}}]]]]]]$ Georgian: 2PL>1SG

d. $[_{\text{PRS}}[_{\text{ADD}_\text{S}}[_{\text{PART}_\text{S}}[_{\text{REF}_\text{S}}\overbrace{[_{\text{PART}_\text{O}}[_{\text{REF}_\text{O}}]]]}^{m\text{-}}]]]]]$ Georgian: 2SG>1SG

The remaining properties of the prefixal alternations, however, remain the same as what we saw in Laz, with third person objects being spelled out with higher material, and local objects spelled out on their own. I now turn to the discussion of these cycles in Georgian.

2.4.3.2 The Suffixes – Plural Spellout in Georgian Continued

Against the background of the analysis of Laz given above, an analysis of the Georgian suffixal system faces multiple empirical points of difference that need to be accounted for: First, we need to show that the inclusion of *gv-* into the system does indeed account for the suffixal differences that accompany the prefixal ones. Secondly, the variation in the distribution of *-t* needs to be accounted for, since it is – unlike its Laz counterpart – not strictly limited to local subjects, but occurs in 3SG>2PL contexts. Finally, we need to account for the fact that the third person singular suffix *-s* does not occur in the present tense of 3SG>2PL *g- (*-s)-t*. Beginning with the plural suffixes, I will argue that Georgian *-t* is in fact identical to its Laz counterpart, but that the third

person subject suffixes *-s* and *-en* are not, giving rise to the variation.

Like Laz, Georgian has an omnivorous suffix *-t*. Unlike its Laz counterpart, however, we do not see its omnivorous distribution across the local subject corners of the paradigm. Instead, we find the omnivorous pattern of *-t* only in the 1>2 corner of the paradigm, as in Table 2.13.

Table 2.13: Georgian – Omnivorous *-t*

	Subject	
Object	1Sg	1Pl
2Sg	g- -Ø	g- -Ø-t
2Pl	g- -Ø-t	g- -Ø-t

As was the case for Laz, this shows us that *-t* can spell out PL_S and/or PL_O. As before, we conclude that PL_O is thus lexicalized under a pointer. In Laz, we furthermore concluded that it must also lexicalize a pointer between PL_S and the subject person structure, $[\text{ADD}_\text{S}[\text{PART}_\text{S}[\text{REF}_\text{S}]]]$, since it occurs with both first and second person subjects. The same argument applies in Georgian, as we can see from the 1>2 corner in Table 2.13 on the one hand, and the fact that it marks subject plurality in *m- -Ø-t* (2PL>1SG) and *gv- -Ø-t* (2PL>1PL) on the other. That is to say, like its Laz counterpart, it is able to spell out a first or second person subject structure, as well as the subject's and/or the object's person agreement structure – it is the same as its Laz counterpart.

(46) *-t* ⇔ '[PL_S ⟶ [ADD_S [PART_S [REF_S ⟶ [PL_O]]]]]'

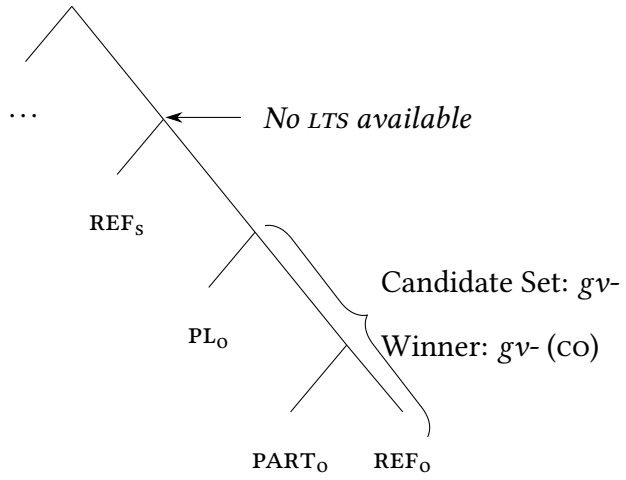
Let me illustrate that this does indeed derive the correct results for the 2>1 corner of the paradigm. Note that among cells with local subjects, only 2SG>1PL and 2PL>1PL are different from their Laz counterparts, and that we can now show how this follows purely from the presence of *gv-* in the Vocabulary. To demonstrate the full paradigms, I will also assume that, like Laz, Georgian has a zero affix for PRS, i.e., that (47) is part of the Georgian vocabulary, paralleling the Laz vocabulary item I introduced in (41b).

(47) -∅ ⇔ '[PRS]'

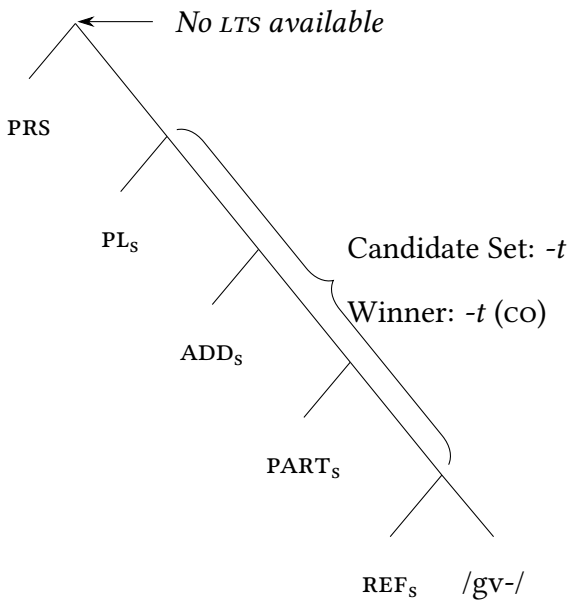
A structure in which both arguments mark their plurality independently, 2PL>1PL, is illustrated in example (48). In a first cycle, *gv-* spells out the structure headed by PL_O, since it is the biggest structure for which a corresponding VI can be found. The remaining structure to undergo spellout is the same as the one we would find with a first person singular object in Laz, and spellout proceeds accordingly, i.e., the subject structure is spelled out by *-t*. This immediately derives the fact that there is a suffixal syncretism for all X>1 forms, relative to any given subject X, i.e., that the number of a first person object does not bear on the suffixes: Given that both first person singular, and first person plural object agreement is spelled out in the first cycle, the second cycle is sensitive only to the subject's features and Tense, and therefore a first person object does not influence the suffixal agreement morphology (quite unlike Laz).

(48) Georgian: 2PL>1PL

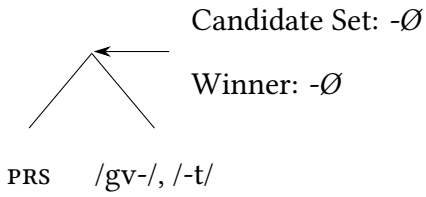
a. *First Effective Cycle*



b. *Second Effective Cycle*



c. *Third Effective Cycle*



The second cycle spells out the subject's person and number structure with the suffix -*t*, identical in specification to the one found in Laz. Completing the presentation in (45), the remainder of the 2>1 corner of Georgian is shown in (49):

(48') *Number Marking with* gv- *(continued)*

a. $[\overbrace{\text{PRS}}^{\text{-Ø}}\ \overbrace{[\text{ADD}_S[\text{PART}_S[\text{REF}_S}^{\text{Ø-}}\overbrace{[\text{PL}_O[\text{PART}_O[\text{REF}_O]]]]]]}^{\text{gv-}}]$ Georgian: 2SG>1PL

b. $[\overbrace{\text{PRS}}^{\text{Ø}}\ \overbrace{[\text{PL}_S[\text{ADD}_S[\text{PART}_S[\text{REF}_S}^{\text{-t}}\overbrace{[\text{PART}_O[\text{REF}_O]]]]]]}^{\text{m-}}]$ Georgian: 2PL>1SG

c. $[\overbrace{\text{PRS}}^{\text{-Ø}}\ \overbrace{[\text{ADD}_S[\text{PART}_S[\text{REF}_S}^{\text{Ø-}}\overbrace{[\text{PART}_O[\text{REF}_O]]]]]}^{\text{m-}}]$ Georgian: 2SG>1SG

Note again the dual work that the second person subject marker *Ø-* performs within the current system. It is important both in the first explanandum, i.e, the account of the prefixal alternation, and in the competition with *-t*.

All other cells of Georgian that have local subjects are identical to the Laz paradigm. We therefore turn to the second difference between the two languages, i.e., the fact that *-t* occurs in the 3SG>2PL cell in Georgian, but not Laz (Table 2.14). Note that this is in fact a predicted target of spellout for *-t*: Once again, it is the person containment that gives rise to this possibility – since *-t* lexicalizes a second person subject structure, and both first and third person are proper subsets of the second person, *-t* was always a candidate for the spellout of [REF_S[PL_O]]. In Laz, however, it is overwritten by *-an*/*-es* in all cases where the subject was third person.

Table 2.14: Second Person Objects

Laz			**Georgian**		
	Subject			Subject	
Object	3Sg	3Pl	Object	3Sg	3Pl
2Sg	g- -s	g- -an	2Sg	g- -s	g- -en
2Pl	g- -an	g- -an	2Pl	g- -t	g- -en

We infer that while *-an* in Laz is able to spell out the head PL_O in this structure, Georgian spells it out using *-t*. It is not the case, however, that Georgian *-en* simply cannot spell out such a head, as we see from the fact that there is a total syncretism between 3PL>2PL and 3PL>2SG, i.e., from the fact that there is no form **g- -en-t*, in which both arguments would mark their plurality independently, with *-t* marking the object's plurality, as in 3SG>2PL, and *-en* marking the subject's plurality, as it does in 3PL>2SG. I called this a *conditional omnivorous number effect*: Georgian *-en* can spell out a low plural feature only in case it *also* spells out a high one. Such a contiguity effect is in fact precisely the kind we expect under the approach with Pointers. The contrast between the two languages arises, because Laz *-an* requires weak contiguity, i.e., it contains a pointer under the Tense structure, whereas Georgian *-en* requires strict contiguity, i.e., it does not contain a

pointer, cf. (49) and (50a). That is to say, the contrast is exactly of the kind we expect, given that VIs may or may not have a pointer – a language specific property of their vocabulary.

(49) *Laz* -an *(repeated)*

-an ⇔ '[PRS → [PL_S[REF_S → [PL_O → [REF_O]]]]]'

(50) *Georgian*

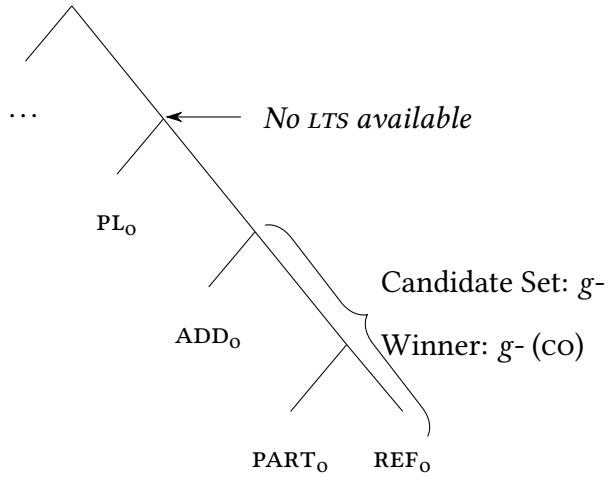
a. *-en* ⇔ '[PRS [PL_S[REF_S → [PL_O → [REF_O]]]]]'

b. *-dnen* ⇔ '[IMPF [PL_S[REF_S → [PL_O → [REF_O]]]]]'

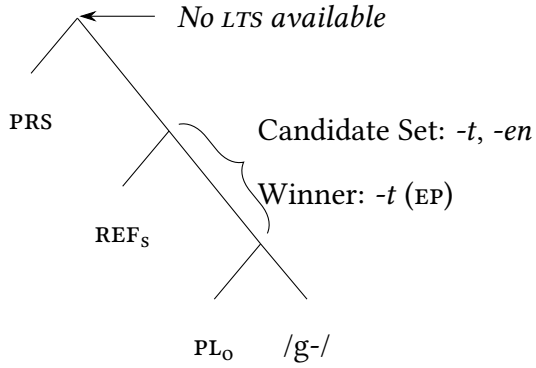
The system that accounted for the way omnivorous plural spellout is Tense dependent in Laz now gives us a handle on Georgian: The omnivorous number effect we found with Laz *-an* is the result of a pointer above the high plural feature, which allows the affix to spell out Tense with either one or both of the arguments being plural. Absent such a pointer, we find a system as the Georgian one, which can spell out the span containing the low plural head and Tense, only if there is strict contiguity between the Tense structure and the subject's phi structure, i.e., if the subject is plural, giving rise to the conditional omnivorous number effect. To see how this contrast between Laz and Georgian plays out, compare (51) and (52).

(51) Georgian 3SG>2PL

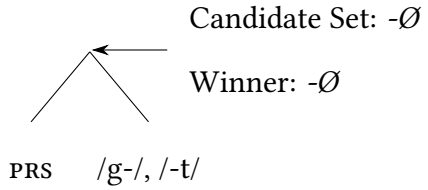
a. *First Effective Cycle*



b. *Second Effective Cycle*



c. *Third Effective Cycle*



(52) $\overbrace{[\text{PRS}[\text{REF}_S[\text{PL}_O}^{\textit{-an}}\overbrace{[\text{ADD}_O[\text{PART}_O[\text{REF}_O]]]]]]}^{\textit{g-}}$ Laz: 3SG>2PL

Both languages spell out the object person span in the first effective cycle, since *g-* is identical

in the two languages, and neither one has a competing or larger vocabulary item. In the second effective cycle, however, the two languages diverge. While Laz *-an* can spell out the complete remaining span, including the Tense node, and therefore overwrite a previous cycle in which *-t* was in the candidate set, Georgian *-en* cannot do so, as it requires contiguity between PRS and PL_S in order to be able to spell out PRS. It is therefore only a candidate for the structure [REF_S [PL_O]]. At this point it competes with *-t*, and loses out to it due to the elsewhere principle.[18] Unlike Laz, then, Georgian requires a third effective cycle of spellout to interpret the PRS structure. At this point, the PRS affix *-Ø* wins, completing the spellout. Note that this immediately explains the absence of *-s*, which occurs in all other cells with third person singular subjects: Its target is spelled out by *-t*.

In 3PL>2SG/PL, by contrast, PL_S and PRS are contiguous, and thus the spellout proceeds as it did in Laz, as indicated in (53a,b). Because *-en* can spell out the whole remainder of the structure if the subject is plural, including the Tense node, it cyclically overwrites *-t*. We therefore derive the *conditional omnivorous number effect*, i.e., the fact that omnivorous spellout of the object's plurality may depend on the subject also being plural, and account for the relevant difference between Laz and Georgian simply in terms of the presence vs absence of a Pointer in the lexical specification of a VI.[19]

[18] Note that in order for *-en* to lose out to *-t* due to the EP, we need to assume that *-t* is less specific than *-en*. I tacitly assume that this is due to the internal structure of the placeholder PRS.

[19] An interesting point about the occurrence of *-t* was brought to my attention by Thomas Wier (p.c.). Third person plural objects can exceptionally trigger plural agreement if they are focused and the subject is inanimate, as shown in (53). Note that a focused inanimate plural object appears to block a non-focused inanimate subject that would normally trigger plural agreement "optionally" from doing so (regardless of the object triggering plural agreement). In contrast, a human third person plural subject triggers *-en*, blocking *-t* (Léa Nash, p.c.).

(i) *mesame seri-is nak'tv-eb-s saerto punkcia a-ertianeb-(*-s)-t.*
third series-GEN form-PL-DAT common function.NOM PRV-unite-(-*3.SG.PRS)-PL

'A common function unites the forms of the third series.' Thomas Wier, Léa Nash (p.c.)

As with second person plural objects, we see that *-s* disappears, and *-t* occurs, when there is exceptional plural agreement with third person objects. From the current perspective, it is expected that [REF_S[PL_O]] receive spellout by *-t*, but it raises the question what spells out the object's person feature. They need to be spelled out independently, in order to block *-en* from being anchored before *-t* can be. At the present moment I can only give an ad-hoc stipulation, suggesting that the system can be made to work by employing a zero affix that spells out a structure along the lines [FOC[REF_O]], and that the element FOC bleeds *-en* (cf. a similar ad hoc solution in Halle and Marantz (1993) fn. 6,

(53) a. $\overbrace{[\text{PRS}[\text{PL}_\text{S}[\text{REF}_\text{S}[\text{PL}_\text{O}}^{\textit{-en}}\overbrace{[\text{ADD}_\text{O}[\text{PART}_\text{O}[\text{REF}_\text{O}]]]]]]]}^{\textit{g-}}$ Georgian: 3PL>2PL

b. $\overbrace{[\text{PRS}[\text{PL}_\text{S}[\text{REF}_\text{S}}^{\textit{-en}}\overbrace{[\text{ADD}_\text{O}[\text{PART}_\text{O}[\text{REF}_\text{O}]]]]]}^{\textit{g-}}$ Georgian: 3PL>2SG

c. $\underbrace{[\overbrace{\text{PRS}[}^{\textit{-Ø}}\ \underbrace{\text{REF}_\text{S}[\text{PL}_\text{O}}_{n/a}}^{\textit{-t}}\overbrace{[\text{ADD}_\text{O}[\text{PART}_\text{O}[\text{REF}_\text{O}]]]]]]}^{\textit{g-}}$ Georgian: 3SG>2PL

The analysis also shows a further divergence between the two languages. Note that the Laz 3SG>2PL case in (52), and likewise the Georgian cases in (53a,b) receive a bi-morphemic spellout. In contrast, the Georgian spellout of 3SG>2PL in (51)/(53c) is tri-morphemic. That is to say, the zero-affix *-Ø* ⇔ '[PRS]', which occurs only with local subjects in Laz (since there, PRS is spelled out with the subject structure if the subject is third person), "spreads" into the 3SG>2PL cell in Georgian, under the current analysis, giving rise to a total syncretism between 3SG>2PL and the omnivorous plural forms of 1>2, including the zero affix. I now turn to showing that this is indeed the correct analysis by looking at all overt counterparts to *-Ø*.

Consider, for instance, the optative, in which *-o* marks the TAM structure with local objects. As can be seen in (54), this suffix does indeed occur in both forms – i.e., its distribution is parallel to the one argued to be correct for *-Ø*. I conclude that the analysis of the absence of *-s* in the 3SG>2PL form as due to *-t* spelling out the subject person structure is on the right track, since it correctly predicts the presence/distribution of overt affixes from the analysis of the distribution of a zero affix, i.e., an affix whose distribution was accounted for by the theory without prior evidence.

(54) a. *g-nax-o-t* 3SG>2PL

2.OBJ-see-OPT-PL

'S/he intends to see you$_{pl}$.'

which suggests that third person arguments can occasionally pattern with participants due to an additional feature).

b. *g-nax-o-t* PL: 1>2

2.OBJ-see-OPT-PL

i. 'I intend to see you$_{pl}$.'

ii. 'We intend to see you$_{pl}$.

iii. 'We intend to see you$_{sg}$.

In fact, after developing the analysis of said zero affix, I went on to check whether this effect generalizes to all Georgian paradigms that include *-s*, and indeed it does – but with an interesting caveat: The zero affix and its overt counterparts do not only spread into the relevant 3SG>2PL cell, but occur with *-s* as well, i.e., they occur in the remainder of the relevant cells with third person singular subjects. Consider the data in Table 2.15, which shows that the suffix *-s* that is absent in the Georgian 3SG>2PL cases, does not – unlike its Laz counterparts – compete with the Tense marker, but instead, they systematically co-occur.[20]

Table 2.15: Georgian Suffix Triplets with *-s*

Local	3rd SG	3rd PL	Context	Source
-∅	-s	-en	Present/Future Tense	p. 42
-i	-i-s	-ian	Present/Future (Conj. 2 Verbs ending in *-am*)	p. 63
-de	-de-s	-dnen	Conjunctive (Conj. 1, 3)	p. 86
-ode	-ode-s	-odnen	Conjunctive (Conj. 2)	p. 86
-o	-o-s	-on	Optative (Conj. 1, Conj. 2 in *-i*)	p. 142
-e	-e-s	-nen	Optative (Conj. 2 in *-d*)	p. 142
-a	-a-s	-an	Optative (irregular verb *tkma* 'say'))	p. 210
-X	-X-s	-Y	Generalization	

Data from Aronson (1990)

This is easily accounted for under our theory by arguing that *-s* simply does not lexicalize Tense,

[20] In fact, some of the third person plural subject affixes appear to contain the TAM marker as well, e.g. the optative affixes *-o*, *-os* and *-on*. I will tentatively suggest that in these cases the *-n* spreads minimally into the TAM domain, but does not spell out all of it. If the vocabulary item *-o* contains a pointer to the same TAM structure *-n* spells out, below the remainder of the TAM domain, we derive the desired result. As this is not an investigation into the Georgian TAM structure, I will leave the precise formulation of this to future research.

i.e., it is a smaller affix than its Laz counterpart:[21]

(55) *-s* ⇔ '[REF_S ⟶ [REF_O]]'

We therefore conclude our analysis of the Georgian 3>2 corner with the following, tri-morphemic analysis.

(56) $\overbrace{[\ \text{PRS}}^{\textit{-Ø}}\ \overbrace{[\ \text{REF}_\text{S}}^{\textit{-s}}\ \overbrace{[\text{ADD}_\text{O}[\text{PART}_\text{O}[\text{REF}_\text{O}]]]]]}^{\textit{g-}}$ Georgian: 3SG>2SG

This analysis, then, concludes the account of explanandum number three: We have seen how the paradigmatic distribution of omnivorous number effects can be modeled in terms of cycles of spellout that are structured by the vocabulary items. I have given a precise characterization of the differences between Laz and Georgian with the identical abstract notions of PF-interpretation, but with a variation in the language specific vocabulary items: Georgian has an additional vocabulary item *gv-* that can spell out PL_O in the first cycle, in case the object is first person, and it lacks a pointer in *-en* that blocks it from overwriting *-t*, unless the subject is third person plural. Both of these result in a breakdown of the omnivorous spellout of number, in the first case by causing PL_O and PL_S to be spelled out in different cycles, in the second case by causing PL_O to be spelled out in spans of different sizes, depending on the presence/absence of PL_S.

I turn to the Imperfect paradigm. While the paradigmatic distribution of the third person plural suffix *-dnen* and the occurrence of *-t* is identical to the previously discussed data, the "bleeding" of *-s* by *-t* in 3SG>2PL has no counterpart in the imperfect, as indicated in the repeated Table 2.16.

Given the discussion so far, the current system runs into an issue: Considering the fact that we just captured the present tense syncretism between 1SG>2PL and 3SG>2PL by arguing that it arises from the fact that *-t* spells out REF_S in this case, this cannot be modeled in the precise terms

[21]Note that this variation is closely linked to another point of variation, namely the Pointer in Laz *-an*, and its respective absence in Georgian *-en*: It is only because *-en* cannot spell out the tense structure unless it also spells out PL_S that allows a tenseless *-s* to surface. Were the Georgian *-en* like the Laz *-an*, it would always overwrite a tenseless *-s*, even in the absence of any plurality. The tenseless *-s*, i.e., the *-s* that always co-occurs with the kind of tense marker that is limited to local subjects in Laz, is only possible with the pointerless *-en*, and the resulting conditional omnivorous number effect.

Table 2.16: Georgian Verbal Agreement, Imperfect (based on Aronson (1990, p. 171))

	Subject					
Object	1sg	1pl	2sg	2pl	3sg	3pl
1sg	—	—	m- -di	m- -di-t	m- -da	m- -dnen
1pl	—	—	gv- -di	gv- -di-t	gv- -da	gv- -dnen
2sg	g- -di	g- -di-t	—	—	g- -da	g- -dnen
2pl	g- -di-t	g- -di-t	—	—	g- -da-t	g- -dnen
3	v- -di	v- -di-t	-di	-di-t	-da	-dnen

developed so far. In the present tense, the system derived the fact that *-s* is present in all 3sg>x contexts, except for the 3sg>2pl case, and in doing so derived the sycretism with 1sg>2pl as in (57).

(57) *Present Tense*

a. $\overbrace{[\text{PRS}}^{\text{-Ø}}\ \overbrace{[\text{PART}_\text{S}[\text{REF}_\text{S}[\text{PL}_\text{O}}^{\text{-t}}\ \overbrace{[\text{ADD}_\text{S}[\text{PART}_\text{O}[\text{REF}_\text{O}]]]]]]}^{\text{g-}}$ Georgian: 1sg>2pl

b. $\overbrace{[\ \text{PRS}[}^{\text{-Ø}}\ \overbrace{\text{REF}_\text{S}[\text{PL}_\text{O}}^{\text{-t}}\ \overbrace{[\text{ADD}_\text{O}[\text{PART}_\text{O}[\text{REF}_\text{O}]]]]]]}^{\text{g-}}$ Georgian: 3sg>2pl

In the Imperfect, however, we do *not* find the same syncretism, instead we find *g- -di-t* in 1sg>2pl, and *g- -da-t* in 3sg>2pl, and *-da* across all cases with third person singular subject. Since we still find *g-* and *-t* in both cases, we may conclude that the first two effective cycles are identical between the present and the imperfective, as expected, given the bottom-up nature of spellout. However, as (58) indicates, this leaves the same structure for spellout in both cases, and yet we find different surface forms.

(58) *Imperfect*

a. $[\text{IMPF}\overbrace{[\text{PART}_\text{S}[\text{REF}_\text{S}[\text{PL}_\text{O}}^{\text{-t}}\ \overbrace{[\text{ADD}_\text{S}[\text{PART}_\text{O}[\text{REF}_\text{O}]]]]]]}^{\text{g-}}$ Georgian: 1sg>2pl

b. $[\text{IMPF}\overbrace{[\text{REF}_\text{S}[\text{PL}_\text{O}}^{\text{-t}}\ \overbrace{[\text{ADD}_\text{O}[\text{PART}_\text{O}[\text{REF}_\text{O}]]]]]]}^{\text{g-}}$ Georgian: 3sg>2pl

Pending a more detailed analysis of the TAM system, I propose that *-di* and *-da* are two contextually determined phonological forms of the same vocabulary item. In particular, I suggest here that a contextual allomorphy rule exists that makes reference to the sister of IMPF, as in (59).[22] Under such a proposal, the resulting form will be *-di* whenever the subject is local, but *-da*, whenever it is third person singular, i.e., whenever the sister of IMPF is REF_S, i.e., exactly the distribution of *-di*/*-da* that we find in Table 2.16. Note that *-s* is thus analysed as a smaller affix than the Laz counterpart, but that *-di*/*-da* does include Tense.

(59) *-di* ⇔ '[IMPF → [REF_S →[REF_O]]]' | /i/→/a/, if the sister of IMPF is headed by REF_S

Crucially, this analysis suggests that the spellout of third person subjects with objects other than second person plural will result in a tri/bi-morphemic spellout (depending on the person of the object) in the present tense type paradigms, but a bi/mono-morphemic spellout in imperfect-type paradigms. This is indicated in (60):

(60) a. PRS-type, *bi-morphemic*

$$\overbrace{[\text{PRS}}^{\text{-Ø}}\ \overbrace{[\text{REF}_\text{S}[\text{REF}_\text{O}]]]]}^{\text{-s}}$$

Georgian: 3SG>3

b. IMPF-type, *mono-morphemic*

$$\overbrace{[\text{IMPF}[\text{REF}_\text{S}[\text{REF}_\text{O}]]]]}^{\text{-da}}$$

Georgian: 3SG>3

(61) a. PRS-type, *tri-morphemic*

$$\overbrace{[\text{PRS}}^{\text{-Ø}}\overbrace{[\ \text{REF}_\text{S}}^{\text{-s}}\overbrace{[\ \text{PART}_\text{O}[\text{REF}_\text{O}]]]]]]]}^{\text{m-}}$$

Georgian: 3SG>1SG

b. IMPF-type, *bi-morphemic*

$$\overbrace{[\text{IMPF}[\text{REF}_\text{S}}^{\text{-da}}\overbrace{[\text{PART}_\text{O}[\text{REF}_\text{O}]]]]]]}^{\text{m-}}$$

Georgian: 3SG>1SG

[22] Presumably, *-di* itself is bimorphemic, with *-d* spelling out some higher structure. I once again abstract away from the details of TAM.

We saw earlier, from the data in Table 2.15 that in the PRS-type paradigms that include *-s*, the third person subject forms always include the same Tense marker as the cells with local subjects, and that the overt counterparts of *-Ø* thus provided an argument in favor of the "small" analysis of *-s*. Looking at the complement set of those paradigms, we can see that this is never true in the other triplets/screeves, as shown in Table 2.17.

Table 2.17: Georgian Suffix Triplets without *-s*

Local	3rd SG	3rd PL	Context	Source
-i	-a	-ian	Present/Future Tense (Conj. 2)	p. 61
-d-i	-d-a	-d-nen	Conditional/Imperfect	p. 45
-od-i	-od-a	-od-nen	Conditional/Imperfect (Conj. 2)	p. 65
-e	-a	-es	Aorist (Conj. 1)	p. 113
-e	-a	-nen	Aorist (Conj. 2 ending in *-i*)	p. 113
-i	-a	-nen	Aorist (Conj. 2 ending in *-d*)	p. 115
-i	-a	-es	Aorist (irregular verb *tkma* 'say')	p. 210
-X-e/-i	X-a	X-Y	Generalization	

Data from Aronson (1990)

In fact, we always find the same pattern, in which /e/ or /i/ alternate with /a/. While no theory of the TAM system will be offered here, this suggests that it might be possible to capture all of these paradigms with a single allomorphic variation, in which a [-low] feature is changed to [+low] under the relevant sisterhood relation. Crucially, for the present purposes, this analysis binds two things together. First, the fact that all paradigms that include the third person subject affix *-s* show the complementary distribution of *-t* and *-s* that we analysed as a result of *-t* spelling out REF_S, *and* the structure *X-s* for these suffixes, where X is the same TAM suffix we find with local subjects. Second, the paradigms that do *not* show the complementary distribution effect with *-t* also never have this bi-morphemic suffixal structure.[23]

[23]Note that the selection of the set of suffixes appears to depend on conjugation class, the phonological form of the stem, or irregular verbs. This does not seem to lend itself to an analysis in terms of spans, as advanced here. Whether an extension of an allomorphy approach (a notion that most Nanosyntax eschews), is a plausible way of accounting for these further details remains to be seen.

2.4.4 Interim Summary & Discussion

In interpreting a five-dimensional system in a spanwise (i.e., one-dimensional) fashion, a significant reduction in complexity was achieved. In pursuing this restrictive hypothesis, we derived various complex interactions between cycles of spellout, and accounted for the four explananda, repeated in (62).

(62) *Explananda* (repeated)

a. Dedicated local object exponents, but subject-dependent third person object exponents (Prefixal Alternation)

b. Omnivorous plural exponents co-vary with Tense, if the subject is third person, but are independent of Tense with local subjects

c. Distribution of omnivorous number effects

d. Laz/Georgian variation

I have shown that the "person-driven" subject/object alternation of the prefix is a possibility that arises because objects are low, and local persons structurally include the third person: An affix that can spell out a span of subject and object features may be blocked from doing so, if the object is "too large". It is the object whose person determines this, since its agreement is low, i.e., subject to spellout early. Thus, a system where a local object may bi-uniquely determine a certain affix, while a third person does not is a structurally determined possibility that a language-specific PF-interpretation exploits.

Omnivorous number effects were interpreted as the effect of a VI spanning multiple arguments' plural agreement structure. This allowed for an account of both the third and fourth explanandum, and linked it to the first one: When a vocabulary item spells out PL_{O} in the first effective cycle, it "bleeds" omnivorous number effects. The analysis also provided an account of the second explanandum in terms that are parallel to the first explanandum: In both cases, higher

material (subject features in one case, Tense in the other) is spelled out with third person material, but is blocked from doing so with local person, due to PART blocking the required contiguity. The person-effects on spellout thus receive a uniform explanation, and are tied into the spellout of number by means of cyclicity.

Having thus accounted for the *paradigmatic* distribution of the agreement affixes, and the various explananda about person asymmetries, and number spellout that arose, I now turn to the question how this system might derive the *linear order* of these affixes, showing that a single statement can characterize the linear distribution of the agreement affixes I have discussed.

2.5 Morpheme Order

I now turn from paradigmatic distribution to linear order. Prefixal placement can be characterized as an effect of phrasal movement and antisymmetry (Kayne, 1994), eschewing affix-by-affix stipulation. Taking the structure in Section 2.4 as part of the functional sequence dominating VoiceP delivers not only the correct linear order, but also several other positions where additional higher material occurs.

Of particular interest, for the obvious reason that they are non-trivial, are cases with two suffixes, and the prefixes. As a first step, let us consider the set of prefixes in (63). Two observations characterize these: They all spell PART, whenever they occur (otherwise they lose the competition to other affixes), and they all can spell out REF_{o}. In fact, the only one of these that can be anchored at something other than REF_{o} is *Ø-*, that is to say (63a-d) are necessarily spelling out the first effective cycle, whenever they occur.

(63) *Prefixes*

a. *gv-* ⇔ ‘[PL_{o} [PART_{o} [REF_{o}]]]’ G

b. *m-* ⇔ ‘[PART_{o} [REF_{o}]]’ L/G

c. *g-* ⇔ '[ADD_O [PART_O [REF_O]]]' L/G

d. *v-* ⇔ '[PART_S [REF_S[REF_O]]]' L/G

e. *Ø-* ⇔ '[ADD_S [PART_S [REF_S →[REF_O]]]]' L/G

The only other affix that can ever spell out a PART head is *-t*, repeated in (64). Note, however, that *t* cannot be a first effective cycle spellout, since it does not lexicalize REF_O which is uniformly the first anchor in all contexts.

(64) *-t* ⇔ '[PL_S → [ADD_S [PART_S [REF_S → [PL_O]]]]]' L/G

Note further that *t* spells out PART_S only in cases where PART_O has already been spelled out by *m-* or *g-* (i.e., with local objects), and does not spell out PART_S with third person objects either, since *Ø* or *v* do so in these cases . That is to say, *t* never spells out the lowest PART head, since it never is a first effective cycle spellout – and any cycle that contains at least one PART head will have one spelled out in the first cycle. That is to say, we can characterize the prefixal position for both Laz and Georgian with a simple generalization, as in (65):

(65) The (if any) affix that spells out REF_O and a PART head in the structure is a prefix.

To see that this does indeed characterize the prefix, consider a few relevant examples. The spell-outs in (66) exemplify that structures without a PART head do not have a prefix. The examples in (67) show spellouts with a prefix. In particular, (67a,b) show the spellout of PART_O being prefixal, whereas the examples in (67c,d) show the spellout of PART_S being prefixal in case there is no PART_O.

(66) *Mono-Morphemic without Prefix*

a. $\overbrace{[\text{PST}[\text{REF}_\text{S}[\text{REF}_\text{O}]]]}^{\textit{-da}}$ Georgian: 3SG>3

b. $\overbrace{[\text{PST}[\text{PL}_\text{S}[\text{REF}_\text{S}[\text{REF}_\text{O}]]]]}^{\textit{-dnen}}$ Georgian: 3PL>3

(67) *Bi-Morphemic with Prefix*

a. $\overbrace{[\textsc{pst}[\textsc{ref}_{s}}^{\textit{-da}}\overbrace{[\textsc{add}_{o}[\textsc{part}_{o}[\textsc{ref}_{o}]]]]]}^{\textit{g-}}$ Georgian: 3SG>2SG

b. $\overbrace{[\textsc{pst}[\textsc{ref}_{s}}^{\textit{-da}}\overbrace{[\textsc{pl}_{o}[\textsc{part}_{o}[\textsc{ref}_{o}]]]]}^{\textit{gv-}}$ Georgian: 3SG>1PL

c. $\overbrace{[\ \textsc{pst}[}^{\textit{-di}}\overbrace{\ \textsc{part}_{s}[\textsc{ref}_{s}[\textsc{ref}_{o}]]]]}^{\textit{v-}}$ Georgian: 1SG>3

d. $\overbrace{[\ \textsc{pst}[}^{\textit{-di}}\overbrace{\ \textsc{add}_{s}[\textsc{part}_{s}[\textsc{ref}_{s}[\textsc{ref}_{o}]]]]}^{\textit{Ø-}}$ Georgian: 2SG>3

Next, let us consider the order of suffixes in tri-morphemic spellouts, in (68). We see that it is the spellout of the structurally higher span that precedes the spellout of the lower one, i.e., we find *-di-t*, not **-t-di*.

(68) *Tri-Morphemic with Prefix*

a. *m- -di-t*

$\overbrace{[\ \textsc{pst}[}^{\textit{-di}}\overbrace{\ \textsc{pl}_{s}[\textsc{add}_{s}[\textsc{part}_{s}[\textsc{ref}_{s}}^{\textit{-t}}\overbrace{[\textsc{part}_{o}[\textsc{ref}_{o}]]]]]]}^{\textit{m-}}$ Georgian: 2PL>1SG

b. *gv- -di-t*

$\overbrace{[\ \textsc{pst}[}^{\textit{-di}}\overbrace{\ \textsc{pl}_{s}[\textsc{add}_{s}[\textsc{part}_{s}[\textsc{ref}_{s}}^{\textit{-t}}\overbrace{[\textsc{pl}_{o}[\textsc{part}_{o}[\textsc{ref}_{o}]]]]]]}^{\textit{gv-}}$ Georgian: 2PL>1PL

c. *g- -da-t*

$\overbrace{[\ \textsc{pst}[}^{\textit{-da}}\overbrace{\ \textsc{ref}_{s}[\textsc{pl}_{o}}^{\textit{-t}}\overbrace{[\textsc{add}_{o}[\textsc{part}_{o}[\textsc{ref}_{o}]]]]]}^{\textit{g-}}$ Georgian: 3SG>2PL

That is to say, we can see that the lowest span can become a prefix, if it contains a PART head, while the suffixal order is correlated with height: The higher of the two spans is to the left of the lower one, with both of them suffixal. I would like to propose an interpretation in terms of

Antisymmetry, such that affix order is determined by phrasal movement of the lower structure: A prefix becomes a prefix by virtue of having been pied-piped to a higher position, while the suffix order reflects structural height of stranded material. The structure discussed above is thus originally built on top of VoiceP, along the lines of (69).

(69) TENSE > PL_S > PERSON_S > PL_O > PERSON_O > VoiceP

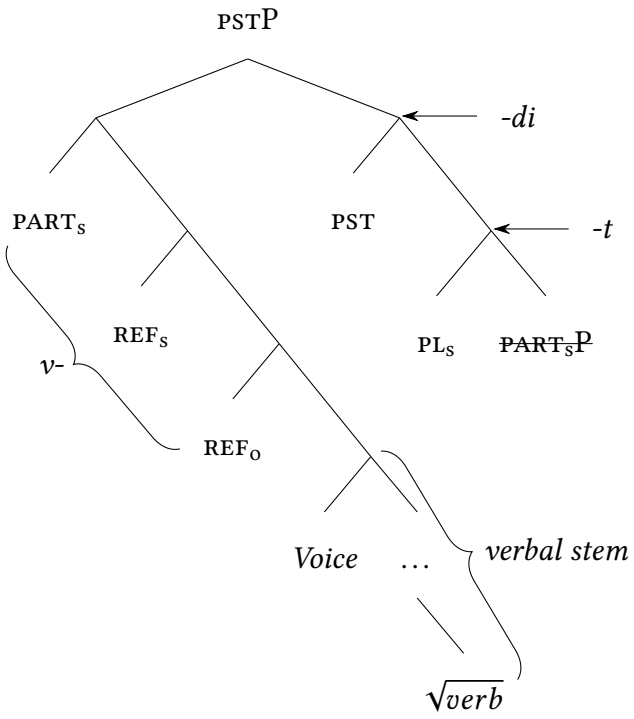
From this perspective, we can now re-characterize the generalization in (65) as a requirement of the voice head to pied-pipe a PART head, if possible, as in (70a), and assume that this structure moves to the specifier of Tense (70b). Admittedly, (70a) is itself in need of an explanation, but the fact that we can capture the linear distribution within the rather complex paradigms discussed above with a single statement is itself rather interesting.

(70) a. Voice pied-pipes the span that contains the structurally closest PART head, if such a span exists.

b. Tense attracts VoiceP

That is to say, we derive suffixation of Tense by movement of the VoiceP to a specifier of Tense, and the lowest span is pied-piped by the VoiceP, if it contains a PART head (recall that if there is at least one PART head, the lowest one is always spelled out in the first cycle). Consider one such derivation in (71). When PST attracts the VoiceP, the VoiceP pied-pipes its adjacent span [PART_S[REF_S [REF_O]]], because it contains a PART head. Antisymmetry now gives us the correct characterization of the linear order: *v-* immediately precedes the verbal stem, since all heads it spells out c-command the VoiceP. The whole PART_S phrase itself is a specifier, and thus precedes the head and complement of PSTP. The spellout of PST, *di*, precedes the spellout of the stranded PL_S head, *t*, and therefore, we correctly arrive at the form *v*-$\sqrt{stem}$-*di*-*t*.

(71) Georgian (1PL>3, Past Tense): *v-stem-di-t*



Such a proposal makes a set of predictions in case we merge material on top of the tense structures we have considered. If we apply the same decompositional logic to Tense as we did for person, we might expect certain Tense structures to structurally contain others. In case we have morphological evidence for such a containment relationship, we thus expect a form that contains another to also show a certain affixation pattern. In particular, if we follow Cinque (2005), Koopman (2017a), and Starke (2018) i.a., in allowing for three types of movement (i.e., all movement must contain the root, and movement out of a moved constituent is blocked, due to Freezing), we expect the following three types of correlating affixation for a structure that merges a head X on top of one of the structures under discussion:

(72) *Affixation of a higher head/structure X*

a. No movement

→ Prefixation before the agreement prefixes

b. Successive Cyclic Movement: Move VoiceP (and pied-piped span, where applicable) to Spec,XP

   → Suffixation immediately after the verbal root, preceding the other suffixes

c. Snowball Movement: Move complement of X to Spec,XP

   → Suffixation to the right, following the other suffixes

We find preliminary evidence for all three strategies; there are however, some caveats which show that, while this might be a promising first step, more in-depth research on this is necessary before drawing firmer conclusions about the antisymmetry interpretation of the spanning account.[24]

As an example for the *no movement* condition, consider the Georgian future tense. It is built from the present tense forms, by means of a set of so called *preverbs*, as we can observe in (73).

(73) *Georgian Future Tense*

a. *v-ṭqer-Ø-t*	b. ***da**-v-ṭqer-Ø-t*
1.SU-write-PRS-PL	PV-1.SU-write-PRS-PL
'We are writing it.'	'We will write it.' Aronson (1990, p. 44)

The crucial point for the current analysis is that there is a morphological containment relation between the future tense and the present tense, that shows a predicted linear order, the caveat being that the preverb appears to be partly lexically determined; a fact that the current analysis does not yet offer a perspective on, given that the relevant head appears to be both structurally

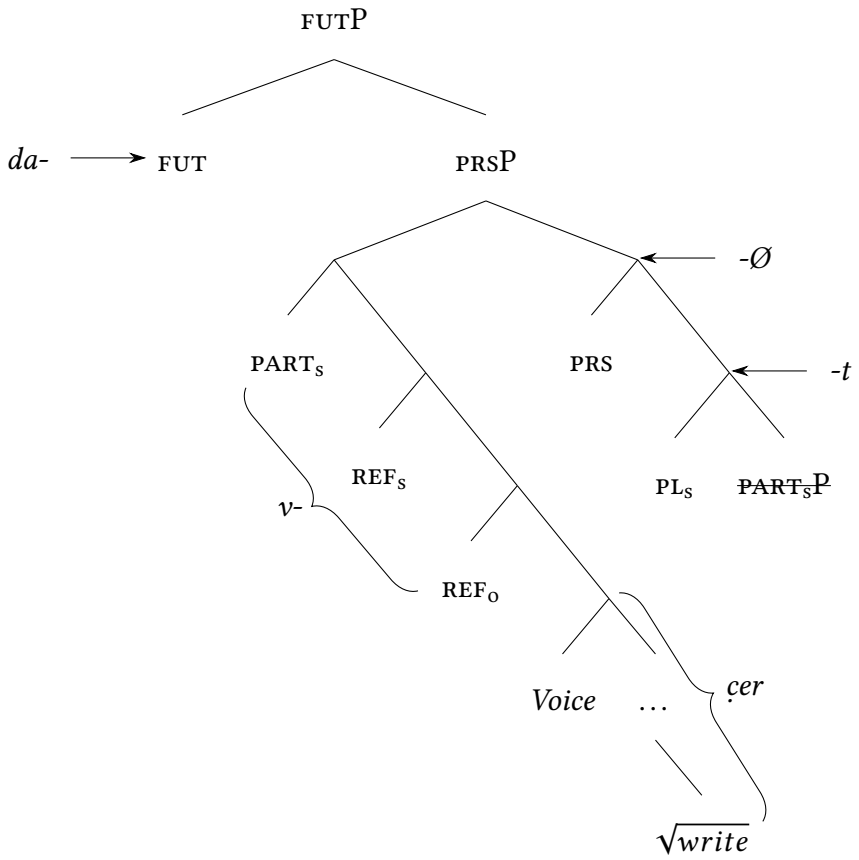
[24] As an anonymous reviewer pointed out, the argument extends to material that is lower than the agreement structure, i.e., VoiceP-internal material: Given the claim in (70) that movement of VoiceP is the relevant operation, stranding of VoiceP-internal material via successive cyclic movement of a smaller phrase is ruled out. We thus predict two positions for VoiceP-internal material, in between the prefix and the stem (no movement/pied-piping), and after the stem but before the suffixes (snowball movement). As the reviewer points out, both positions are attested in, a.o., Laz causative morphology (their example):

(i) *Ma bere-s mektubi do-v-**o**-nç'ar-**ap**-i*
I.ERG child-DAT letter.NOM PV-1.SU-CAUS-write-$\text{CAUS}_{\text{trans}}$-PST

'I made the child write the letter.'

and linearly distant from the root.[25] Consider the derivation in (74):

(74) Georgian: *da-v-ċer-Ø-t* 'We will write it'



Under the assumption that the morphological containment reflects a structural one, the linear position is exactly the one we expect under the *no movement* condition. Since FUT is structurally higher than PRS, and does not trigger movement, it will be prefixed before the agreement prefix, as indicated. Movement proceeds as it did before, and the higher head, FUT, c-commanding the remainder of the heads, precedes the whole structure.

Now contrast this with a structure in Laz that is comparable insofar as it, too, is built from the

[25] The preverbs serve a variety of functions beyond the future marking, such as spatial relations, or perfectivity. They are also present in the past Tense (aorist). Insofar as their function of future marking is concerned, they are "usually unpredictable" (Hewitt, 1995, pp. 148-169), i.e., specific to the verb, but selected from a small class. See ibid., pp. 148-169 for details. Note also that certain verbs do always come with their respective preverb, and do not distinguish the present from the future, i.e., they exhibit a syncretism (Aronson, 1990, pp. 42ff, 61f).

present Tense, namely the subjunctive, before we turn to Laz's future Tense. Descriptively speaking, the subjunctive is built by using the present Tense morphology, but affixing the subjunctive marker *-a*.[26] Note that while the zero affix in (75a) is uninformative with respect to the linear order of the suffixes, we can see that the present/non-past plus agreement morphology in (75b) is to the right of this higher head, i.e., the order we find is prefix-stem-SBJV-PRS. That is to say, this is the linear order we would expect under successive cyclic movement: The verbal structure, once again pied-pipes the lowest PART containing span, and moving through the specifier of PRS, it succesive cyclically moves to the specifier of SBJV.

(75) *Subjunctive*

a. *p'-t'ax-**a**-Ø-t*
1.SU-break-SBJV-PRS-PL
'Let us break it.'

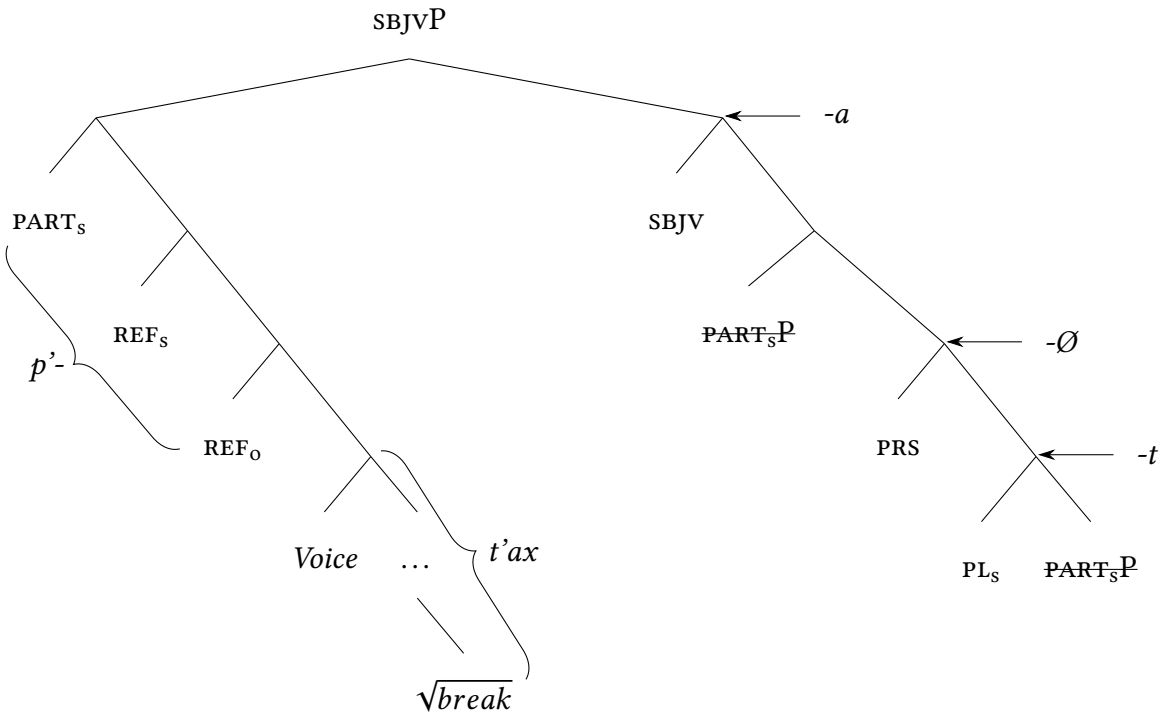
b. *t'ax-**a**-s*
break-SBJV-3SG.PRS
'Let her/him break it.' LAZ

Öztürk and Pöchtrager (2011: 76)

Such a successive cyclic derivation is exemplified in (76).

[26] A second caveat: Laz shows a set of thematic suffixes in certain configurations that denote properties of the event and argument structure. These are obligatory in the present tense Öztürk and Pöchtrager (2011, p. 69), but absent in the subjunctive forms, and again it is not a priori clear what the right analysis of this is under the present account. See Öztürk and Erguvanlı Taylan (2017) for a detailed analysis of the argument structure, event type and aspectual functions of the thematic suffixes.

(76) Laz: *p'-t'ax-a-Ø-t* 'We will break it'



We find that the additional suffix *-a* that spells out the SBJV head (or structure) is suffixed immediately after the verbal structure, since the PART$_S$P moved successive cylically through the PRS head's specifier. Since it is now the specifier of SBJV, it precedes its spellout, which in turn precedes the lower cycles of spellout that were stranded when the VoiceP pied-piped the structure spelled out by *p'* (the phonologically conditioned variant of *v-*).

Finally, consider the Future tense of Laz, which morphologically contains the subjunctive form. As evident from a comparison of (75) and (77), a suffix *-(e)re*, (where the occurrence of /e/ is morpho-phonologically conditioned) is suffixed to the subjunctive form to derive the future tense. Crucially, this containment relation shows the additional affix in the final position, following the agreement affixes.

(77) *Future Tense*

a. *p'-t'ax-a-Ø-**re***

1.SU-break-SBJV-NPST-FUT

'I will break it.'

b. *t'ax-a-s-**ere***

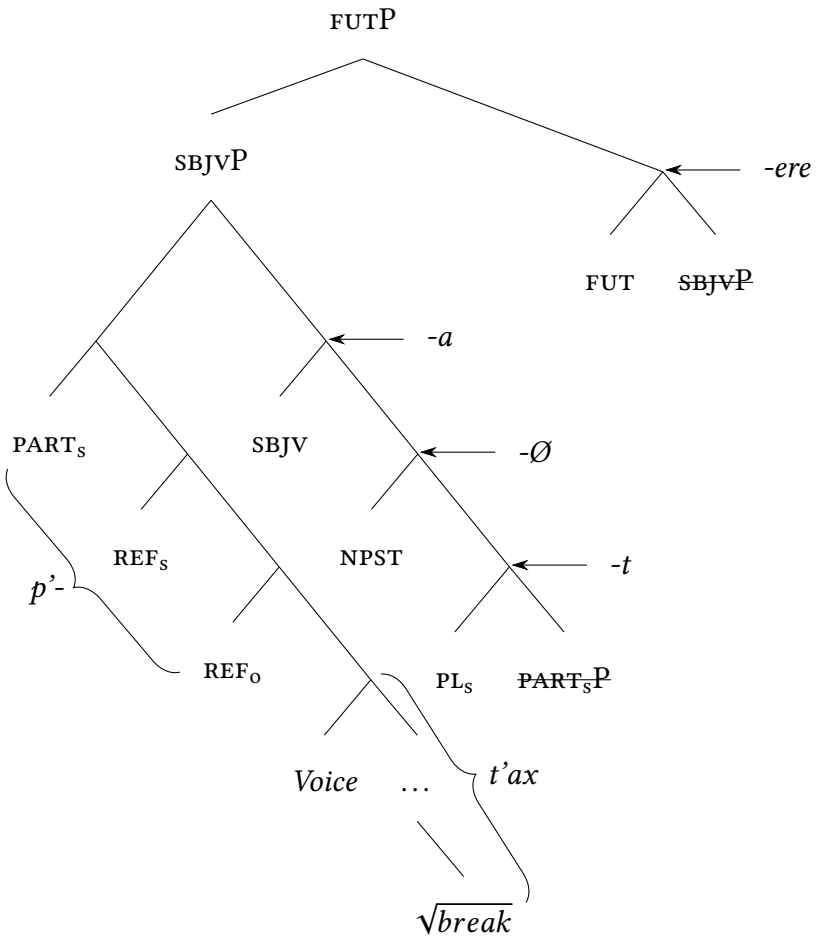
break-SBJV-3SG.NPST-FUT

'She/he will break it.' LAZ

Öztürk and Pöchtrager (2011: 70)

Under a Cinque-style derivation of affixal order, this is indeed the final locus of higher material that we expect, namely the locus we expect if the higher head induces snowball movement, i.e., phrasal movement of the subjunctive phrase into the specifier of the future head. The order is thus derived as in (78). As every element spelled out by *p'* (the phonologically conditioned allomorph of *v*-) c-commands the verbal root, it must precede it. Specifiers precede heads and complements, therefore *p'-t'ax* precedes *a-Ø-t*, and the whole structure in the specifier of FUT precedes *-ere*.

(78) Laz: *p'-t'ax-a-Ø-t-ere* 'We will break it'



We thus find preliminary evidence for all three affixation types that are expected under a phrasal movement account of the affix order. The primary reason for pursuing such an approach, of course, was that the spanning account itself offered a simple characterization of affix order in these terms, providing simple description of the linear order facts in terms of pied-piping. I take the additional data presented in this section to be preliminary evidence in favor of such a view of affixation, but of course such a proposal faces a number of theoretical and empirical challenges if it is to be integrated with the general syntax: The respective structure that receives spell-out needs to be evacuated, i.e., all arguments must leave the TP.[27] The integration of adjuncts into

[27]See Svenonius (2016) for an alternative to evacuation that relies on specifier-complement asymmetries for the definition of a span, which might provide an alternative to evacuation analyses, and might offer a direction of re-

such a structure is not a priori clear, and it raises interesting questions on the derivational timing of e.g., Late Merger (Lebeaux, 1991; Stepanov, 2001). Heads might have to allow for two specifiers, to allow for said evacuation. Spellout must be able to feed back into syntax (see also Starke's (2018) notion of *spellout driven movement*), given that it is the Vocabulary that determines the size of the pied-piped structure. The individual movements are in need of an explanation, and so is the pied-piping description in (70a). Possibly last but definitely not least, the complex remainder of the verbal structure would have to be shown to be derivable under such assumptions. As this is far outside the scope of a single paper, I will leave it at these speculative remarks, suggesting that the current research might open the door for a *radically atemplatic* account not just of the paradigmatic distribution of the relevant affixes discussed above, but also of their linear order, as well as verbal structure beyond these. What I have shown, then, is that a phrasal interpretation of affix order is both possible and simple under the spanning approach: The spanning approach to *paradigmatic* distribution carries over to the *linear* distribution with minimal additional assumptions.

2.6 Discussion

Any theory of the Georgian system of verbal agreement will have to justify its existence, given the number of theories already in existence. I will offer a brief comparison with two competing analyses from the literature, namely the subset-based Distributed Morphology account from Halle and Marantz (1993), as well as the Cyclic Agree account in Béjar (2003), Béjar and Rezac (2009). In many ways the current system builds on insights from those previous analyses, and thus a comparison is of particular interest. Concretely, the current approach takes the idea that the person asymmetries between participant arguments and third person ones should be accounted for in terms of cyclicity and complex person from Béjar (2003), Béjar and Rezac (2009),

search, under which the derivational timing as well as the construction of phonological words from syntactic structure might be understood in terms of PF-instructions relativized to heads.

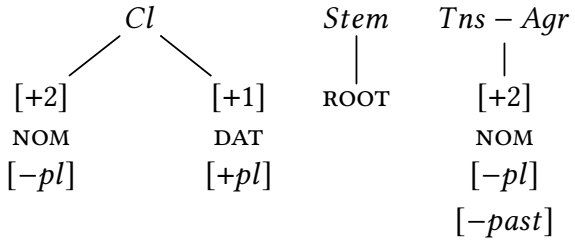
yet crucially retains the idea that this is ultimately is a morpho-syntactic *effect* of the spellout mechanism that exploits structural asymmetries, rather than a property of the syntactic operation Agree as in their approach. In this sense it also takes up crucial insights from Distributed Morphology (Halle & Marantz, 1993) about the way that the morphological component *interprets* syntactic structure. By combining these two insights and applying them to a bottom-up, span based notion of interpretative spellout, it offers several improvements over both.

2.6.1 The Subset Approach – Distributed Morphology

Like the current model, DM assumes morphology to interpret the syntactic structure at PF. For DM, however, the input is a syntactic structure in which heads have sets of features as their internal, often lexically determined internal structure. DM is thus *templatic*, insofar as a head provides the locus of spellout, i.e., a position of exponence, but various processes, such as fissioning or fusing nodes, or impoverishing features interfere with this mapping from syntactic structure to morpho-phonological structure.[28]

In their analysis of Georgian, Halle and Marantz (1993) follow Nash (1992) in assuming that the prefixes of Georgian are the result of a clitic cluster that collects the features of up to two participant arguments. There is a Tense-Agreement node with features of the subject and Tense. The structure that is subject to spellout to be like the one exemplified in (79), with the subject's features represented twice, if the subject is local.

(79) DM representation: 2SG>1PL



Halle and Marantz (1993: 119)

[28] For a general overview of DM see Harley and Noyer (1999), Embick and Noyer (2007).

Such an input representation is subject to some rule-based manipulation, e.g., a [+plural] feature is fissioned off the clitic cluster, unless the plural feature is part of a first person object (to account for *gv-*, which occurs only with 1PL objects), as in (80).

(80) *Fission*

Cl (with [+pl]) + Stem → [+pl] + Cl + Stem (linear order irrelevant)

Unless the [+pl] is part of a [+1], DAT argument. Halle and Marantz (1993: 118)

This modified clitic cluster node is fused, and then subject to vocabulary insertion, with the vocabulary item that matches the largest set of features being inserted; the items that compete for the clitic cluster are shown in (81). Similarly, three vocabulary items compete for the Tns-Agr node, and *-t* spells out the fissioned plural feature.

(81) *Clitic Cluster*

a. [+1], DAT, [+pl] ↔ /gv-/
b. [+1], DAT ↔ /m-/
c. [+2], DAT ↔ /g-/
d. [+1] ↔ /v-/
e. [+2] ↔ Ø

(82) *Suffixes*

a. [+3], [-pl], [-past] ↔ /-s/
b. [+3], [+pl], [-past] ↔ /-en/
c. [-past] ↔ -Ø

Since this does not derive what I dubbed the *conditional omnivorous number effect*, i.e., the fact that we get *g- -en(*-t)* in the 3PL>2PL context, an impoverishment rule is postulated that deletes the fissioned plural terminal in this context. Presumably, a second impoverishment rule would apply in the omnivorous number context 1PL>2PL, to avoid two independent exponents of plurality. Similarly a readjustment rule deletes the suffix *-s* again, in the 3SG>2PL context, where we find *g- (*-s)-t*. Note also that despite the fact that local subjects are represented twice, they only ever trigger one number marking.

What I intend to argue here – given that DM account is descriptively adequate –, is that the structural assumptions that the DM account makes, and that are partly necessitated by the notion of a subset-based insertion into terminals do not carry appropriate explanatory load, but instead require post-syntactic manipulation that make the system so powerful that it does not provide adequate explanations for the phenomena that this paper discussed at length.

In the current approach, two asymmetries between participants and third person arguments have been linked to size: Both the absence of dedicated third person object exponents, and the fact that omnivorous number is Tense dependent only with third person subjects, receive explanations in the same terms. The agreement structure of third person is spelled out in a span with higher material, but with participant arguments, PART blocks the relevant affixes from spanning a structure that includes the higher material (subject agreement, or tense, respectively). Crucially, this results in an additional cycle of spellout, i.e., the number of affixes we find for a particular agreement context is linked to these asymmetries, and consequently we find only one agreement affix, *-u* in 3SG>3 contexts, but three affixes in, say, 1PL>2SG *g- -i-t*.

In a system where spellout is restricted to terminals, we do a priori expect to find the same number of terminals, regardless of the person specifications that such terminals bear, and thus we expect to find a uniform number of morphological positions of exponence across a paradigm, i.e., DM remains fundamentally a *templatic* approach to morphology. Since this is not in fact uniformly the case, these deviations from a uniform template have to be implemented by postulating additional mechanisms of structure manipulation, such as Fusion or Fission that are intertwined with vocabulary insertion, or by stipulating that the information of local subjects is represented twice. This creates redundancy and fails to capture both the person asymmetries, and the relation between person asymmetries and the number of surface morphemes we find. It is redundant, on the one hand because the same features are sometimes represented twice, and on the other hand, because the rules that are necessary to derive the correct results create precisely the kind of contexts that vocabulary items are sensitive to: Consider the exception in the Fission rule in (80),

and the vocabulary item *gv-* that is sensitive to precisely the context this exception includes. For Fission, this has in fact been noted early on, by Trommer (1999), who proposes that the Fissioning of [+pl] in the non-*gv-* contexts is an effect of the other VIs not spelling out plural, i.e., leftover material that has not been matched may be fissioned automatically. However, even such a system fails to capture the fusional effects, and their corresponding person asymmetries. Consider the second third/participant asymmetry, i.e., the fact that we find omnivorous number being spelled out with Tense only if the subject is third person. In the Laz paradigms, we need to make reference to the number of the object, the number of the subject, and the person of the subject to describe the distribution of *-s*, *-an*, and *-t*. We thus find the kind of effects that Fusion is supposed to capture, but we find them only with third person. As I have shown, it is impossible to encode the opposite effect in the current system, but there is nothing in the notion of Fusion that stops us from postulating that a fissioned object [+pl] feature is fused with the Tns-Agr node only in case the subject is a participant, and there is nothing stopping us from fissioning Tense from said node only in case the subject is third person, i.e., the person asymmetry is purely accidental under the DM account, as is the parallelism between the two asymmetries.[29]

In contrast to DM, the current system is radically *atemplatic*, and dispenses with notions such as Fusion and Fission. Morphological positions of exponence are merely an effect of the cyclic nature of spellout, which targets contiguous spans, thus linking the number of positions of exponence to the person asymmetries – we find dedicated exponents for local objects, and the absence of fusion with Tense with local subjects, because in both cases, the presence of a PART head blocks the relevant contiguity, thus leading to additional effective cycles of spellout, i.e., positions of ex-

[29]Trommer's approach is thus quite similar in spirit to the current one, insofar as it is an attempt at reduction to vocabulary insertion. For Trommer, Impoverishment reduces to insertion of a zero VI, and Fusion itself should also be eliminated from the system, and replaced with mutually conditioned contextual allomorphy. That is to say, rather than fusing two nodes, say the object's [+pl] feature, and the Tns-Agr node, we would postulate two contextual allomorphy rules: [+pl] is spelled out as zero in the context of a third person subject, and a third person subject Tns-Agr is spelled out as *-an* in the context of a [+pl] node. Since we find ourselves in a system with omnivorous number, this might necessitate two rules for *-an*, one for [+pl] originating in the Tns-Agr node, and one for the contextual allomorphy, thus creating even more redundancy. It still fails to account for the person asymmetry.

ponence. Similarly, the fact that PL_o may receive a prefixal spellout (*gv-*), or a suffixal one (e.g., *-t*) follows naturally from the way the Vocabulary interprets the structure. In contrast, Fusion and Fission (or their alternative implementations in Trommer's sense) are attempts to repair such divergences of the surface facts from templatic predictions, and fail to capture their relation to the person asymmetries. Neither does the DM system capture the relation these cycles have to omnivorous number, which is merely an effect of deleting (or leaving uninterpreted) up to two [+pl] features, in the DM account. The current system is thus more explanatory, precisely because it is more restrained: The effects of spanning can be derived by Fusion rules, but they cannot be linked to the specifications of vocabulary items in a systematic way; and in fact, they fail to capture generalizations over the data.

I thus conclude that the current system is an improvement over the subset-based analysis. It captures person asymmetries and the distribution of omnivorous number as simply effects of matching vocabulary items to syntactic structure, without additional structure manipulation – by not being subject to manipulation, but only interpretation, the structure is given more explanatory load. My proposal dispenses with the notion of a template, capturing the relation of person asymmetries to the number of surface morphemes we find. Unlike the DM system, which has local subjects represented twice, despite no clear evidence of such multiple exponence, the current system manages to model the five-dimensional paradigms (person and number of subject and object, as well as tense) as fundamentally a one-dimensional structure, i.e., spans are computationally speaking merely ordered lists, and matching reduces to sets of tails (and, given pointers, linked lists of sets of tails), a computationally simple mechanism. At the same time, the spanning approach is a much stronger hypothesis than a Fusion & Fission based one. By making the system less powerful, we have increased its explanatory power. Note, however, that the crucial insights of DM remain intact: We are still dealing with a system in which there is a post-syntactic, interpretative vocabulary. In fact, one might argue that in getting rid of the idea that heads are pre-syntactic bundles of features that create morphological positions of expo-

nence, the current approach takes the DM notion of *syntactic structure all the way down* radically serious. The current approach is therefore in no any way "anti-DM", but rather an attempt at taking DM's proposals extremely seriously, and showing that doing so may lead to better explanations. By elaborating on the structure, and removing operations from the system, the current approach opens a perspective in which the only PF-interface mechanism is Matching, in a form that subsumes the effects of Fusion as a Matching effect, rather than a dedicated rule.

2.6.2 Cyclic Agree

The Cylic Agree approach to Georgian (Béjar, 2003; Béjar & Rezac, 2009) focuses largely on the prefixal alternation, arguing that the DM-style accounts do not capture the subject/object, third/local asymmetries. They argue that the language can be characterized as "having a single core agreement slot, for the control of which multiple arguments compete" Béjar and Rezac (2009: 35). They take this morphological slot (here: the prefix) to be a direct correlate of a syntactic probe, i.e., like the DM approach, their perspective is templatic. From this perspective they suggest that the controller of the Georgian prefix, i.e. the argument that enters the agree relation, is best described as in (83).

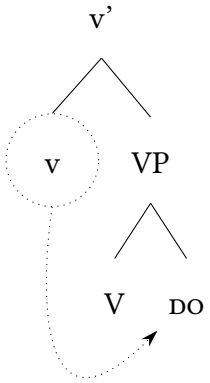
(83) Local Object > Local Subject > Third Person

They interpret this as evidence for a markedness structure on a probe, and adopt the same person structures from Harley and Ritter (2002a) that this paper argued for, with either first or second person as the more marked one.

These features occur as interpretable features on arguments, and uninterpretable ones on probes. Crucially, they revise the matching requirement of Chomsky (2001), such that a given argument matches the probe if it carries a *subset* of the probe's uninterpretable features: A probe specified only for [$u\pi$] will result in the behavior familiar from e.g. Romance or Germanic lan-

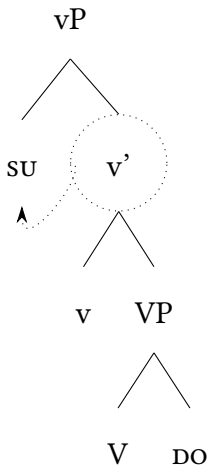
guages, namely agreement with the first argument in the domain.[30] If a probe is specified for a more complex structure, and the first goal it encounters matches only a proper subset of its uninterpretable features, however, the probe remains active, and upon merge of a specifier projects to the bar level. Since it now c-commands the specifier, it engages in a second cycle of probing. For example, a probe specified for [$u\pi$, uParticipant] will agree with the first argument it c-commands as well, as in (84). If, however, this argument matches only [$u\pi$], [uParticipant] will remain active, and engage in a second-cycle probing (85).

(84) First Probing



[30] As they note on page 45 this means divorcing the feature valuation from the resulting morphological expression, as a probe specified for only [$u\pi$] (German, Romance etc) can clearly result in morphological expression of the full feature structure.

(85) Second Probing



This derives the prefixal alternation: A local object will always value both features on the probe, and therefore no second cycle ensues. A third person object, however, will result in a projected [*u*Participant], and thus a second cycle may agree with the subject. This first/second cycle distinction between *m-* (first person, first cycle) and *v-* (first person, second cycle) is then accounted for by a contextual allomorphy rule: If spellout targets the bar level, we get second cycle morphology, conditioned by the first cycle, otherwise we get first cycle morphology.[31] Recall, however, that Georgian always shows agreement with both arguments, even if the object is local. As (86) shows once again, varying either argument's person features results in co-varying morphology, i.e., there is agreement with the subject, even if the object is local, counter the Cyclic Agree predictions.

(86)	a.	*m-xedav-Ø*	2>1	b.	*m-xedav-s*	3>1
		1.OBJ-see-L.SBJ			1.OBJ-see-3.SBJ	
		'You$_{sg}$ see me.'			'He sees me.'	

[31] Note that their approach thus breaks with the DM notion of spellout targeting terminals as well, while building a similar notion into the system as I did: *v-* is essentially the spellout of first person in the context of a third person probing.

c. *g-xedav-Ø* 1>2
2.OBJ-see-1.SBJ
'I see you$_{sg}$.'

d. *g-xedav-s* 3>2
1.OBJ-see-3.SBJ
'He sees you$_{sg}$.' GEORGIAN

While this might potentially be saved by arguing that in addition to the probe on little *v*, Georgian has another probe on T that always agrees with the subject (thus essentially recapitulating the local subject doubling approach of Halle and Marantz 1993), the case is more dire with number agreement. While Béjar and Rezac (2009) do not actually treat this, Béjar (2003) – again, like the current proposal – suggests that singular is total underspecification, and plural is not. The number probe is high, and finds the closest plural argument, since singular subjects cannot intervene. This offers a potential account of omnivorous number (in fact, such a system would quite elegantly derive Laz, though it offers no account of the person-asymmetry of omnivorous number spellout), but fails to account for its breakdown, as in Table 2.18. If *gv-* is involved, first person plural objects have their number marked independently of the subject, even if the subject is plural.[32]

Table 2.18: Georgian First Person Objects

	Subject			
Object	2Sg	2Pl	3Sg	3Pl
1Sg	m- -Ø	m- -Ø-t	m- -s	m- -en
1Pl	gv- -Ø	gv- -Ø-t	gv- -s	gv- -en

The cyclic agree perspective therefore simply does not seem to offer an empirically adequate

[32]The same point applies to Nevins (2011), which derives omnivorous number from Multiple Agree. For Nevins' approach, the prefixes are clitic arguments that can be the targets of Multiple Agree; it is unclear why Multiple Agree would be blocked exactly in case the clitic itself has vocabulary items that make reference to the number distinction. Under said approach we would expect the Georgian first person objects to trigger omnivorous number the same way as second person subjects/objects and first person subjects do, and the proposal does not seem to offer a clear way of distinguishing these two cases. In addition, Nevins (2011: 962) also seems to suggest that Omnivorous Number and co-variance with Tense are in complementary distribution, but Laz clearly provides a counterexample to such a generalization.

account of the Georgian agreement data that it purports to explain. It wrongly predicts certain observable agreement facts – the co-occurrence of subject agreement with local objects, and of object number agreement with plural subjects – to be impossible. It misses crucial parts of the interaction between the prefixal and suffixal spellout of number, and only accounts for what appears to be an arbitrary subset of the data. While it offers an account of the person asymmetry of the prefix, it offers no potential insights into the fact the same person asymmetry appears to be at play in the spellout of number: Both the prefixal position and the suffixal position exhibit local/third asymmetries, and the fact that portmanteau morphology for Tense/Number seems to be subject to similar generalizations is not captured in the cyclic agree account.

2.7 Conclusion

In this paper, I have argued in favor of a novel perspective on the Georgian agreement system, and provided new data from Laz that shed light on the distribution of omnivorous number effects within Georgian. In this analysis, a variety of phenomena have been interpreted as the effects of a generalized portmanteauhood, *spanning*, which is driven by the available vocabulary. I have shown that the five-dimensional system can be interpreted as strictly ordered (by the order of Merge), and that the distribution of all morphemes can be linked to them spelling out a contiguous subpart of a fixed hierarchy. In doing so, I provided an analysis of the prefixal alternation, as well as prefix-suffix interactions, which did not find appropriate explanations in previous accounts. In particular, I have shown that two person-based splits can receive parallel explanations in terms of third person agreement being spelled out with higher material, and local agreement blocking the relevant contiguity due to their larger size. I have also shown that the same perspective offers an account of the distribution of omnivorous number effects in Georgian.

In many ways, this approach builds on previous analyses. The first one is the Cylic Agree approach of Béjar (2003), Béjar and Rezac (2009), with which it shares the notion that the prefixal

facts should be accounted for in terms of derivational timing (objects are low), and size asymmetries of person (participants contain third person). The current approach differs sharply from theirs in arguing that these person asymmetries are not genuine syntactic effects of a relativized probing mechanism, but PF-effects that arise since the language's vocabulary inventory exploits these structural asymmetries. In doing so, the current approach achieved an empirically accurate picture of all agreement facts, rather than just the singular prefixes.

The second approach is the Distributed Morphology account of Halle and Marantz (1993), from which it takes the general perspective that PF interprets abstract syntactic structure composed of minimal units. It differs from the original DM account not in general architectural questions, but in a highly specific ones, namely the nature of the matching algorithm/the Vocabulary. Correspondingly the two approaches differ with respect to terminals provide templatic slots, and the question of how minimal the minimal syntactic units are. In dispensing with head-internal feature structures, and the corresponding templatic notion of insertion into terminals only, I have shown that it may be possible to reduce these operations to a form of Fusion that is itself subsumed under the notion of Matching. In doing so, I hope to have shown that taking the DM notion of *hierarchical syntactic structure all the way down* more serious than even the original DM paper did, we may be able to provide new and interesting explanations for morpho-syntactic phenomena. In arguing that spanning is what matches abstract syntactic structure with phonological material, I have shown that – for the case of Laz and Georgian, at least – we may dispense with a variety of post-syntactic operations that this approach necessitated for what appeared to be syntax-morphology mismatches under the terminal insertion perspective. The current system contrasts with theirs in being radically atemplatic, deriving all effects from a simple restriction, contiguity, but otherwise it remains committed to its larger architectural notions.

Of course, there is a great variety of morphological phenomena that have not been touched upon, and unlike research in core syntax, research on inflectional morphology in particular is limited by the finitude of forms that one can find in any given language, so in these senses this

paper's proposals must be taken with (at least) a grain of salt. Despite that, however, I hope to have demonstrated that a Minimalist attempt that reduces all structure building to the narrow syntax and having the interpretative component of PF drive the bundling of heads into exponents may indeed lead to new and interesting insights.

3 Arabic Agreement

3.1 Introduction

Narrowly, this paper discusses the linear and paradigmatic distribution of the Classical/Modern Standard Arabic (MSA) agreement affix *t*. Paradigmatically, the suffix's distribution is of interest, because its distribution has to be described disjunctively under a subset approach to the post-syntactic matching of a vocabulary item to a morpho-syntactic context. Linearly, it is of interest, because it can occupy three distinct morphological positions, as a prefix, or in two different suffixal positions.

Broadly, the approach proposed here argues that the hierarchy Noyer (1992) describes as governing various morphological effects in Afro-Asiatic languages, such as Discontinuous Bleeding, the order of spellout, or Impoverishment, really describes the functional sequence of Tense and Agreement, and that a bottom-up, span-based spellout can derive these effects in MSA without recourse to a template, an independently existing hierarchy, or Impoverishment. I argue that the span-based approach to the spellout of agreement advocated in Blix (2021b) can derive effects that cannot be stated in a non-contradictory or arbitrary fashion in a subset principle approach to matching a vocabulary item to its contexts, and link the paradigmatic distribution to the linear one in a systematic way.

The first aspect, i.e., the paradigmatic distribution of *t* bears on the nature of natural classes that morphemes are sensitive to, i.e., potentially able to spell out at PF. It is standardly assumed

within Distributed Morphology (DM) that natural classes are characterized by an SPE-style subset principle: A set/structure of features [A,B] defines the classes [A], [B], and [A,B], and – possibly – [ ], i.e., there is a structural symmetry to (certain) features that is resolved by other means, such as an Elsewhere Principle based competition, or markedness. The distribution of *t*, and the contrasts it shows with *y*, however, cannot be captured in such a system. To see this, consider the data in Table 3.1: First consider the subjunctive contrast between *y-aktub-a* (3MS) and *t-aktub-a* (3FS): In a subset-based approach, minimally, *t* spells out feminine gender, or *y* spells out masculine gender, given that these verb forms form a minimal pair for gender.

Table 3.1: Classical Arabic *t* and *y* with *k-t-b* 'to write'

	SUBJUNCTIVE	
PERSON, NUMBER	FEMININE	MASCULINE
3SG	***t**-aktub-a*	***y**-aktub-a*
3PL	***y**-aktub-na*	***y**-aktub-uu*
2PL	***t**-aktub-na*	***t**-aktub-uu*

Next, consider the same contrast in the third person plural of the subjunctive: Here we find *y-aktub-uu* (3MP) and *y-aktub-na* (3FP). Since *y* occurs in both genders, it cannot be specified for gender, leading to the conclusion that it must be *t* that is responsible for the first contrast, i.e., specified for feminine gender. In the second person, however, we are led to the opposite conclusion: In both the masculine *t-aktub-uu* (2MP), and the feminine *t-aktub-na* (2FP) we find *t*, rather than *y*, but with the same suffixes as in the third person plural. If the subset principle leads us to conclude that both *t* and *y* are unspecified for gender, however, there is no explanation for the initial contrast in the third person singular.[1] Instead, we arrive at a descriptive disjunction: From the contrast within the third person singular, we see that the *y/t* contrast encodes gender, but in the third/second person plural, we see the same *y/t* contrast encoding person.

[1]Note that the point holds under any kind of standardly taken assumptions about feature structure, such as *i*) one gender is specified and the other is underspecified, *ii*) one is a subset of the other, or *iii*) both are specified (privatively or as opposite values of a binary feature), as long as there is a structural symmetry within the AGR node.

The second aspect of interest with respect to the agreement morpheme *t* pertains to its linear distribution: In addition to the prefixal paradigms, MSA also has a purely suffixal agreement paradigm called the Perfect. As shown in Table 3.2, some of the affixes that occurred as prefixes above, surface as suffixes in these paradigms. The first person plural form *n-aktub-a* has *katab-n-aa* is its Perfect counterpart, and every form that exhibits prefixal *t* also shows a suffixal *t* — with an additional occurrence of *t* in the first person singular.

Table 3.2: Prefix/Suffix Alternations with *k-t-b* 'to write'

Subject	Subjunctive	Perfect
1sg	ʔ*-aktub-a*	*katab-t-u*
1pl	***n***-*aktub-a*	*katab-**n**-aa*
2fs	***t***-*aktub-ii*	*katab-**t**-i*
2ms	***t***-*aktub-a*	*katab-**t**-a*
3fs	***t***-*aktub-a*	*katab-a-**t***
3ms	*y-aktub-a*	*katab-a*

Curiously, however, *t* can differ in its relative position as a suffix, as seen in the contrast between *katab-a-t* (3ms) and *katab-t-a* (3fs), both of which have *t-aktub-a* as their subjunctive counterpart. We thus find *t* to be involved in the spellout of two paradigmatic contrasts, gender *or* person, as well as occurring in three distinct linear positions: In the non-perfect paradigms, *t* is always prefixal, but in the perfective, the gender contrast is encoded by *t* at a suffixal position that is different from the position of first/second person *t*.

To my knowledge, there are two different perspectives in the DM literature that have been taken on the paradigmatic distribution of *t* in the prefixal paradigms, and none that explicitly link it to the suffixal paradigms. Halle (2000) proposes that the prefixal *t-* is a pure elsewhere marker, devoid of features. It is inserted in the second person, because there are no competing prefixal vocabulary items. In the third person feminine singular, an Impoverishment rule deletes the third person feature [−participant], bleeding the insertion of the third person affix *y-*, and feeding the insertion of *t-*. The first issue with this account is the reliance on competition among

prefixes only, which does not extend to the suffixal paradigms; the second issue is the reasoning that lead Noyer (1992) to argue for the other perspective in the literature, namely a homophony account: As Noyer shows in great detail, there appears to be a general hierarchy that governs Impoverishment in the Afro-Asiatic language family's prefixal conjugations, as in (1). According to his account, a co-occurrence filter on two features can trigger Impoverishment of the feature that is lower on the hierarchy, but never vice versa — the exact opposite of the Impoverishment rule proposed by Halle.

(1) *Noyer's Impoverishment Hierarchy*

Person > Number > Gender

In fact, Noyer (1992: 93) makes a further claim about the hierarchy: It does not only govern Impoverishment, but also the order of spellout — if the relevant vocabulary items are disjoint or overlapping, then person is spelled out before number, and number is spelled out before gender. Taking Bobaljik's (2000) position that spellout operates in a bottom-up fashion, I argue that we can account for both the linear and the paradigmatic distribution of Classical Arabic agreement by taking Noyer's hierarchy to represent an insight into the syntactic structure of MSA agreement, rather than an independent hierarchy governing morphological operations: It represents an inversion of the syntactic hierarchy, and person is spelled out first, precisely because it is at the bottom of the agreement structure. Taking the Distributed Morphology notion of *syntactic hierarchical structure all the way down*, radically serious, I will assume that all syntactic hierarchy is composed of maximally simple heads, not feature bundles (Bobaljik, 2012; Caha, 2009; Starke, 2009), and that the PF-interpretation of syntactic structure by vocabulary items targets contiguous spans of such maximally simple heads. In particular, I assume that person is an internally complex region of the agreement structure, and that a containment relationship as in (2) holds (following e.g., Béjar & Rezac, 2009; Blix, 2021b; Harley & Ritter, 2002a).

(2) Internal Structure of Person

a. 3^{rd} [3]

b. 1^{st} [1[3]]

c. 2^{nd} [2[1[3]]]

Concretely, I adopt the following Nanosyntactic assumptions: Firstly, a vocabulary item lexicalizes a span, or a span of spans (by means of a Pointer, Blix, 2021b; Caha & Pantcheva, 2012). Secondly, vocabulary items match all their subspans (the *Superset Principle*). Thirdly, spellout of such spans targets the largest span that can be matched by a vocabulary item (*Cyclic Overwrite*). Fourthly, the commonly shared Elsewhere Principle assumption that the more specific vocabulary item takes precedence over a less specific one in case they match the same span.

From that perspective, *t* is argued to be a first person singular feminine marker, lexizalizing a pointer between number and person, as in (3a). Pointers have been argued to account for Blansitt's generalization (Pantcheva & Caha, 2012), as well as certain omnivorous number and person hierarchy effects (Blix, 2021b). A vocabulary item may, under this perspective, lexicalize two spans X, Y with a pointer [X→Y]. Each of these two spans characterizes a set of matching subspans, and the vocabulary item then matches any span that is formed by contiguity between a subspan of X and a subspan of Y. Under inclusion of the empty span as a trivial subspan, this gives rise to the possibility of a disjunctive surface distribution, since the vocabulary item in question can spell out a subspan X or Y, or a combination thereof, since a subspan X_s of X together with the empty subspan is identical to X_s by definition. Applying this previously developed theory to the data at hand, I argue that the structure in (3a) allows us to account for the disjunctive properties of the *y/t* contrast in these spanning terms.

(3) a. */t/* ⇔ '[FEM[SG→[1[3]]]]'

b. */y/* ⇔ '[SG[3]]'

c. */Ø/* ⇔ '[SG[2]]'

d. */na/* ⇔ '[FEM [PL]]'

Crucially, the pointer, denoting the fact that this vocabulary item lexicalizes a *span of spans*, allows for the item to spell out the two internally contiguous spans [FEM[SG]] and [3] together, as [FEM[SG[3]]], i.e. a third person singular feminine, as in (4a).

This is possible, because [3] is a subspan of [1[3]], i.e., matched by the lower span below the pointer. The [FEM[SG]] is matched trivially by the higher part, due to their identity. Finally, the whole structure is contiguously matched by *t* at large, since no head that is not lexicalized by *t* intervenes between the two spans, and thus *t* can spell out all of [FEM[SG[3]]].

(4) a. [TENSE $\overbrace{\text{[FEM [SG [3 }}^{t}$ [aktub]]]]] 3FS

b. [TENSE [FEM [PL [2 $\overbrace{\text{[1 [3}}^{t}$ [aktub]]]]]] 2FP

c. [TENSE [FEM [PL $\overbrace{\text{[3}}^{y}$ [aktub]]]]] 3FP

d. [TENSE$\overbrace{\text{[SG [3}}^{y}$ [aktub]]]] 3MS

In contrast, in the second person in (4b), the affix *t* can spell out part of the second person's structure, namely [1[3]], since this is a trivial subspan of [1[3]], and contiguity with the *empty* subspan of [FEM[SG]] is trivially fulfilled. However, since the person head [2] blocks contiguity with gender, it cannot spell out person and number/gender together in the second person, thus accounting for its presence in both the masculine and the feminine second person.

In the third person feminine plural, [FEM[PL[3]]] in (4c), where we find *y*, the plural head below gender blocks contiguity with [FEM], and therefore *t* cannot spell out the whole structure that includes Person, Number and Gender. Instead, both *y* and *t* are candidates for the spellout of [3], and the smaller, more specific item *y* wins due to the Elsewhere Principle. In a parallel fashion,

the third person singular masculine in (4d), where no FEM head is in the structure, has both *t* and *y* as candidates for the spellout of [SG[3]], and *y* wins the competition for the same reason.
We thus derive the distribution of *y* and *t* as natural classes as characterized by the superset principle. The remainder of the material is then spelled out in further cycles, deriving what Noyer called *discontinuous bleeding*: When *t* is able to spell out feminine gender in the first (lowest) cycle, and thus *y-aktub-a* (3MS) and *t-aktub-a* (3FS) are syncretic at the suffix. In contrast, when *t* is unable to spell out gender at the prefix, as in *t-aktub-na* (2FP), or *y-aktub-na* (3FP) we see gender expressed at the suffix instead — a sycretism that the zero affix for [2] derives, since the third cycle of 2FP and the second cycle of 3FP are identical, as shown in (5):

(5) a. [TENSE [FEM [PL [2 [1 [3 [aktub]]]]]] — *na* *Ø* *t* — 2FP

b. [TENSE [FEM [PL [3 [aktub]]]]] — *na* *y* — 3FP

The linear order of these affixes will then be derived by suggesting that the phrase headed by [3] moves to the specifier of the Tense head, pied-piping all material that is spelled out within the same span as [3], and stranding the remaining material; under an Anti-Symmetry approach (see Kayne 1994, 2017 broadly, as well as more narrowly for morpho-syntactic issues Blix 2021b; Caha 2009; Koopman 2017a; Starke 2009), this derives the correct order. Unlike previous approaches, which stipulate the prefix/suffix distinction as a property of the vocabulary item, and thus miss a generalization that holds for them *as a class*, this approach provides an explanatory account of the set of prefixes as a derivational result: Whatever affix spells out [3], may become a prefix.

The same approach is then shown to extend naturally to the suffixal paradigm. In particular, I show that the morphological evidence points to the Perfect being characterized by a head PERF between Person and Number. I then argue that the same constraint for linear adjacency between the affix that spells out [3] and the verbal stem will holds in both prefixal and suffixal paradigms,

and allows for a uniform characterization of the paradigmatic and linear distribution of *t*. As Table 3.3 shows, there is no suffixal counterpart to the smaller third person affix, *y*, but instead the contrast between the third person singular masculine/feminine is encoded by *addition* of *t*, rather than by the *y/t* contrast from the prefixal paradigms.

Table 3.3: Prefix/Suffix Alternations and Their Absence

P/G/N	Subjunctive	Perfect
3ms	y-aktub-a	katab-a
3fs	t-aktub-a	katab-a-t
2ms	t-aktub-a	katab-t-a

I take this to show the presence of a zero affix that spells out a larger structure than *y*, namely *Ø* ⇔ '[perf[3]]'. From that perspective, it follows, that *t* cannot spell out [3] in the third person singular feminine of the Perfect, since *Ø* can spell out a larger span [perf[3]] that *y* and *t* do not match. However, this first cycle leaves fem uninterpreted, i.e., subject to a later cycle, for which *t* is the matching candidate: The span [fem [sg]] is spelled out by *t*, as in (6).

(6) [tense $\overbrace{[\ \text{fem}\ [\ \text{sg}}^{t}$ $\overbrace{[\ \text{perf}\ [\ 3}^{Ø}$ [aktub]]]]] 3fs

Under this perspective, the correlation between the affix encoding a gender contrast vs a person contrast with its linear position as a prefix or a suffix receives an explanation that is unified with the account of the location of the other affixes, by simply stating that the affix that spells out [3] must be adjacent to the root, a more precise implementation of which will be given below, in terms of anti-symmetry and Heck's (2008) edge condition on pied-piping. The current approach advances a fully *atemplatic* notion of morphology, where linear order is derived purely in syntactic terms, with movement targeting phrases, and pied-piping being determined by spans; the lesson from the variation in linear order that a single affix may exhibit, show that a lexical specification for prefixhood/suffixhood cannot be empirically correct, *and* misses a generalization

about the set of prefixes. Instead of assuming templatic effects, I show that it is precisely the flexibility of matching that the pointer theory provides that corresponds to the paradigmatic and linear distribution of this affix, and the contrasts that it can encode; the linear and paradigmatic peculiarities of the distribution of *t* find a uniform explanation in the terms developed in Blix (2021b). I further show that from this perspective, not only does Noyer's (1992) hierarchy find a well-defined place in the syntactic theory, as part of the functional sequence, but also that the syncretism effects in Standard Arabic that were characterized by Noyer as Impoverishment can be captured purely as syncretism effects of vocabulary insertion without a second mechanism such as Filters.[2]

The paper is organized as follows: Section 2 will provide the theoretical machinery that I am employing. Section 3 will discuss the prefixal paradigms, focussing first on the paradigmatic distribution, and then deriving the linear effects, and their derivation. Section 4 will extend this approach to the suffixal Perfect paradigm, and show how certian affixes can occur as either a prefix or a suffix. Section 5 will offer a brief discussion of the current approach against Halle's and Noyer's perspectives, and Section 6 concludes.

3.2 Theoretical Background

This section introduces the technical framework in which this work is couched, i.e., Nanosyntax. For the basic notions of span based spellout, it follows Starke (2009), Caha (2009), with some technical details adopted from Svenonius (2018); the discussion of the Pointer is based on

[2]Note that in the kinds of cases I discuss here, what it at issue is whether syncretisms, partial or total, should be captured by multiple mechanisms — the empirical fact that two morpho-syntactic contexts find (partially) identical expression is in principle a possible effect of the matching mechanics (e.g., the subset principle, the superset principle): The contexts [A,B] and [A] can be syncretic, because an item [A] matches both of them, or because an Impoverishment rule deletes, say, [B] in the context of A. While I argue here that the matching mechanism should, if plausible and possible, be the only mechanism to achieve this goal, there is a more general argument in favor of Impoverishment, given for example by Bobaljik (2002), who argues that trans-paradigmatic syncretisms are not properly captured by reducing them to vocabulary insertion, since it would be an accident of the set of items, when in fact it appears to be a property of the language. While the boundary between these two cases is not always clear-cut, I leave this issue aside here.

Pantcheva and Caha (2012), as well as Blix (2021b).

The general tenants of Nanosyntax lie in the idea that morphological items interpret contiguous "chunks" of syntactic structure. Lexical access occurs after every instance of Merge, and cyclically overwrites the previous spellout, until no vocabulary item can be found that is able to interpret the newly built chunk — in this case, the previous cycle of spellout is finalized, and a new cycle of spellout begins, starting with the head that could not be spelled out in the previous cycle.

(7) *Cyclic Overwrite*

Spellout proceeds in a cyclical, bottom-up fashion, paralleling the syntax. Every cycle of spellout overwrites the previous cycle, until no vocabulary item can be found for a span, making the previous cycle effective. The next cycle is anchored at the head for which no span could be found.

Closely intertwined with this conceptualization, is the notion that head-movement and the Distributed Morphology process of Fusion can be replaced by vocabulary items lexicalizing spans of heads. By doing so, the fact that morphological co-exponence of two or more heads by a single morpheme/vocabulary item occurs, is derived from the content of the vocabulary item itself, i.e., Fusion is *vocabulary driven*. The notion of a span itself is a contiguous n-tuple of heads in a syntactic structure, such that each head has the following head as the one heading its complement. This is formalized as in (8).

(8) *Span*

An n-tuple of heads $< X_n, \ldots, X_1 >$ is a span in a structure S, if and only if $X_{n-1}P$ is the complement of X_n

Adapted from Svenonius (2018: 90)

Based on this notion, we can construct a matching condition, as in (9): A vocabulary item lexicalizes a span, which characterizes a set of subspans that it matches, i.e., that it can potentially spell out (and will spell out unless it is overwritten cyclically, or a better candidate exists).

(9) *Matching*

A Vocabulary Item that lexicalizes a span $< X_n, \ldots, X_1 >$ matches any syntactic span $< Y_m, \ldots, Y_1 >$, s.t.:

(i) $X_1 = Y_1$, and *Anchoring*

(ii) for any Y_p, s.t. $p < m$: $X_p = Y_p \rightarrow X_{p+1} = Y_{p+1}$ *Contiguity*

Matching thus derives what has been dubbed the *Superset Principle*: A Vocabulary item matches all its sub-spans, subject to an identity condition for the lowest element of the respective spans, called *Anchoring*, as well as one on the order of items, called *Contiguity*. A Vocabulary item that lexicalizes a span [A[B[C]]] will thus match all spans that are identical with respect to the bottommost element, [C], and show the same contiguity, i.e., it will match the spans [A[B[C]]], [B[C]], as well as [C].[3] A more simple way of thinking of this is in terms of subspans, as in (10).

(10) *Subspan*

Any vocabulary item $< X_n, \ldots, X_1 >$ characterizes a set of contiguous subspans $\{< X_m, \ldots, X_1 > \mid m \leq n\}$ that it matches.

That is to say a span $< X, Y, Z >$ characterizes the set of subspans $\{< X, Y, Z >, < Y, Z >, < Z >\}$, i.e., its matching spans. In this paper, I defend an extension of this basic matching principle that has been advocated in Caha and Pantcheva (2012) and Pantcheva and Caha (2012) and Blix (2021b), in order to account for certain regularities in "multidimensional paradigms", i.e., in cases where a vocabulary item may spell out heads that belong to adjacent syntactic regions, in our case regions like Person and Number. Formally speaking, a region may be defined by an internal hierarchy, the bottom-element of which itself selects for a region, rather than a specific head. To pick up the example of Person and Number, as touched upon in the Introduction: While

[3] I continue to use the familiar bracketing notation for spans here, despite the fact that a span is not necessarily a constituent, or vice versa. For similar approaches in terms of constituency rather than spans, however, see Svenonius (2018), Starke (2018).

the participant head [1] always subcategorizes the third person structure [3P], the singular head may subcategorize any person structure, thus characterizing Person and Number as two distinct regions.

Under the current proposal, a vocabulary item may lexicalize parts of two adjacent regions rigidly, i.e., without a pointer, in which case Subspans are defined as before. In such a case, exemplified in (11) the only subspan of a first person singular [SG[1[3]]] would be the first person [1[3]], or the third person [3], but crucially not the third person singular span [SG[3]], because this is not a *contiguous* subspan of the first person singular:

(11) a. VI: $\alpha \Leftrightarrow$ '[SG[1[3]]]'

b. Matching Subspans:

i. [SG[1[3]]]

ii. [1[3]]

iii. [3]

It may also, however, lexicalize the two regions as individual spans, with a pointer, as per (12).

(12) *Pointers*

A Vocabulary Item that lexicalizes two spans $< X_n, \ldots, X_1 >, < Y_m, \ldots, Y_1 >$ by means of a pointer X⟶Y matches any syntactic span that is formed by contiguity between a subspan characterized by X and a subspan characterized by Y (including the empty ones):

$\{< X_q, \ldots, X_1 >, < Y_r, \ldots, Y_1 >, < X_q, \ldots, X_1, Y_r, \ldots, Y_1 > \mid q \leq n \ \wedge r \leq m\}$

Under these circumstances, the contiguity requirement between the two spans is relaxed, in that it is required to hold between the spans, rather than between the individual heads. Transposing the previous example, this allows such a corresponding vocabulary item β that differs from the previous example only in the Pointer to spell out [SG[3]] as a subspan, as shown in (13). In contrast to the pointerless version, this affix can spell out two additional spans, namely [SG[3]],

and [SG]. The former can be matched by β, because contiguity between a subspan of [SG] (i.e., the span itself) and a subspan of [1[3]], namely [3] holds, the latter, because the same is true if we take the empty span as a trivial subspan of [1[3]].

(13) a. VI: $\beta \Leftrightarrow$ '[SG$\rightarrow$[1[3]]]'

b. Matching Subspans:

i. [SG[1[3]]]

ii. [1[3]]

iii. [3]

iv. [SG[3]]

v. [SG]

I will show below that the flexibilities that this interpretation allows, can be motivated by showing that it allows for a unification of diverse facts about the distribution of certain affixes that are otherwise treated as stipulative distributions of an elsewhere marker, with conditions for insertion being derived by a set of independent means such as lists of Impoverishment rules.

Finally, the system adopts the usual Elsewhere Principle, by which competition among multiple matching items is resolved in favor of the more specific item, as in (14).

(14) *Elsewhere Principle*

If two vocabulary items match the same span, the item that lexicalizes the fewest heads wins.

In the next section, I detail the individual regions that we can assume to be relevant for the problem at hand, and derive the prefixal paradigms along these lines, focussing on the paradigmatic distribution of the affix *t*.

3.3 The Non-Perfect Paradigms

MSA has four basic inflectional paradigms for finite verbs, the Jussive, the Subjunctive, the Imperfect and the Perfect. The Perfect differs from the other three in two basic regards: While the Perfect is exclusively suffixal, the other three exhibit both prefixes and suffixes for agreement; the same three also share are a vocalic melody that is not shared by the Perfect.[4] This is exemplified in (15).

(15) a. *y-aktub* (3MS.JUSS)

b. *y-aktub-a* (3MS.SBJV)

c. *y-aktub-u* (3MS.IMPF)

d. *katab-a* (3MS.PERF)

While I have no intention of developing an elaborated theory of the internal structure of Tense/Aspect/Mood in Arabic here, there is some evidence for a containment relation among the Jussive, the Subjunctive, and the Imperfect, and I will assume that this is reflected in their internal structure. The contrast in (15a,b) provides some evidence that the Subjunctive might properly contain the Jussive, given that one is derived from the other by adding the suffix *-a*. The same can, of course be said about the Imperfect, which exhibits the suffix *-u* in the same place; there is, however, another contrast that suggests that the Imperfect contains both the Subjunctive and the Jussive, as in (16).

(16) a. *t-aktub-ii* (2FS.JUSS, 2FS.SBJV)

b. *t-aktub-ii-na* (2FS.IMPF)

[4] I follow Kastner's (2018) analysis for Hebrew here, in assuming that the melody is in fact a contextually conditioned spellout of *Voice* (or, from the current perspective, possibly a span of such argument structure related heads). This raises interesting questions from the current perspective, but I will only be able to touch on this briefly, when I return to the Perfect below.

I therefore assume that these three can be characterized by a containment relation as in (17): Since *-ii* can spell out the subjunctive, it can, by the superset principle, also spell out the jussive, accounting for the syncretism; *-na* spells out the IMPF head in this case.

(17) a. *Jussive*: [JUSS]

b. *Subjunctive*: [SBJV[JUSS]]

c. *Imperfect*: [PRS[SBJV[JUSS]]]

As it is evidently the simplest case, I will begin with the Jussive, as it will allow me to focus mostly on the agreement structure itself, in particular the distribution of *y* and *t*. I will then derive the Subjunctive and Imperfect from it. For now, I abstract away from the linear order of the relevant affixes as well, returning to it only once I have accounted for the paradigmatic distribution of the relevant affixes, showing that the paradigmatic account extends to the linear account. I later link it to the distribution in the Perfect, showing how a single head can account for the differences between the prefixal and suffixal paradigms, both with respect to the linear and the paradigmatic aspects, under the assumption that the agreement structure is built on top of the verbal domain, and movement targets various sub-constituents within this syntactic structure in Cinque (2005) U20 style approach to linear order.

3.3.1 *Y*, *T* AND THE JUSSIVE

Table 3.4 shows the complete paradigm of the Jussive, with the distribution of *y* and *t* highlighted. We can see that *t* occurs in all second person cells, as well as the third person feminine singular and dual cells; not, however, in the third person feminine plural one, where *y* is found in the prefixal position instead. The prefix *y* is also found in all third person masculine cells. Crucially for the present purpose, this distribution is also found in all other paradigms, modulo the Perfect, which shows a related distribution that I will return to below.

Table 3.4: Arabic — Jussiv

	Number		
Person/Gender	Singular	Dual	Plural
1st	ʔ-aktub		n-aktub
2nd m.	t-aktub	t-aktub-aa	t-aktub-uu
2nd f.	t-aktub-ii	t-aktub-aa	t-aktub-na
3rd m.	y-aktub	y-aktub-aa	y-aktub-uu
3rd f.	t-aktub	t-aktub-aa	y-aktub-na

Source: Ryding (2005: 616)

In this analysis, I adopt the same structures for person I adopted in Blix (2021b), in turn based on ideas in Béjar (2003), Béjar and Rezac (2009), and Harley and Ritter (2002a): Third person forms simplest case, containing only a REF head, merged in the clausal spine on top of other structure.[5,6] The first person is composed by merging a PART(icipant) node on top of this, and the second person is in turn composed by adding ADD(ressee) on top of the first person.[7]

(18) The Internal Structure of Person

a. Third Person [REF [...]]

b. First Person [PART[REF [...]]]

c. Second Person [ADD[PART[REF [...]]]]

[5]I adopt this notation from Harley and Ritter (2002a), but REF(erential) is a slight misnomer, in particular when it comes to agreement systems, since not all third person elements are referential. While third person may thus well have a complex internal structure, it is sufficient for the current purposes to assume these structures, however.

[6]One way of conceiving of agreement in the current system where it is part of the functional spine, is in terms of Sportiche (2005, 2006) determiners; this perspective raises serious questions about clauses without agreement, as well as those with multiple agreeing verbs, however. Since there is no general consensus on the nature of agreement in the larger generative world either, I leave these questions aside for further research.

[7]Many languages have some syncretisms between second and third person to the exclusion of first person. I generally assume that a SPEAKER node is also available as an option of UG, and that languages may vary parametrically with respect to which participant is marked (cf. Harley and Ritter 2002a). Crucially for the current study, I will assume that in Arabic, this partial syncretism with respect to *t* comes about as a *ABA violation induced by higher material. In a recent paper, Vanden Wyngaerd (2018) takes this kind of possibility to be an argument against the Pointer-based approach, which I argue here to be wrong — I return to this question in more detail below.

As I noted in the introduction, I take number to be the region above person, in line with the idea of re-interpreting Noyer's (1992) hierarchy as reflecting syntactic structure. I do, however, not assume that gender is a distinct region, but rather that gender is part of the number hierarchy, following Ritter (1993), who argued that gender can be *on* number in certain languages. The assumptions I make are laid out in (19), with the brackets around FEM denoting its presence in the feminine, and absence in the masculine forms, as motivated by pairs such as *katab-a-t* (3FS.PERF) / *katab-a* (3MS.PERF). In empirical terms, this assumption is grounded in the fact that we find *t* in the third person feminine only in the singular and the dual: Under these structural assumptions, we can account for the fact that FEM is spelled out by *t* in the third person feminine singular and dual, but suffixally, by *na*, in the third person feminine plural: FEM is contiguous with SG only in the singular and the dual, but not in the plural. Further arguments for this containment structure will be given below, when I turn to the Perfect.

(19) The Internal Structure of Number+Gender

a. Singular [(FEM) [SG [...]]]

b. Plural [(FEM) [PL [SG [...]]]]

c. Dual [DU [PL [(FEM) [SG [...]]]]]

With these assumptions in place, we are ready to account for the paradigmatic distribution of affixes in the Jussive, assuming the larger order of the "regions" Tense, Number/Gender, and Person, as in (20).

(20) [Tense [Number/Gender [Person]]]

In particular, we can solve the issue of *y* and *t* by assuming the following specifications for the person-sensitive affixes.

(21) *REF Lexicalizing Affixes*

a. /*n*/ ⇔ '[FEM[PL[SG[PART[REF]]]]]'

b. /ʔ/ ⇔ '[JUSS →[FEM[SG[PART[REF]]]]]'

c. /*t*/ ⇔ '[FEM[SG → [PART[REF]]]]'

d. /*y*/ ⇔ '[SG[REF]]'

Since the person-sensitive affixes are sensitive to the lowest part of the structure, they necessarily determine the earlier cycles of spellout, given the assumptions about bottom-up, cyclic spellout that were laid out above. Spellout in the first person is rather trivial, since both the singular and the plural have their complete person/number/gender structure spelled out immediately in the first cycle, by ʔ and *n*, respectively (21a,b). In the second person, however, no mono-morphemic spellout of all the phi-structure is possible. As the examples in (22) show, there is no vocabulary item available to spell out the whole structure [ADD[PART[REF]]], and therefore only [PART[REF]] receives spellout in the first effective cycle. Three vocabulary items, *t*, ʔ, and *n* all compete for insertion, since they all match the span [PART[REF]]. The Elsewhere Principle determines the outcome in this case, and decides in favor of *t*, since the other two have a higher number of "unused" heads.[8] Since the first effective cycle of spellout occurs within the person region, this holds for all person/number combinations of the second person, explaining why it is not *t* that encodes these differences, but other affixes, i.e., later cycles.

[8]Note that this is partly due to the stipulation that ʔ spells out JUSS. The structure that *t* lexicalizes will be argued to be somewhat more extended, once we get to the perfect, to the point where I need to extend this stipulation somewhat. This is easily solved by assuming that JUSS is in fact a span that consists of two or more heads. Since nothing about this is surprising, or, from the current perspective, particularly interesting, I gloss over this fact here.

(22) *Second Person – First Cycle*

a. 2MS: [JUSS[SG[ADD[PART[REF]]]]]
t
n/a

b. 2FS: [JUSS[FEM[SG[ADD[PART[REF]]]]]]
t
n/a

In the third person, we derive the fact that the distribution of gender encoding is not uniformly prefixal or suffixal, in contrast to the second person: In the singular and the dual, gender is spelled out in the first cycle (and therefore, as I will show below, prefixal), it is encoded suffixally in the plural. Consider first the examples in (23), where we see that in both the singular and the dual, the presence of FEM leads to its spellout by *t*, due to cyclic overwrite. As the structure [FEM [SG [REF]]] is the largest structure that can be spelled out contiguously in (23b,d), with *t* as the solve candidate. In contrast, examples (23a,c) show that in the absence of FEM (i.e., in the third person masculine singular/dual) leads to the spellout of [SG[REF]] as the largest contiguous structure that finds a match in the vocabulary. Both *y* and *t* are candidates that match said structure — and since *y* has zero superfluous heads, while *t* has two, the EP decides in favor of *y*.

(23) *Third Person – Singular & Dual*

a. 3MS: [JUSS [SG [REF]]]
y
n/a

b. 3FS: [JUSS [FEM [SG [REF]]]]
t
n/a

c. 3MD: [JUSS [DU [PL [SG [REF]]]]]
y
n/a

d. 3FD: [JUSS [DU [PL [FEM [SG [REF]]]]]]
t
n/a

Next, consider the plural of both feminine and masculine third person, as in (24): Neither *t*, nor *y* can spell out the span [PL[SG[REF]]]. As the largest span that can be spelled out in the masculine as well as the feminine is therefore [REF[SG]], the EP returns the same verdict as in the parallel (i.e., third person masculine singular/dual) case above: *y* wins over *t* due to the elsewhere principle.

(24) *Third Person Plural*

a. 3MP: [JUSS [PL [SG [REF]]]]
y (SG [REF]); *n/a* (PL [SG [REF]])

b. 3FP: [JUSS [FEM [PL [SG [REF]]]]]
y (PL [SG [REF]]); *n/a* (FEM [PL [SG [REF]]])

The system thus successfully derives the distribution of *y* and *t* without resorting to any means but the principles governing the matching of vocabulary items to a syntactic context that receives spellout, and an Elsewhere Principle governing competition between multiple matches. Crucial to this is the fact that natural classes arise in terms of (partial) containment relationships, rather than SPE-style subsets; no disjunctive set of rules needs to be invoked to derive the fact that *t* can spell out FEM in the third person (non-plural), but not the second person, or the fact that *y* surfaces in the third person plural; instead the apparent disjunction follows as a possibility from the idea that an item may lexicalize a span of spans. These facts can now be interpreted as following from contiguity within a syntactic interpretation of Noyer's (1992) hierarchy.

Note that, in the third person plural, as in the second person, we are not dealing with a case which induces a syncretism (by spelling out a span that contains FEM in the same way as one that does not), but rather one in which the spellout of FEM has not happened in the *first* cycle. Now that we have dealt with the first cycle of all person/number/gender combinations, we can provide an analysis of the remaining cycles, and show how the variation in the locus of gender spellout that Noyer termed *Discontinuous Bleeding* comes about as an effect of the cyclical nature of spellout. Recall that a second cycle will now begin, spelling out the next span, and again,

in accordance with cyclic overwrite, it will target the largest span that a matching vocabulary item can be found for. Continuing with the third person, consider first the singular: Both in the masculine and the feminine case, the only remaining head to be spelled out is JUSS. I thus simply follow Noyer (1992) in assuming that there is a zero affix (25) that spells out the Jussive head.

(25) */Ø/* ⇔ '[JUSS]'

The third person singular forms, *t-aktub* (3FS) and *y-aktub* (3MS), are thus spelled out cyclically as shown in (26):

(26) *Third Person – Singular*

a. 3MS: $\overbrace{[\ \text{JUSS}}^{Ø}\ \overbrace{[\ \text{SG}\ [\ \text{REF}\ ]]]}^{y}$

b. 3FS: $\overbrace{[\ \text{JUSS}}^{Ø}\ \overbrace{[\ \text{FEM}\ [\ \text{SG}\ [\ \text{REF}\ ]]]]}^{t}$

Under the assumption that the dual affix *aa* lexicalizes the structure in (27), the third person dual forms, *y-aktub-aa* (3MD) and *t-aktub-aa* (3FD) , are derived in the same fashion, i.e., with the remaining higher material spelled out fully in the second cycle, as in (28).

(27) */aa/* ⇔ '[JUSS[DU[PL]]]'

(28) *Third Person Dual*

a. 3MD: $\overbrace{[\ \text{JUSS}\ [\ \text{DU}\ [\ \text{PL}}^{aa}\ \overbrace{[\ \text{SG}\ [\ \text{REF}\ ]]]]]}^{y}$

b. 3FD: $\overbrace{[\ \text{JUSS}\ [\ \text{DU}\ [\ \text{PL}}^{aa}\ \overbrace{[\ \text{FEM}\ [\ \text{SG}\ [\ \text{REF}\ ]]]]]]}^{t}$

Finally, consider the third person plural forms, where FEM had not been spelled out in the first cycle. Consequently, we do not (necessarily) expect a partial syncretism (i.e., a syncretism

for the second cycle) that would parallel the way the second cycle was identical for the masculine and the feminine forms of the singular and the dual respectively: In these cases, the first cycle of the feminine and the singular spelled out different structures, with different affixes, but the second cycle spelled out the same structure. The result was a first cycle difference, and a second cycle syncretism. The opposite effect holds in the plural, where the first cycle was identical, but the second cycle results in a difference, namely *y-aktub-uu* (3MP) and *y-aktub-na* (3FP). We can simply assume that the second cycle spells out the remaining material with the affixes in (29), as in (30). This way, the present system derives Noyer's *discontinuous bleeding*, without reference to autonomous morphological structure, or templates of any kind.

(29) a. */uu/* ⇔ '[JUSS[PL]]'

b. */na/* ⇔ '[JUSS[FEM[PL]]]'

(30) *Third Person Plural*

a. 3MP: $\overbrace{[\ \text{JUSS}\ [\ \text{PL}}^{uu}\ \overbrace{[\ \text{SG}\ [\text{REF}]]]]}^{y}$

b. 3FP: $\overbrace{[\ \text{JUSS}\ [\ \text{FEM}\ [\ \text{PL}}^{na}\ \overbrace{[\ \text{SG}\ [\text{REF}]]]]]}^{y}$

Returning to the second person, and its subsequent cycles of spellouts, it is worth partly repeating the paradigm to point out some further interesting distributional facts. Table 3.5 shows the second and third person of the Jussive.

Table 3.5: Arabic — Jussiv Second/Third Person (repeated)

	Number		
Person/Gender	Singular	Dual	Plural
2nd m.	t-aktub	t-aktub-aa	t-aktub-uu
2nd f.	t-aktub-ii	t-aktub-aa	t-aktub-na
3rd m.	y-aktub	y-aktub-aa	y-aktub-uu
3rd f.	t-aktub	t-aktub-aa	y-aktub-na

Source: Ryding (2005: 616)

Notice that, with the exception of *t-aktub-ii* (2FS), there is a general suffixal syncretism between the second and the third person: Whatever suffix a number/gender combination in the third person shows, is also found in the corresponding second person. That is to say, we find *Ø* in the masculine singular, *aa* in the dual, *uu* in the masculine plural, and *na* in the feminine plural, both in the second, and in the third person. Notice further that there is no overt gender contrast in the second person dual, with both the masculine and the feminine form being *t-aktub-aa*.
The sole exception is easily derived by assuming that *ii*, once again, simply spells out the remainder of the structure (i.e., *ii* lexicalizes the structure in (31a)), after *t* has spelled out [PART[REF]] in the first cycle, as in (22b). The remaining facts are captured simply by postulating a zero spellout for ADD, as in (31b).

(31) *ADD lexicalizing affixes*

a. */ii/* ⇔ '[JUSS [FEM [SG [ADD]]]]'

b. */Ø/* ⇔ '[FEM [SG [ADD]]]'

In the singular and plural, spellout is fairly straightforward: Only in the second person singular feminine does *ii* spell out the whole remaining structure. In the second person singular masculine, as well as the plural, *Ø* spells out [SG[ADD] — *ii* cannot cyclically overwrite *Ø*, due to contiguity. In particular, in the masculine singular, [JUSS] cannot be spelled out, because *ii* does

not lexicalize [JUSS[SG]], only [JUSS[FEM[SG]]], of which [JUSS[SG]] is not a contiguous sub-span. In the plural, it is PL that blocks the contiguity with JUSS. In these cases, both *ii* and *Ø* are candidates for insertion, but since *Ø* is the smaller affix, the EP decides in its favor. Since *Ø* spells out [SG[ADD]] in both cases, the remaining structure to be spelled out is identical to the second cycles in the third person plural, and consequently we derive the suffixal syncretism.

(32) *Second Person – Singular & Plural*

a. 2MS: [JUSS [SG [ADD [PART [REF]]]]] — Ø Ø t

b. 2FS: [JUSS [FEM [SG [ADD [PART [REF]]]]]] — *n/a* ii t

c. 2MP: [JUSS [PL [SG [ADD [PART [REF]]]]]] — uu Ø t

d. 2FP: [JUSS [FEM [PL [SG [ADD [PART [REF]]]]]]] — na Ø t

Note that I postulated that *Ø* in (31b) lexicalizes FEM, but is overwritten by *ii* in the second person singular feminine, as shown in (32b). This accounts for the fact that the second person exhibits a gender syncretism in the dual, as can be seen in (33).

(33) *Second Person – Dual*

a. 2MD: [JUSS [DU [PL [SG [ADD [PART [REF]]]]]]] — aa Ø t

b. 2FD: [JUSS [DU [PL [FEM [SG [ADD [PART [REF]]]]]]]] — aa Ø t

Here, the second cycle shows a simple superset effect, resulting in a simple syncretism: There is no affix that is anchored at ADD that spans all the heads up to PL, and therefore the cycles [SG[ADD]] (33a) and [FEM[SG[ADD]]] (33b) receive spellout. The smallest vocabulary item that can lexicalize this span is, in both cases, *Ø*. Note that this is similar to the situation we saw in the second person feminine singular (32b), where *Ø* also matched [FEM[SG[ADD]]], but was overwritten by *ii*. Here, however, we find the heads PL and DU blocking the contiguity between JUSS and FEM that are required for *ii* to match the whole span. Thus it competes with *Ø* for the spellout of [FEM[SG[ADD]]], and loses due to the EP, and this is what produces the syncretism.

We have thus derived the paradigmatic distribution of the affixes relevant to the paradigm of the Jussive. In assuming that Noyer's (1992) hierarchy reflects properties of a syntactic structure that is spelled out cyclically, in a bottom-up manner, and span by span, the current system still can account for the effects modeled as Impoverishment, both hierarchy governed, such as the loss of overt gender marking in 2FS, and arbitrary, such as the distribution of *t* and *y* (in the analysis of Halle, 2000).

In the next steps I will very briefly discuss the subjunctive and the imperfect, and then show how we can derive the linear order of the affixes from the simple fact that it is, in all cases, the first cycle of spellout that becomes the prefix.

3.3.2 The Imperfect and the Subjunctive

Both the Imperfect and the Subjunctive are fairly trivial extensions of the Jussive. Insofar as the main concern here is the agreement structure, and insofar as the relatively small amount of additional morphology is not particularly informative with respect to their possible internal structures, I will only provide a very brief account for the sake of completeness. Table 3.6 provides the subjunctive paradigm. As highlighted, the only obvious difference between the Subjunctive and the Jussive is the fact that the former shows a suffix *a* in all and only those cases where the Jussive has no overt suffix.

Table 3.6: Arabic — Subjunctive

	Number		
Person/Gender	Singular	Dual	Plural
1st	ʔ-aktub-a		n-aktub-a
2nd m.	t-aktub-a	t-aktub-aa	t-aktub-uu
2nd f.	t-aktub-ii	t-aktub-aa	t-aktub-na
3rd m.	y-aktub-a	y-aktub-aa	y-aktub-uu
3rd f.	t-aktub-a	t-aktub-aa	y-aktub-na

Source: Ryding (2005: 609)

The Imperfect, in Table 3.6, is characterized partly by the same description as the Subjunctive, in that the same context that show no overt suffix in the Jussive, and *a* in the Subjunctive, show *u* in the Imperfect. In addition, however, we find *na/ni* suffixed to those forms ending in *aa*, *ii*, or *uu*; the only forms that are identical across all three paradigms are the ones ending in *na*, i.e., the second and third person feminine plural.

Table 3.7: Arabic — Imperfect

	Number		
Person/Gender	Singular	Dual	Plural
1st	ʔ-aktub-u		n-aktub-u
2nd m.	t-aktub-u	t-aktub-aa-ni	t-aktub-uu-na
2nd f.	t-aktub-ii-na	t-aktub-aa-ni	t-aktub-na
3rd m.	y-aktub-u	y-aktub-aa-ni	y-aktub-uu-na
3rd f.	t-aktub-u	t-aktub-aa-ni	y-aktub-na

Source: Ryding (2005, p.441) , Noyer (1992, p. 60)

We can account for the absence of any additional affixes with *-na* by simply assuming that it is in fact capable of spelling out the whole Tense structure, and revising (29b) as (34). Given that it is anchored at PL, and spells out the remaining structure in all cases, we account for the syncretism in the usual way, that is, by containment.

(34) */na/* ⇔ '[PRS[SBJV[JUSS[FEM[PL]]]]]'

It has long been recognized, of course, that the distribution of the suffixes *a*, *u*, and *na/ni* in the remaining can be described well in morpho-phonological terms: We get *a* and *u* at the edge of the verbal stem, and *na/ni* after long vowels. I follow this line of thought by simply assuming that *a* spells out the SBJV head, but deletes after long vowels, for phonological reasons[9], i.e., to avoid an illicit structure /V:V/. To keep the phonological characterization of the distribution of *u* and *nV* in place, I will simply assume a phonologically driven allomorph selection for *u* / *na* / *ni*, and assume that the difference between *na* and *ni* is also phonologically resolved.[10]

(35) a. */a/* ⇔ '[SBJV]' (*with* /a/ → /Ø/ / V:__)

b. $\left\{ \begin{array}{l} /u/, \\ /nV/ \,/\, \text{V:}\_\_ \end{array} \right\}$ ⇔ '[PRS[SBJV]]'

Insofar as the differences between the Imperfect, the Subjunctive and the Jussive are not of great interest for the distribution of the agreement affixes, which show otherwise identical distribution across these three paradigms, I leave it at this brief description.

3.3.3 Deriving a Prefix

The discussion so far has abstracted away from the linear order of affixes, both with respect to each other, and with respect to the verbal stem. The perspective developed so far does, however, offer the possibility for a simple characterization of the distribution of the relevant affixes as prefixes. In every case of the prefixal paradigms, it is the *first cycle* that becomes a prefix, and

[9] See e.g.Brame (1970), Rosenthall (2006) for details on the deletion of short vowels in the context of long ones.

[10] A characterization of these distributions in the terms advanced so far is, of course, also plausible, with the varying affixes just below SBJV and IMPF spelling out different sizes of the Tense region. However, insofar as the internal structure of Tense is of limited importance, given that it does not affect the distribution of the agreement markers themselves, at least insofar as limited to the contrasts between Imperfect, Subjunctive, and Jussive. It would also lose the insight that the distribution of these extra affixes appears to be characterized by phonological criteria. However, nothing crucial hinges on either take here.

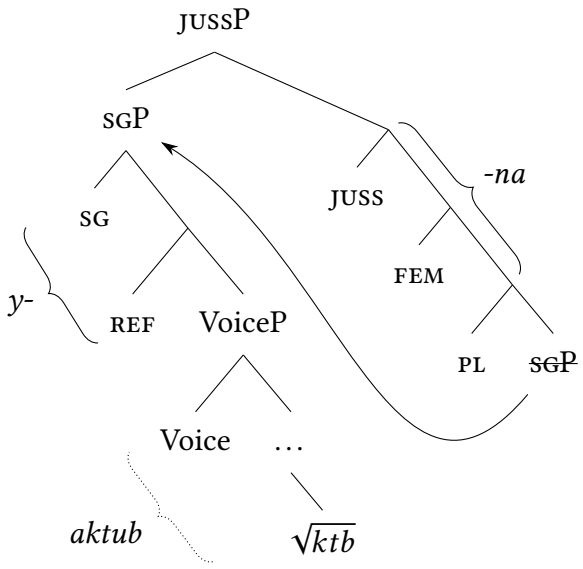
the first cycle is uniquely characterized by spelling out REF, as this is the first cycle's anchor in every agreement configuration. In this subsection I show how this correlation may be captured under an Antisymmetry perspective (Kayne, 1994; Koopman, 2017b) of linear order, and derived by phrasal movement. In particular, I assume that the JUSS head carries an EPP feature that attracts REF, and that movement is constraint so as to not break up a span that received spellout, along the lines of (36).

(36) *Span-based Pied-Piping — No Breakup*

EPP-driven movement of HP targets the phrase headed by be highest head in the span that spells out H.

Consider, for example, the third person feminine plural form *y-aktub-na*, as in (37). The highest head, JUSS, has an EPP property, attracting REF. Since REF is spelled out in a span [SG[REF]], the whole phrase headed by SG moves to the specifier of JUSS, as per the span-based pied-piping introduced above. Antisymmetry now gives us the correct result: Since every element of the span spelled out by *y* c-commands the VoiceP, *y* must linearly precede it. As the whole SGP c-commands the elements of the span spelled out by *na*, it must linearly precede it as well, i.e. *y* must precede *aktub*, and they together must precede *na*, resulting in *y-aktub-na*.

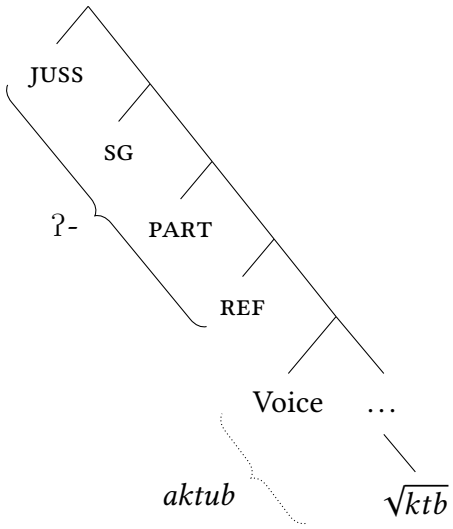
(37) 3FP: '*y-aktub-na*'



The same applies to most other prefixes: Since REF pied-pipes the part of the structure that is contained in the span it is spelled out with, we derive *t* as a prefix in a parallel way. However, ʔ has been argued to spell out JUSS, i.e., JUSS and REF receive spellout in the same span. Since movement out of the span that includes both JUSS and REF is blocked, nothing happens in this case.[11]

[11]If an unchecked EPP property always crashes a derivation, as some approaches hold, this would suggest that being interpreted within the same span is a different way of providing the correct locality conditions. If (at least some) EPP properties are instructions to the syntactic algorithms that apply only if they can, we do not need to say anything additional.

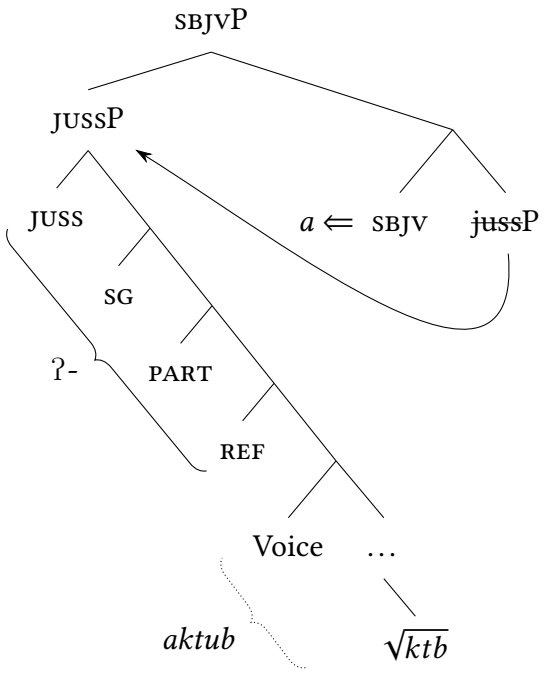
(38) 1s: 'ʔ-*aktub*'(Jussive)



Since no movement occurs in this structure, we still derive a prefix; every element of the span that ʔ spells out, c-commands the VoiceP, and thus ʔ must precede *aktub*. We do, however, see a suffix in conjunction with ʔ in the Subjunctive and the Imperfect. Koopman (2017b) argues that suffixes may be marked with an EPP feature that moves the whole complement to its specifier — in the current approach, we can translate this into a property of a head, rather than an affix, and suggest that SBJV bears such an EPP feature.[12] Once again, under the assumption that movement may not break apart a span (i.e., move neither out of a span, nor into a span), we derive a system under which such an EPP feature triggers movement of the complement of the lowest element of the span to the specifier of the highest element. In the simplest case, such as *a*, which spells out a single head, we derive the Subjunctive counterpart ʔ-*aktub-a* to the Jussive ʔ-*aktub* simply via movement of the SBJV's complement to its specifier position, as in (39).

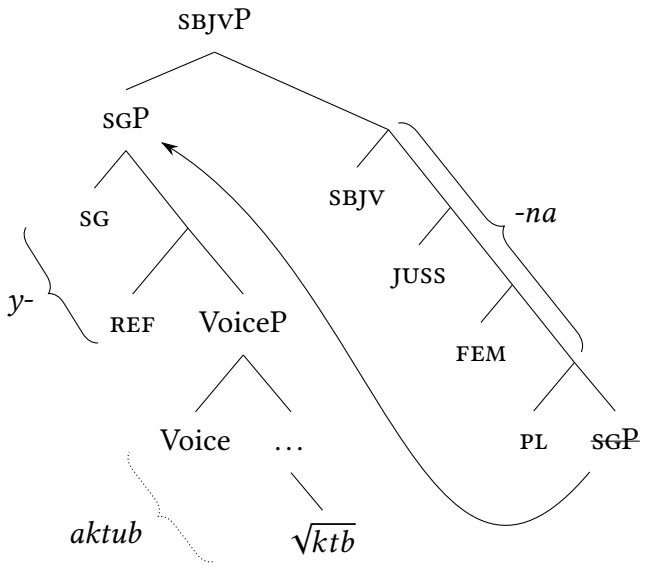
[12]Starke (2018) develops some ideas for a more principled account for this kind of movement, frequently dubbed "snowball movement" in the cartographic literature, by deriving it from a constituent based spellout, rather than a phrase-based one. In that perspective, prefixes are specifier-like structures, i.e., complex left branches with a binary structure [A B] at its bottom, whereas suffixes are structures with a unary bottom, [A [B [X]]], that cause movement of the complement XP to form a constituent that can receive spellout.

(39) 1s: '*ʔ-aktub-a*' (Subjunctive)



In the case of affixes like *na* that can spell out *sbjv* together with juss, the movement of ref to the specifier of sbjv may satisfy the requirements of both juss and sbjv at the same time. In (40), the movement of sgP is simultaneously the movement of the phrase pied-piped by ref into the nearest specifier available to juss that does not break up a span, and the movement of the highest phrase below sbjv that is not spelled out within the same span as sbjv.

(40) 3FP: '*y-aktub-na*' (Subjunctive)



What is particularly interesting about this approach is, of course, not merely that it can derive the linear order of the affixes, but that it does so by linking it once again to the structural interpretation of the hierarchy discovered by Noyer (1992). In doing so, the current approach provides an explanatory account of a number of interrelated properties of the Arabic agreement system: It accounts for the *y/t* contrast purely in terms of matching, rather than stipulating accidental homophony, or arbitrary Impoverishment rules; in fact all relevant Impoverishment is reduced to matching as well. It provides an account of discontinuous bleeding, i.e., the fact that a feature that is spelled out by the prefix is not spelled out at a suffix position, that is derived in purely from the syntactic order of merge. Finally, it accounts for the set of prefixes in an explanatory manner, rather than just stipulating certain affixes as prefixes, in the same manner: The first cycle must necessarily spell out [REF], since it is the lowest head, and it is the cycle that spells out REF that becomes a prefix.

As I have thus shown in this section, the structural approach to the spellout of agreement can

explain facts that are merely a list of accidents (such as limited competition among prefixes only) in competing approaches. In the next section, I discuss the suffixal paradigm, which shows that the listing approach is not only less elegant, but in fact contradictory: The affixes in question *cannot* be prefixes by stipulation, since they occur as suffixes in other paradigms.

3.4 The Perfect

We can now consider the Perfect, i.e., the suffixal paradigm. As in the previous section, I begin by focussing on the paradigmatic distribution, and discuss the linear facts after that. Table 3.8 shows an interesting pattern to the distribution of *t*: In every case (i.e., person/gender/number combination) that had a prefixal *t* in the non-perfect paradigms discussed above, we also see a suffixal *t* in the perfect, highlighted in light gray. In particular, we see that *t* occurs in the third person feminine non-plural, as well as all second person cells. I take this identity in distribution to strongly suggest that they are indeed the same object. In addition, however, we find that the first person singular, highlighted in a darker shade, also shows a suffixal *t* — a fact not entirely unsurprising, given the fact that I argued *t* to lexicalize the structure of first person.

Table 3.8: Arabic — Past Tense Indicative (Perfect)

	Number		
Person/Gender	Singular	Dual	Plural
1st	katab-t-u		katab-naa
2nd m.	katab-t-a	katab-t-um-aa	katab-t-um
2nd f.	katab-t-i	katab-t-um-aa	katab-t-un-na
3rd m.	katab-a	katab-aa	katab-uu
3rd f.	katab-a-t	katab-a-t-aa	katab-na

Source: Ryding (2005, p.443)

There are a number of facts that need to be accounted for: First and foremost, the fact that *t* is a prefix in the non-perfect paradigms, but a suffix here. In addition, suffixal *t* varies in its position,

preceding *a* in the second person singular masculine, but following it in the third person singular feminine, differing in a highly interesting manner from the prefixal paradigms, where these two are always syncretic, as shown in Table 3.9. I will relate this to the pointer in the structure of *t*, and suggest that the difference in position is a result of *t* spelling out either the top or the bottom span that it lexicalizes.

Table 3.9: The Loci of *t*

P/G/N	Subjunctive	Perfect
2ms	t-aktub-a	katab-t-a
3fs	t-aktub-a	katab-a-t

Secondly, the Dual shows a curios pattern of containment: As Table 3.10 shows, the Dual properly contains the Plural form in the second person masculine. In the third person singular feminine, however, the Dual properly contains the Singular form, yet the same suffix *aa* is added in both cases.

Table 3.10: Dual Containment in the Perfect

P/G	sg	du	pl
2m	katab-t-a	katab-t-um-aa	katab-t-um
3f	katab-a-t	katab-a-t-aa	katab-na

Note in particular, that the suffix *um* in the second person masculine dual/plural is strictly limited to a highly specific context: It does only occur in the second person non-singular, and only in the Perfect, i.e., not in any of the prefixal paradigms. It is thus clearly reflective of both the second person structure, and the structure of the perfect, but cannot spell out the complete number structure, since the dual affix *aa* appears to still spell out du. However, the Dual clearly breaks contiguity of Person and the Tense structure — I therefore conclude that the Perfect is

marked lower than the Tense structure we have seen in the prefixal paradigms, namely between Number and Person. We thus extend Noyer's hierarchy slighty, as follows:[13]

(41) *Revised Hierarchy*

Tense > Number/Gender > Perfect > Person

As for the Tense structure that the Perfect is co-occuring with, I take pairs such as *katab-t-a* (2SM.PERF), *t-aktub-a* (2SM.SBJV) from Table 3.9, to suggest that the Perfect shares the higher structure with the Subjunctive.[14]

From that perspective, then, we have additional evidence for the head ADD that has been spelled out by a zero affix in most cases seen so far; we can revise the previous instance of the zero affix, as in (42a), and add the two new affixes *i* (42b) and *um* (42c) to our lexicon.[15]

(42) *ADD Lexicalizing Affixes (Perfect)*

a. */Ø/* ⇔ '[SG[PERF[ADD]]]'

[13] Arguing that the Arabic morphological facts point towards a mixed Tense/Aspect system has some (albeit contested) support from other works on the matter: (Bahloul 2008: 51-57) makes the point that the Imperfect denotes a relation between utterance time and topic time, i.e., it behaves like Tense proper, whereas the Perfect can actually occur with past, present and future meanings in different contexts. Similarly, Comrie (1998: §4.4) argues that the Imperfective/Perfective contrast in Standard and Classical Arabic is a mixed Tense/Aspect system. There appears to be no general consensus on the matter, however, on the question of the precise meaning and structure of the categories that these forms express. It is my hope that the investigation undertaken here might be fruitful in further exploring the intricacies of the meaning and structure of these categories, but for now I leave the more semantic side of the question aside, and point the reader to Bahloul (2008: §3) for a general overview of stances that have been taken in the literature.

[14] A more plausible take might be to suggest that the structure of the imperfect is more complex, embedding the subjunctive first under an indicative structure which is then in turn embedded under the present tense structure, as in (i). From that perspective, affixes like *a* might lexicalize [IND[SBJV]], leading to the desired syncretism without implying that the Perfect is a form of the Subjunctive. I leave that issue aside both for expository reasons, and since I do not have much evidence to bear on the internal structure of the Tense/Aspect/Mood configurations themselves, beyond claims about their containment.

(i) [PRS[IND[SBJV[JUSS]]]]

[15] This obviously does not account for the similarity between *i* and *ii* that both occur only in the second person singular feminine. A plausible alternative might be to suggest that these are allomorphic variants of the same affix (i.e., lexicalizing a pointer above PERF), with lengthening (or shortening) determined on other grounds.

b. */i/* ⇔ '[SBJV[JUSS[F[SG[PERF[ADD]]]]]]'

c. */um/* ⇔ '[SBJV[JUSS[PL→[F[SG[PERF[ADD]]]]]]]'

These specifications derive the containment facts that hold for the second person in the Perfect; in particular it accounts for the fact that the second person masculine dual contains the second person masculine plural form.

Consider first the second person singular. Its spellout proceeds essentially in a way that is parallel to the spellout we have seen in the prefixal paradigms, modulo the difference between *i* and *ii*, as can be seen in (43).

(43) *Second Person Singular – Perfect*

a. 2MS: [SBJV [JUSS [SG [PERF [ADD [PART [REF]]]]] a Ø Ø t

b. 2FS: [SBJV [JUSS [FEM [SG [PERF [ADD] [PART [REF]]]]] i t Ø

Next, we derive the second person plural masculine, in (44). Again, spellout proceeds in the usual way, with *t* spelling out [PART[REF]], and *um* spelling out the remaining structure.

(44) *Second Person Plural Masculine – Perfect*

2MP: [SBJV [JUSS [PL [SG [PERF [ADD [PART [REF]]]]]] um t

The second person plural feminine, however, has a curious property. In the form *katab-t-un-na*, we can clearly identify *na*, which spells out the feminine plural in the second and third person, across all paradigms. The form appears to properly contain the masculine plural, *katab-t-um*, however, presumably with a process of nasal assimilation changing *um* to *un*.[16] There thus

[16]If the reader is opposed to such an analysis, suggesting that *unna* is a simple morpheme does, of course, pose

appears to be a doubly marked plural, expressed by both *um* (*un*) and *na*, as indicated in (45).

(45) *Second Person Plural Feminine – Perfect*

2FP: [SBJV [JUSS [FEM [PL] [SG [PERF [ADD [PART [REF]]]]]]]

(SBJV [JUSS [FEM [PL] = na; PL] [SG] = um (→ [un]); PART [REF]]]]]]] = t)

I propose that the reason for this double marking lies in the fact that there is no Vocabulary item that can be anchored at FEM. That is, at the point at which *um* has spelled out the structure up to PL, there is no applicable vocabulary item for the next cycle of spellout. I suggest that a "last resort" mechanism may lead to lower material being part of the target of spellout, along the lines of (46).[17]

no further problems, it would be solved rather trivially by suggesting that *unna* lexicalizes the same structure as *um*, modulo a FEM head above PL. The fact, however that a string *um* appears in all other second person perfect non-singular forms, coupled with evidence for *na* being a morpheme in its own right, suggests to me that analyzing the string *unna* as bimorphemic is a more plausible path to take.

[17]Note that more obvious cases of double marking do exist within Semitic, and that they can occur discontinuously as well, for example in Modern Hebrew (Berman 1997: 318):

(i) a. *te-saper* 'she will tell'
b. *ye-saper* 'he will tell
c. *te-sapr-na* 'they (fem) will tell'
d. *ye-saper-u* 'they (masc) will tell'

Under the assumption that FEM in Hebrew is always located just above SG, we could account for this by suggesting that both forms involve this kind of rewinding, e.g. as follows:

(ii) *Hebrew (Toy) Lexicon*
a. */y/* ⇔ '[SG[REF]]'
b. */t/* ⇔ '[FEM[SG[REF]]]'
c. */u/* ⇔ '[PL[SG]]'
d. */na/* ⇔ '[PL[FEM]]'

The spellout would thus proceed as in (iii), with rewinding causing the dual marking in both cases.

(iii) a. 3MP [PL [SG] [REF]]] (*no anchor*; SG + [REF] = y; PL = u)

b. 3FP [PL [FEM [SG [REF]]]] (*no anchor*; FEM = na; SG [REF] = t)

(46) *Rewind*

If no Vocabulary Item can be anchored at the current cycle, the target of spellout includes the head of the (previous) anchor's complement.

In fact, the same mechanism allows us to explain the containment facts that hold for the dual. Recall that *aa* occurs across all paradigms in all the Dual cells, and that we postulated the structure in (27). Crucially, the fact that it occurs in *all* dual cells (and only there) leads us to naturally assume that it is also the only affix to lexicalize DU.

(27, *repeated*) */aa/* ⇔ '[JUSS[DU[PL]]]'

Since there is no other affix that can spell out DU, we expect the same kind of rewinding to occur if the previous cycle spelled out PL, in order to find an anchor for *aa*. As (47) shows, this assumption immediately derives the correct set of affixes, assuming, as before, that *aa* and *a* merge phonologically. Note that we derive the same gender syncretism as in the prefixal paradigms in an identical way, i.e., by suggesting that *um*, like *Ø* (31b) lexicalizes the low FEM.

(47) *Second Person Dual – Perfect*

a. 2MD: [SBJV [JUSS [DU [PL] [SG [PERF [ADD [PART [REF]]]]]]]

a: [SBJV
aa: [JUSS [DU [PL]
um: [SG [PERF [ADD
t: [PART [REF]]]]]]]

This approach would suggest that cases of double marking that do not involve actual doubling in the syntax (such as negative concord) would involve structurally overlapping vocabulary items in a fairly limited set of contexts, namely the absence of an anchorable vocabulary item, and that they should occur only when the doubly expressed element is expressed within structures that are both contiguous with the doubled element, i.e., multiple exponence should be an extremely locally determined phenomenon. Obviously, this hypothesis needs to be checked against a much broader set of data; as a preliminary heuristic, however, I took a look at the set of analyses of prefixal Afro-Asiatic paradigms that Noyer (1992) provides. In his approach, affixes can be secondary exponents of a feature, i.e., apply only in those contexts that bear the relevant feature, but without blocking application of other affixes. All the affixes that exhibit secondary exponence that he provides refer either to a context of number and gender, such as "pl (f)" (i.e., plural in the context of feminine), like the one provided in (45), or to second person as the secondary feature, a case that the current approach deals with by suggesting that application of *t* leaves ADD as the anchor of the next cycle. It appears, then, that the current interpretation of his hierarchy might be extendable to account for the distribution of multiple exponence as well, but further research is needed to ensure that this is indeed a correct characterization of the facts.

b. 2FD: [SBJV [JUSS [DU [PL [FEM [SG [PERF [ADD [PART [REF]]]]]]]]]]

(a: SBJV; aa: JUSS; um: PL; t: ADD)

Turning to the first person forms, we find *katab-t-u* in the singular, and *katab-naa* in the plural. For the plural, we thus have to account for the fact that *n* occurs across paradigms, and appears to include both person and number. Given that we determined PERF to be structurally located between these two regions, we thus have to conclude that *n* can in fact spell out PERF as well, given that it must span across both PL and the first person structure. Furthermore, given that *n* can spell out PL without spelling out PERF, as we saw in the previous discussion of the prefixal paradigms, we can conclude that PERF must be under a pointer, i.e., lexicalized without requiring strict contiguity. The distributional facts thus lead to the specification in (48). As for the fact that the suffixal structures result in *naa*, rather than *na*, I simply follow Noyer (1992: 98) in assuming that the affix' form is underlyingly /*na*/, but that a truncation process applies in the prefixal position.

(48) */n(a)/* ⇔ '[FEM [PL [SG → [PERF [PART [REF]]]]]]' (*revised (21a)*)

(49) *First Person Plural – Perfect*

1PL: [SBJV [JUSS [PL [SG [PERF [PART [REF]]]]]]

(a: SBJV; Ø: JUSS; n(a): PL SG PERF PART REF)

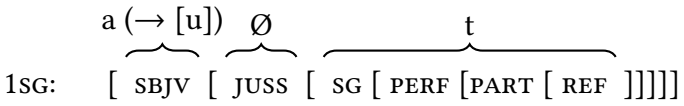
This analysis raises an issue for the first person singular: Since *n(a)* can spell out the whole span from SG to REF in the first cycle, *t* must at least be able to spell out the same span, so as to not be overwritten. That is, we need to revise *t* to include PERF as well, as in (50).

(50) */t/* ⇔ '[FEM[SG → [PERF [PART [REF]]]]]' (*revised (21c)*)

This accounts for the fact that *t* spells out the first person singular in the perfect, while ʔ does so in the prefixal paradigms: While the latter can overwrite *t* in the absence of PERF, it

is *t* that overwrites ʔ in its presence: Since ʔ does not lexicalize PERF, it is merely a candidate for [PART[REF]], and thus overwritten cyclically by *t*. The analysis cannot account naturally for the presence of a suffix *u*, however, given that all person/number/gender material will receive spellout by *t*. Comparing the first cycle in (51) to the first two cycles in (43a), we find that the next cycle should be identical; we do however find *katab-t-u* in the first person singular, and *katab-t-a* in the second person singular masculine. Given that situation, I am currently left to stipulate that *a* has an allomorphic variant [u], conditioned by the first person singular perfect structure.[18]

(51) *First Person Singular – Perfect*



Turning to the third person, a number of aspects are of interest. Firstly, the comparison between the third person singular feminine and masculine across the prefixal and the suffixal paradigm, as given in Table 3.11, shows that there is no overt suffixal counterpart to *y*: While the prefixal paradigms mark the difference in gender with a different prefix, the suffixal paradigms show that the feminine form *katab-a-t* (3FS.PERF) properly contains the masculine form *katab-a*

[18]The issue would not arise, if number on the first person is marked not in the number region, but rather in the person region; we could then argue that *t* does not lexicalize PERF at all, but cannot be overwritten by *n(a)*, because the relevant head, say, MULTISPEAKER, would break contiguity between PERF and PART in a vocabulary item like (i), and thus not be able to overwrite *t* for [PERF[PART[REF]]]. If we further assume, that a first person structure only encodes presence or absence of a PL head, the following specification would resolve two issues:

(i) */n(a)/* ⇔ '[FEM [PL → [PERF [MULTISPEAKER [PART [REF]]]]]]'

Firstly, from that perspective, *t* would not lexicalize PERF. Insteadl, *u* could spell out the relevant structure, anchored at PERF, thus avoiding the contextual allomorphy rule. This is compatible with the proposed spellout for the second and third person, where PERF is spelled out with ADD or REF respectively. Secondly, this affix points towards a unification of *n(a)* and *na* as a plausible analysis, since it now matches the feminine plural structure. I have not pursued this path here, insofar as I mostly focused on contrasts between and within second and third person, but Ritter (1997) provides some arguments along those lines, arguing that Arabic does in fact mark the first person plural in the person domain, rather than the number domain. Harley and Ritter (2002a) provide similar points, suggesting that in some languages that distinguish number only in the first person, the relevant difference is found in the person domain. A future refinement of the current analysis that is based on a larger, crosslinguistic dataset, might thus take this as a plausible direction to take.

(3MS.PERF).

Table 3.11: Third Person Singular

P/G/N	SUBJUNCTIVE	PERFECT
3MS	y-aktub-a	katab-a
3FS	t-aktub-a	katab-a-t

The same absence of *y* extends to the plural forms, where we find *y* in the prefixal paradigms: While the prefixal subjunctive forms are *y-aktub-uu* (3MP.SBJV) and *y-aktub-na* (3FP.SBJV), the corresponding perfect forms are *katab-uu* and *katab-na*, i.e. we find the same suffixes, but no corresponding *y*. In contrast, we find *t* in the third person singular/dual feminine, i.e., we find suffixal *t* in precisely those contexts where we find prefixal *t*, and no suffixal counterpart to *y*. These facts follow immediately, if we assume another zero affix, as in (52), that can cyclically overwrite *y* in the perfect.[19]

(52) /Ø/ ⇔ '[SG[PERF[REF]]]'

In the third person singular, this derives the surface containment facts as in (53): In both the masculine and the feminine, Ø spells out the structure [SG[PERF[REF]]], since no other vocabulary item can spell out a bigger structure. In particular, *t* requires contiguity between PERF and PART in order to spell out PERF and is thus not a candidate for insertion. The first cycle is thus identical between the feminine and the masculine in the third person singular perfect, contrasting with the same cells in the prefixal paradigms, where the contrast was established here. In masculine, the remaining structure to be spelled out is the Tense structure, and spellout proceeds in the usual fashion. In the feminine, however, we find the same issue as before, namely that no vocabulary

[19] A plausible alternative is that Ø actually is the suffixal counterpart of /*y*/ ⇔ '[SG → [PERF[REF]]]', but that *y* is phonologically deleted in suffixal position. According to Rosenthall's (2006) analysis of the verbal phonology of Classical Arabic, glides are deleted unless they occur in a position where they are the only available onset, or in certain positions where they are the final element of a root. Since neither would protect *y* against deletion in a suffix position, this would therefore result in a zero realization of *y* in the cases under discussion in this section.

item can be anchored at FEM, and we thus rewind, and anchor the next step at SG again. The same affix *t* then spells out [FEM[SG]] in the second cycle, i.e., it spells out only the structure that it lexicalizes above the pointer. We will see below that this same fact will also allow us to account for the difference in linear order between the forms *katab-a-t* (3FS.PERF) and *katab-t-a* (2MS.PERF). After that, the spellout of Tense proceeds in the same way as it does in the masculine form, i.e., the presence of FEM causes an additional cycle of spellout, but the following cycle is anchored at the same element as in the masculine, thus deriving the formal containment between the two forms.

(53) *Third Person Singular – Perfect*

a. 3MS: [a SBJV [Ø JUSS [Ø SG [PERF [REF]]]]]

b. 3FS: [a SBJV [Ø JUSS [FEM [SG] [Ø PERF [REF]]]]]]
(FEM [SG] = t)

The facts of the dual containment, i.e., that the dual form contains the plural form in the second person masculine, but the singular form in the third person feminine, receives an explanation as well: Because the cycle that spells out PERF in the second person, anchored at ADD, spells out the span up to (and including) PL, the dual contains that form. We see this in (54), where the spellout proceeds largely parallel to the singular, modulo *aa* spelling out the span [JUSS[DU[PL]]]. Because there is no affix that could be anchored at REF and spell out a structure larger than the span up to SG, *Ø*

(54) *Third Person Dual – Perfect*

a. 3MD: [a SBJV [aa JUSS [Ø DU [PL [SG [PERF [REF]]]]]]]

b. 3FD: [SBJV [JUSS [DU [PL [FEM [SG [PERF [REF]]]]]]]]

(a: SBJV; aa: JUSS DU PL; t: FEM; Ø: SG PERF REF)

The *Rewind* proposal thus accounts for the containment facts that hold for the dual forms, i.e., the pairs *katab-t-um* (2MP.PERF) vs. *katab-t-um-aa* (2MD.PERF) and *katab-a-t* (3FS.PERF) vs. *katab-a-t-aa* (3FD.PERF): In the second person masculine, plural is spelled out together with the cycle that spells out ADD, but the fact that no item can be anchored at DU causes rewinding. In the third person feminine, in contrast, *t* does not spell out PL, but only the structure just below that. Both of these, however, lead to the next cycle targeting [JUSS[DU[PL]]], and therefore lead to the partial syncretism with respect to *aa*. Crucially, the same is not true in the prefixal paradigms, because *um* is not applicable there, since the absence of PERF blocks it from matching the span between SG and ADD, i.e., the system derives the (a)symmetry between second and third person, as well as the fact that it is restricted to the suffixal perfect paradigm simply from the matching algorithm.

This leaves the third person plural to be accounted for. In particular, we see that the suffixes here are identical between the suffixal and the prefixal paradigms, while the contrast involves *y* vs Ø, as shown in Table 3.12.

Table 3.12: Second/Third Person Plural

P/G/N	SUBJUNCTIVE	PERFECT
3MP	y-aktub-uu	katab-uu
2MP	t-aktub-uu	katab-t-um
3FP	y-aktub-na	katab-na
2FP	t-aktub-na	katab-t-um-na

Like the dual asymmetries discussed just above, this follows from the specification of Ø and the fact that the perfect differs from the other paradigms in being marked lower: Because Ø wins

out against *y*, which cannot spell out PERF, but both affixes do not spell out the structure above SG, we derive the partial syncretism as in (55), with the second cycle identical to the one in the prefixal paradigms, while the first cycle varies. Crucially, this differs from the facts about the second person (which, in the prefixal paradigms also exhibits *uu*/*na*) that we saw early, insofar as *um* disrupts the remaining parallelism between the spellout of the prefixal and the suffixal structures by spelling out PL. Thus, in the prefixal paradigms, both the second and the third person plural show *uu* and *na*, but in the second person that syncretism breaks down partly, for reasons we saw above, and that I showed to be linked to the containment facts about the dual as well. In the third person, however, the partial syncretism under discussion remains intact, and this follows immediately from the system:

(55) *Third Person Plural – Perfect*

a. 3MP: $\overbrace{[\ \text{SBJV}\ [\ \text{JUSS}\ [\ \text{PL}}^{\text{uu}}\ \overbrace{[\ \text{SG}\ [\ \text{PERF}\ [\ \text{REF}\ ]]]]]]]}^{\text{Ø}}$

b. 3FP: $\overbrace{[\ \text{SBJV}\ [\ \text{JUSS}\ [\ \text{FEM}\ [\ \text{PL}}^{\text{na}}\ \overbrace{[\ \text{SG}\ [\ \text{PERF}\ [\ \text{REF}\ ]]]]]]]}^{\text{Ø}}$

As I have shown then, the nanosyntactic interpretation of Noyer's (1992) hierarchy advanced here can be extended to the suffixal perfect paradigm of Standard Arabic by assuming that the Perfect is at a lower position in the same hierarchy, namely between Number and Person. In doing so, a variety of facts about the ways these paradigms differ have been shown to be implementable naturally: First, we saw that *t* extends to the first person singular in the perfect, while having an otherwise identical distribution, a fact that followed from assuming that *t* is underlyingly a first person marker. Secondly, we saw that the containment facts about the dual are different between the prefixal and the suffixal paradigms, and that the dual containing the plural form only in the second person masculine of the perfect followed from the presence of a specific suffix

um that was non-applicable in the prefixal paradigms. Thirdly, we saw the parallelism between the second and the third person with respect to the plural marking that the prefixal paradigms exhibit is absent due to the same suffix, while the same suffixes are retained in the third person where said affix does not apply.

In the next section I will lay out how the linear distribution of these affixes follows from the same presence of PERF in a low position, focussing in particular on the fact that certain affixes alternate between prefixes and suffixes, and the fact that certain

3.4.1 Deriving a Suffix

It should be clear that the perspective on morpheme order in current approach is in a sense preliminary, insofar as the the Cinque style Universal 20 / Antisymmetry perspective is a typological one that needs to be tested against a range of languages, insofar as specific ways to implement this make subtly different predictions, and insofar as the predictions for possible postverbal orders are much weaker within a Universal 20 approach than those for preverbal orders. That being said, the current approach does provide a clear intuition for how to characterize the linear distribution in both the prefixal and the suffixal paradigms: REF needs to be local to the verb in both of them. In particular, this intuition links the fact that whenever *t* spells out REF, it is adjacent to the verb, to the Pointer approach developed to account for its paradigmatic distribution: In the repeated Table 3.13, we see that *t* is non-adjacent to the verb in the third person singular feminine of the perfect only, i.e., in the context where I argued that it spells out [FEM[SG]] only, but not the REF-containing part, i.e., the part below the pointer.

Table 3.13: The Loci of *t* (repeated)

P/G/N	Subjunctive	Perfect
2MS	t-aktub-a	katab-t-a
3FS	t-aktub-a	katab-a-t

This is shown in the repeated spellouts of the relevant forms in (56). In all three forms where *t* is adjacent to the verb, it spells out REF, but in (56d) it does not. The same holds for the other alternating affix, in the first person plural forms, *n-aktub-a* and *katab-na-a*, where we find *n(a)* adjacent to the verbal stem, paralleling the behavior of *t* in those cases where it spells out REF.

(56) t *in* 3FS *vs* 2MS

a. 2MS.SBJV: [SBJV [JUSS [SG [ADD [PART [REF]]]]]]
   (*a* Ø Ø t)

b. 3FS.SBJV: [SBJV [JUSS [FEM [SG [REF]]]]
   (*a* Ø *t*)

c. 2MS.PERF: [SBJV [JUSS [SG [PERF [ADD [PART [REF]]]]]
   (a Ø Ø t)

d. 3FS.PERF: [SBJV [JUSS [FEM [SG] [PERF [REF]]]]]
   (a Ø Ø; FEM [SG] = t)

With this in mind, the differences between the prefixal and the suffixal paradigms can be accounted for by postulating two things: Firstly, the difference is characterized purely by PERF attracting the VoiceP.[20] Secondly, REF requires the VoiceP to be local to it on the surface. In the prefixal paradigms, PERF is absent, and the locality of REF and VoiceP is trivially given. In contrast, the suffixal paradigms show the verbal structure (i.e., VoiceP) moving to the edge of the span that spells out REF (in line with the landing sites being relativized to spans, as in the discussion of the derivation of a prefix), and pied-piping it. This type of obligatory movement to an edge followed

[20]Note again that Kastner (2018) argues that the vowels spell out the voice structure in Hebrew, contextually conditioned by Tense. Benmamoun (1999) argues that in Arabic, the vowels occurring in the prefixal paradigms are the default, whereas the ones occurring in the suffixal Perfect paradigm are special. While I do not intend to advance a theory of how Kastner's theory could be implemented under the current view here, I would like to tentatively point out the fact that the fact that Voice receives a special spellout in the presence of PERF, and that PERF attracts the VoiceP to its specifier might provide a hint towards the mechanism of and reason for this relation.

by pied-piping is, of course, attested elsewhere, and discussed as *The Edge Generalization* by Heck (2008, 2009), as well as Cable (2012), for examples such as (57).

(57) a. Bill would never buy [[that big] a car]

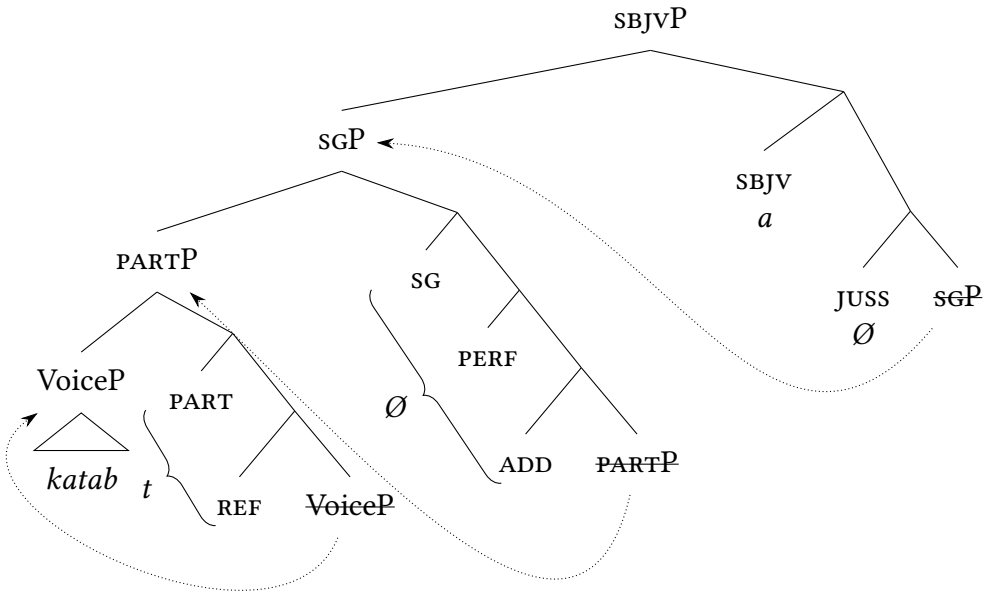
b. Bill would never buy [[a car] that big]

c. [[How big] a car] did Bill buy?

d. *[A car [how big]] did Bill buy? *Source:* Cable (2012)

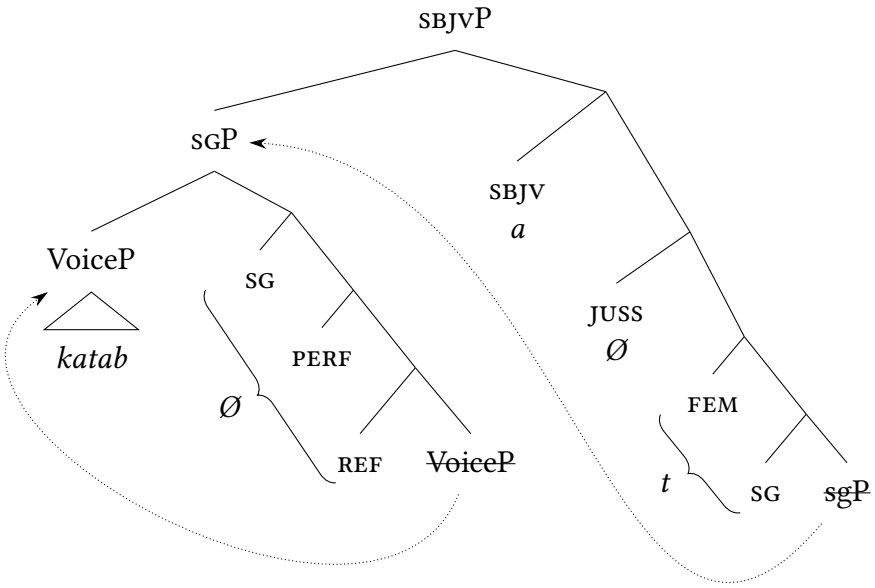
If the locality requirements of REF with respect to Voice blocks movement of the VoiceP to the specifier of PERF in the same manner, with the same repair strategy, namely movement to the edge plus pied-piping, we derive the desired facts: In the perfect, whatever affix spells out REF is right-adjacent to the verbal stem. To see this, consider the two trees for the second person singular masculine and the third person singular feminine in (58) and (59) respectively. In the second person singular masculine, we see that *t* is pied-piped by the verb, since the first movement targets VoiceP, and the VoiceP moves to the edge of the span that spells out REF, i.e. it moves to the specifier of PARTP. The whole phrase headed by PART then moves to the specifier of the span that spells out PERF, i.e., the head that attracted Voice. Since the Tense structure, too attracts REF, as argued in Section 3.3.3, the whole structure moves to the specifier of SBJVP. Note in passing that I left out the movement to the specifier of JUSS; I follow Koopman (2017a), in assuming that the typology of EPP features makes available a type that requires surface true movement (i.e., movement that requires subsequent pied-piping if another EPP feature targets a piece of the structure), as those indicated in the tree, as well as a weaker type that only requires successive cyclic movement. Independent evidence for this is provided by Cinque's (2005) approach where both successive cyclic movement and pied-piping inducing movement are required to derive the various postnominal orders of nominal modifiers.

(58) 2MS.PERF *katab-t-a*



In contrast, the same EPP features will derive a different structure in the third person singular feminine, crucially stranding the structure that *t* spells out, and thus deriving the desired result. Since it is *Ø*, not *t* that spells out REF and PERF in this structure, a single movement operation suffices to fulfill both head's locality requirements. As in the previous derivation, the whole structure then continues to move to the Tense region; the crucial difference lying in the fact that [FEM[SG]] is stranded. Note that the Rewind approach suggests that SG is doubled in the structure, presumably due to the fact that the whole structure is copied, but SG cannot be deleted, due to its presence in the span spelled out by *t*.

(59) 3FS.PERF *katab-a-t*



The same mechanism accounts trivially for all forms that have two overt suffixes: It is always the one expressing REF that is adjacent to the verb. It also accounts for the trimorphemic 2FP *katab-t-um-na*: Since *t* spells out REF, it is linearly adjacent to the verb; and since *um* spells out PERF, it moves PARTP into its specifier, followed by *na* moving the whole structure to the specifier of SBJV, paralleling the derivation of 2SM *katab-t-Ø-a*, modulo *um* being in the position of the zero element, as expected, since *um* and *Ø* spell out PERF. The dual forms, however, need an additional comment: While the second person dual *katab-t-um-aa* could be derived by merely stranding *aa* (which would cause its adjacency and thus phonological fusion with *a*), the third person feminine dual *katab-a-t-aa* posses a problem: If *t* is stranded, we would expect *aa* to precede it. There are various possible ways of addressing this: We might assume that the re-ordering of *t* and *aa* is really a morpho-*phonological* phenomenon, triggered in order to avoid the fusion of *a* and *aa*,

along the lines of a RealizeMorpheme constraint. Alternatively, DU might have an EPP feature for SG that is not demanding a surface specifier; under this approach, *t* would be smuggled above *aa* in the sense of Collins (2005): The whole FEMP moves to the specifier of DUP, but only the SGP moves out after that. This approach would not interfere with the explanation for the second person's form *katab-t-um-aa*, as *um* independently needs to be pied-piped. It would hold for the prefixal paradigms as well, but since all smuggled material is phonologically zero, there is currently no way for me to test this analysis; it does, however, weaken the predictions of the Universal 20 approach, with respect to the possible set of orders — a proposal that has independently been advanced (see Abels 2011, but also Cinque in p.c. to Koopman 2017a). As of right now, however, I do not see a possibility to test these hypotheses within MSA, and thus future work on a larger set of languages that fall under Noyer's (1992) hierarchy, and their postverbal morpheme orders will have to decide between these types of analyses, as both are compatible with the larger argument in this paper.

Crucially, what I have shown in this section is that the the perfect can be properly integrated into the f-seq interpretation of Noyer's (1992) hierarchy. Doing so allows for a natural integration, both of the distribution of *t* in paradigmatic terms, with its occurrence in the first person singular of the perfect, but not the prefixal paradigms, and its linear distribution: The pointer approach that was developed above to account for the fact that *t* occurs in the third person feminine non-plural as well as the second person, but that it cannot spell out both second person and feminine at the same time, was shown to lend itself to a natural characterization of the linear distribution of *t*: The suffixal nature of the perfect in general is captured by a single EPP feature on the single head that distinguishes the two paradigms, while the fact that *t* occupies varying positions follows from the perfect head as well: When PERF blocks the contiguity between REF and [FEM[SG]], *t* must be anchored at SG instead of REF, which leads to its being stranded, rather than pied-piped, and thus not be adjacent to the root. The pointer approach does thus not only account for the paradigmatic distribution of *t*, but also for its linear distribution, unifying them under a single

lexical specification.

3.5 Discussion

In this paper I have argued for a structural interpretation of Noyer's (1992) hierarchy for Modern Standard Arabic, and extended it to the Perfect. Under the current theoretical perspective, a variety of phenomena reduce to the notion that the vocabulary drives the morpho-syntactic *Fusion* (in DM parlance) of maximally simple heads, and independently motivated syntactic operations. It is worth briefly comparing this system to the approaches of both Noyer (1992) and Halle (2000), who employ a large set of independent mechanisms to derive these facts (alongside many similar phenomena in other languages that the current theory still needs to be tested against), among which are Fission, Impoverishment, an independently existing hierarchy, Fusion, a subset principle, as well as a lexical specification of affixes as prefixes or suffixes.

On the theoretical side, it is surely worth noting that a smaller set of mechanisms ought to be preferable, and if we can get by with only a Fusion-like mechanism, such as spanning, that is itself an interesting result, offering not only fewer mechanisms, but also a perspective on the theoretical status on Noyer's hierarchy: It is simply a part of the functional sequence. It is also worth noting that in the current perspective, the hierarchy extends to governing Fusion effects as well, i.e., since an exponed span is a contiguous part of the hierarchy, no vocabulary item can expone non-contiguous parts of the hierarchy, thus making the theory more restrictive — while the paradigmatic distribution of affixes can be understood purely in terms of a linear system, the f-seq, reducing a five dimensional system (Tense, Aspect, Person, Number, Gender) to a more restrictive one-dimensional one (modulo the effects of pointers), their linear distribution is subject to the usual syntactic mechanisms of feature driven displacement. The system proposed here offers an account in which the effects that Noyer attributed to Impoverishment are reduced to

vocabulary insertion,[21] the notion of discontinuous bleeding (the fact that certain prefixes block certain suffixes) has been accounted for without a template, and linked to the "continuous bleeding" counterparts in the perfect paradigm, and the order of the affixes has been accounted for in the same terms, by suggesting that affixes that spell out REF must retain the local relation to the verbal stem.

More crucial, however, are the empirical facts: Firstly, both Noyer and Halle rely on the notion that it is the vocabulary items themselves that are prefixes or suffixes, a notion that runs into obvious trouble given the prefix/suffix alternations of *t* and *n(a)* that we have seen.[22] Insofar as Noyer does not deal with the suffixal paradigms, it is not obvious whether his system could be amended to deal with these paradigms, but it has to retreat to treating the distribution of *t* as an accidental homophony of two different affixes, an approach that Halle takes issue with. In addition to a prefixal Arabic paradigm, Halle's approach does treat both the prefixal and suffixal paradigms of Biblical Hebrew, and the issues that would arise in extending it to his analysis of Classical Arabic are obvious: Halle treats *t* as an elsewhere prefix that encodes nothing, and argues that all prefixal paradigms have a requirement to exhibit at least one prefix. Since no other prefix encodes a subset of the second person specification [+PART, -AUTHOR], *t* is inserted in these contexts. As for its occurrence in the third person feminine singular/dual, he posits an impoverishment rule for these contexts that deletes the feature [-PART] that *y* spells out, and they are thus subject to the insertion of elsewhere *t* as well.[23] There are two major issues with this proposal: Firstly, it is clearly a violation of exactly the hierarchy-based restriction on Impoverishment that Noyer discovered: Given his hierarchy of *Person* > *Gender*, it should be gender that is impoverished here, not person, i.e., Halle's approach is incompatible with Noyer's insights. Secondly,

[21] See Trommer (1999) for a similar notion, where Impoverishment is reduced to the insertion of zero affixes.

[22] Though Noyer somewhat contradicts himself: On page 92, he argues that affixes bear a polarity property [±prefix] that determines which templatic slot an affix will be competing for, but on page 98 he notes that *n* appears as a suffix.

[23] Halle (2000: 140) actually posits an Impoverishment rule that deletes [-Pse] (his name for participant) in the context of [+fem,-sg], but insofar as that would derive the wrong distribution of *t*, I assume this is an error, and the rule is meant to apply in the context [+fem, -pl], i.e., the singular and the dual.

this mechanism of limiting competition to a set of prefixes, does clearly not extend to the perfect paradigm — *t* does occur as a suffix as well, and thus it cannot be a lexical property of the affix to be a prefix. Even if *t* were underspecified with respect to [±prefix], and thus able to compete in both paradigms, we run into trouble: There are no obvious "positions" (as in a fissioned terminal) where *t* could apply, and nothing that could distinguish 2SM *katab-t-a* from 3FS *katab-a-t* with respect to the linear order of these two affixes. His analysis of Biblical Hebrew posits no Fission in the perfect, and five monomorphemic suffixes that all have *t* as their first segment; clearly an analysis that does not capture the Arabic paradigms discussed here, where every cell that has *t* as a prefix has *t* as a suffix in its perfect counterpart. In contrast, the current analysis gives an account of the paradigmatic distribution of *t* that accounts for the prefixal as well as the suffixal paradigms, and provides an account of its linear distribution that derives from the same specification: We can capture the variation in linear distribution by a single descriptive statement, namely that REF needs to be adjacent to the verb; the prefix/suffix distinction is a derived generalization over different spellouts, rather than a list of individual stipulations that combines with an autonomous morphological template. The radically atemplatic approach advanced here accounts for the variation in linear distribution by linking it to the paradigmatic distribution purely in terms of matching, a possibility that is excluded in the other approaches.

That is not to say that the current approach is without issues: Unlike Noyer's and Halle's approach, it has not yet been tested against a wide range of languages. The movement approach to affix order within a broader Antisymmetry approach necessitates syntactic movement of arguments to evacuate the Tense structure, if it is to fit into the wider analysis of the language. More crucial is the fact that I have tacitly assumed that there is no first person dual, while simultaneously assuming that the first person does bear gender that is always neutralized (in order to allow for the specification of *t* as a first person singular feminine affix); a treatment of two identical surface properties in entirely different ways. Ritter's (1997) approach to number in the first person as part of the person domain (through a MULTISPEAKER head above the PART head)

that I mentioned in footnote 18 might provide a perspective on this that solves that problem, but for now it remains an open question. Lastly, and possibly most importantly, the proposal introduced here basically suggests that the peculiar distribution of *t* is part of an *ABA violation in the prefixal paradigms, where the largest and the smallest structures, third and second person, are spelled out by *t*, while the medium sized first person is instead spelled out by ʔ, which capable of overwriting *t* since it lexicalizes additional higher material. This is a fairly unique property of the pointer approach, and the reason Vanden Wyngaerd (2018) rejects it. To my knowledge, no parallel effects are known for cases of root suppletion, and if this is indeed a difference that characterizes morphology at least in cases like the one discussed here, it begs the question of its boundaries and the reasons these differences would arise for.

Finally, let me offer a brief note on predictions: The approach proposed here captures the Impoverishment effects Noyer described for MSA (modulo the absence of the first person Dual, to which Ritter's proposal might provide a solution, as just discussed), but it does make subtly different predictions with respect to the possible effects we might find elsewhere. While Noyer's proposal allows any type of filter to impoverish a lower ranked feature, a strikter locality is required under the current approach, such that person should not be able to create a syncretism with respect to gender, unless it is also expressing part of the number structure, person should not be able to create a syncretism with respect to Tense, unless it also expresses Number and Gender, etc. As of now, I have to leave it to future research to test these predictions, but it provides clear and testable predictions, as any theoretical stance should.

3.6 Conclusion

In this paper I applied the Nanosyntax system for the spellout of agreement that was developed in Blix (2021b) to a completely different dataset, arguing that certain peculiar predictions of the pointer approach account for effects that cannot be captured properly in a subset-based ap-

proach to matching. In arguing that Noyer's hierarchy is a description of the functional sequence, and that the vocabulary items spell out differently sized "chunks" of this functional sequence, I have shown that an account of the distribution of *t* can be given that does provide an explanation for an apparent disjunction: The affix in question spells out the feminine, or the second person, but not both. I have shown that the pointer account, by which *t* lexicalizes two contiguous spans, derives this effect, and I have shown that it makes available an implementation of of affix order that relies on a coherent set of requirements, in particular linking it to the fact that *t* can, but does not have to spell out the person structure REF in order to spell out FEM, and that it is non-adjacent to the root only in case it is not part of this pied-piping inducing head. I have shown that the perfect can be understood within this implementation of Noyer's hierarchy by arguing that it is a single head within the hierarchy that derives both the distribution of affixes, and the fact that these paradigms are suffixal. In doing so, I have shown that Impoverishment can be unified with the syncretisms introduced by the fact that vocabulary items match a set of structures determined by the matching algorithm, I have derived discontiguous bleeding in Arabic from a purely local relation of (spans of) heads, and provided a hypothesis about the nature of Noyer's hierarchy that implements it in the narrow syntax. Finally, I believe that I have shown that the atemplatic approach that derives all affix order from syntactic movement is empirically superior to approaches that restrict competition semi-templatically by lexically specifying affixes as prefixes and suffixes; an approach that the data itself shows to be untenable, but that also misses interesting generalizations about *why* a group of affixes can be prefixal, and why an affix may turn up in a variety of positions.

While I think these are very promising results, it must be said that the current approach has not yet been shown to have the same empirical scope as the many DM approaches to agreement have. It is my hope, however, that the results shown here can inspire future research into and refinements of the span-based perspective on the syntax-morphology interface.

4 | Movement-Containing Trees

4.1 Introduction

In Nanosyntax (Starke 2018, Caha 2019), vocabulary items (VIs) are hypothesized to interpret phrasal nodes only. Mediated by a superset-based matching algorithm (a vocabulary item matches all trees that the tree it lexicalizes contains), this leads to the vocabulary driving (parts of) the syntactic derivation: In order for a derivation to converge, it must consist solely of trees with matching vocabulary items. To ensure such convergence upon merging a feature F, the derivational algorithm 'attempts' a variety of operations in an ordered fashion until it finds a candidate that can successfully spell out, as in (1):

(1) Merge(F, XP), then:

a. Try: Spell out FP

b. If it fails, try: Move(Spec of XP) to Spec FP, Spell out FP

c. If it fails, try: Move(XP) to Spec FP, Spell out FP

d. If it fails: Go to the previous cycle that merged X, and try the next option for that cycle

Lightly adapted from Starke (2018), Caha (2019)

Upon merging a feature F, the spellout algorithm will first (1a) attempt to spell out the resulting

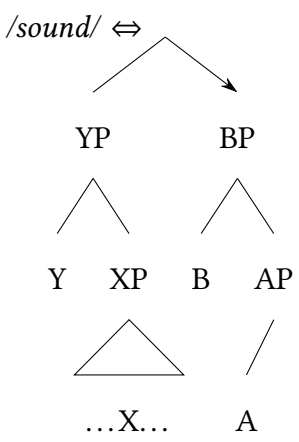
FP (resulting in F being spelled out with the current 'root' structure). Should that fail (1b), it will attempt spec-to-spec movement (generally resulting in F being spelled out with the current suffix structure). If that, too, fails (1c), F moves its complement to its specifier – roughly corresponding to the "construction" of a new suffixal position. These three steps create a preference for spellout targets that are as large as possible. Should they fail, however, option (1d), the so-called *backtracking* option, returns to the previous cycle, changing the structure of XP itself.

Note however that *backtracking* differs categorically from the other three steps: Step (1d) calls the spellout algorithm itself, i.e., its addition turns spellout into a recursively defined algorithm. The corollary, in terms of computational cost, is that it turns the linear algorithm that is defined by (1a-c), into an exponential algorithm. That is to say, for a derivation of length n (or $n + 2$, if we consider first merge), an algorithm without backtracking would consider no more than on the order of $3n$ derivational stages (in the sense of Collins and Stabler 2016). Since backtracking makes the function recursive, however, every feature triples the set of possible derivational stages that may need to be considered in the full algorithm, i.e., the worst-case scenario is a comparatively costly 3^n.[1]

Given this computational cost, it is worth asking if backtracking is a necessary part of the spellout algorithm, or if we can do without a recursive extension of the theory. In this paper, I explore the possibility that the desired effects of backtracking can in fact be implemented in a linear spellout algorithm, if we employ branching vocabulary items and pointers, as in (2) – with A, B and Y part of the extended projection of X.

[1]Technically, the algorithm is a depth-first search over a decision tree in which any node has three daughters, representing the three ways to introduce and spell out a feature. However, since any node with three daughters that are all impossible to spell out will not have its daughters' daughters explored, the worst-case scenario is, of course, an unlikely edge case. This does not change the fact that the algorithm is exponential with backtracking and linear without backtracking. Section 4.2 will provide examples and a comparison of decision trees for particular derivations.

(2)



A pointer (indicated by the arrow on the right branch) is a common nanosyntactic device of conjunction: In this case, it conjoins YP with any subtree of BP (here: BP or AP). Hence, the vocabulary item in (2) could be described as follows: It matches any subtree of the set of trees that is formed by conjoining YP with a subtree of BP.

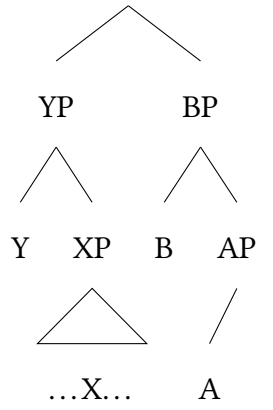
Such branching vocabulary items with a pointer have two properties that will turn out to be crucial for our purposes: First, they are able to drive movement even when the root node is the target of spell out. That is to say, when A is merged with YP, the result of comp-to-spec movement of YP can be matched by the vocabulary item at the root, as can the result of subsequent spec-to-spec movement upon merging B. Second, such vocabulary items can be subject to *partial overwrite*, an effect that turns out to look highly similar to aspects of backtracking.

Both properties are illustrated in (3): After merging A, simple spellout fails. There is no relevant specifier that spec-to-spec movement could target, and thus comp-to-spec is attempted. The resulting AP is matched by */sound/* at the root. Merging of B triggers spec-to-spec movement of YP to Spec BP, and */sound/* again matches the result at the root, and thus spells out the whole BP.[2] Upon merging C, simple spellout fails, and thus spec-to-spec movement is attempted. The

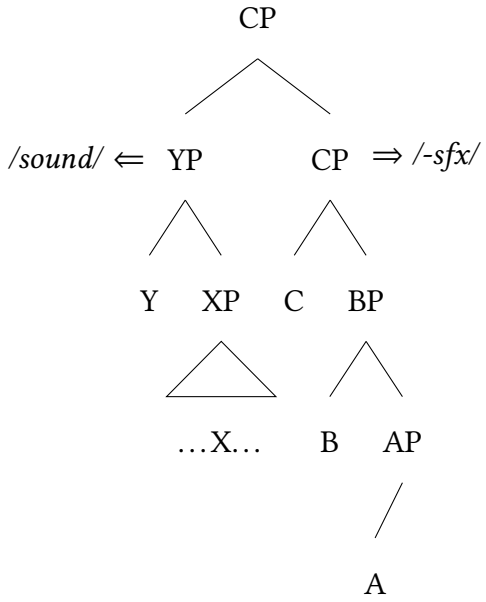
[2]The suffix */-sfx/* in (3b) that spells out [C[B[A]]], matches [B[A]] at this stage as well, yet [B[A]] is spelled out by */sound/*. I assume that the preference for */sound/* follows from general Nanosyntactic principles such as the preference for large targets over small ones. Concretely, this could be accomplished simply by having the spellout algorithm check whether the root can be matched before checking whether the right branch can be matched. (A left branch must have spelled out successfully prior to being merged/moved to the specifier position, hence checking whether the right branch can be matched is sufficient.)

structure spells out successfully. However, instead of spelling out the whole BP, */sound/* now targets only the left branch YP – it has been partially overwritten.[3]

(3) a. */sound/* ⇐ BP



b.



If we consider the effects of partial overwrite in terms of a lexicalization table, we see that our branching vocabulary items are able to create an effect that looks similar to those created by backtracking; they do, however, create this effect without necessitating any undoing of the derivation, or a recursive spellout algorithm.

Table 4.1: Lexicalization table for partial overwrite

	XP	Y	A	B	C
AP	*sound*				
BP		*sound*			
CP	*sound*			*-sfx*	

Both branching vocabulary items and pointers are already part of standard Nanosyntax, and neither generates the computational complexity involved with backtracking.[4] In the interest of

[3]This idea thus exploits the fact that Nanosyntax provides two ways to successfully spell out a phrasal node X: Either spell out X directly, or spell out both daughters of X.

[4]The lexicon contains only well-formed syntactic expressions, and branching vocabulary items correspond to

a minimalist theory of spellout, both in the sense of having few components, and in the sense of computational complexity, it is thus worth exploring whether the theory can get by without the fourth clause of the spellout algorithm. In what follows, I explore a few cases from Caha's (2019) recent analysis of case competition, as well as a relevant example of Pseudo-ABA patterns. I chose this work because it represents a recent state-of-the-art view of Nanosyntax, and because it contains an explicit argument in favor of backtracking as well as a variety of analyses that make crucial use of this technology.

Section 4.2 lays out an analysis of the Iron Ossetic pronoun data that Caha uses to motivate backtracking, and shows that a partial overwrite analysis can capture the facts. Section 4.3 discusses the case of the Digor Ossetic augment, arguing that the backtracking analysis does not, in fact, capture the data perfectly, but needs additional tools. Once such tools are in place, however, an analysis without backtracking is possible. Section 4.4 takes a look at Caha's (2019, 2020) size-based theory of declension class, and sketches a modified configurational theory of declension class. I suggest that, in addition to capturing the relevant data, branching vocabulary items offer an empirical advantage: Since they 'split' the functional sequence (f-seq) into a lower left and a higher right branch, a configurational theory may model different classes by locating their f-seq split in different positions. Section 4.5 takes a look at Pseudo-ABA patterns (H. J. Middleton, 2020), and shows that we can similarly model these as subextraction without backtracking. Section 4.6 concludes.

4.2 Iron Ossetic pronouns

Caha motivates the backtracking component on the basis of a comparison between the Iron Ossetic first person plural pronoun *max* and the nominal paradigm of *fyd* 'father' in Table 4.2.

expressions that can in principle be derived by the spellout algorithm. Pointers are needed independently for idioms, as well as certain types of cross-categorial syncretisms where a vocabulary item applies to the conjunction of sets (Caha and Pantcheva 2012, Blix 2021b).

While the latter is affixed with *-y* in the accusative and genitive, the pronoun shows a syncretism between nominative, accusative, and genitive. That is to say, under standard nanosyntactic assumptions, *max* lexicalizes a larger structure than *fyd* does – one that is at least as big as the genitive (or the phrase headed by the feature which builds the genitive from the accusative, F3 in Caha's terminology).

Table 4.2: Iron Ossetic pronoun vs noun (Caha 2019: 74f, 119)

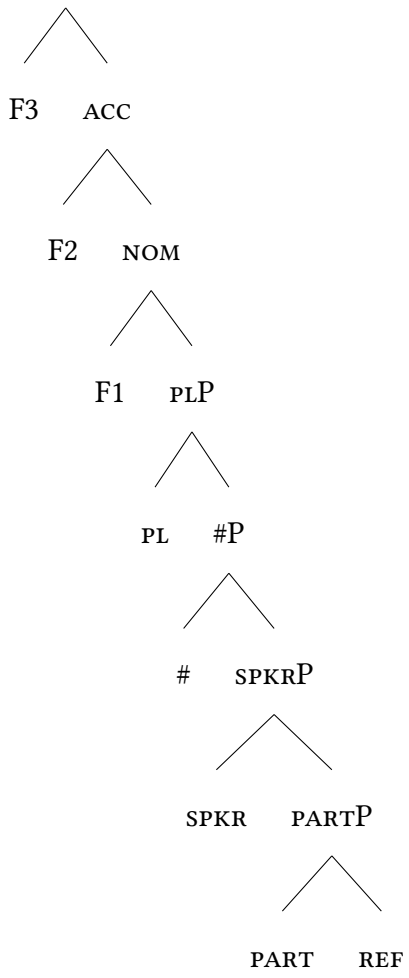
	1PL	father, SG
NOM	max	fyd-∅
ACC	max	fyd-y
GEN	max	fyd-y
INS (ABL)	max-æj	fyd-æj
DAT	max-æn	fyd-æn

Note that both share the same case suffixes in the larger cases, such as the instrumental/ablative, or the dative. Since the instrumental *-æj* does not appear to co-occur with *-y*, we assume that they compete in some sense, i.e., under the standard assumptions, *-æj* is able to *overwrite -y* in the instrumental. For overwriting to take place, *-æj* must therefore lexicalize a superset of the features that *-y* lexicalizes, and *lexicalize the same bottommost feature* (or foot).

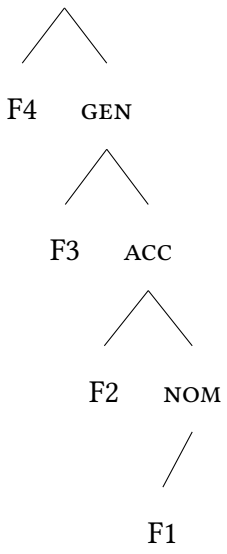
Under Caha's proposal, the conjunction of these facts leads to the necessity of backtracking: i) *max* lexicalizes a full GENP, as in (4a). ii) The suffix *-y* that we observe with *fyd* 'father' lexicalizes, minimally, the features F2, F3 that build the genitive from the nominative structure, and iii) *-æj* lexicalizes a superset of the features lexicalized by *-y*, up to the instrumental, as in (4b). Therefore, the foot of *-æj* is lower than the largest structure *max* can spell out – once we reach F4 on top of *max*, we must backtrack in order to anchor *-æj* (say, at F1).

(4) The vocabulary (backtracking hypothesis)

a. *max* ⇔ GEN

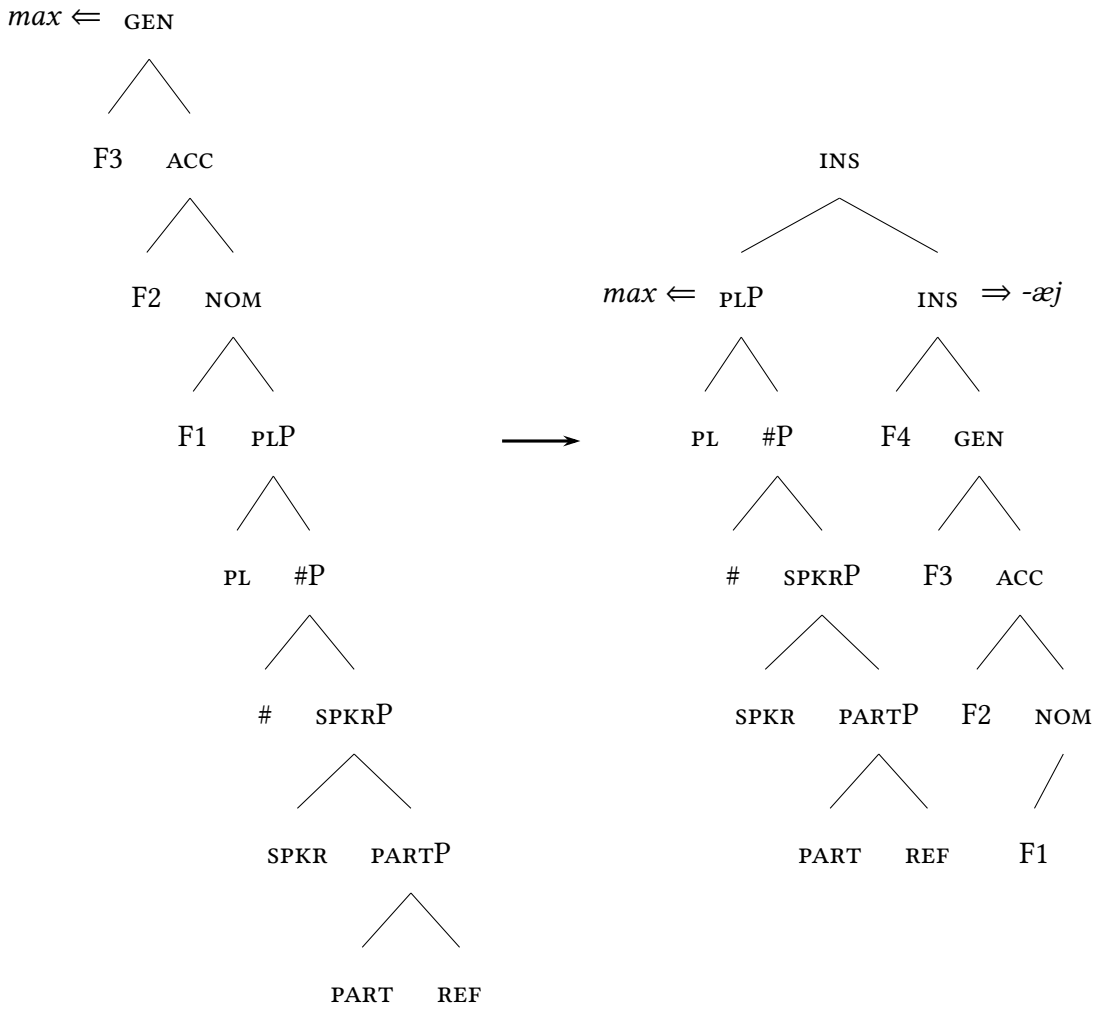


b. INST ⇔ *-æj*



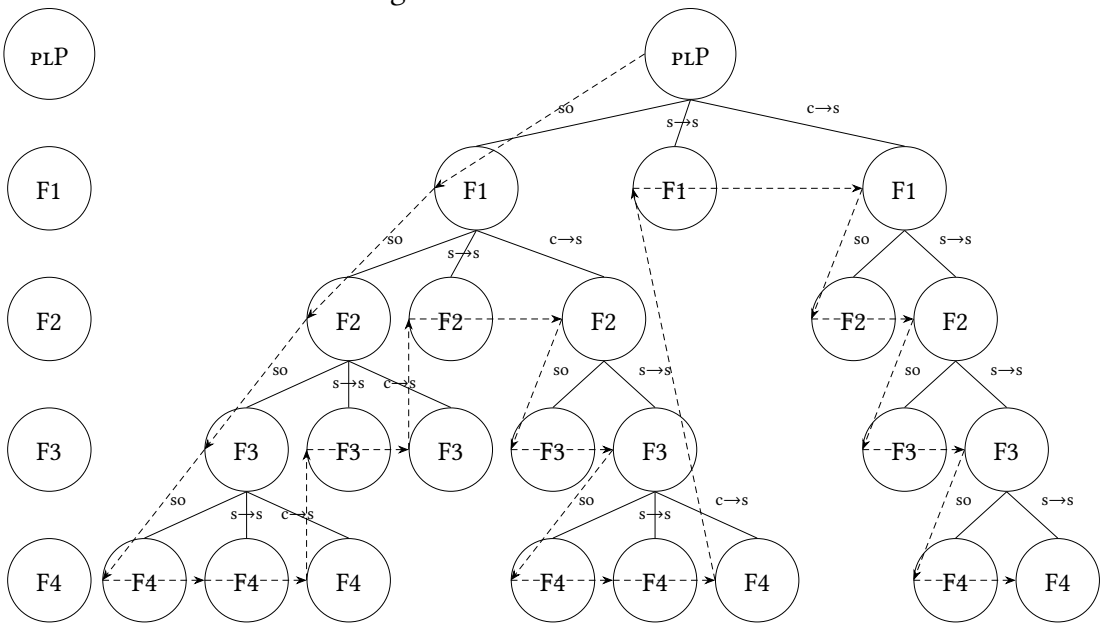
That is to say, backtracking is unnecessary until we reach F4 – at this point, the backtracking analysis suggests that the derivational path taken is incompatible with the spellout of F4, and it walks back step by step, attempting different derivations and checking whether they are compatible with the spellout of F4. The result of backtracking is indicated in (5) – though backtracking itself is not an operation on the tree, its results look just like standard phrasal movement, because it effects a series of standard movement operations.

(5) Backtracking at F4



The series of these operations can be visualized in the form of the decision tree in (6). Every node of the tree describes a derivational state, with the feature in the label being the most recently introduced one. The label of the edge connecting it to the mother node describes which of the three basic steps led to the derivational state, i.e., spellout (SO), spec-to-spec movement (s→s), or comp-to-spec movement (c→s). The dashed arrow describes the sequence in which these derivational states are built/explored.

(6) Decision tree with backtracking



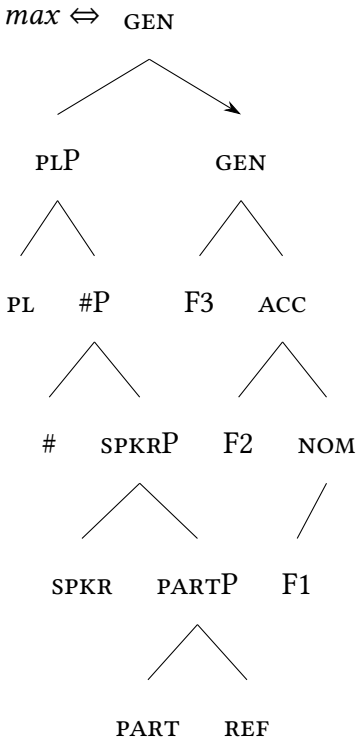
Essentially, the algorithm first merges F1–F3 on top of the PLP without any movement (SO). However, since this does not lead to a derivation that can successfully spell out F4 (i.e., all three options fail), it traces back its steps, first attempting to change the configuration of F3 where neither alternative to the SO route can be matched. It then reconsiders the configuration of F2. Here, the algorithm is forced to check whether a derivation with *-y* (which lexicalizes [F3[F2]]) may succeed, since F2 with F1P in its specifier can be spelled out by the hypothetical *max-y*. However, this derivation, too, fails to spell out F4, so a different configuration for F1 is explored next. Here, iteration of spec-to-spec movement finally succeeds and successfully derives *max-æj*.

The core idea of the alternative I would like to advance is that *max* is a branching lexical item, as in (7). The left branch corresponds to the phi-structure, and the right branch to the case structure. Crucially, the latter is embedded under a pointer, i.e., the vocabulary item can spell out any subtree of the tree that is formed by conjunction of the PLP with a subtree of the case structure.[5] Let us consider the stage at which we have built the PLP, and not yet merged any case

[5]See Caha 2019: Chapter 4.6 for a relevant discussion of mixing grammatical features with pointers to lexical entries. Note that pointed to material is sometimes thought to be obligatorily present in the syntactic structure to be matched, whereas I take it to be optional here.

features. Since a vocabulary item matches all subtrees of the tree it lexicalizes, and the PLP is a subtree of the tree in (7), *max* matches PLP, and thus simple SO is the path for the PL feature. After merging F1, spec-to-spec movement is unavailable (since PL does not have a specifier), but comp-to-spec movement is successful and *max* self-overwrites. F1 now has a specifier, PLP, and subsequently merging F2 triggers spec-to-spec movement, as does merging F3. This is due to the fact that *max* can match the respective resulting trees. That is to say, we retain the ability of *max* to spell out F1, F2, F3 that the backtracking approach offered, but we derive a configuration where, for spellout to be successful, the PLP must be in the specifier of F1/F2/F3 respectively.

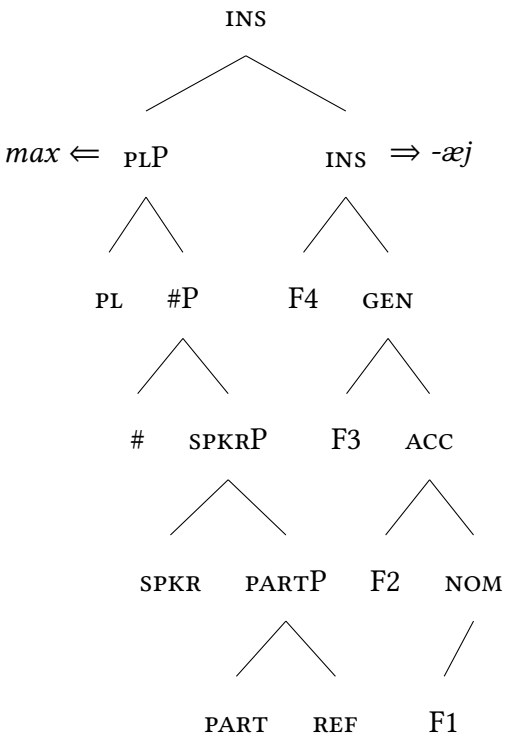
(7)



The crucial difference between this analysis and the backtracking one comes about upon merging F4, as in (8). In both cases, simple spellout after merging F4 fails, and spec-to-spec movement is attempted. In the backtracking analysis, this fails, and previous structure building operations are now undone. However, under the current hypothesis, the derivation does spell out successfully after spec-to-spec movement. PLP – now in the specifier of F4 – can be spelled out by *max*, because

it is a subtree of the one *max* lexicalizes. The right branch is spelled out successfully by a suffix *-æj* that is identical to the one from the backtracking analysis. This *partial overwrite* analysis takes advantage of a theoretical ambiguity: A node is considered interpreted successfully if a) it was spelled out, or b) both its daughters were spelled out. That is to say, the structure [F4[F3[F2[F1]]]] can now be spelled out by an affix anchored at F1, with no backtracking being necessary: *max* is effectively overwritten with respect to spelling out F1, F2, and F3, but it remains the winning candidate for spelling out the PLP. Note in passing that the latter part is an accident of the lexicon (there is no smaller competitor for the spellout of PLP). I return to cases where partial overwrite leads to the emergence of a smaller competitor for the spellout of the left branch in section 4.5, which discusses Pseudo-ABA patterns of this type.

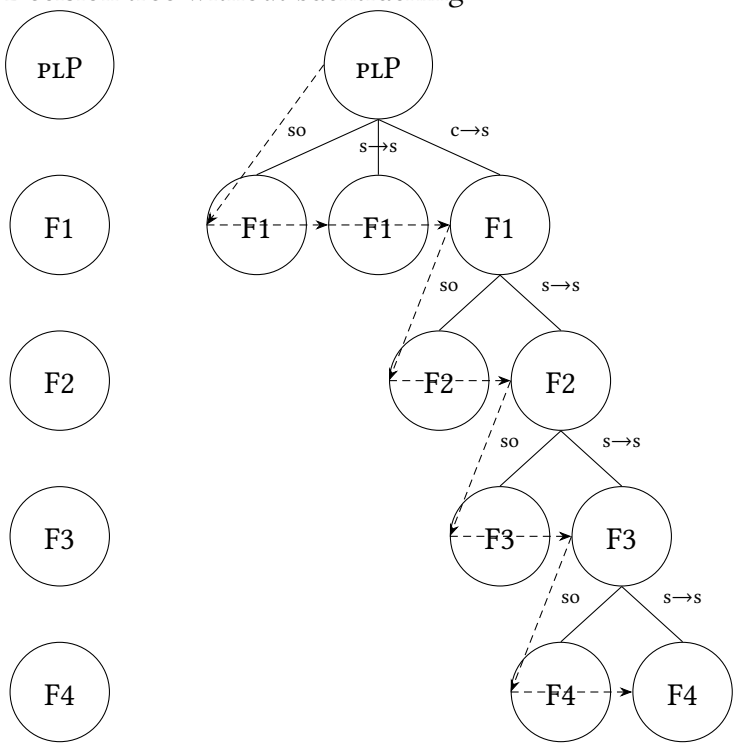
(8)



Both analyses derive the data successfully, and we can thus compare them in terms the number of derivational stages that had to be considered. The operation of the spellout algorithm in the analysis without backtracking can be visualized as in (9). Crucially, no backtracking means that

there is no walking back up the tree and opening a new branch: At every level, there is a maximum of three nodes to explore, and consequently the derivational stages form a proper subset of those that the backtracking algorithm needs to explore. The maximum number of derivational stages that need to be explored simply grows linearly with the number of features in our tree.

(9) Decision tree without backtracking



Though the algorithms differ considerably in the way they achieve their aim, a look at a lexicalization table that describes which features are interpreted by which vocabulary item, as in Table 4.3, is a useful comparison to see similarities: The table describes *both* hypotheses equally well. In the backtracking analysis, the spellout of F1, F2, F3 by *max* was undone by returning to previous stages of the derivation and attempting alternative operations. However, the same result was achieved by the backtracking-less analysis by means of *partial overwrite* – just with a less costly algorithm.

Table 4.3: The lexicalization table for *max*

	PLP	X1	F1	F2	F3	F4
NOM	max					
ACC		max				
GEN			max			
INS	max			-æj		

4.3 The Digor Ossetic augment

Digor Ossetic numerals show an augment in the oblique cases, but not the structural cases. As the comparison in table 4.4 shows, however, the numerals take the same oblique case suffixes as a regular noun.

Table 4.4: Digor Ossetic augment (partial paradigm)

	two	horse
NOM	duuæ-∅	bæx-∅
ACC	duuæ-∅	bæx-i
INS (ABL)	duu-em-æj	bæx-æj

The core of Caha's analysis is represented in the Table 4.5. The augment *-em* lexicalizes a low head [X], but a zero affix that lexicalizes [F2[F1[X]]] overwrites it in the structural cases. Any affix that can spell out F3 and subsequent case features, however, must be anchored at F1 – i.e., backtracking occurs. In this case, backtracking has the further effect of uncovering *-em*: The cycles of spellout that overwrote *-em* are undone by backtracking, and thus the augment (re-)surfaces in the larger cases.

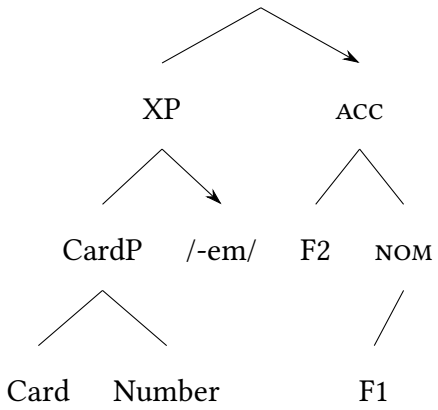
Table 4.5: The backtracking analysis

	CardP	X	F1	F2	F3–F6
NOM	duu(æ)	-Ø			
ACC	duu(æ)		-Ø		
INS	duu(æ)	-em		-æj	

This analysis elegantly captures the distribution of the augment, and I cannot see a way of capturing the same data solely with a backtracking-less algorithm for spellout-driven movement. That is to say, an algorithm without backtracking may need to stipulate feature-driven movement to derive such lexicalization tables. Using such an auxiliary tool/assumption, we could once again postulate a branching vocabulary item with a pointer, as in (10):

(10) The vocabulary (to be revised)

a. */duu(æ)/* ⇔ ACC



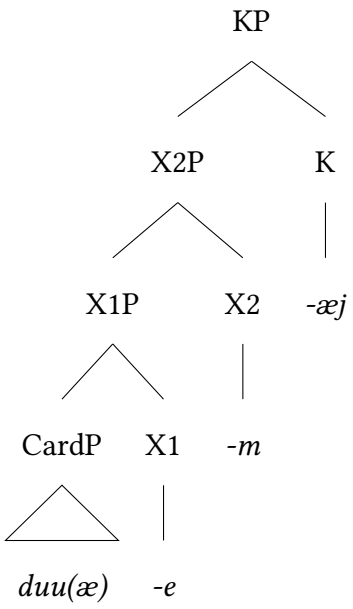
b. */-em/* ⇔ XP

The derivation continues in the usual way, and upon merging F2, we derive the full tree that *duu(æ)* lexicalizes. The core idea is the following: If, by stipulation, F3 has an EPP feature for CardP, then such a feature may lead to stranding of the structure above CardP, namely X. Once CardP has stranded X, the interpretation by *-em* that was hitherto overwritten by the idiom, resurfaces. As in the cases above, however, *duu(æ)* remains the spellout for the moved CardP, and thus we derive the augment as a resurfacing effect, and capture the data in a way reminiscent of Caha's proposal, but without backtracking.

There is reason to believe that an operation not driven by spellout is needed to derive the Ossetic augment on independent grounds, even for the backtracking analysis, and I will lay out in more detail both why and how to derive the relevant facts with vocabulary entries such as (10) in the remainder of this section.

Let me begin by elaborating on an issue with Caha's analysis that cannot be captured with backtracking alone: Nanosyntax generally takes zero-affixes to be reason for suspicion (see in particular the arguments about zero-distribution in Caha et al. 2019), but they are certainly not reason enough to reject an analysis. There is, however, a deeper problem that is laid bare by Caha's subsequent comparison with the augment in the pronominal and demonstrative system. In particular, Caha (2019: 142ff) argues highly convincingly that the augment must in fact be analyzed as a bimorphemic structure *-e-m*. He revises the structure as represented in the simplified representation in (11) (his 24, p. 144).

(11)



A bimorphemic augment however, cannot be the result *solely* of backtracking undoing the work of a zero affix overwriting the augment. To see why, consider the original analysis in Table 4.5, in which a zero affix overwrites the augment in the nominative and accusative cases, but backtracking undoes this in the oblique cases. The augment (re-)surfaces in the larger cases, since the spellout cycles that overwrote it were undone. Table 4.6 shows why this solution is impossible for a bi-morphemic augment: The zero affix would continue to be a better candidate for the spellout of X1, X2, regardless of backtracking to F1, and no augment would re-surface. This limitation

follows from core principles of Nanosyntax (the preference for few large morphemes over many small ones, as encoded in the preference for spec-to-spec movement, or the biggest wins theorem): Backtracking that results in the re-emergence of a smaller vocabulary item in place of a bigger one is *necessarily* limited to the re-emergence of a single affix.

Table 4.6: The problem of the bimorphemic augment

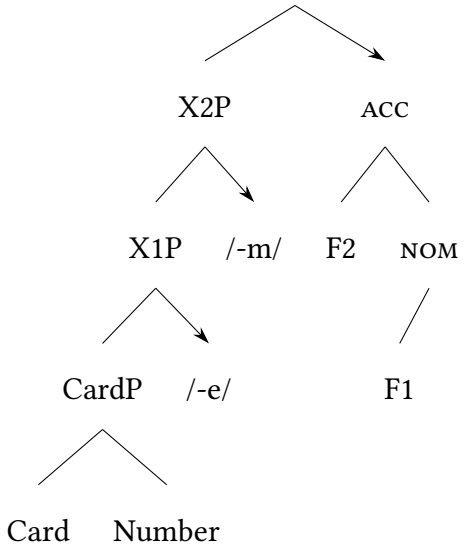
		CardP	X1	X2	F1	F2	F3–F6
	NOM	duu(æ)		-Ø			
	ACC	duu(æ)			-Ø		
*	INS	duu(æ)	-e	-m		-æj	
☞	INS	duu(æ)		-Ø		-æj	

The backtracking analysis thus needs some tool(s) beyond backtracking itself to derive the data. One possible analysis employing standard syntactic tools would be to augment the backtracking analysis with feature driven movement, as above. If F3 triggers the displacement of X1P, it would block the ability of *-Ø* to overwrite *-e* and *-m*, because they are no longer in a phrasal configuration that could be matched by *-Ø*. Once such feature driven movement is brought into the account, however, we can make due without backtracking, and the analysis sketched above can easily be modified to deal with a bi-morphemic augment.[6]

The non-backtracking analysis I would like to propose builds on Caha's analysis in (11), but includes the fact that there is no overt case marking in the structural cases. Both the augments and the case structure can be analyzed as being spelled out by right branches of the vocabulary item i.e., we simply iterate the approach from before, as in (12).

[6]That being said, the backtracking analysis still has the advantage of not requiring a stipulation of feature-driven movement in cases where a monomorphemic augment appears. For instance, Caha's (2019:304) analysis of the declension of Russian *žen-a* 'woman' also features the re-surfacing of a single smaller vocabulary item in the instrumental. If this is the correct analysis, the backtracking-less story will need to stipulate a feature-driven movement where the backtracking analysis does not, and would operate in a manner parallel to the one advanced here. An anonymous reviewer helpfully points out that parts of the Finnish case paradigm also feature the resurfacing of a single morpheme, as discussed recently in Kloudová (2020).

(12) a. /*duu(æ)*/ ⇔ ACC



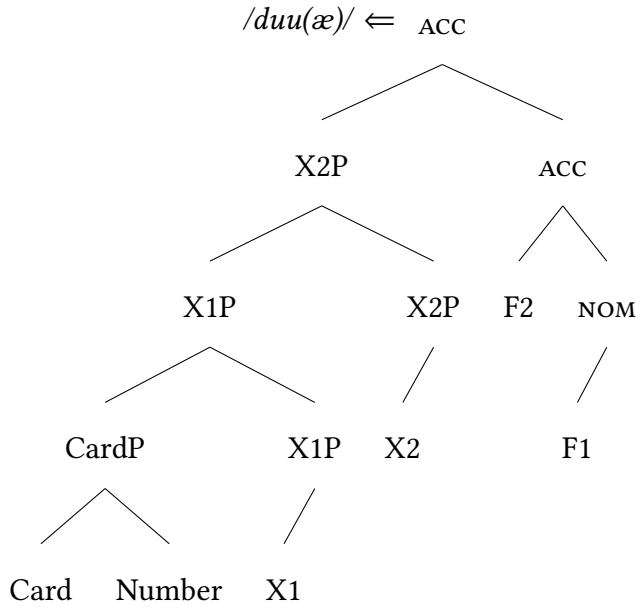
b. /*-e*/ ⇔ X1P

c. /*-m*/ ⇔ X2P

Simply put, this vocabulary item is a complex idiom that can overwrite a phrase that contains the augments */-e/* and *-m/*, as well as any subtree of [F2[F1]]. Since Nanosyntax maximizes targets, it will therefore result in the corresponding derivation after the merger of F2, as in (13):[7]

[7]The derivation occurs as follows: 1. Merge(Card, Number), spell out 2. Merge(X1,CardP), c→s 3. Merge(X2,X1P), c→s 4. Merge(F1,X2P), c→s 5. Merge(F2,F1P), s→s.
The vocabulary item *duu(æ)* matches the whole tree at the root at every single step.

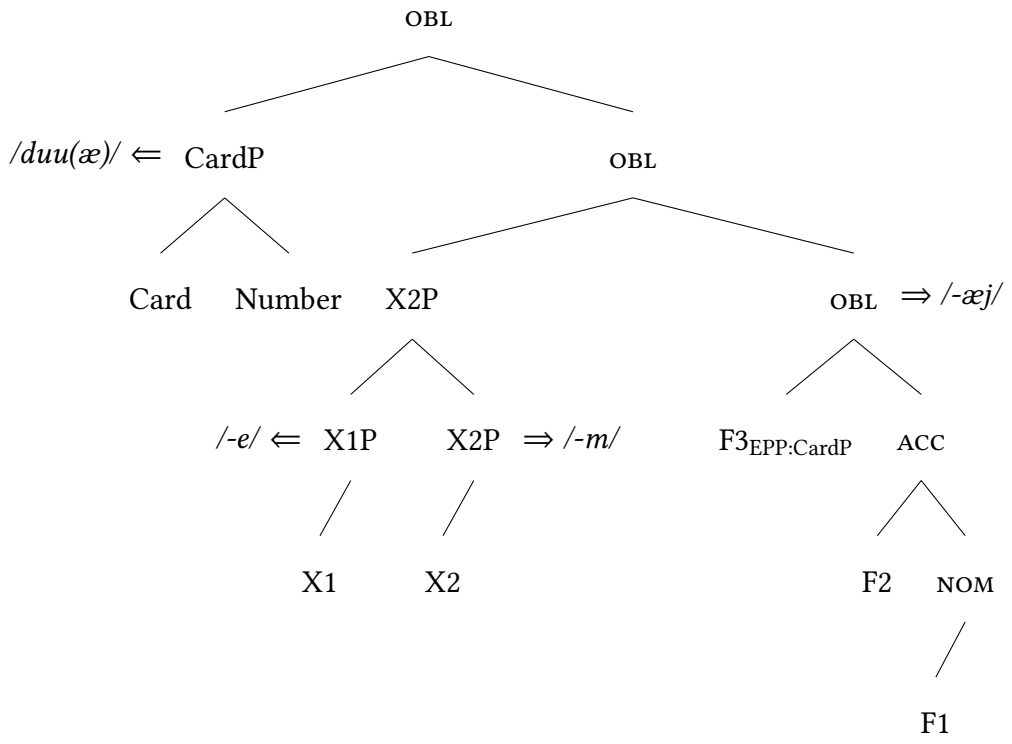
(13)



Next, let us again assume that there is an EPP-feature associated with F3 that attracts CardP. This results in feature-driven movement of CardP to the specifier of F3. However, the remainder cannot spell out the newly merged F3, and hence the spellout algorithm continues operating as usual: spec-to-spec movement targets the remnant X2P, and moves it from the specifier of F2 into the specifier of F3. This allows for the spellout of F3 in a constituent [F3[F2[F1]]], essentially parallel to the case with *max* above.[8]

[8]It is of course essential that CardP end up in the outer specifier position in order to derive the proper linear order. There are two obvious paths to ensure this: Either feature-driven movement precedes spellout driven movement and the latter is subject to *tucking in*, or spellout-driven movement precedes feature-driven movement, and we subextract from the inner specifier position to an outer specifier position. In either case, subsequent spellout driven movement may have to move *both* specifiers to create viable spellout targets (and presumably do so in an order preserving fashion). To my knowledge, the relation between feature driven movement and spellout driven movement has not been discussed at much depth (at least publicly), so I will leave these details aside, merely noting that there are options to ensure the desired results.

(14)



The case structure thus spells out the same way as above; however, the feature-driven movement has additionally extracted CardP from a larger idiom. Because *duu(æ)* can no longer match X1P and X2P (since they are no longer forming a phrasal node that can be matched by *duu(æ)*), the *pointed to* vocabulary items, i.e., the overwritten idiom chunks re-appear. The analysis is thus quite similar in spirit to Caha's, but derives it with different means.

None of this is to suggest that this ought to taken as a proper analysis of the Digor Ossetic Augment. The point was to show that a) the backtracking analysis runs into an issue that will require a further stipulation, such as feature driven movement, and that b) once such an allowance is made, an analysis without backtracking becomes possible as well. That is to say, the augment case does not offer a good case for the *necessity* of backtracking.

4.4 Size and configuration: declension class

Caha (2019, 2020) employs backtracking to develop a *size theory of declension classes*: A declension class is determined by the size of the root, and the consequences for anchoring affixes and the point at which affixes need to start backtracking. This is a highly attractive theoretical development, for a variety of reasons. In featural theories of declension class, the theoretical status of the features generally remains unclear – there is no reason to suspect that declension class represents syntactically active features (for instance, gender may be the target of agreement/concord processes, but declension class is not). Yet, a featural conception in a late insertion model requires post-syntactic reference to them, i.e., they seem to violate the inclusiveness condition (Chomsky, 1995). In contrast, the size theory does not have declension class features, and models declension class purely in terms of the cyclicity of exponence targeting feature sets (trees) of varying sizes.

A simplified example of Caha's theory is provided in Table 4.7 (I refer the reader to Caha's work for the actual details; the overview here is meant to lay out the spirit of the proposal). Both *zavód* 'factory' and *mést-o* 'place' share an identical f-seq, but the former lexicalizes are larger part of this structure, reaching all the way up to F2, while the latter lexicalizes only #P. This leads to *mést-o* 'place' having affixes that are anchored at the F1 node in the nominative (F1P) and accusative (F2P), while *zavód* 'factory' is able to spell these out as part of the root. However, for both items, subsequent affixes must be anchored at the F1 level, i.e., the derivation of *zavód-a* (GEN) involves backtracking.

I believe, however, that the featureless conception of declension class can not only be derived without backtracking, but that such a theory may in fact improve upon the size theory of declension class.[9] Consider first how we can replicate the basic aspect of these lexicalization tables, as in

[9]Technically, both conceptions of declension class make slightly different predictions. Insofar as the backtracking-less theory of spellout is a sub-theory if the one with backtracking, however, the tools of the backtracking-less theory can, of course, be employed to create a theory that can encompass the predictions of both theories.

Table 4.7: Simplified lexicalization tables for Russian *zavód* 'factory' and *mést-o* 'place'

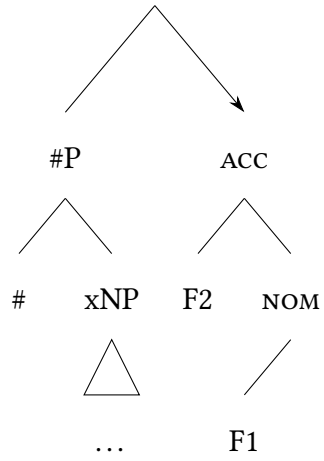
	xNP	#	F1	F2	F3	F4	F5	F6
NOM	zavód							
ACC	zavód							
GEN	zavód			a				
LOC	zavód				e			
DAT	zavód					u		
INS	zavód						om	

	xNP	#	F1	F2	F3	F4	F5	F6
NOM	mést		o					
ACC	mést			o				
GEN	mést			a				
LOC	mést				e			
DAT	mést					u		
INS	mést						om	

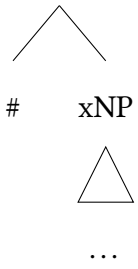
(15). As in the previous discussion, this branching conception of *zavód* can spell out nominative and accusative (hence the absence of dedicated surface morphemes vis-a-vis the *-o* in *mést-o*). The fact that it is a branching vocabulary item with a pointer, however, allows for *partial overwrite*, i.e., it is compatible with the lexicalization tables introduced above without any need for backtracking.

(15) An alternative without backtracking

a. *zavód* ⇔ ACC



b. *mést* ⇔ #P



By employing branching vocabulary items to model declension classes, we can explore the possibility of a *configurational theory of declension class*: If vocabulary items fall into declension classes, they vary by size (and possibly f-seq, in case they have different gender, but I leave this aside for now), as in the size theory. However, they may *additionally* vary in the point at which the f-seq is split into a left and a right branch (and if it is split at all). There is data that suggests that such an approach is necessary, even for a theory that does adopt backtracking. Consider for instance the two strong feminine consonantal inflection classes from Icelandic (Müller 2005, p. 232), or the Greek declension classes VII and VIII (Alexiadou and Müller 2008, p. 120).

Table 4.8: Strong feminine consonantal declension classes (Icelandic)

	geit 'goat' (FC1)	*vík* 'bay'(FC2)
NOM SG	geit	vík
ACC SG	geit	vík
DAT SG	geit	vík
GEN SG	geit-ar	vík-ur

In both the Greek and the Icelandic case we can observe the following: For two declension classes, nouns of the same gender show zero morphology for the lower part of the case hierarchy. As

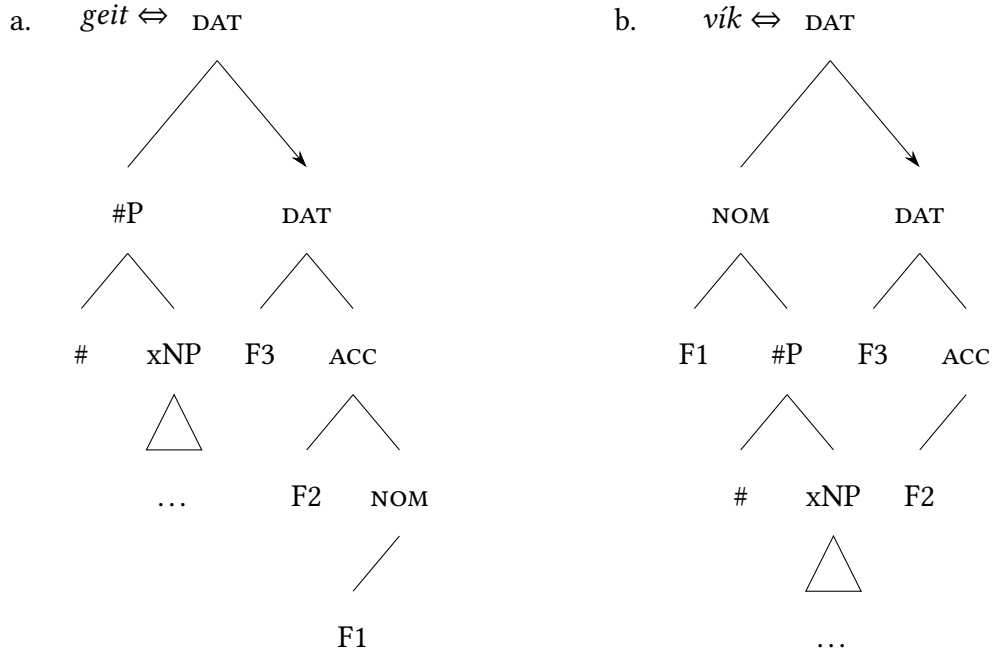
Table 4.9: Neuter declension classes VII and VIII (Greek)

	spiti 'house' (VII_n)	*soma(t)* 'body'($VIII_n$)
NOM SG	spiti	soma
ACC SG	spiti	soma
GEN SG	spitj-u	somat-os

usual, we conclude that Icelandic *geit* 'goat' and *vík* 'bay' are able to spell out the whole phrase corresponding to a dative. In parallel fashion we conclude that Greek *spiti* 'house' and *soma* 'body' lexicalize the whole accusative structure. They do, however, take different suffixes for the larger cases. Under an analysis that relies solely on backtracking, this is unexpected: They spell out structures of the same size, and presumably share the same f-seq.[10] Under the configurational perspective laid out above, however, two vocabulary items may lexicalize the same set of features, but in a different syntactic configuration. This is not the place to develop a serious analysis of these inflectional systems, but let me illustrate how one can account for these type of data in principle with vocabulary items such as those in (16):

[10]The latter is, of course, not necessarily a given, but examples like these are easy to find, and reducing all such cases to differences in the f-seq risks being indistinguishable from returning to a theory that employs declension class features.

(16) Configurational declension classes



Assuming two suffixes *-ar* [F4[F3[F2[F1]]]] and *-ur* [F4[F3[F2]]] now explains how these two classes differ besides sharing an identical f-seq and being of equal size, as in the lexicalization tables in Table 4.10: The partial overwrite of the right branch requires a different anchor. While *geit* restricts overwriting of the right branch to affixes anchored at F1, *vík* restricts it to affixes anchored at F2. Crucially, backtracking on its own is unable to derive lexicalization tables such as those in Table 4.10: Given the existence of *-ur*, backtracking would never result in backtracking all the way to F1. In a configurational theory, however, identity in size does not necessarily imply identity of inflection class.

I believe that I have shown that a backtracking-less theory can in principle underly a featureless theory of declension classes. Obviously, this configurational theory of declension classes will need to be put to a broader empirical test. Independently of the question of backtracking, however, I believe to have shown that thinking of declension class not only in terms of size but also in terms of configuration is likely be necessary to account for relations between declension classes – whether a theory deploys backtracking or not.

Table 4.10: (Toy) lexicalization tables for Icelandic *geit* 'goat' and *vík* 'bay'

	xNP	#	F1	F2	F3	F4
NOM	geit					
ACC		geit				
DAT			geit			
GEN	geit			ar		

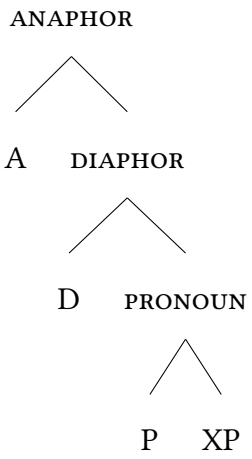
	xNP	#	F1	F2	F3	F4
NOM	vík					
ACC		vík				
DAT			vík			
GEN	vík			ur		

4.5 Pseudo-ABA

H. J. Middleton (2020) and J. Middleton (2021) provides us with a final case of interest: Pseudo-ABA patterns. In her investigation, Middleton uncovers strong cross-linguistic evidence for a general *ABA restriction for pronouns/diaphors/anaphors, i.e., a ban on using the same form for anaphors and pronouns to the exclusion of diaphors. She derives the restriction from a structure along the lines of (17).[11]

[11] H. J. Middleton (2020) and J. Middleton (2021) defines an anaphor as a locally bound variable, a diaphor as a non-locally bound variable, and a pronoun as a free variable. The lexical entries for English pronouns such as *they/she/he* etc are systematically ambiguous between diaphors and anaphors (an AAB pattern) – or, in terms of the theory adopted here, they lexicalize [D[P[XP]]] and English lacks a smaller competitor.

(17)



While there are no *ABA patterns, however, there are Pseudo-ABA patterns, A-B-A+x, as shown in (18), from H. J. Middleton (2020).[12]

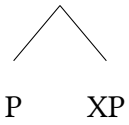
(18)

	PRONOUN	DIAPHOR	ANAPHOR	
a.	wén	jì	ə̀wénə̀ wén	(Babanki)
b.	avan	tan	avanavan	(Malayalam)
c.	rẹ́	òun	ara rẹ́	(Yoruba)

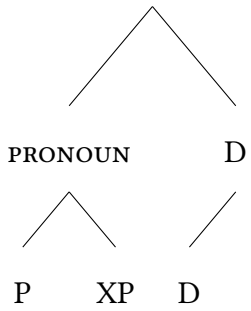
In order to account for Pseudo-ABA patterns, we extend our approach slightly. As before, the effect can be captured as a partial overwriting; however, this time a smaller candidate re-emerges. Consider the vocabulary items for Malayalam, as in (19). In the same spirit as the analyses above, the key to making this system work lies in the fact that (19b) enforces comp-to-spec movement subsequent to merging D. In this way – and unlike a vocabulary item that lexicalizes [D[P[XP]]] in a strict head-complement sequence – *tan* makes this derived specifier available for subsequent spec-to-spec extraction.

[12]While comparable cases are rare in the degrees of adjectives, Bulgarian offers a similar pattern with *mnogo* 'much/many': *mngogo* (POS) – po-veče (CMPR) – naj-mnogo (SPRL), see Bobaljik (2012: 126).

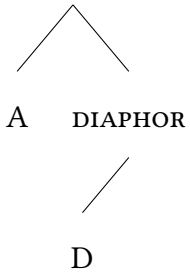
(19) a. *avan* ⇔ PRONOUN



b. *tan* ⇔ DIAPHOR



c. /REDUPL/ ⇔ ANAPHOR

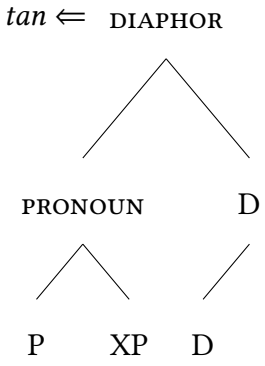


That is to say, *tan* can spell out the whole diaphor, but subsequent merger of the A feature that derives an anaphor leads to sub-extraction of the pronoun. This is shown in (20): In (20a), *tan* is able to spell out the root after the spellout algorithm attempts comp-to-spec movement (after simple spellout and spec-to-spec movement were unsuccessful). Upon merging A (20b), the spellout algorithm can now move this derived specifier by spec-to-spec movement. This movement results in a structure that can be spelled out, with the right branch being matched by the reduplicating vocabulary item REDUPL (19c). On the left branch, the previously overwritten *avan* re-emerges as the spellout of the PRONOUN structure.[13] The only real difference to the previous analysis is

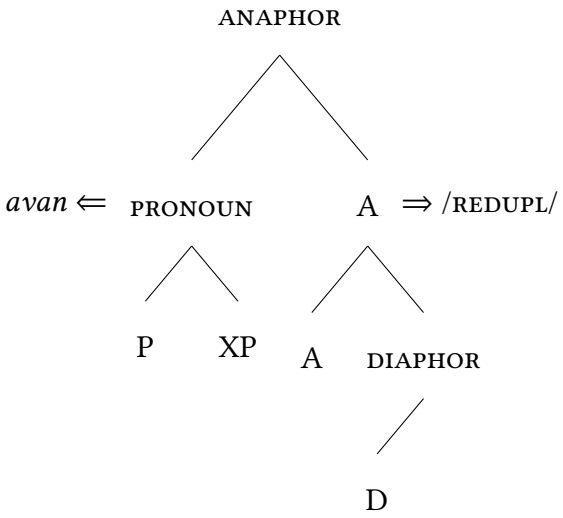
[13]Note that I am abstracting away from linear order here; the astute reader may notice that this assumption isn't altogether innocent, as the re-emerging smaller element has to be to the right, rather than to the left in Babanki as

that the left branch of *tan* has an independent candidate for its spellout, but the mechanisms are otherwise identical.

(20) a.



b.



Once again, an analysis in which a vocabulary item can force movement prior to spellout, despite targeting the root, allows for an account that can model Pseudo-ABA effects without backtracking.

well as Yoruba – i.e., the analysis here might require the suffix to undergo subsequent movement above the root.

4.6 Conclusion

In this paper, I have offered an analysis of a variety of phenomena that eschews the notion of *backtracking*. Instead, vocabulary items were argued to take a highly active role in structuring the course of a derivation, frequently enforcing movement even when the spellout target was the root node – a possibility that has not been explored at great depth within Nanosyntax, but one whose availability follows from standard assumptions. I have shown that such vocabulary items can result in *partial overwrite*, and that a variety of empirical phenomena that appeared to necessitate backtracking can in fact be analyzed without backtracking, once such vocabulary driven movement is fully exploited.

I have further argued that regardless of the adoption of *backtracking*, a configurational extension of the size theory of declension class is likely to be necessary, and that such an extension requires the tools I employed as an alternative to backtracking. That is to say, a theory with backtracking will still need pointers, branching vocabulary items, and partial overwrite, but a theory with these features does not necessarily need backtracking.

These results are of core importance, insofar as the proposed analyses all manage to keep the computation of spellout in the realm of a linear algorithm. I believe that this should be a welcome simplification of the theory, and any data that might suggest the need for the adoption of backtracking should be scrutinized carefully before we conclude that the realm of linear algorithms is one we have to abandon.

5 | Kipsigis Singulatives

5.1 Introduction

Nominal classification systems that have morphological effects, but no discernible syntactic ones are crosslinguistically common, but their nature has remained theoretically problematic. In this paper, I develop a novel account of one such system, the Kipsigis (Kalenjin) number-based noun classification system described and analysed in Kouneli (2019, 2020). I explore the idea that systems of nominal classification arise as the effect of *interface legibility conditions* that are imposed by the nature of the vocabulary, and its role in translating syntactic structure at PF (Halle and Marantz 1993, Starke 2018): Under a superset-based approach to spellout, vocabulary items (VIs) determine the configuration in which number is PF-interpretable, and thus give rise to different classes, depending on the structure they are able to interpret. Building on Caha (2020), I argue that declension classes do not correspond to classificatory features (primitives), but rather to different syntactic configurations of the same features (derived structural properties), that arise in response to interface legibility requirements. Like Kouneli's (2020) analysis of the system, the one presented here derives the classes from uninterpretable number features – but rather than stipulating this property, the (un)interpretability of number is an interface effect derived from general principles of the Nanosyntactic theory of PF interpretation.

The Endo-Marakwet (Kalenjin, Nilo-Saharan) data in (1) reflects the core of the system of nominal classification we will explore. Kalenjin nouns come in three classes with respect to

number marking: There are those that mark only the plural, those that mark only the singular, and those that mark both singular and plural.

(1)

	SG	PL	
a.	kipaw	kipaw-**tiin**	'rhino'
b.	peel-**yaan**	peel	'elephant'
c.	pata-**yaan**	pat-**een**	'duck'

Endo-Marakwet (Kalenjin), Zwarts (2001) via Kouneli (2020: 2)

Note that we find a bi-directional pattern of morphological containment in the surface form: In some cases, the plural form morphologically contains the singular form (1a), but in others the singular form morphologically contains the plural form (1b), while a third case marks both forms independently (1c). The choice between these three situations is determined by the root. In her detailed analysis of Kipsigis (Kalenjin), Kouneli (2020) shows that count nouns from all three classes behave like ordinary singular/plural count nouns, regardless of which class they belong to – that is, there are no known syntactic effects that would distinguish between count nouns from these classes: The effect is purely morphological.[1]

In the morphological part of Kouneli's (2020) analysis, the three classes correspond to the value of an uninterpretable binary number feature *u*SG (singular) on the categorizing head little *n*, i.e., a *classificatory feature*. This uninterpretable feature can take three values, [+SG], [−SG], or underspecified, and affects the marking of number (only number values that differ from the classificatory feature get marked), see Section 5.5 for details.

Kouneli's (2020) analysis is of special interest for a nanosyntactic account for nominal classification, because it constitutes a twofold challenge to such an approach. First, it makes a sustained

[1] See also van Urk and Sun (2021), who apply these insights to the number marking system of another Nilotic language, Dinka (Nilo-Saharan), and show that the challenges the system appears to pose to item-based approaches to morphology can be tackled if the system is in fact based in similar number-based nominal classes. Erschler (2022) investigates a dedicated singular marker in Digor Ossetic.

argument that *binary* number features are necessary for an account of the data, while Nanosyntax eschews binary features. The empirical core of this challenge is the bi-directional pattern of morphological containment: For one class, the morphological plural form contains the morphological singular form, yet for another class, it is the other way around. Secondly, Kouneli's (2020) analysis employs these features as uninterpretable classificatory features on little *n* (inspired by the approach to gender developed in Kramer 2015, 2016). The (un)interpretability of a feature, however, should follow from the theory of interface interpretation, rather than simply be stipulated, and hence this poses a further challenge.

The first challenge is essentially an empirical one: To answer it, one must show that the data can be derived in a theory that makes no use of binary features. The second challenge, however, is related to a deeper conceptual issues. First, a theory of the PF interface should offer an explanation of why a certain feature is uninterpretable at PF. Secondly, a classificatory *feature* poses issues regarding the way morphological classes are introduced in the Y-model. Alexiadou and Müller (2008) identify the problem roughly as follows: Certain inflectional classes appear to be relevant solely to the morphology, but not the syntax. Such classes are frequently modelled as features, such as the uninterpretable [±SG] feature in Kouneli's analysis. In a Y-model with late insertion (i.e., a model with an interpretative vocabulary that is distinct from the set of formants, cf. Halle and Marantz 1993, 1994), the status of such "morphomic features" is unclear: Insofar as such features are essentially morphological in nature, their presence in the syntax would seem to violate the Legibility Condition, which restricts the syntax to operating on objects that are syntactic in nature – such a violation raises the question of why these particular non-syntactic objects can be featuralized and dragged through a derivation. If, on the other hand, they are introduced post-syntactically, their presence (as features) would violate the Inclusiveness Condition – a post-syntactic system that can insert features would itself appear generative, rather than interpretative.

At the heart of the proposal I develop here is the idea that declension classes correspond

not to features, but to spellout configurations, as articulated in Caha's (2020) size-based theory of declension classes, and its configurational extension in Blix (2021a). In a nutshell: Under a superset approach to spellout (Starke, 2009), a vocabulary item (VI) effectively characterizes a finite set of PF-legible syntactic structures (in crucial contrast to VIs in a subset approach), namely the set of trees contained in the lexicalized tree. For a derivation to converge, all its parts must be legible at the interfaces (Chomsky, 1995).[2] Hence, a vocabulary item may impose extremely local legibility restrictions. To ensure convergence/legibility, then, the syntax may perform a variety of operations – say, moving the complement of a feature to it's specifier position (*spellout-driven movement*). In other words: Particular vocabulary items induce specific legibility conditions, which cause highly local movement. Declension classes correspond to the resulting different spellout configurations.

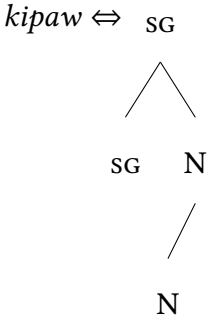
Let me first illustrate this for nouns that mark plural only, versus nouns that mark both singular and plural. Assuming a functional sequence (f-seq) that determines a merge-order N ≻ SG ≻ PL, we simply account for the unmarked singular with *kipaw* 'rhino', by postulating that it lexicalizes SG, and for the marked singular with *pata* 'duck' by not doing so. That is, the former lexicalizes a set {SG,{N}}, while the latter simply lexicalizes {N}, as in (2), which provides both the sets and their representations in tree form.[3]

[2]We might consider Nanosyntax itself as a movement towards a *theory of uninterpretable features*. In current minimalist reasoning, the binary distinction between interpretable and uninterpretable features is conceptually well motivated, but practically uninterpretability is generally simply stipulated – that is to say, it does not follow from well understood general properties of the interfaces that such and such a feature cannot be interpreted at PF or LF, and that certain syntactic operations may resolve the legibility issue. While by no means comprehensive, the Nanosyntactic reasoning that I apply here, is ultimately a first stab at developing a theory that derives PF-uninterpretability from the bare necessity that the PF interface contains pairings of syntactic and phonological information, and a basic set of operations that may resolve such uninterpretability.

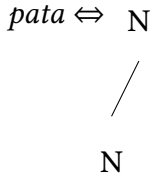
[3]As per the minimalist set-theoretic interpretation of linguistic structure building (Chomsky, 1995), I assume trees to be representations of sets. The terminals/features are elements, the edges (top to bottom) denote set membership in a labeled set. The difference between the sets (i-a) {α,{β}}, (i-b) {α,β}, and (i-c) {{β},{α}} is represented as follows:

(2) *The Vocabulary Items*

a. < {SG,{N}}, /kipaw/ >

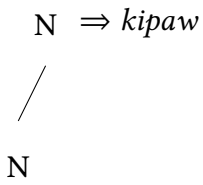


b. < {N}, /pata/ >

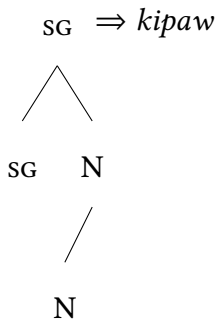


In the case of *kipaw* 'rhino', the derivation of a singular will proceed as in (3): After selecting *kipaw* as the spellout of {N} (assuming that choice between roots – the VIs that spell out {N} – is free, cf. Marantz 1996, Caha et al. 2021a), we simply merge SG with this set, and *kipaw* continues to match it. Hence, no further syntactic operation is required, and no morphological marking is associated with the singular, since it gets spelled out with {N}.

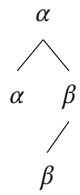
(3) a.



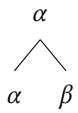
b.



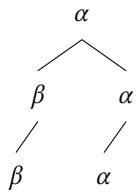
(i) a.



b.



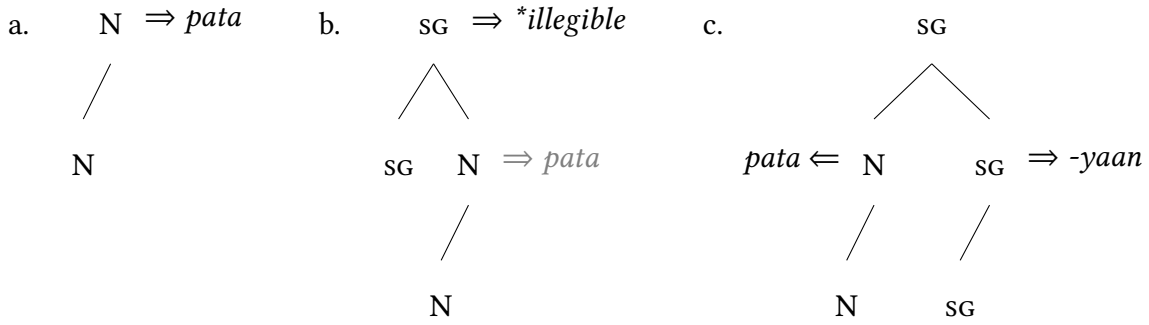
c.



By hypothesis, only sets (but not terminals/heads/non-set elements) are possible targets of spellout. That is, if α, β are heads, then the singleton set containing β, $\{\beta\}$, is a possible target of spellout in (i-a,c), but there is no corresponding way to spell out β to the exclusion of α in (i-b). In parallel fashion, only (i-c) makes α available as a target of spellout to the exclusion of β, since (i-a,b) do not contain a set $\{\alpha\}$.

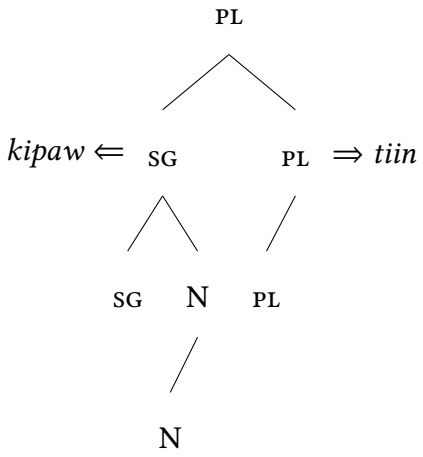
In contrast, merging SG in the case of *pata* 'duck' does not produce a PF-interpretable structure: Since *pata* does not lexicalize SG, the resulting syntactic structure cannot be matched – SG is uninterpretable in this configuration. In order to create an interface-legible structure, the syntax transforms the set {SG,{N}} into the set {{N},{SG}}, i.e., it moves {N} into a specifier position. Since spellout targets phrasal objects only, this transformation is crucial: Now SG is itself a phrasal object in its own right, and can thus be spelled out on its own, in this case by the suffix *-yaan* that lexicalizes the relevant structure, thus rendering the right branch interpretable. Note that in both cases, we have built a singular structure, i,e., a phrasal object labeled SG that obeys the functional sequence. The difference in morphology comes about, because the syntax has to change the configuration to ensure the PF-legibility of the singular feature.

(4)



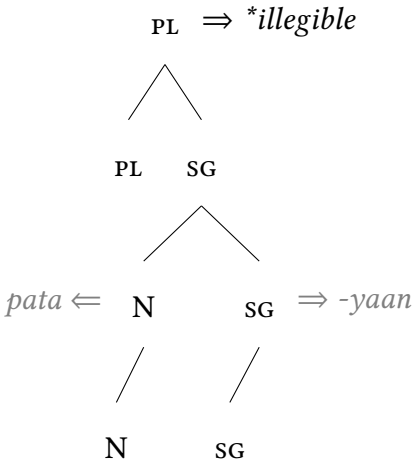
To build a plural structure, we continue by merging PL with the singular structure. In neither case can the immediate result be spelled out, and hence the syntax once again attempts repairs. In the case of *kipaw*, this repair takes the same form as the one we saw for the singular of *pata*: Move the complement to the specifier position of PL, i.e., make PL into a phrasal node that can be a target of spellout on its own, as in (5). The left branch continues to be matched and spelled out by *kipaw*, and the new right branch, {PL} can be spelled out by the plural suffix *-tiin*.

(5)

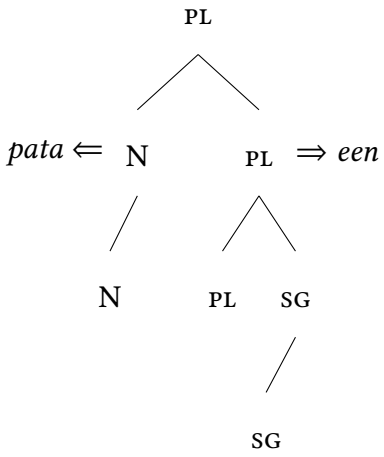


With *pata*, too, simply merging PL with the singular structure does not result in a PF-legible structure (6a). In this case however, there is a more economical repair strategy available: Rather than transforming the newly merged feature PL into its own set that can receive its own spellout, the syntax will first attempt to move the specifier of SG out of the way, i.e., so-called spec-to-spec movement: After merging PL with the singular structure, we move the specifier {N} out of the way and into the specifier position of PL, as in (6b). This is an operation that maximizes the potential target for spellout, since it avoids the spawning of new suffixal positions that is generally associated with movement of the complement, just as we saw above. Instead, it makes available a single structure {PL,{SG}} as the right branch, which can then be spelled out by a single VI, *-een*, which overwrites the previous spellout of {SG} (i.e., it keeps the number of surface morphemes identical, rather than increasing it by one, as comp-to-spec movement usually does).

(6) a.



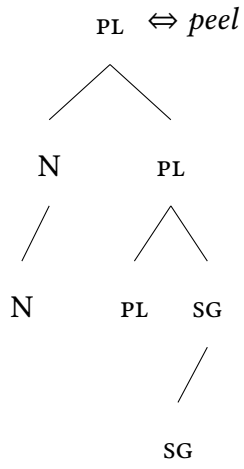
b.



Note that the differences in the distribution of number markers and the form of the plural markers (*-tiin* vs. *-een*) both arise from the way the syntax reacts to the legibility conditions imposed by the vocabulary at the PF interface, i.e., the two classes correspond to interface-induced configurational differences.

With this basic mode of generating declension classes as effects of legibility restrictions in mind, let me sketch how we will address the general empirical picture of Kouneli's (2020) challenge: We need to derive the three classes without reference to binary features, and accounting for the fact that some nouns show a morphological pattern in which the plural form contains the singular form (such as *kipaw* 'rhino', above), while for others, it is the singular form that morphologically contains the plural form, such as *peel* 'elephant' in (1b). The answer is already contained in the sketch above, in particular in the derivation of *pata-een* in (6): If a vocabulary item can lexicalize a well-formed tree (Starke, 2009), then it should be able to lexicalize the particular structure that we derived in (6), i.e., a lexical entry such as (7).

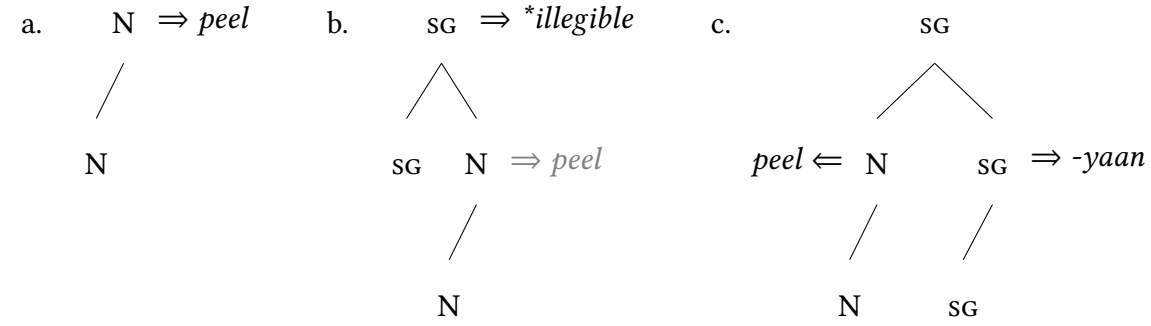
(7)



Such a vocabulary item has the curious property that it contains the singular feature SG, without containing a singular configuration – that is to say, there is no sub-tree of (7) that would contain N and SG to the exclusion of PL.

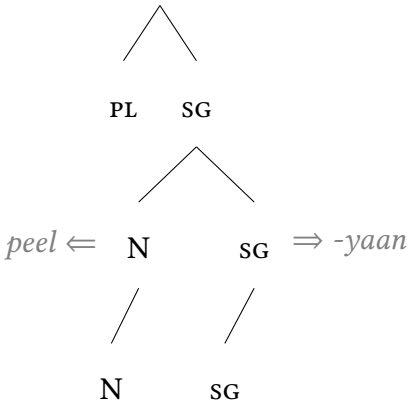
The effect is that such a vocabulary item triggers the same initial syntactic response as *pata* 'duck', i.e., nouns that mark both numbers, as shown in (8). Initially, {N} is spelled out by *peel*, a possibility that is given, because {N} is a tree contained in the one lexicalized by *pata*, (8a). After merging SG with {N}, however, the resulting structure is illegible, since *peel* does not contain this tree as a sub-tree. Consequently, the syntax transforms the set in the same fashion as above, i.e., by transforming {N} into a specifier, deriving {{N},{SG}}. The resulting set is legible at PF, since both its elements, {N} and {SG}, can be spelled out independently, resulting in singular-marked *peel-yaan* '*elephant-SG*', (8c).

(8)

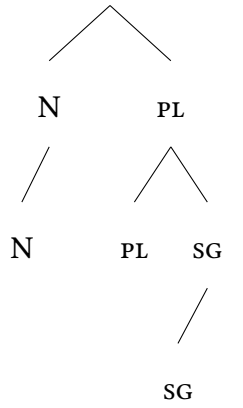


Turning to the plural form. As with *pata* in (6), simply merging PL with the singular structure does not result in a legible structure (9a), and hence spec-to-spec movement is attempted as a repair strategy (9b), exactly as before. However, in contrast to *pata-een*, where spellout matched the left and right branches independently, *peel* can actually match the whole resulting structure. Since Nanosyntax is target-maximizing, *peel* overwrites both *-yaan* as the spellout of {SG}, and itself as the spellout of {N}. The result is a return to a mono-morphemic spellout structure, without surface marking of number.

(9) a. PL ⇒ **illegible*



b. PL ⇒ *peel*



This type of vocabulary item thus constitutes the third number-marking class of Kipsigis, marking the singular, but receiving no dedicated marker for the plural, as a result of the particular properties of containment that pertain to lexical items with complex left branches.

In addition to the three basic classes, I explore further details of number in Kipsigis, focusing in particular on allomorphy in the number domain, as well as the interaction between number class and number allomorphy on the one hand, and the so-called thematic affixes on the other. In particular, I focus on the fact that Kouneli (2020) unearthes an interesting property of the thematic suffix that appears descriptively disjunctive: The allomorph is determined by the root, in case number is unmarked, but by the number-allomorph in case number is marked suffixally. I show that under the current system, this falls out as a simple case of locality without any dis-

junction: Whatever vocabulary item spells out number – whether it is a root, or a number suffix – is maximally local to the thematic structure, and can thus influence its spellout.

The paper provides four contributions: First, it addresses the immediate challenges raised by Kouneli's (2020) argument for binary classificatory features. By shifting nominal classification from syntactic features into the vocabulary, we derive the number-based noun class system (and the bi-directional containment effects) without binary features. Second, rather than simply stipulating that there are uninterpretable number features associated with the different classes, we derive the configurational (un)interpretability of number features from general principles of spellout. That is, we provide an explanatory account of the (un)interpretability of the features postulated by Kouneli (2020). Third, in doing so, it provides a conceptually sound locus for purely morphological classes, such as the Kipsigis number-based nominal classification system: They arise as interface-effects that are imposed by the vocabulary. If this is the right path to take, it suggests that syntacticians should not treat PF-interpretation as a purely interpretative system of translating syntactic structure into morpho-phonological structure, but rather one that imposes highly local legibility conditions – that is to say, the vocabulary may play an active role at the interface. This issue is also at the heart of two competing conceptions of matching vocabulary items and syntactic structure at the interface: Only under a superset-based approach does the vocabulary generate serious legibility restrictions, while a subset-based approach suggests no such mechanisms. Thus, this paper is a part of a bigger argument about the role of the vocabulary at the PF-interface, and the role of the PF-interface in syntax.

Finally, the paper makes an argument that is internal to Nanosyntax: I lay out a novel argument in favor of recent developments of the notion of phrasal spellout, in particular the algorithm developed in Starke (2018), Caha (2019). I will argue that the details of the bi-directional containment effects are a serious issue for older, de facto span-based versions of Nanosyntax (such as Caha 2009), while they do not constitute a problem for the analysis sketched above. In the nanosyntactic literature that is compatible with both spans and phrases as targets of spellout,

gapping has been suggested as the appropriate analysis for bi-directional morphological containment (see for example Márkus 2015, Wyngaerd et al. 2020, Caha et al. 2021b). Similar to the proposal sketched above, a gapping analysis also derives the "unexpected" case of nouns marking only the singular, but not the plural, from a lack of contiguity.[4] Under such a perspective, however, singulative marking really concerns the spellout of higher material. Suppose we have the two lexical items in (10), with the root lexicalizing the whole plural structure plus a higher head X.

(10) a. *sg-marking-root* ⇔ [X [PL [SG [N]]]]

b. *singulative-suffix* ⇔ [X]

In the plural, a root such as (10a) would simply spell out the whole XP. In the singular, however, the absence of PL leads to an inability of the root to spell out X, since the root can match X only if it is contiguous with PL. Hence, the spellout of X involves a suffix only in the singular, but not in the plural – the appearance of singulative marking arises. As I will argue below, however, such a conceptualization does not provide us with sufficient degrees of freedom to account for all the data, once we take the related domains of number allomorphy and thematic marking into account.

The paper is organized as follows: Section 5.2 provides an extremely brief introduction to the Nanosyntactic spellout algorithm, and the particular assumptions I make here. Section 5.3 lays out the basics of the analysis of number spellout in Kipsigis, starting with an account of the number-based classes, and then building an account of number-allomorphy. Section 5.4 provides evidence in favor of the proposed analysis of number spellout from the properties of the thematic suffix. I show that the phrasal spellout analysis for number immediately provides a straightforward, non-disjunctive characterization of the triggers of thematic allomorphy, in that

[4] I would like to thank two anonymous reviewers for bringing up the possibility of deriving the number facts with a gapping analysis rather than highly structured vocabulary items.

the element that determines the thematic allomorph is always the element that spells out the number structure. Section 5.5 offers a discussion of the proposal and its implications, as well as a brief comparison with the subset-based approach of Kouneli (2020), and an argument against a gapping analysis of the facts. Finally, Section 5.6 concludes.

Before we start, please note that all Kipsigis data in this paper comes directly from Kouneli's (2020) excellent study of the Kipsigis number system without which this paper would have been impossible.[5]

5.2 Background – The Spellout Algorithm

I adopt the relevant aspects the phrasal spellout algorithm from Starke (2018).[6] I lay out the ideas in a highly abbreviated form here, and refer the reader to Caha (2019) for a more detailed version.

In Nanosyntax (as in DM), the vocabulary is interpretative, i.e., it translates abstract syntactic structure into (morpho-)phonology. The pieces that undergo interpretation, however, are phrasal nodes rather than internally complex heads. Vocabulary items (VIs) – usually called lexical items – are tuples of semantic information (left aside here), phonological information, and a well-formed lexicalized syntactic tree that characterizes the set of syntactic structures it can interpret/that it matches: A VI matches those trees that are contained in the tree it lexicalizes (Superset Principle).[7] Multiple matches are resolved in the usual way, i.e., by an Elsewhere Prin-

[5]All data in this paper is presented in accordance with the conventions introduced by Kouneli (2020): Double vowels indicate long vowels. Underlined vowels indicate [−Atr] vowels. A low tone is indicated by V̀, a high tone by V́, and a HL contour tone by V̂. Data is generally presented in its underlying morpho-phonological form on the left, and the surface form on the right of an arrow. See also Kouneli (2019) for further discussion of the phonology of Kipsigis.

[6]Since neither backtracking nor prefixes play a role in the current analysis, I leave aside those technical aspects.

[7]Note that the Superset Principle approach to matching is essential to reconceptualizing VIs as interface legibility conditions. In an approach based on a Subset Principle, it is the context that defines a finite set of possible VIs that would match that context, namely its powerset. Any VI whose feature set is not an element of the powerset of the target will not be a subset, and hence won't match the context. The corollary is that VIs themselves match an infinite number of contexts (in principle, if not in practice); the set of sets whose powerset contains the VI's set is not inherently finite. In other words, illegible contexts arise only under extreme edge conditions, and any necessary

ciple (the item with the fewest 'unused' nodes wins).

Binary Merge operates on individual privative features (and sets built from such features), in an order that is constrained by a functional sequence (f-seq). Every operation of Merge is followed by an attempt at spellout. Crucially, since matching is defined over phrasal nodes, not every operation of Merge trivially results in an interpretable structure. That is to say, after merging a feature F with an XP, there may not be an appropriate VI to spell out the resulting FP. In these situations, the spellout algorithm may attempt 'repairs', i.e., extremely local movement operations driven by the requirement of interface interpretabilty (and thus by the available vocabulary).[8,9]

(11) Merge(F, XP), then:

a. Try: Spell out FP

b. If it fails, try: Move(Spec of XP) to Spec FP, Spell out

c. If it fails, try: Move(XP) to Spec FP, Spell out

Lightly adapted from Starke (2018), Caha (2019)

The first step in (11) is just simple spellout of the resulting FP, as in (12). By definition, XP must have spelled out successfully at the previous cycle, so a successful spellout of FP is said to

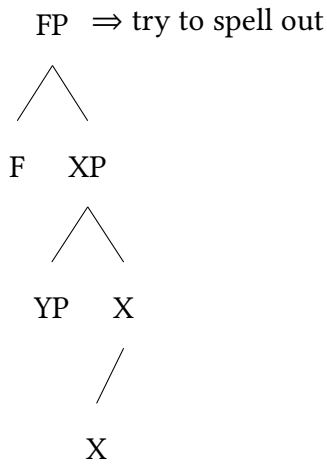
restriction must be imposed in core syntax, rather than at the interface.
In a Superset-based approach, the issue is reversed: The set of VIs that can match a particular context is in principle infinitely large, but the set of contexts that a VI can match is finite – hence, VIs effectively impose interface legibility restrictions on contexts, and thus provide us with a means of stating restrictions as extremely local interface conditions.

[8]Instead of considering these as movement operations in the sense of internal merge, it might be more fruitful to view these repair strategies as different attempts at constructing interpretable sets (by employing external merge) that obey the restrictions imposed by the functional sequence. For instance, we might merge F with an XP {X,Y} by constructing a set {F, {X,Y}}, or by constructing a set {{F}, {X,Y}}, with F being 'wrapped' in a set itself before being merged (i.e., subject to what we might call unary Merge). In both cases, the fact that F projects is determined by the f-seq, but the latter set corresponds the result of comp-to-spec "movement". This constitutes a way to account for the extremely local properties of spellout-driven movement, and its difference from internal merge and feature-driven movement.

[9]There is a similarity with locally optimizing approaches to syntax, such as Heck and Müller's (2007), in that there is an algorithm that selects between a variety of derivational options, at any relevant derivational step. In the current conceptualization, however, selection is based on absolute interface conditions, rather than violable syntactic constraints.

overwrite the previous spellout of XP. That is to say, the spellout algorithm as laid out above is generally *target maximizing*: It spells out structures that are as big as possible, given the available vocabulary and the conditions on matching.

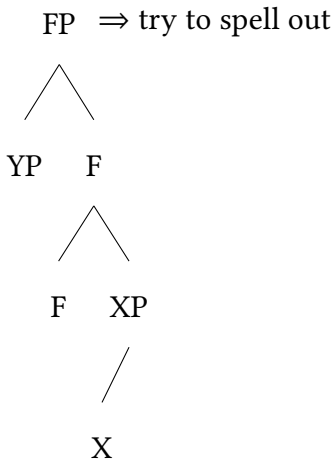
(12)



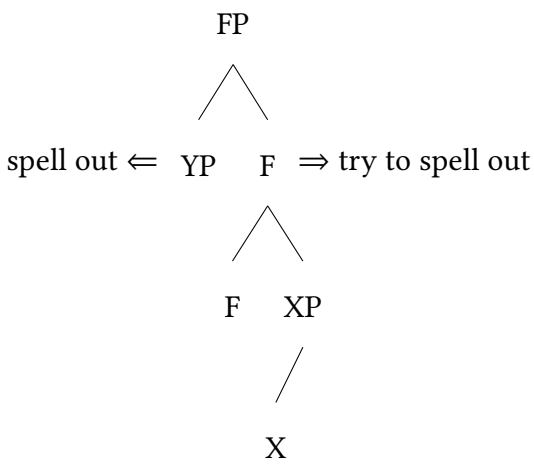
If that fails, spec-to-spec movement is attempted (if applicable), as per (11b). Concretely, I will assume that the spellout algorithm will first attempt to spell out the resulting structure at the root node (13a). If that fails (i.e., if no vocabulary item lexicalizes a tree that contains the root node), the spellout algorithm will attempt to spell out the two daughters of the root node instead, as in (13b).[10] Note that YP must have been able to spell out successfully at some earlier derivational stage, and will therefore never be the issue. In those cases where the root is not the target of spellout, spec-to-spec movement thus creates the condition for overwriting a suffix: If they can be matched successfully, F and X now get spelled out together as a suffix to the spellout of YP.

[10] Alternatively, the spellout algorithm might first target the two daughters, and then attempt to spell out the root node (regardless of whether the previous attempt was successful). If successful, this would overwrite the previous cycles, as usual. As an anonymous reviewer points out, De Clercq (2020: 6.3.2), De Clercq and Wyngaerd (2019) propose precisely such an algorithm. Either way, we arrive at a preference for spelling out the root node, if possible.

(13) a.

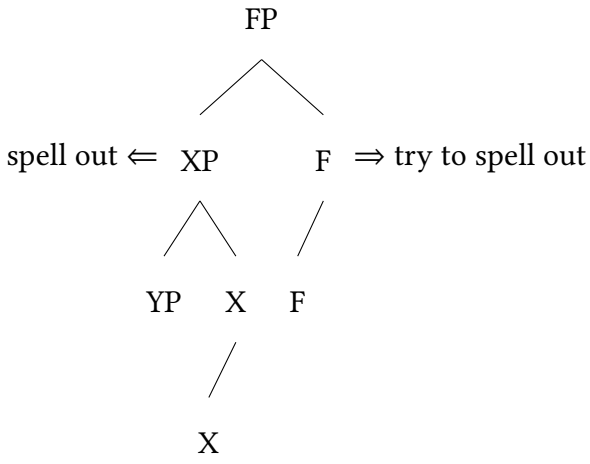


b.



Should the resulting structure also be uninterpretable, comp-to-spec movement (11c) is attempted instead. Again, the algorithm will attempt to spell out the resulting structure at the root. As before, in case that does not succeed, it will then attempt to spell out both daughters, as in (14), generally creating a new suffixal position.

(14)



Crucial for our purposes here is the fact that spellout – and thereby indirectly the available vocabulary – may trigger extremely local movements, such as spec-to-spec or comp-to-spec, and that spellout may be able to target the root following such an operation. I now turn to describing how vocabulary items driving such movement can be used to derive the Kipsigis system of number

based noun classification, and the corresponding number and thematic affixes.

5.3 Spelling out Number

Kipsigis count nouns fall into three classes with respect to their number-marking properties: First, nouns that mark only the plural (15). Second, nouns that mark only the singular (16). And third, nouns that mark both singular and plural (17). In addition to the number suffix, nouns generally show a theme sign (TH), and an obligatory secondary suffix (SEC). I set the latter two aside for the purpose of this section, and return to them below.

(15) a. peet-u-it → pêetúut
day-TH-SEC
'day (SG)'

b. peet-**uus**-ya-ik → pêetùusyék
day-**PL**-TH-SEC
'days (PL)'

(16) a. ngeend-**yaan**-ta-it
bean-**SG**-TH-SEC
→ ngéendyáat
'bean (SG)'

b. ngeend-a-ik → ngéendéek
bean-TH-SEC
'beans (PL)'

(17) a. sigis-**yaan**-ta-it → sìgìsyáat
sock-**SG**-TH-SEC
'sock (SG)'

b. sigis-**iin**-ik → sìgìsìiník
sock-**PL**-SEC
'socks (PL)'

Let us begin with the derivation of (17), which shows overt marking for both singular and plural. As a notational device, I will use bidirectional ⇔ for lexicalization and unidirectional ⇐ and ⇒ for the actual spellout of a particular syntactic structure. The relevant vocabulary necessary for the derivation of (17) is given in (18):

(18) a. xNP ⇔ *sigis*

...

b. SG ⇔ *-yaan*

SG

c. PL ⇔ *-iin*

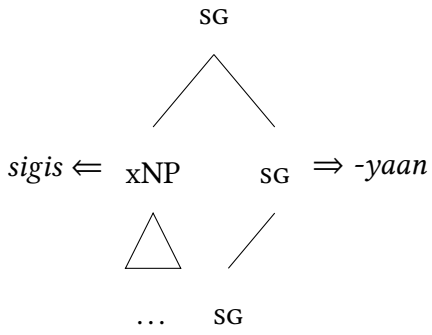
PL SG

SG

We'll operate under the standard assumption that roots that correspond to different concepts do not compete, and that choice among them is free (cf. Marantz 1996). Put in a somewhat simplified nanosyntactic framing: The choice of the lexical item spellout out the bottom is not regulated by the elsewhere principle, and roots that lexicalize different concepts cannot overwrite one another. I abstract away here from the particular implementation here, and restrict myself to merely pointing out that there are a variety of reasons and technical methodologies for doing so. One possibility would be to implement a variant of Harley (2014) by assuming an indexed $\sqrt{i}$ node at the bottom of the structure (with corresponding variation in the bottom of lexical items). Alternatively – and more in line with the current proposal – one might follow Caha et al. (2019, 2021a), Wyngaerd et al. (2021) in assuming that competition between roots is limited by general restrictions on overwriting. Insofar as the particular mode of implementing this non-compete clause is irrelevant to the points at hand, I will simply assume that such a restriction exists for the remainder of the paper.

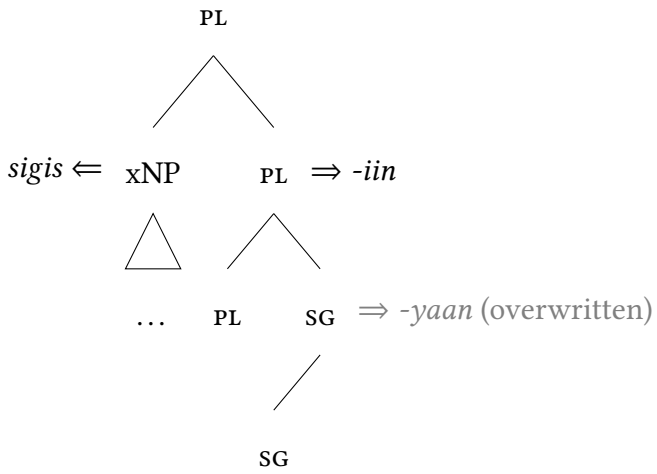
Let us take as a starting point the moment where xNP – the relevant part of the extended projection of the noun, just below number – is built, with *sigis* as its spellout. Once we we merge SG with this xNP, there is no candidate for spelling out the resulting SGP. Spec-to-spec movement is not an option, since xNP does not have a specifier. We therefore attempt comp-to-spec movement, and the resulting structure spells out successfully as in (19).

(19)



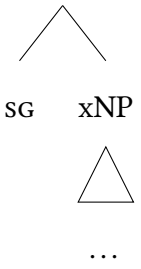
In the next step, we merge PL on top of the phrase we just built. Once again, there is no candidate that can spell out the whole structure – this time, however, the complement of PL does have a specifier, namely xNP. We thus attempt spec-to-spec movement, as in (20), moving xNP from the specifier of SG to the specifier of PL. The left branch continues to be spelled out by *sigis*, but the right branch is spelled out by the plural suffix *-iin*, which overwrites the singular suffix *-yaan*.

(20)



Next, we turn to the other common case, nouns that mark only the plural, like *pêetúut* 'day' in (15). In standard Nanosyntax fashion, we will assume that the absence of singular marking is due to root-size, i.e., the relevant vocabulary is as in (21).

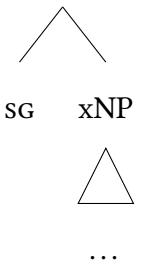
(21) a. SG ⇔ *peet*



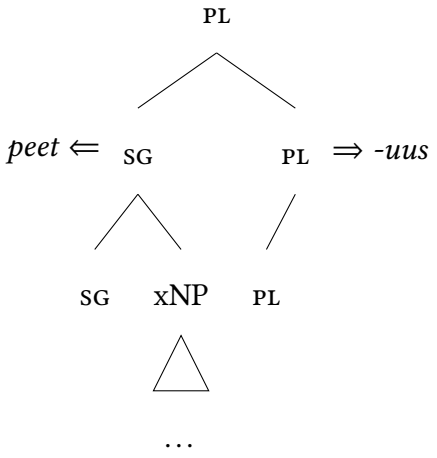
b. PL ⇔ *-uus*

That is to say, the singular spells out successfully after merging SG, and no movement operation is required, as in (22-a). Upon merging PL, however, no matching tree is available. Once again, there is no relevant specifier, and thus comp-to-spec movement is attempted, resulting in the creation of a suffixal position for the spellout of PL, as in (22-b).

(22) a. SG ⇒ *peet*



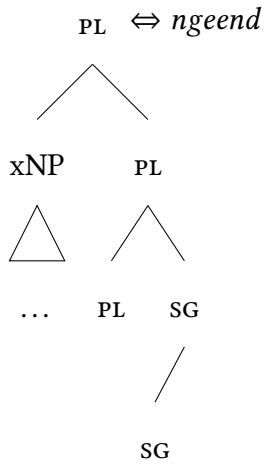
b.



Finally, we'll consider nouns such as *ngéendyáat* 'bean' in (16) that have number marking only in the singular. We have already come across the singular affix *-yaan* (18-c) in the derivation of nouns that mark both singular and plural; we thus have to only provide a structure for the root. If, as assumed here and elsewhere in the Nanosyntactic literature, plural and singular stand in a superset relation, this root must have the curious property of being able to spell out both PL and SG together, but not SG on its own. That is exactly the property of a vocabulary item like

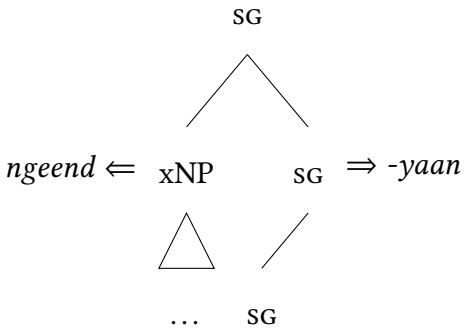
(23): It contains a tree in which xNP, PL, and SG form a constituent. There is, however, no sub-tree contained in this tree that would form a constituent containing xNP and SG to the exclusion of PL. That it is a well-formed tree nonetheless is obvious from a quick comparison with the derivation in (20), where this precise constituent structure was built.

(23)



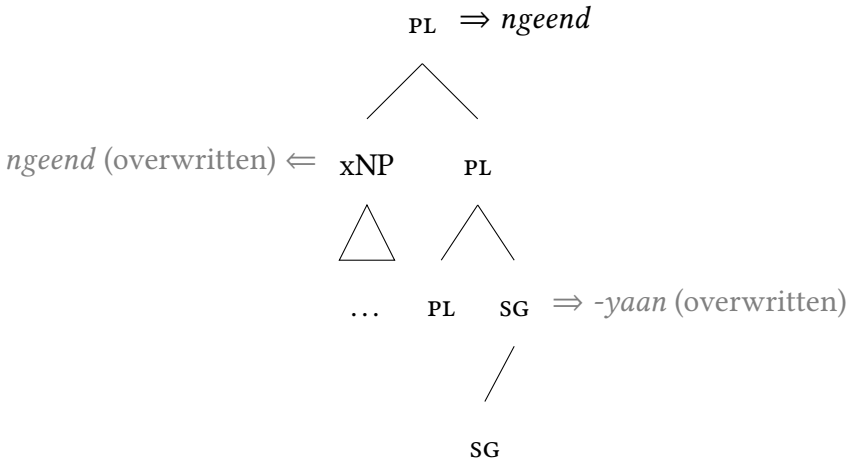
Let us consider then, how the derivation proceeds. As before, we start by merging the relevant xNP with SG. As just discussed, the resulting tree cannot be matched by *ngeend*, and as before repair strategies are attempted. Once again, there is no relevant specifier, and thus comp-to-spec movement is the relevant repair strategy, as in (24). Crucially, the xNP itself is a tree that is (properly) contained in the tree lexicalized by *ngeend* 'bean', and it can thus be spelled out by the VI in (23). The right branch is identical to the one we built previously, in the case of *peel* 'elephant', which marked both singular and plural, and as in that case, it is matched and spelled out by the singular suffix *-yaan*.

(24)



We continue as before, by merging PL, and there is, once again, no immediate match. The result of subsequent spec-to-spec movement, however, can be matched. Indeed, it can be matched at the root node, i.e., *ngeend* can simultaneously overwrite itself (as the spellout of xNP) and the singular suffix *-yaan* (as the spellout of {SG}) – we return to a monomorphemic spellout of the tree, and there is no dedicated exponent for number.

(25)



We have thus shown that the three classes of nominal behaviour with respect to number marking in Kipsigis can in fact be captured by a system based on privative features (pace Kouneli 2020), provided that spellout targets phrasal nodes. The way this noun classification system is implemented relies solely on different sizes and configurations of lexicalization. In eschewing uninterpretable classificatory features, we have thus arrived at an analysis that treats these classes

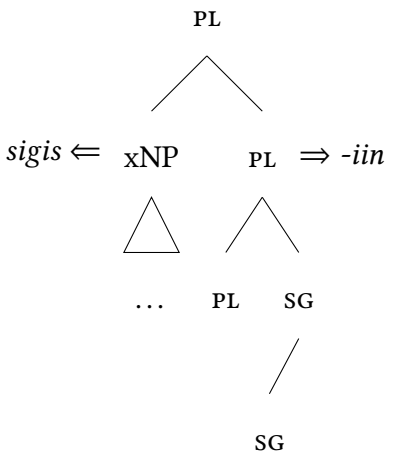
as interface conditions imposed by the lexicon. What gives rise to the particular system, is the (probably unusual, perhaps comparatively difficult to acquire) shape of the vocabulary items.

Having derived the classes of number marking, we now turn to the allomorphy that overt number marking exhibits.

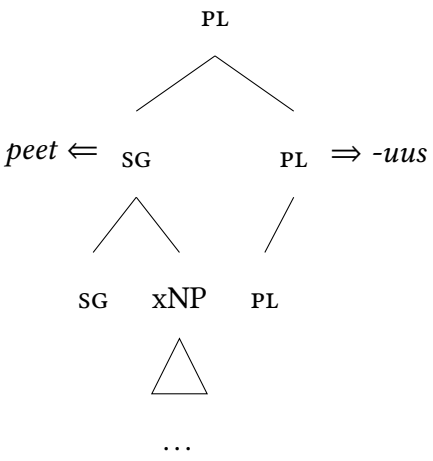
5.3.1 Number Allomorphy and Partial Overwrite

In addition to the threeway distinction in number classes just discussed, Kipsigis exhibits root-determined allomorphy for the particular number suffix used in the marked case, especially in the marking of plural. The previous section has already tacitly laid out one case of this root-conditioned number allomorphy: While *sigis* 'sock' takes the plural suffix *-iin*, *peet* 'day' takes the plural suffix *-uus*. The way the previous section implemented this, was as in (26). The root determined the structure the right branch of the PLP whose spellout is the plural suffix. In the case of *sigis* 'sock' (26a), the foot (bottommost element) of the right branch is SG, because SG could not be spelled out by the root, and thus triggered comp-to-spec movement in a previous cycle. Subsequent merger of PL and spec-to-spec movement creates a right branch {PL,{SG}}. In the case of *peet* 'day', in contrast, SG was spelled out by the root, and only PL could not be matched, hence triggering comp-to-spec movement. As a result, the right branch consists only of {PL} (26b).

(26) a.



b.



This analysis already lays the base for an account for the allomorphy in the number domain. If the plural head is spelled out together with the singular one, we find one allomorph, if it is spelled out on its own, we find a different allomorph. Such an account is configurational (what configuration number is spelled out in determines its interface interpretation, i.e., its allomorph), and it is root-driven (the lexical item that is the root determines the relevant aspect of the derivational path). This section will offer a generalization of this approach to number allomorphy.

Since the allomorphy of plural suffixes is considerably larger than that of singular suffixes, and thus more interesting and challenging, we will begin with the former, and turn to the latter ones below. Consider the three nouns in examples (27-29). All three nouns are unmarked in the singular (modulo the thematic and secondary suffixes to which we turn below). However, in the plural, *peet* 'day' combines with the plural suffix *-uus*, *oosn* 'forest' combines with the plural suffix *-oos*, and *laak* 'child' combines with the plural suffix *-oy*.

(27) a. peet-u-it → pêetúut
day-TH-SEC
'day (SG)'

b. peet-**uus**-ya-ik → pêetùusyék
day-**PL**-TH-SEC
'days (PL)'

(28) a. oosn-a-it → òosnêet
forest-TH-SEC
'forest (SG)'

b. oosn-**oos**-ya-ik → òosnòosyék
forest-**PL**-TH-SEC
'forests (PL)'

(29) a. laak-wa-it → làakwéet
child-TH-SEC
'child (SG)'

b. laak-**oy**-ik → làagóok
child-**PL**-SEC
'children (PL)'

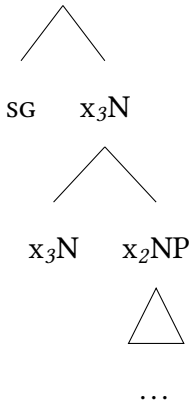
Note that in all cases, the selection of the plural suffix is determined by the root (as is the selection of number allomorphs more generally). Coupled with the size-based approach to allomorphy selection commonly used in Nanosyntax (and as sketched above already), that suggests that the lexicalized tree structure of the root must encode this information more generally: Lexical items for roots must vary in the structure they lexicalize in a way that can determine the anchor (bot-

tommost element) of the tree in which the number heads are spelled out.

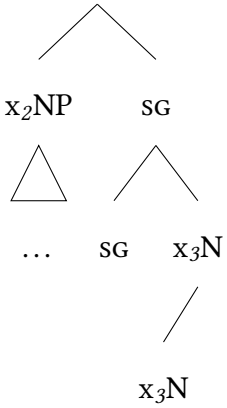
To implement such variation between roots, we need to turn to the material below number, i.e., we need to decompose (the higher end of) xNP into component features. Since I do not have enough data available to argue for particular features, I will agnostically label them x_nN for our present purposes, assuming a basic fseq [x_3N[x_2N[x_1N[...]]]], i.e., the assumption is essentially that there is some structure below number that is common to all count nouns.[11] Since all three roots are plural marking roots, we continue to encode this behavior by having them lexicalize SG, but not PL. However, the configuration in which SG is lexicalized differs: For *peet* 'day' we assume essentially the same structure as before (30a). For *oosn* 'forest', and *laak* 'child', we introduce a slight modification: On the one hand, *oosn* 'forest' lexicalizes the singular and the topmost part of the xNP as a right branch {SG,{x_3N}}, as in (30b). On the other, *laak* 'child' lexicalizes a slightly larger part of the xNP and the singular as a right branch {SG,{x_3N,{x_2N}}}, as in (30c).

[11] It is worth noting that these features are necessary to the analysis only insofar as we do not allow for the option of contextual allomorphy in the way that is commonly assumed in DM, namely as extra information in vocabulary items that is not itself the target of spellout. However, below I will briefly argue that for at least one such feature, HUMAN, there is good reason to believe that it exists and behaves like these other features. While these assumptions obviously call for further research into the nature of these features, the only claim that is necessary for my argument to work is that there is some syntactic structure below the number structure that does not receive independent spellout. I believe that, despite the fact that I put these assumptions to perhaps unusual use here, they are actually uncontroversial assumptions for theories such as DM or Nanosyntax, where such heads/features abound (gender, categorizing heads, animacy related features, etc).
Note also that I am assuming at this point that the functional sequence is essentially identical across count nouns. This is an obvious simplification, but an innocent one for our purposes, insofar as differences in the functional sequence associated with different nouns would only make it easier to account for differences in associated number morphology.

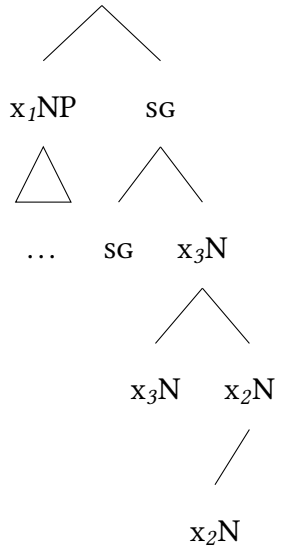
(30) a. *peet* ⇔ SG



b. *oosn* ⇔ SG



c. laak ⇔ SG

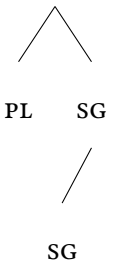


All such patterns result in an unmarked singular, which is unaffected by the configurational variation, since the root node containing all relevant features up to and including SG will be the target of spellout in the singular. That is to say, SG continues to be spelled out by the root, regardless of this configuration variation. They do, however, make a different right branch available for partial overwrite in a subsequent cycle, in this case after the merging of PL – the variation in "shape" induces allomorphy for number spellout.

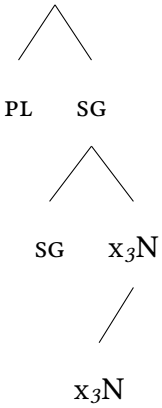
To see how we use these different right branches to account for the plural allomorphy, let us introduce the relevant plural suffixes – with the configurations for *-uus* and *-iin* repeated – in (31).

(31) a. PL ⇔ *-uus*

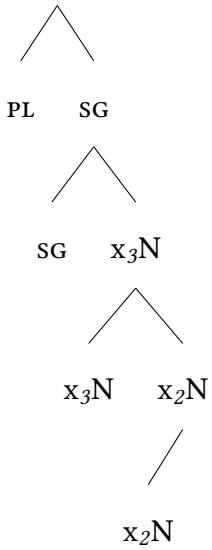
b. PL ⇔ *-iin*



c. PL ⇔ *-oos*



d. PL ⇔ -oy

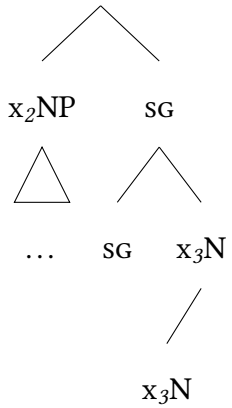


We have already accounted for the plurals *peet-uus* and *sigis-iin*, which remain unaffected by these new elements. The same principles provide us with the derivation of *oosn* 'forest', which combines with the plural suffix *-oos*. In (32a), we see the spellout of the singular structure by the root *oosn*, which lexicalizes this precise structure.[12] Upon merging PL, the resulting structure finds no match, and hence spec-to-spec movement is attempted, as in (32b). No item can spell out

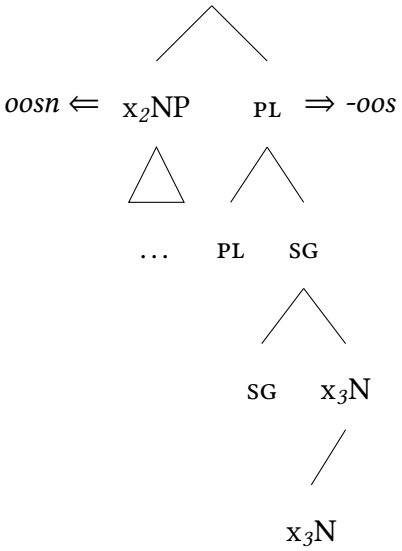
[12]Note in passing the double role of the plural suffix: It allows for the initial movement of the x_2NP interpreted by *oosn* (recall that there is no competition between concepts, and hence a non-suppletive root is selected low) up the tree, through the specifier of x_3N. It is only after the merger of SG, and accompanying spec-to-spec movement that the plural suffix disappears. Hence, lexical items of these shapes predict patterns of type A-x/A/A-x in *ABA domains, in which a suffix may disappear in an intermediate level.

the root node, but *oosn* continues to be a valid spellout of x_2NP, and the right branch {SG,{x_3N}} can be spelled out by *-oos*, partially overwriting *oosn*, and correctly deriving *oosn-oos*.

(32) a. *oosn* ⇐ SG

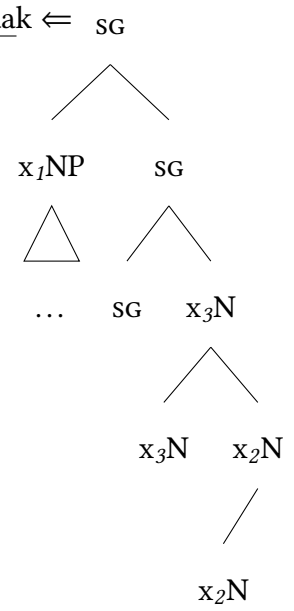


b. PL

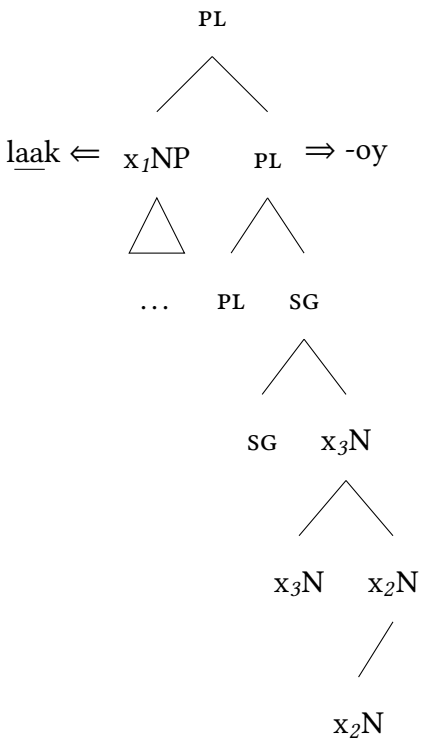


In parallel fashion, we derive *laak* 'child' combining with the plural suffix *-oy*, as in (33): The singular structure is spelled out by *laak*, in the configuration corresponding to the lexicalization pattern in (30c). The contrast to *oosn*, whose lexicalization pattern made an x_2NP available for spec-to-spec movement, comes about because *laak* instead lexicalizes x_1NP as a complex left branch. Consequently, after merging PL, and spec-to-spec movement, the right branch that is subject to spellout/partial overwrite has x_2N as its bottommost element. The right branch is therefore spelled out by the plural suffix *-oy*, which lexicalizes the corresponding structure.

(33) a.



b.



At this point, I have provided a root-driven approach to two aspects of nominal classification: The ifs (number marking class), and hows (number marking allomorphy) of number marking. Pending a more detailed analysis of the lower nominal structure in Kipsigis, I have made use of placeholder heads, x_iN. It is worth noting that depending on the details of the analysis of these nodes – in particular, if certain heads are not present across all nouns –, the approach makes the prediction that particular features in the region below number (say, ANIMATE, HUMAN, etc) are necessary but not sufficient conditions for certain number allomorphs: If such a head is lexicalized in the left branch of the root, it will not affect the selection of the number allomorph, since it is the right branch that is determining the relevant allomorphy (and hence, such an item may pattern with other items that lack the relevant semantics). If it is lexicalized in the right branch, however, spellout of number must also spell out this lower feature. While further study will be needed to ascertain the usefulness of such a prediction for the set of plural suffixes, it does turn out to be a good prediction for the much simpler system of singulative suffixes.

According to Kouneli (2020: 9), "[t]here are two singulative suffixes in the language, *-iin* and *-yaan*, both followed by the thematic suffix *-ta*. The suffix *-yaan* is by far the most productive suffix in the language, while *-iin* is only used with a few human nouns."[13] Crucial for our purpose is the fact that *-iin* occurs only with human denoting nouns, yet does not occur with all of them. That is, HUMAN is *necessary but not sufficient* for the appearance of *-iin*, i.e., it shows exactly the behavior we expect, as shown in (34).

(34) a. ngeend-**yaan**-ta-it → ngéendyáat
bean-**SG**-TH-SEC
'bean (SG)'

b. puun-**yaan**-ta-it → pùunyáat
enemy-SG-TH-SEC
'enemy (SG)

c. sig-**iin**-ta-it → sìgìindét
parent-SG-TH-SEC
'parent (SG)'

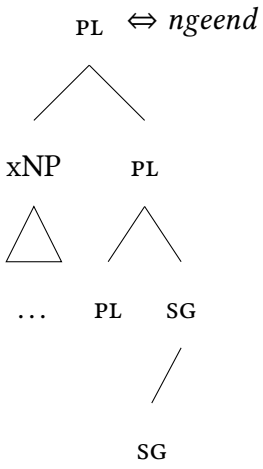
In extending the analysis of plural allomorphy to the singulative suffixes, we thus have to capture three cases: Nonhuman nouns combining with *-yaan* (34a), human nouns combining with *-yaan* (34b) and human nouns combining with *-iin* (34c), while excluding the possibility of *-iin* combining with non-human nouns. Once again, the process is root-specific, and hence we want to capture it in the lexicalization pattern.

We have already discussed the lexicalization pattern associated with *ngeend* 'bean', which is repeated below for the reader's convenience. The trees in (35a,b) provide the lexical items for *ngeend* and *-yaan* respectively. Upon merging of SG with the xNP, *ngeend* does not contain the resulting tree, hence comp-to-spec movement was attempted, which resulted in singulative

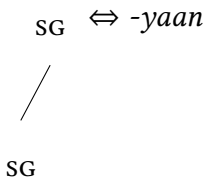
[13] I follow Kouneli (2020) in assuming that the singular *-iin* here is a different vocabulary item from the plural *-iin* introduced above. There is independent reason to believe that this is correct: Number morphemes always determine the thematic suffix that follows them, and individual thematic suffixes occur both in the singular and the plural, i.e., they are not sensitive to number, but to the identity of the number suffix (as will be discussed in more detail in the next section). Singular *-iin* occurs with the thematic suffix *-ta*, while plural *-iin* has a zero (or no) thematic suffix.

marking, since {SG} could be matched by *-yaan*, and *ngeend* continued to spell out the xNP.

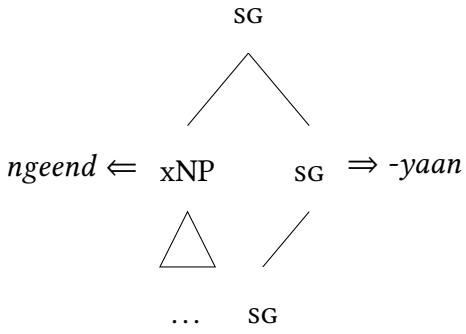
(35) a.



b.



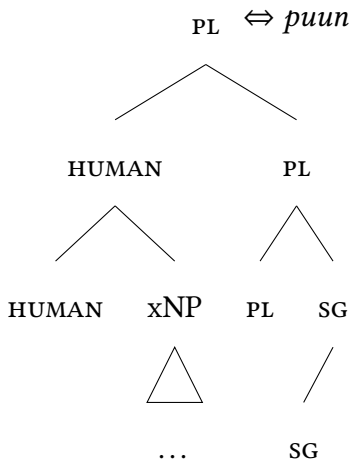
c.



What we want the system to capture, is the fact that *-iin* only combines with some human nouns, while the remaining human nouns pattern with non-human nouns with respect to singulative marking. That is, we want to capture the fact that the feature HUMAN is necessary, but not sufficient for the occurrence of *-iin*. The mechanism we have used for allomorph-selection was the splitting point of the functional sequence into a right branch and a left branch. Crucially for our purpose, only elements lexicalized in the right branch are available as targets for spellout by the number suffix. In this case, we can actually identify one of the heads labeled x_iN as HUMAN, and it shows exactly the expected behavior: On the one hand, *puun* 'enemy' type nouns lexicalize HUMAN as part of the left branch, thus making it irrelevant to the spellout of the right branch, as in (36).[14] Crucially, this makes their *right branch* identical to the one of *ngeend* type non-human nouns.

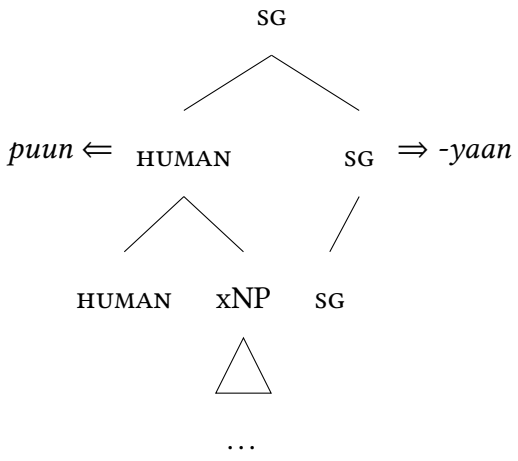
[14]As long as we are not concerned with the interaction between multiple such features, the particular place of HUMAN in the f-seq is not essential to the argument, provided that it be merged somewhere below the number structure, but for expository purposes it is most convenient to assume that is immediately below number.

(36)



Suppose, thus, that we have an xNP that was spelled out by *puun*. Upon merging HUMAN to create {HUMAN,xNP}, we find that *puun* continues to match this structure, and hence no movement is required. As with the singular-marking nouns before, however, merging SG with this structure requires comp-to-spec movement in order to be matched. As shown in (37), the right branch can be spelled out by *-yaan*, just as it was with non-human nouns. Because *puun* lexicalizes HUMAN in the left branch, it simply gets spelled out by the root even in the singular, and hence does not affect the spellout of number.

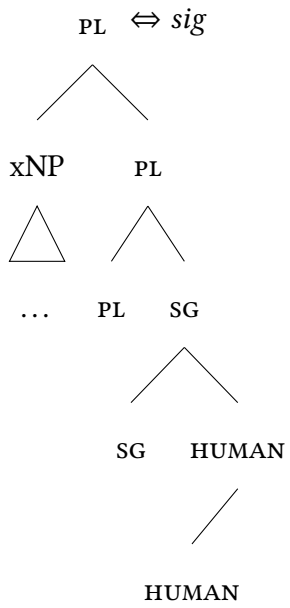
(37)



On the other hand, if a noun lexicalizes the feature in the right branch, as in (38), we get a

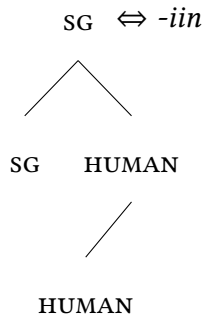
divergence from non-human nouns: The right branch now features HUMAN as the bottommost feature for any potential spellout of the right branch.

(38)



Hence, we need a singulative suffix capable of spelling out such a structure {SG,{HUMAN}}, as shown in (39).

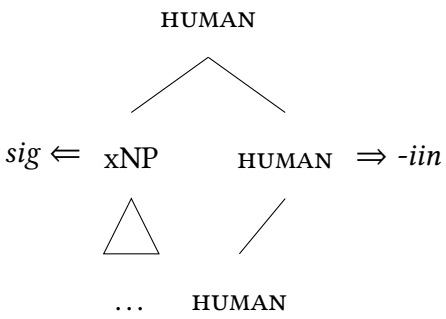
(39)



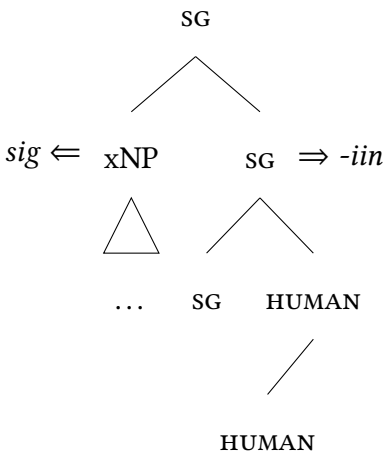
Let us assume, as before, that we have built an xNP spelled out by the relevant root, in this case *sig* 'parent'. Upon merging HUMAN, the divergence arises: Since {HUMAN,xNP} is not contained in the lexical item, and hence cannot be matched, we must attempt comp-to-spec movement (no spec is available for spec-to-spec movement) at this point in the derivation, as in (40a). The root

sig trivially continues to match the xNP, and the right branch can be spelled out by *-iin*. Next, we merge SG, which again fails to match at the root, hence spec-to-spec movement is attempted. As shown in (40b), this structure can be matched: Again, the left branch continues to be spelled out by *sig*, and the right branch can be spelled out by *-iin*, which self-overwrites.

(40) a.



b.



We have thus derived the morphological pattern of singulative allomorphy from the same mechanisms as number allomorphy more generally. I have shown that the lexical variation between different human-denoting nouns follows naturally from the ways in which lexicalization patterns are parameterized in Nanosyntax.[15] There is a singulative suffix that depends on the feature HU-

[15]Note in passing that this Nanosyntactic implementation of a parameter in Nanosyntax as a lexical property builds on the Borer conjecture (Borer, 1984) – occasionally called Borer-Chomsky conjecture – in assuming that lexical items are the locus of variation. However, in line with other arguments from the Minimalist Program, it provides us with a theory in which the *interface* items – essentially language-specific legibility conditions – give rise to variation, rather than (solely) a set of pre-syntactic objects (the heads, or the list of formatives). This variation may occur between languages, or between items, as in the current case.

MAN for its occurrence, and one that does not. A particular "root" may lexicalize the feature HUMAN, in the left branch or the right branch. Only in the latter case does the feature come to bear on the suffixal spellout, predicting the kind of distribution we find: The feature is a necessary, but not a sufficient condition for the occurrence of singulative *-iin*. Or, to put it the other way around: Only human nouns can create the conditions to the occurrence of singulative *-iin*, but not all of them do so.

To summarize: I have laid out a system that answers Kouneli's (2020) challenge: It can derive the number-based system of noun classification that we find in Kipsigis, and it does so by imposing legibility conditions at the PF-interface that determine how number is spelled out. It does so without binary features, and it does so without requiring classificatory features. I have then extended the same mechanisms to an account of number-allomorphy, showing not just that it can be implemented, but also that it makes interesting predictions regarding the role of particular features in determining allomorphy as necessary but not sufficient.

I now turn to the thematic suffixes. I will show that the locality conditions necessary to account for the way their allomorphy is driven follow immediately from the above system if we assume that their order reflects their structural height in the usual mirror principle way.

5.4 Thematic Suffixes

In addition to the number suffix, Kipsigis nouns generally also come with a *thematic* suffix and a *secondary* suffix, and they each interact with the number system. According to Kouneli (2020), the secondary suffixes are historically related to specificity markers, and she analyses them as D heads that agree for number. The form they take is *-it* in the singular and *-ik* in the plural (except with a small class of athematic roots where they take the forms *-ta* and *-ka* respectively when immediately following the root). Since the details for the secondary suffix appear to be

See also Caha (2021) for a reflection on the relation between Nanosyntax and the Borer conjecture.

fairly trivial morphologically, I will simply assume that Kouneli's analysis is correct. If this is true, a Nanosyntactic adaptation of the spellout of D is trivial, and we do not need to concern ourselves further with them.[16] The thematic suffixes, on the other hand, interact with the system of number marking in more intricate ways. In this section I show that the particularities of this interaction follow immediately from the locality conditions arrived at in the account above.

As the repeated examples (41-43) – now with the thematic suffixes highlighted – show, the thematic suffix can take a variety of shapes. If there is a number suffix, the thematic suffix follows the number suffix, otherwise it immediately follows the root. Note also that the plural form in (43b) does not have a thematic suffix.

(41) a. peet-**u**-it → pêetúut
day-**TH**-SEC
'day (SG)'

b. peet-uus-**ya**-ik → pêetùusyék
day-PL-**TH**-SEC
'days (PL)'

(42) a. ngeend-yaan-**ta**-it →
bean-SG-**TH**-SEC
ngéendyáat
'bean (SG)'

b. ngeend-**a**-ik → ngéendéek
bean-**TH**-SEC
'beans (PL)'

(43) a. sigis-yaan-**ta**-it → sìgìsyáat
sock-SG-**TH**-SEC
'sock (SG)'

b. sigis-iin-ik → sìgìsìiník
bean-PL-SEC
'socks (PL)'

Fortunately, Kouneli (2020) has already done the hard work of extracting interesting generalizations about the thematic suffix. I quote:

(44) "[T]hematic suffixes in Kipsigis are placed after the singulative or plural suffix if they are present. In this case, the thematic suffix of the noun in its unmarked form is absent;

[16]The fact that we get a special form only when athematic roots do not take a number suffix either, is easily interpreted as roots spelling out all of the thematic domain, and then – and only then – being local enough to the higher D domain, as will be clear by the end of this section. In that sense, this would iterate the argument to be made here, but insofar as there are no other triggers of secondary suffix allomorphy, nothing of much interest beyond could be said.

only one thematic suffix per noun is overt at any time. Moreover, the thematic suffix of a noun in its unmarked form is dependent on the root (or nominalizing suffix in the case of nominalizations), but the form of the thematic suffix present after number suffixes is predictable by the number suffix, not the root" (Kouneli 2020: 25)

That is to say, the form of the suffix may be dependent on the particular *number suffix*, but is itself not immediately determined by *number*. In the above examples we have seen particulars noun each combining with different thematic suffixes in the singular and the plural. In (45-47), we see that a particular thematic suffix can occur both in the singular and the plural, in this case in the singular of an plural-marking singular noun, and the plural of an singular-marking noun. That is, the thematic suffix in the unmarked case is determined by the root, and not by number, or by the number class of the root.

(45) a. or-**a**-it → óorêet
road-**TH**-SEC.SG
'road/clan (SG)'

b. saram-**a**-ik → sárámêek
twin-**TH**-SEC.PL
'twins [animals] (PL) '

(46) a. kar-**i**-it → kàríit
car-**TH**-SEC.SG
'car (SG)'

b. sig-**i**-ik → sìgíik
parent-**TH**-SEC.PL
'parents (PL)'

(47) a. peet-**u**-it → pêetúut
day-**TH**-SEC.sg
'day (SG)'

b. sugar-**u**-ik → sùgàrúuk
sugar-**TH**-SEC.PL
'sugar (PL)'

Essentially, the form the thematic suffix takes is determined in one of two ways: In case the root is unmarked for number, the root directly determines the thematic suffix, as we just saw in (45-47). If there is a number suffix, however, the root cannot determine the form of the thematic suffix. Instead, it is now the number suffix that determines the form of the thematic suffix. That is to say, according to Kouneli, both roots and number suffixes are able to determine a thematic suffix.

For instance, the singulative suffix *-yaan* always combines with the thematic suffix *-ta* (48), and the plural suffix *-uus* always combines with the thematic suffix *-ya* (49).

(48) a. ngeend-**yaan-ta**-it
bean-SG-TH-SEC
→ ngéendyáat

'bean (SG)'

b. puun-**yaan-ta**-it → pùunyáat
enemy-SG-TH-SEC
'enemy (SG)

c. peel-**yaan-ta**-it → pèelyáat
elephant-SG-TH-SEC
'elephant (SG)

d. sigis-**yaan-ta**-it → sùgàryáat
sock-SG-TH-SEC
'sock (SG)

(49) a. peet-**uus-ya**-ik → pêetùusyék
day-PL-TH-SEC
'days (PL)'

b. pug-**uus-ya**-ik→ sìgìsyáat
book-PL-TH-SEC
'books (PL)'

Both roots and suffixes are potential determinants of the form of the thematic suffix, and in fact, they may may even trigger the same form„ as we see in the comparison between (49) and (50): Both the plural suffix *-uus* (49) and the root *tariit* 'bird' (50) require the form of the thematic suffix to be *-ya*.

(50) tariit-**ya**-it → tàrìityét
bird-**TH**-SEC.SG

'bird (SG)'

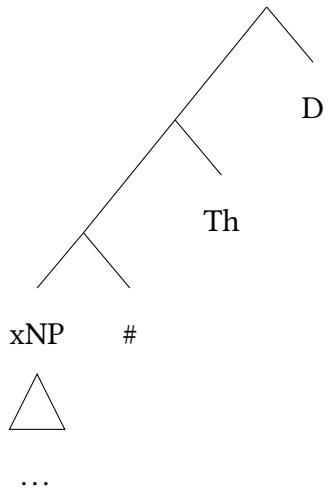
In summary: The allomorphy of the thematic suffix is independent of number, and of a root's number class. The element that determines the particular thematic allomorph is either the root (in case number is unmarked) or the number allomorph (in case number is marked): On the descriptive level, the statement describing the trigger of thematic allomorphy is disjunctive.

5.4.1 Determining the Thematic Suffix

I will now show that these facts follow quite naturally from the analysis laid out above. While they call for further study, we are concerned here with the number system and its interaction with the thematic suffixes. As such I am content here to merely sketch an account of the principles that lead to the alternation between a root determined thematic suffix vs. a number-morphology determined one. Beyond that, I will leave a more detailed analysis of the thematic suffixes to future work.

Let us assume, in the absence of evidence to the contrary, that the linear order, *root-(#)-*TH*-*SEC, arises in accordance with the mirror principle (Baker, 1985; Muysken, 1981). Then, the linear order of the Kipsigis nominal suffixes points towards a structure along the very rough lines of (51) (where D, Th, #, xNP are regions of the noun phrase that each are composed feature by feature as above).

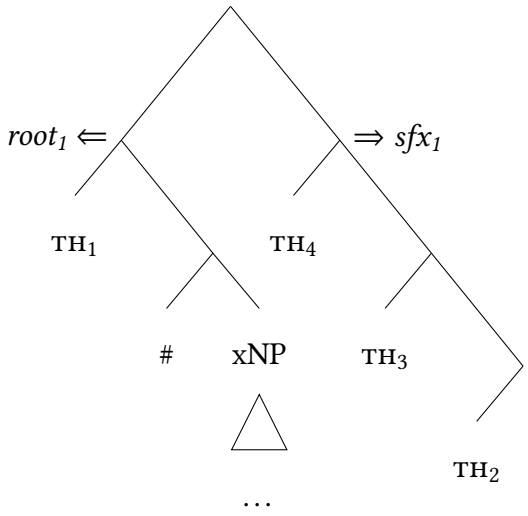
(51)



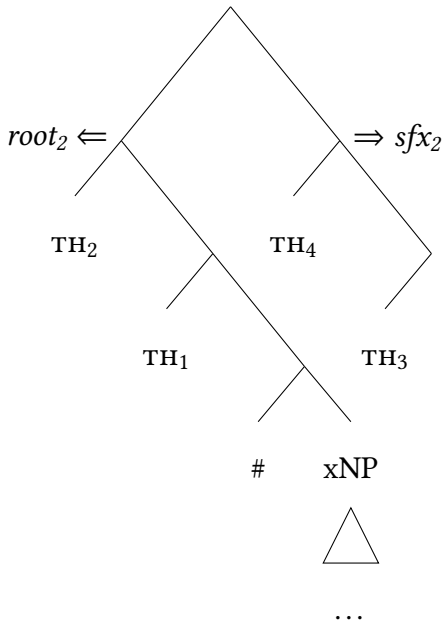
Let us further assume that the thematic region is internally complex, consisting, say, of some heads TH_1, TH_2, TH_3, TH_4. Again, as was the case above, these heads are labeled agnostically, but the content of this assumption boils down to saying that, insofar as there are no correlations

between certain semantic/syntactic features and the thematic affix, we assume that they are the spellout of some structure in the extended projection of the noun that is common to all count nouns, even if we do not know (or currently care about) what exactly that structure is. Allomorphy for the thematic domain is determined in a fashion parallel to what we have seen above for the allomorphy of number: Some previous cycle determines the bottommost feature that the affix spells out, thus determining the allomorph for the thematic domain. For instance, two roots might differ in size as in (52). While root_1 in (52-a) lexicalizes a TH_1P, the slightly larger root_2 in (52-b) lexicalizes a TH_2P. Consequently, root_1 combines with a thematic suffix spelling out a constituent $\{\text{TH}_4,\{\text{TH}_3\{\text{TH}_2\}\}\}$, while root_2 combines with a thematic suffix spelling out a constituent $\{\text{TH}_4,\{\text{TH}_3\}\}$.

(52) a.



b.



There are two basic explananda that we want to follow from our theory. First, we want to capture the fact that there are two basic pathways for determining the thematic allomorph: In pathway (i), number is unmarked, and the root itself determines the thematic allomorph. In pathway (ii), number is marked. In this case, the root determines the number allomorph, and the number allomorph in turn determines the thematic allomorph. That is, there is a disjunction in the description of the determination of the allomorph – it is either determined by the identity of the root (on the first pathway), or the identity of the number allomorph (on the second pathway). We would like our theory to instead provide us with a non-disjunctive characterization.

Second, we want to capture the flipside of the fact that it is the identity of the root or number suffix that determines the thematic suffix – the fact that it is independent of properties like the root's number class or number (though these obviously have a role in determining whether there is a number suffix at all, i.e., they have indirect effects on the determination of the thematic allomorph, but no direct ones).

I will first show how the two pathways arise as immediate consequences of the implementation of number spellout advanced above. After that, I will turn to discussing the interplay of a

noun's number class, a noun's number allomorph, and a noun's thematic suffix in the unmarked case. This discussion will focus on the way they are empirically independent, and how the degrees of freedom that a lexical item is afforded under the present theory can implement this by varying three different factors: The features that are lexicalized, the "breaking point" between the left and right branch, and the height of the left branch in the lexical item.

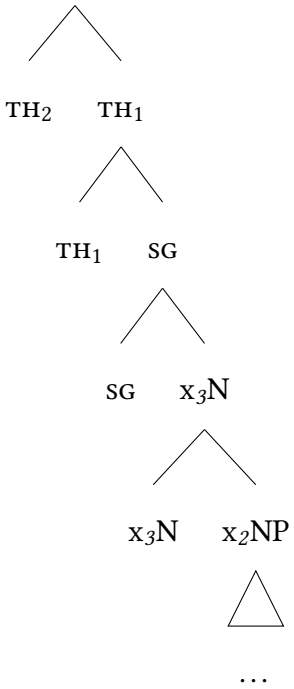
Let us consider first the case of a plural marking root, *peet* 'day', repeated in (53). In the unmarked singular, *peet* triggers the thematic allomorph *-u* (53a). In the marked plural in (53b), however, the root's ability to determine the thematic allomorph directly, is bled. Instead, the root determines the plural allomorph *-uus* (as we saw above), which in turn triggers *-ya* as the form of the thematic suffix.

(53) a. peet-**u**-it → pêetúut
day-**TH**-SEC
'day (SG)'

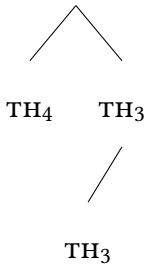
b. peet-uus-**ya**-ik → pêetùusyék
day-PL-**TH**-SEC
'days (PL)'

We have already seen how the lexical items for nominal roots encode number marking class (by varying how much of the number structure the root lexicalizes, and in what configuration it does so), and the number allomorph they combine with in the marked case (by additionally varying the breaking point of the f-seq into a left and a right branch). We now extend this system further by varying how much of the thematic domain a root lexicalizes, i.e., we revise our entry for *peet* 'day' as in (54a), and postulate a corresponding thematic suffix *-u*, as in (54b), such that *peet* and *-u* split up the thematic domain.

(54) a. *peet* ⇔ TH$_2$

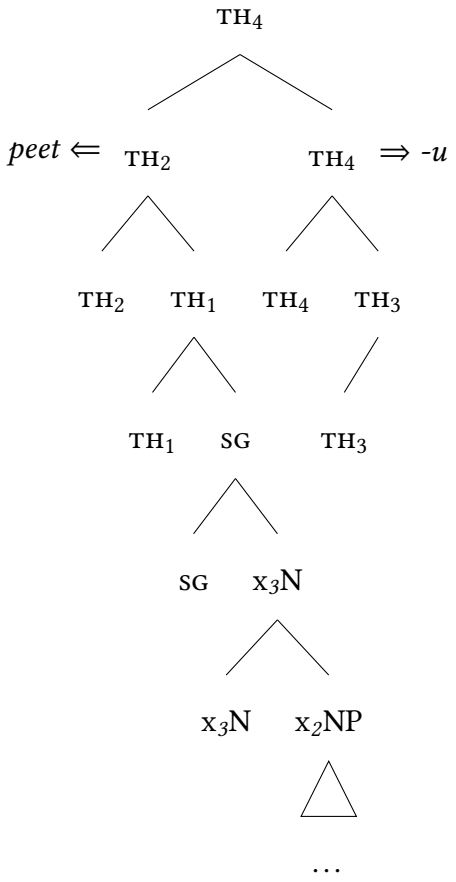


b. TH$_4$ ⇔ *-u*



In the singular, this leads to the derivation in (55): We continue simply merging the relevant features from the nominal domain, the number domain, and the thematic domain, and all the way up to TH$_2$, *peet* is simply able to spell out the resulting structure at the root node, i.e., it continually self-overwrites. Upon merging TH$_3$, however, *peet* is no longer a candidate for the whole structure, and since no specifier is available, comp-to-spec movement is attempted. The lexical item *-u* can spell out the resulting right branch {TH$_3$}. Merging TH$_4$ does not result in a PF-legible structure, and hence spec-to-spec movement is attempted, i.e., the left branch TH$_2$P is moved to the specifier of TH$_4$. The left branch has not changed, and continues to be spelled out by *peet*, and the right branch can now be spelled out by *-u*, which self-overwrites.

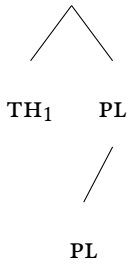
(55)



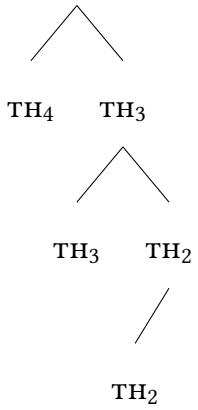
Note that it is *peet* that determines the foot of the thematic affix, and hence the thematic allomorph. It does so simply by spelling out TH_2 and TH_1, i.e., the information is encoded simply via the tree it lexicalizes. Crucially, however, its ability to determine the thematic allomorph is contingent on the right containment relations: It can only spell out a structure in which TH_1 is merged directly on top of SG.

Before we can illustrate how this inability to determine the thematic suffix in the plural will arise, we will need to introduce a slightly revised lexical entry for *-uus*, and a corresponding thematic suffix *-ya*, as in (56).

(56) a. $\text{TH}_1 \Leftrightarrow$ *-uus*



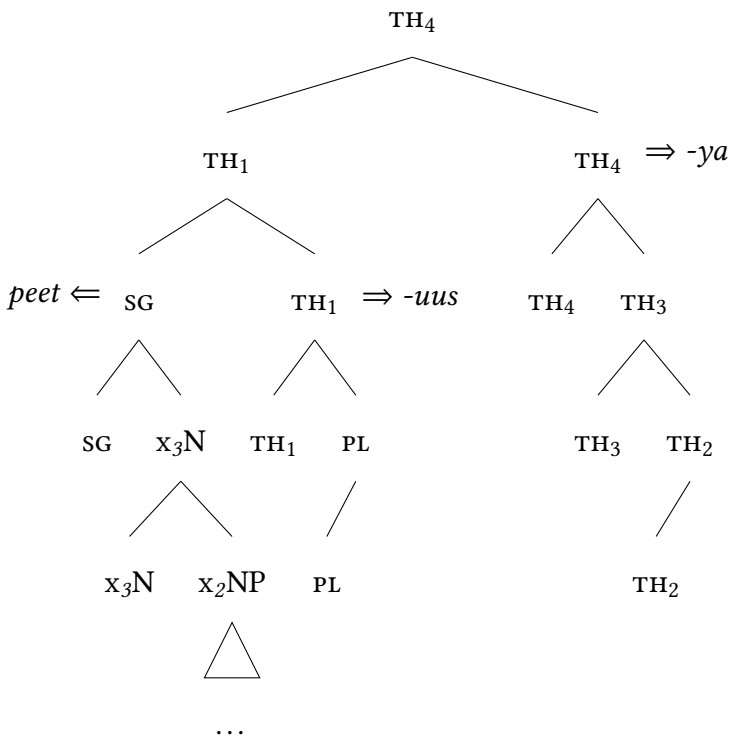
b. $\text{TH}_4 \Leftrightarrow$ *-ya*



Recall that it was the structure of *peet* that determined the bottom of the right branch in the case of number spellout. In this case, since *peet* does not split into a left and right branch, its associated plural spellout was *-uus*, whose foot is PL. We simply retain the mechanisms that we introduced above, but add to *-uus* the same ability to spell out part of the thematic domain, and thus trigger a subsequent thematic allomorph. The derivation of the plural form *peet-uus-ya* thus proceeds as in (57). The first crucial difference with the singular derivation in (55) occurs when we merge PL. Since *peet* is a marked plural noun – i.e., it does not lexicalize PL –, such a tree cannot be spelled out at the root node. Consequently, comp-to-spec movement moves the SGP into a left branch, and the right branch can now be spelled out by a suffix. The bottom-most element of the right branch is determined by the structure of *peet* – in this case, *peet* is not itself branching, and hence, the bottom-most feature of the right branch is PL. The consequence is that *-uus* is the relevant spellout for the right branch {PL}. Next, we merge TH_1. Again, simple spellout fails, so we attempt spec-to-spec movement, to create a right branch {TH_1,{PL}}, which continues to be matched by *-uus*. Crucially, merging PL has bled *peet*'s ability to spell out part of the thematic domain – the fact that *peet* lexicalizes TH_2 and TH_1 is irrelevant to the derivation, because there is no configuration that would include the thematic heads to the exclusion of PL, which *peet* does not lexicalize. Consequently, the identity of the root is irrelevant to the selection

of the thematic allomorph now, which is instead determined by the identity (and structure) of the plural suffix. This becomes evident once we merge TH_2: The set {TH_2, TH_1P} cannot be matched at the root node. Spec-to-spec movement does not result in a legible structure either, since *-uus* does not lexicalize TH_2 (and there is no other candidate for the branch). Consequently, we turn once again to comp-to-spec movement, creating a new suffixal position whose bottom-most feature is TH_2. That is to say the plural suffix *-uus* has determined (the anchor for) the thematic allomorph. Subsequent introduction of TH_3 and TH_4 each trigger spec-to-spec movement, and *-ya* ultimately spells out the set {TH_4,{TH_3,{TH_2}}}.

(57)



The same approach immediately extends to the other relevant case of a root-suffix alternation in the determination of the thematic allomorph, namely singular-marking roots (roots that mark both numbers never get to determine the thematic suffix directly). Let us return to the case of singular-marking *ngeend* 'bean', repeated in (58). In the singular (58a), *ngeend* combines with *-yaan*, which in turn determines the thematic allomorph *-ta*. In contrast, the plural is unmarked,

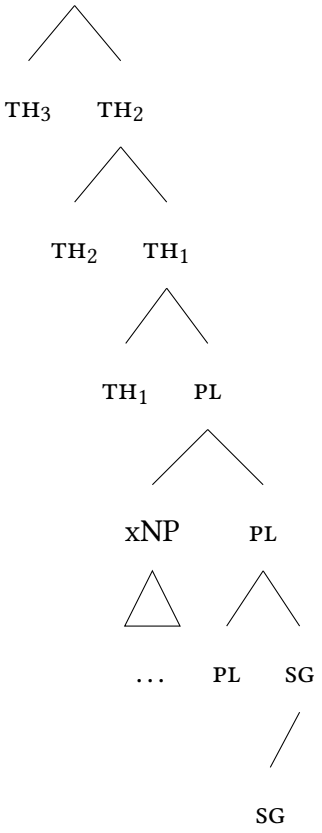
and the root determines the thematic allomorph *-a* directly (58b).

(58) a. ngeend-yaan-**ta**-it → ngéendyáat
bean-SG-**TH**-SEC
'bean (SG)'

b. ngeend-**a**-ik → ngéendéek
bean-**TH**-SEC
'beans (PL)'

Given the analysis of number marking and allomorphy that we introduced above, these facts follow in parallel fashion to the plural-marking case of *peet*: The root can only directly determine the thematic allomorph it is capable of spelling out parts of the thematic domain, and it is capable of doing so only in case it first spells out the number structure. Concretely, we implement this – as before – by revising the lexical entry for *ngeend*. We simply add thematic structure on top of our existing lexical entry, so that *ngeend* also lexicalizes part of the thematic domain (59a), and add a corresponding thematic allomorph *-a* to our inventory (59b).

(59) a. TH_3 ⇔ *ngeend*



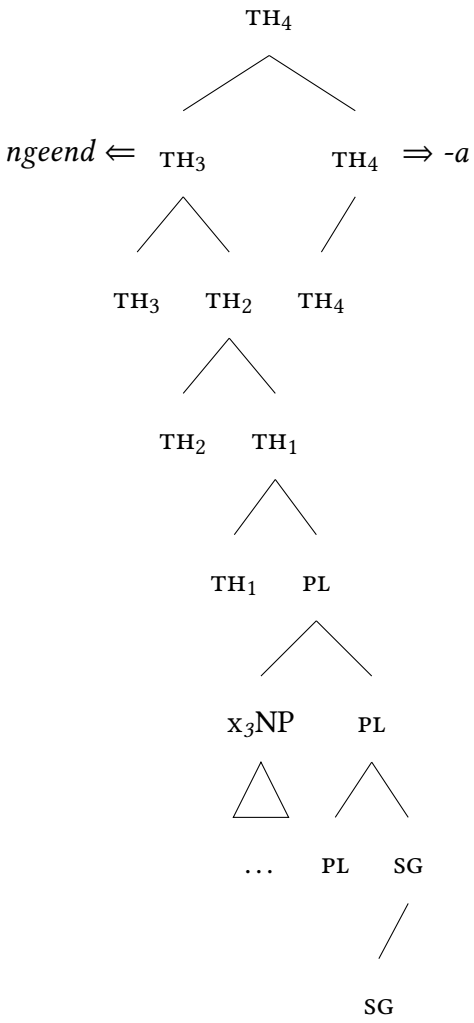
b. TH_4 ⇔ *-a*

In the plural, the derivation initially proceeds just like described above in section 5.3, example (25): Merging SG with x_3NP requires in comp-to-spec movement, where the right branch is spelled out by *-yaan*. Merging PL with this structure triggers spec-to-spec movement, which in turn allows *ngeend* to overwrite both itself as the spellout of the left branch, and *-yaan* as the spellout of the right branch, since it matches the resulting structure at the root node. Given the revised entry for *ngeend* in (59), subsequent merger of TH_1, TH_2, and TH_3 can all be matched directly by *ngeend*.[17] Only when we turn to TH_4 is *ngeend* no longer a candidate for spelling out the whole structure, and comp-to-spec movement creates a new right branch {TH_4} that can be spelled out by *-a*. That is to say, the size of *ngeend* determines the thematic allomorph in the plural. Note

[17]Note that lexical entries like this essentially trap the xNP, since it can no longer be extracted by spellout-driven movement after TH_1 is merged. This might open up interesting derivational options that I am not exploring here.

that the heads that *ngeend* lexicalizes in the thematic domain are independent of the fact that it is a singular-marking root; the number heads and configuration may vary independently of the thematic structure that a noun may lexicalize. I will return to this fact below.

(60)

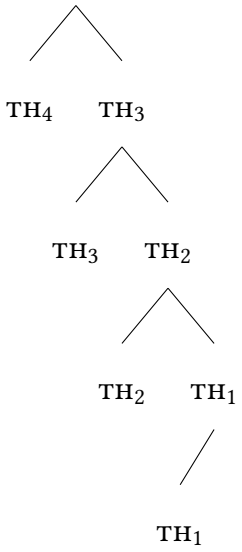


Turning to the singular, recall that *-yaan* always occurs with the thematic allomorph *-ta*. Let us assume that *-yaan* does not lexicalize any part of the thematic domain, and that *-ta* correspondingly lexicalizes all of it, as in (61).[18]

[18]It seems possible that *-ta* and *-ya* are in fact morphologically complex forms *-y-a* and *-t-a*, which contain the thematic suffix *-a*. This would easily be implemented by assuming that *-t* and *-y* only lexicalize the structure up to TH$_3$ (while otherwise being just like the postulated *-ta* and *-ya* respectively). Since *-a* lexicalizes {TH$_4$}, this would immediately derive the desired result. Such multimorphemicitiy in the thematic domain would, in fact, be quite

(61) a. SG ⇔ *-yaan*

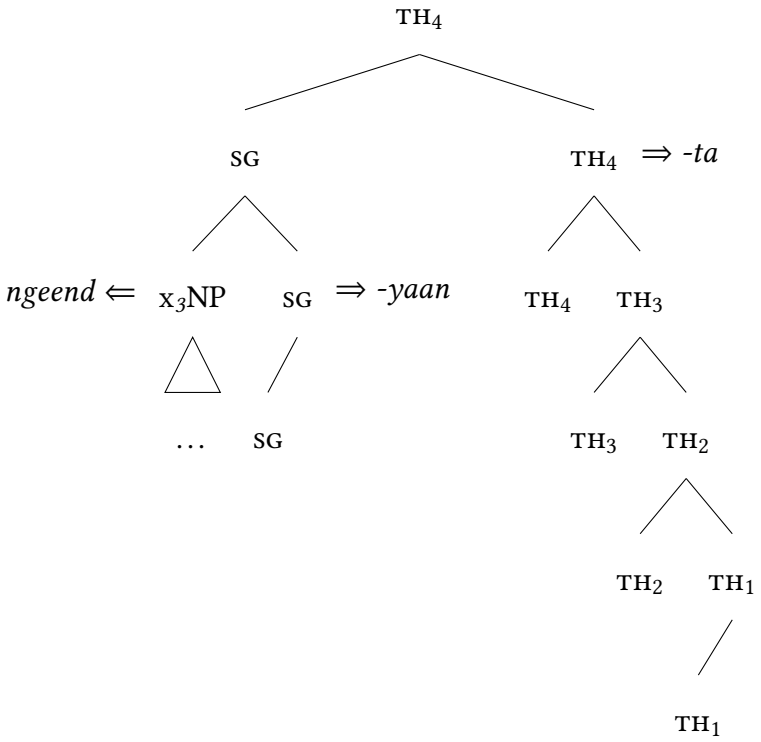
b. TH_4 ⇒ *-ta*



The initial part of the singular derivation continues to operate as introduced above, in example (24): Merging SG with x_3NP requires comp-to-spec movement, and the right branch of the resulting structure is spelled out by *-yaan*. Merging TH_1 with this structure does not result in a PF-legible structure, and spec-to-spec movement does not spell out successfully either, since *-yaan* does not lexicalize TH_1. Hence, comp-to-spec movement is attempted, and the new right branch $\{TH_1\}$ is spelled out by *-ta*. Subsequently merging TH_2, TH_3, and TH_4 all result in spec-to-spec movement, and *-ta* spells out the whole of the thematic domain. As was the case with *peet* 'day' above, it is only when the root spells out the exact number structure that it lexicalizes that it gets to determine the thematic allomorph. If any part of number is spelled out by an affix, however, it is the number affix that determines the thematic allomorph, bleeding a root's ability to do so. (Of course, the root itself selects the number allomorph, and thus it has an indirect effect on thematic selection nonetheless).

expected under the current perspective. I abstract away from such details, insofar as the aim of the current section is not to give a proper analysis of the inherent details of the thematic domain (whose internal complexity this section won't give justice to). Here, I merely aim to capture the interaction with the number system, show the way that the relevant information can be encoded in roots, and how the accessibility of this information follows directly from the theory of number spellout.

(62)



As an interim summary, we have accomplished the first task of accounting for the thematic allomorphy.[19] We have seen that assuming that the linear order of the affixes corresponds to their

[19]Having fully introduced the thematic domain, we now have the ability to tie up a small loose end from the analysis of number allomorphy with nouns that mark both numbers. According to Kouneli (2020: 53), nouns that mark both numbers, and take *-yaan* as their singulative suffix may differ in what plural suffixes they take. For instance, both *kew* 'shoe' and *sigis* 'sock' combine with *-yaan* in the singular, but the former takes the plural suffix *-oos*, while the latter takes *-iin*. However, in either case, the foot of the right branch is supposed to determine the number allomorph, regardless of whether it is a singular or plural allomorph. Given our analysis of the plural allomorph, and given the much larger number of those, vis-a-vis singular allomorphs, we could interpret this as *-yaan* being able to spell out right branches with different feet, for instance {SG} as well as {SG,{x_3N}}. The usual way to implement this is conjunction of subtrees via pointers (Caha and Pantcheva 2012, Blix 2021b, i.e., [SG→[x_3N]]. Such an affix could be anchored either at SG or x_3N, and hence it can be overwritten by different plural affixes. This, however, raises the issue of competition: In section 5.3, we analysed plural *-iin* as {PL,{SG}}, and for the purpose of spelling out only {SG}, it is now unclear whether *-yaan* is indeed the candidate with the least number of "unused" features. To ensure that *-yaan* remains the candidate selected by the elsewhere principle for the spellout of SG in the absence of PL, we are thus pushed to employ the thematic domain: If *-yaan* does not lexicalize any part of the thematic domain, but the plural suffixes do, and hence the plural suffixes have more unmarked features, we continue to derive the facts. In turn, we are forced to predict that *-ta*, which combines with *-yaan*, must be a lexical item for the whole thematic structure. There is, in fact, reason to believe that this is a good prediction:
In addition to the singular-marking nouns that combine with *-yaan* and *-iin*, there are in fact also a few nouns that are unmarked for number in both the singular and the plural. In the plural, they take a lexically determined thematic suffix only. In the singular, these take the thematic suffix *-ta*, but no singulative suffix (Kouneli 2020, fn. 13 & Table A.2). A natural hypothesis is that these are minimally nouns that lexicalize a structure [PL[SG[xNP]]], plus some

structural position along the lines of the mirror principle immediately provides us with a way of characterizing the trigger of thematic allomorphy: Whatever spells out number may also spell out parts of the thematic domain, and that way determine the thematic allomorph. The disjunctive character of the trigger has disappeared, and a simple statement, derived from basic locality arises.

5.4.2 On the independence of the thematic allomorph

We now turn to the second part: Accounting for the fact that thematic allomorphy selection is determined *only* by the statement above. That is, it not immediately determined by other factors (however relevant they may be in a mediated fashion, insofar as they are relevant to the determination of the spellout of number). These other factors include number, a root's number class, and the root/affix distinction. In the remainder of this section I will demonstrate these facts, and show how the kind of lexical entries I have proposed allow for this independence. As I will discuss in section 5.5, this is crucial vis-a-vis simpler nanosyntactic lexical entries and an account of singulatives in terms of *gapping*, which would fail to provide the necessary degrees of freedom.

In essence, this section will discuss three contrasts: First, two elements from the same number class and in the same number showing different thematic allomorphs, showing that the former two do not predict the latter. Second, two elements from different number classes, and in different number showing the same thematic allomorph, showing that the latter does not predict either of the former. Hence, the thematic allomorph is independent of both number and number class. And third, the case of a root and a number affix determining the same thematic allomorph, showing that the thematic allomorph cross-cuts the root/affix distinction.

part of the thematic domain. Since such nouns can spell out both singular and plural structures, they are unmarked for number in both cases, thus constituting a fourth class. However, regardless of any thematic structure they may lexicalize above PL, they will always be bled from spelling out any part of the thematic domain in the singular. Hence, we predict that in the singular they combine with the same thematic suffix as *-yaan* does. Kouneli (2020), fn 13, suggests that it is indeed be the case that these nouns all combine with the same thematic suffix *-ta* just as *-yaan* does, suggesting that the prediction is born out.

Consider first the independence of the thematic allomorph from number marking class. We have already seen *peet* 'day' and *oosn* 'forest', both of which mark only the plural. In section 5.3, I laid out how the two roots "select" different plural allomorphs, due to their difference in branching. As is highlighted in (63-64), they also trigger different thematic allomorphs in the unmarked singular: While *peet* combines with *-u*, *oosn* combines with *-a*.

(63) a. peet-**u**-it → pêetúut
day-**TH**-SEC
'day (SG)'

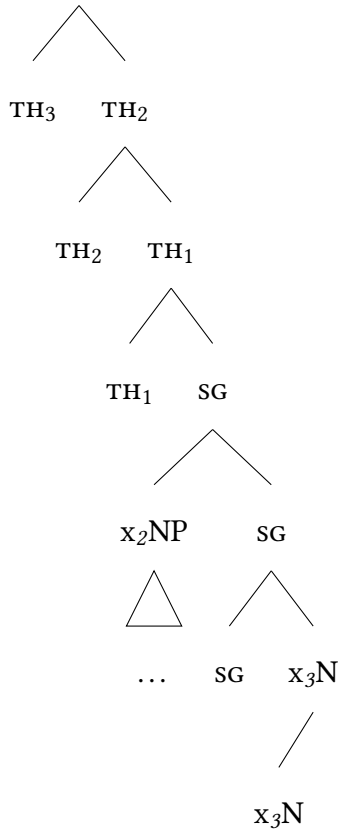
b. peet-uus-**ya**-ik → pêetùusyék
day-PL-**TH**-SEC
'days (PL)'

(64) a. oosn-**a**-it → òosnêet
forest-**TH**-SEC
'forest (SG)'

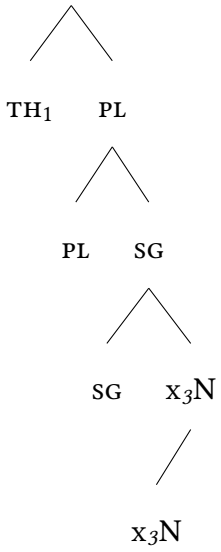
b. oosn-oos-**ya**-ik → òosnòosyék
forest-PL-**TH**-SEC
'forests (PL)'

We can immediately derive the behavior of *oosn* by revising the previous analysis of *oosn* and its associated plural allomorph *-oos* to the structures in (65). In the singular, *oosn* will spell out the whole structure until TH_3, with TH_4 being spelled out by *-a* (59b). In contrast, in the plural, it will fail to determine the thematic allomorph just as we saw with *peet* above. The plural allomorph *-oos* will determine the thematic allomorph in the plural – incidentally (unless one might be driven towards further decomposition), both *-uus* and *-oos* trigger the same thematic allomorph *-ya* (56b), i.e., they both spell out TH_1, but no other material in the thematic domain. Crucially, the fact that *peet* and *oosn* are of the same number-marking class is determined simply by them not lexicalizing PL, and their behavior with respect to thematic allomorph selection (by varying the size in the thematic domain), and their behavior with respect to the plural allomorph (by varying the breaking point of the f-seq below number) are independent of this fact.

(65) a. *oosn* ⇔ TH$_3$



b. TH$_1$ ⇔ *-oos*



We have seen two roots from the same number class triggering different thematic allomorphs, and we now turn to the fact that the opposite is also a possibility, i.e., two roots from different number marking classes may trigger the same thematic allomorph in the unmarked case.[20] We have, in fact, already introduced both the data and the analysis; compare the unmarked singular form *oosn-a-it* 'forest' from (66a) with the unmarked plural form *ngeend-a-ik* 'beans' in (66b):

(66) a. oosn-**a**-it → òosnêet
forest-**TH**-SEC
'forest (SG)'

b. ngeend-**a**-ik → ngéendéek
bean-**TH**-SEC
'beans (PL)'

This illustrates that thematic selection is a root-property that is independent of number-marking

[20] That being said, Kouneli (2020: 8) points out that the set of thematic suffixes that occur with singular-marking nouns is a subset of those that occur with plural-marking nouns. I do not currently have a principled explanation for this; it appears simply as a lexical gap.

class: Above, we saw that two roots from the same number-marking class may determine different thematic affixes in the unmarked case. The same data also illustrates that both a singular-marking and a plural-marking root may determine the same thematic affix. As is evident from comparing the lexical entry for *ngeend* (59) with the one for *oosn* (65a), this is due to the fact that the size of a lexical item with respect to the thematic domain is independent of whether or not a noun lexicalizes the kind of structure that results in marked singulars, or the kind that results in marked plurals: In either case, they can determine the thematic allomorph only in case they spell out the number structure, and nothing blocks them from determining the same thematic allomorph by lexicalizing identical parts of the thematic domain. Hence, thematic allomorphy selection is independent of number-marking class.

The comparison in (66) is also immediately relevant for showing that thematic allomorph selection is not dependent on number itself: Different numbers may co-occur with the same thematic allomorph. We have already seen that the flipside is also true: The same number may occur with different thematic affixes, e.g., *peet-**u**-it* 'day' (63a) vs *oosn-**a**-it* 'forest'(64a).

Finally, let me illustrate that both a root and an affix may in fact select the same thematic allomorph. Consider the comparison in (67) – both the plural suffix *-uus* and the plural-marking root *tariit* 'bird' select for the thematic suffix *-ya*.

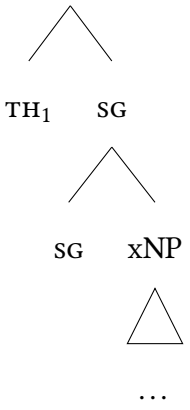
(67) a. peet-uus-**ya**-ik → pêetùusyék
day-PL-**TH**-SEC
'days (PL)'

b. tariit-**ya**-it → tàrìityét
bird-**TH**-SEC.SG
'bird (SG)'

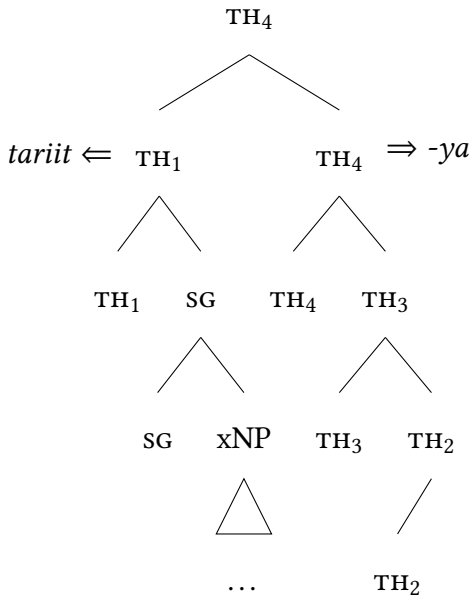
We have already seen how the plural suffix *-uus* determines the thematic suffix *-ya*: By not lexicalizing the relevant part of the thematic region that *-ya* spells out (in this case, all of it). The fact that a root triggers the same thematic suffix when it spells out number, can be modeled in the same way: When a root and an affix lexicalize the same parts of the thematic domain, and they actually get to spell them out, they will trigger the same thematic suffix. In this case, the root *tariit* 'bird' (68) and the plural suffix *-uus* trigger the same thematic suffix *-ya* because neither

lexicalizes any part of the thematic region above TH_1 (68a), (56a). Hence in both cases, TH_2, TH_3 and TH_4 are spelled out in a constituent, by *-ya* (68b), (57).

(68) a. *tariit* ⇔ TH_1



b.



The results of this subsection are summarized in Table 5.1, with trivial cases, such as same class, same thematic affix, or different class, different thematic affix omitted.[21]

Table 5.1: Independence of the Thematic Allomorph: Summary

Domain	S/D	Them.	Data	Analysis
Number Class	same	diff.	(63), (64)	(54-55), (65)
	diff.	same	(66)	(65), (59-60)
Number	same	diff.	(63a), (64a)	(54-55), (65)
	diff.	same	(66)	(65), (59-60)
Root/Affix	diff	same	(67)	(56-57), (68)

This concludes the discussion of the thematic affixes in relation to number. I have shown first

[21] One thing I have not investigated here, is the relation between the thematic suffix in the unmarked number case, and the number suffix in the marked number case. According to Kouneli's (2020) Tables 3 and A.1 there exists a nontrivial degree of correlation between these in plural-marking roots, but neither predicts the other perfectly: For instance, there are pairs of plural-marking roots that take the same thematic suffix in the singular, but different plural suffixes, and there are pairs of roots that take the same plural suffix, but different thematic suffixes in the unmarked singular.

that the analysis I proposed immediately provides the right locality condition to unify an otherwise disjunctive generalization over the trigger of thematic allomorphy. Secondly, I have laid out (following Kouneli 2020) that it is empirically necessary to account for three different facts: Roots must be related to a number class, to the allomorphy for marked number they trigger, and to the allomorphy for the thematic suffix in the unmarked case, and these must be independent. I have then shown that we can vary three properties of lexical items to account for these facts: We vary the features and configuration of number that a root lexicalizes to account for its number marking class. We vary the breaking point (if any) below number to account for number allomorphy, and we vary the size of the lexical item with respect to the thematic domain to account for thematic allomorphy in the unmarked case. Crucially, while the lexical items are able to determine steps of the derivation in a cascading manner (e.g., such that a root selects a number allomorph which in turn selects a thematic allomorph), there is no notion of morphological class that is featuralized in some pseudo-syntactic manner. Instead, the morphological behavior was determined by two things: First, an ordered set of possible ways to introduce a feature into a derivation. Secondly, a set of interpretative lexical items that themselves constitute *interface legibility conditions* that force the syntax to different configurations in which to introduce features, depending on an initial root choice. That is, the apparent "morphomic" character of declension class is not featuralized, but rather arises as a consequence of the necessity to build interface-legible syntactic objects.

5.5 Discussion

Above, I have provided an analysis of the Kipsigis system of nominal classification, and the root-specific effects of number marking (both whether it is marked, and if yes, how), and thematic allomorphy. I have shown that we can conceive of these morphological effects as arising as syntactic operations that ensure PF-interpretability, i.e., in reaction to interface legibility conditions imposed by the vocabulary. The effort was motivated conceptually: Narrowly, by Kouneli's

(2020) argument that binary features are necessary for the derivation of the Kipsigis system, and broadly by issues raised by the common idea of implementing morphological classes in terms of features. In addressing these conceptual points, we have made empirical progress as well: I have shown that a system where morphological containment is bi-directional – the plural form containing the singular form in one case, and vice versa in another – can be captured by lexical items with complex left branches. From this basic notion, I developed an account of the relevant allomorphy, and a novel way of deriving conditions on the occurrence of a vocabulary item that are necessary but not sufficient. I then showed that the theory leads to a natural account of the allomorphy in the thematic domain, and how the richly structured lexical items I have proposed, allow us to account for both the locality conditions of thematic allomorphy, and the independence of that allomorphy from other factors, such as a noun's number class, number, or the root/affix distinction.

At this point, I will offer a brief discussion of this analysis against the background of two alternatives: On the one hand, Kouneli's (2020) analysis of the system discussed above, and on the other a potential alternative account in Nanosyntax, namely one that would involve gapping. With respect to the former, I will argue that what I have done here captures the insights and intuitions that Kouneli (2020) has advanced, but that it overcomes some conceptual issues that arise from the subset principle she employed. With respect to the gapping analysis, I will argue that it is not rich enough to account for the observable data.

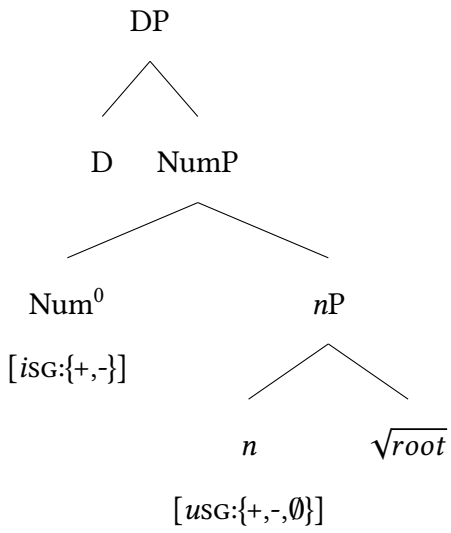
5.5.1 Kouneli's 2020 analysis

For the most part, this paper has simply adopted general insights from Kouneli's (2020) work. First, I take as the explanandum the empirical insights uncovered by Kouneli (2020): The paper lays out in great detail that count nouns all share the same syntactic behavior, regardless of the number-marking class they fall in, and in particular that singular-marking nouns do not behave like collectives. That is, the number-class that a root belongs to is an idiosyncratic morphological

property of a root, not a function of its meaning.[22]

Secondly, the current paper can be read as a re-interpretation of the analytical core of Kouneli's (2020) analysis of the morphological classes. The basic aspect of her analysis is a classificatory feature on the categorizing head *n*. Since the analysis aims to capture the fact that Kipsigis nominal classification is number-based, this uninterpretable classificatory feature is *u*SG, and it can take three values, plus, minus, or underspecified, as indicated in (69). Immediately above *n*, we find the interpretable number feature *i*SG, which differs from the interpretable one in lacking an underspecified option.[23]

(69)



The intuition behind this approach is that roots (through their associated *n* and its uninterpretable number feature) cause some particular number to be unmarked. Kouneli implements this through a kind of OCP rule: Whenever the number head and the categorizing head *n* have the same value, the number head gets obliterated at PF. Coupled with a stipulation that only the number feature

[22]It is worth noting that Kouneli (2020) also offers an account of mass nouns, arguing that the uninterpretable feature on *n* can be agreed with in mass nouns, where Num is absent. This is an intriguing idea. Seen from the current perspective, a proper response would require its own Nanosyntactic account, presumable related to a more general account of mass nouns, pluralia tanta, and general coercion phenomena. For now, I will have to leave this issues aside.

[23]That is to say, the system is modelling the number classes on the theory of gender advanced by Kramer (2015, 2016).

on Num^0 receive a PF interpretation, this derives the three classes: Roots that combine with the underspecified flavor of *n* mark both numbers, but roots that combine with the +SG flavor will only mark the plural, and roots that combine with the −SG flavor will only mark the singular (since +SG Num^0 and −SG Num^0 will be deleted, respectively).[24]

In many ways, this paper is not so much an alternative to Kouneli's (2020) analysis, but an attempt at an explanatory account of the tools she employs, particularly the notions of an uninterpretable feature (with the associated doubling of the number feature, one variant interpretable, one uninterpretable), and of obliteration. What I have proposed here, is that we do not need to create an uninterpretable shadow of the number feature, if we assume that *configurational interpretability* of the feature itself is at stake: It is the roots lexicalized tree that determines which features it can interpret, and in what configuration. This bypasses the necessity to first double the feature and then obliterate it, and it derives the notion of *uninterpretable* number features from general and independently motivated principles of PF-interpretation, namely the superset principle and phrasal spellout. Since a feature being uninterpretable by the root means that it must be moved into a configuration where it is interpretable by a suffix, the need for obliteration also disappears, since the question of whether it receives a suffixal interpretation is immediately reduced to the very same principles of phrasal spellout. That is, we arrive at an explanatory account of the notion of uninterpretable number features, all the while reducing the necessity for stipulated PF operations such as obliteration.

[24]Note that this system generates exactly three classes: Nouns that mark both numbers, nouns that mark only the singular, and nouns that mark only the plural. The Nanosyntactic account I developed here leads to the possibility of a fourth class, namely nouns that mark neither singular nor plural (i.e., for instance, because they simply lexicalize [PL[SG[xNP]]]), while Kouneli's (2020) account does not. At present, it is not clear to me if there is decisive evidence in either direction. On the one hand, the tripartite system seems to classify most Kipsigis nouns, and has a broad validity in the language family at large. On the other hand, there are at least a handful of nouns that do not mark singular or plural, such as (i) according to according to Kouneli (2020), fn. 13.

(i) a. pool-**ta**-it → pôoldét
cloud-TH-SEC.SG
'cloud (SG)'

b. pool-**i**-ik → pôolík
cloud-TH-SEC.PL
'cloud (SG)'

In doing so, we have also bypassed the issues raised by Alexiadou and Müller (2008) regarding the nature of declension class features, and their locus in the Y-model: No class-features are introduced at any point. Morphological classes arise as *interface legibility conditions*, i.e., precisely because different root VIs make the same features interpretable in different configurations. Declension classes are configurational effects arising in response to the general properties of VIs and spellout, not primitive features of the system.

5.5.2 Against a gapping analysis

In the Nanosyntactic literature, bi-directional morphological containment has previously been analyzed as the result of *gapping* (see e.g., Márkus 2015, Wyngaerd et al. 2020, Caha et al. 2021b). Crucially, gapping analyses operate without the need for lexical items with complex left branches. Here, I will argue that while such lexical items may allow for an account of the distribution of number marking, they are not sufficiently rich in structure to simultaneously offer an account of the form of number, i.e., the allomorphy discussed in section 5.3.1. The argument I will advance here, essentially boils down to this: A theory of allomorphy under which all allomorphy is due to differences in spellout targets (i.e., one in which there is no true non-phonologically conditioned allomorphy) cannot be a successful theory of spellout unless it includes lexical items with complex left branches.[25]

A gapping analysis of the basic three classes would look roughly like the following. First, assume that the functional sequence is minimally enriched by some head X, as in (70).

(70) xNP ≻ SG ≻ (PL) ≻ X ≻ TH

[25]Note that the argument here is essentially internal to Nanosyntax. I will assume that a gapping analysis can make reference to *backtracking*, an last resort operation that undoes previous operations in case no match can be found, and recursively attempts the next derivational option for the previous step (e.g., undoing spec to spec movement on the previous cycle and attempting comp to spec movement instead). Such an operation allows higher heads to change the configuration in which lower heads are spelled out. See Starke (2018) for details.

Second, assume that the three classes correspond to vocabulary items along the lines of (71).

(71) a. [xNP] ⇔ $\text{root}_{\text{sg-pl-marking}}$

b. [X [SG [xNP]]] ⇔ $\text{root}_{\text{pl-marking}}$

c. [X [PL [SG [xNP]]]] ⇔ $\text{root}_{\text{sg-marking}}$

In standard fashion, roots that mark both singular and plural do so because they lexicalize neither SG nor PL (71a), and roots that mark only the plural do so, because they lexicalize only SG, but not PL (71b). In each case, some affix has to spell out the features not lexicalized by the root VI. The *gapping* part comes in at (71c), the singular-marking roots: In the plural, such a VI can spell out the whole XP. However, in the singular, where the PL head is absent, the same VI can not spell out X, because [X[SG[xNP]]] is not a tree that is contained in the lexicalized tree. Note however, that such a root VI can still spell out the singular structure [SG[xNP]]. The head X, however, must be spelled out by an affix in the singular, even though it is spelled out by the root VI in the plural: The appearance of singular marking arises.

Such an analysis can account for the presence or absence of number marking. However, it runs into trouble when it comes to simultaneously determining the form that number marking takes, i.e., the allomorphy facts discussed in section 5.3.1. For the singulative marking, I will argue that it does not offer the ability to state that HUMAN is a necessary condition for the appearance of *-iin*, while not being sufficient. For the plural marking, I will argue that a gapping account does not provide sufficient degrees of freedom in the vocabulary item itself to encode the plural allomorph, forcing an account in terms of different functional sequences, even where there is independent reason to believe that such an account is unwarranted.

Let me elaborate in the singulative first: As we saw in section 5.3.1, there are two singulative suffixes, *-yaan* and *-iin* (72-73).

(72) a. puun-**yaan**-ta-it → pùunyáat
enemy-SG-TH-SEC
'enemy (SG)

b. puun-i-ik → pûuníik
enemy-TH-SEC
'enemies (PL)

(73) a. sig-**iin**-ta-it → sìgìindét
parent-SG-TH-SEC
'parent (SG)'

b. sig-i-ik → sìgíik
parent-TH-SEC
'parents (PL)'

While the *-yaan* has a kind of elsewhere distribution, the latter occurs only with human-denoting nouns. Crucially, however, it does not occur with all human-denoting nouns – human-denoting is a necessary condition for the occurrence of *-iin*, but not a sufficient one. I have shown that this is exactly the expected behavior if a root VI can lexicalize HUMAN in either the left branch or the right branch: Only in the latter case is it available for partial overwrite by the singulative suffix.

A gapping analysis without lexical items with complex left branches makes no such device available, since the lexical items cannot vary along a dimension such as the f-seq "breaking point", instead it would have to advance a theory where HUMAN is sometimes spelled out by the root VI, and sometimes by the singular affix. While such a system is a possibility in principle, it does not lend itself to an extension of the gapping analysis of singulatives. To see why, let us consider the possible place of our feature HUMAN in the f-seq (with the obvious constraint that it must be sufficiently local to X to be spelled out together with it). The first possibility is that it is above X, in which case we expect *all* human-denoting nouns that combine with a singulative suffix to take *-iin*: The premise of the gapping analysis is that X, not being matched by the root VI in the singular, is the actual target of the singulative suffix – that is, it necessarily separates the structure spelled out by the root VI from higher material, and thus the root VI cannot be responsible for the difference. If HUMAN is merged in between the number structure and X, the same logic applies: Since it is the absence of PL that bleeds the singulative-marking roots from spelling out material above number, such roots would always fail to spell out HUMAN in the singular, and consequently, we'd expect it to be relevant to the spellout of X in the singular across the board. This leaves us

with the option that HUMAN is merged in a place lower than the number structure. From the absence of plural marking, we know that the relevant nouns can spell out all features below PL, and hence, the roots would be able to spell out HUMAN in both the singular and the plural, and hence it should never be relevant to singular-marking. To sum up, the analysis I proposed above gives the root the ability to determine whether HUMAN is interpreted by an affix or the root, even if the root *principally* has the ability to spell out HUMAN on its own (as evidenced by the unmarked plural). In contrast, a gapping analysis without lexical items with complex left branches offers no such way to encode the information as a structural property of the root VI, and any consistent place for a feature HUMAN on the f-seq results in the false prediction that singular marking should be sensitive to HUMAN either in all cases or no cases.

A similar issue arises in the allomorphy of plural marking. As above, the theory I advanced allows the root to induce particular configurations of the features below PL through its f-seq "breaking point", and it is the root-induced configuration of these features that conditions the plural allomorphy. However, no such claim is possible in a gapping analysis, which cannot link such allomorphy to lower features, or a particular root VI. Consider the nouns that mark only the plural in Kipsigis: Since they spell out the singular without an affix, they must be able to spell out the singular structure, and in case of a gapping analysis, the hypothetical X (since not being able to spell out X is how singulative marking arises). Hence, all allomorphs of plural should, in principle, be anchored at PL (or, if there is backtracking, the highest head that any PL-lexicalizing suffix can be anchored at). Consequently, it cannot be the structure of the root VI that gives rise to the allomorphy – it must be heads above PL that do so, and these roots are blocked from spelling out any features merged on top of PL. In essence, then, we would have to postulate that every plural allomorph arises because there is a unique functional sequence that gives rise to that particular allomorph, i.e., there would be structural differences between nouns in the region above number. Not only does this strike me as a particularly bad version of an analysis that involves declension class features in the syntax, but it also questionable empirically, insofar as

there are minimal pairs where we have reason to believe that their thematic domain (the domain for which there is independent evidence suggesting that it is immediately above number) is in fact identical, despite them having different plural allomorphs. Consider the examples in (74), which show that both *oosn* 'forest' and *saa* 'buffalo' combine with the thematic suffix *-a* in the unmarked singular, suggesting that they share the same structure in the thematic domain.

(74) a. oosn-**a**-it → òosnêet
forest-TH-SEC
'forest (SG)'

b. saa-**a**-it → sáaêet
buffalo-TH-SEC
'buffalo (SG)'

Nonetheless, their plural forms in (75) show that they do take different plural allomorphs, so whatever variation there is cannot be due to differences in the thematic domain.[26]

(75) a. oosn-**oos**-ya-ik → òosnòosyék
forest-PL-TH-SEC
'forests (PL)'

b. saa-**iin**-ik → sàaèeník
buffalo-PL-SEC
'buffalo (PL)'

In summary, a gapping analysis does not appear to offer a clear account of number allomorphy in Kipsigis. When it comes to singulative suffixes, it does not offer a clear way of stating that HUMAN is necessary but not sufficient for the occurrence of singulative *-iin*. When it comes to the plural suffix, it is forced to account for the allomorphy in terms of features, and it has to push these features above the number structure. Insofar as it would be forced to postulate features for the sole purpose of accounting for allomorphy, it would appear to be subject to the initial conceptual criticism: It can only account for morphological issues by making otherwise unnecessary claims about syntactic heads.

In contrast, the analysis I proposed does not require particular features, it merely requires

[26]There is, of course, the possibility that there is a domain in between the thematic domain and the number structure that can vary independently of either, and that it's this domain that is different between the two nouns. However, we have seen above that the particular item that spells out the number structure determines the form of the thematic allomorph. It would seem to me that, besides being otherwise unmotivated, such an intermediate domain would possibly undermine the general account of the fact that a number allomorph determines the thematic allomorph, since there is now an independently varying domain in between those two regions.

that there be some structure associated with nouns that is lower than number and that can be spelled out by root VIs. Under such an analysis, the root VIs enforce particular configurations that give rise to particular number allomorphs, and the morphological classes were true interface properties.

5.6 Conclusion

In this paper, I have developed an analysis of the Kipsigis system of nominal classification that treats such classes as effects of interface legibility conditions that are imposed by the interpretative vocabulary at the PF interface (Halle & Marantz, 1993; Starke, 2009) that did not require binary features. I have adopted the general intuition of Kouneli's (2020) analysis that these classes are the effects of an uninterpretable number feature, but I have shifted the explanatory load: Rather than stipulating the feature as an uninterpretable "twin" of the interpretable number feature, I have derived the PF (un)interpretablity of the number feature from a general theory of PF-interpretation, Nanosyntax. Under such an approach, the (un)interpretability of number features is configurational: A number feature is uninterpretable when it cannot be matched at PF – since syntax needs to ensure that all features are interpretable, it will react by manipulating the syntactic structure to create PF-legible (interpretable) trees, in response to the matching requirements imposed by the vocabulary. In deriving these effects from general principles of spellout, I hope to have offered not so much an alternative to Kayne's (2020) analysis, but an explanatory account of the tools she makes use of: (Un)interpretability now appears as simply an effect of the superset approach to the PF-interpretation of syntactic objects, rather than a stipulated property of a feature.

In doing so, we have also (following Caha 2020) provided a solution to the conceptual issues raised by Alexiadou and Müller (2008): To model declension classes as syntactic or postsyntactic features is inconsistent with either basic properties of modularity (syntax should operate on

syntactic objects), or it gives rise to a generative morphology that can introduce features. Since we have not modelled declension classes as features, but rather as classes of vocabulary items that each impose particular legibility conditions, the issue does not arise: Declension classes correspond different syntactic configurations that arise in response to the general minimalist requirement that syntactic structures be legible at the interface.

Bibliography

Abels, K. (2011). Hierarchy-order relations in the Germanic verb cluster and in the noun phrase. *GAGL: Groninger Arbeiten zur germanistischen Linguistik*, (53.1), 1–28.

Adger, D., Harbour, D., & Watkins, L. J. (2009). *Mirrors and microparameters: Phrase structure beyond free word order*. Cambridge University Press
OCLC: ocn318871926.

Aissen, J. (1999). Markedness and Subject Choice in Optimality Theory. *Natural Language & Linguistic Theory*, *17*, 673–711. https://doi.org/10/dgc7n4

Aissen, J. (2003). Differential Object Marking: Iconicity vs. Economy. *Natural Language & Linguistic Theory*, *21*(3), 435–483. https://doi.org/10/fdfgcb

Alexiadou, A., & Müller, G. (2008). Class Features as Probes. In A. Bachrach & A. Nevins (Eds.), *Inflectional identity* (pp. 101–155). Oxford University Press.

Anderson, S. R. (1992). *A-Morphous Morphology* (Vol. 62). Cambridge University Press.

Aronson, H. I. (1990). *Georgian: A reading grammar* (Corr.). Slavica Publ.

Bahloul, M. (2008). *Structure and function of the Arabic verb*. Routledge
OCLC: ocm80020201.

Baker, M. (1985). The Mirror Principle and Morphosyntactic Explanation. *Linguistic Inquiry*, *16*(3), 373–415.

Baunaz, L., De Clercq, K., Haegeman, L. M. V., & Lander, E. (Eds.). (2018). *Exploring nanosyntax*. Oxford University Press.

Béjar, S. (2003). *Phi-syntax: A theory of agreement* (Thesis). University of Toronto.

Béjar, S., & Rezac, M. (2009). Cyclic Agree. *Linguistic Inquiry*, *40*(1), 35–73. https://doi.org/10.1162/ling.2009.40.1.35

Benmamoun, E. (1999). Arabic morphology: The central role of the imperfective. *Lingua*, *108*(2-3), 175–201. https://doi.org/10.1016/S0024-3841(98)00045-X

Berman, R. A. (1997). Modern Hebrew. In R. Hetzron (Ed.), *The Semitic Languages*. Routledge. https://doi.org/10.4324/9781315002682

Blansitt, E. (1988). Datives and allatives. In M. Hammond, E. A. Moravcsik, & J. Wirth (Eds.), *Studies in Syntactic Typology* (p. 173). John Benjamins Publishing Company. https://doi.org/10.1075/tsl.17.14bla

Blix, H. (2021a). Phrasal Spellout and Partial Overwrite: On an alternative to backtracking. *Glossa: a journal of general linguistics*, *6*(1), 62. https://doi.org/10/gbmv

Blix, H. (2021b). Spans in South Caucasian agreement: Revisiting the pieces of inflection. *Natural Language & Linguistic Theory*, *39*(1), 1–55. https://doi.org/10.1007/s11049-020-09475-x

Blix, H. (2022). Interface Legibility and Nominal Classification: A Nanosyntactic Account of Kipsigis Singulatives. *Glossa: a journal of general linguistics*, *7*(1). https://doi.org/10.16995/glossa.5825

Bobaljik, J. D. (2000). The ins and outs of contextual allomorphy. *University of Maryland working papers in linguistics*, *10*, 35–71.

Bobaljik, J. D. (2002). Syncretism without paradigms: Remarks on Williams 1981, 1994. In G. Booij, J. van Marle, S. Anderson, M. Aronoff, M. Baker, L. Bauer, R. Botha, J. Bybee, A. Carstairs-McCarthy, G. Corbett, W. Dressler, M. Haspelmath, J. Hoeksema, R. Lieber, P. Matthews, F. Rainer, S. Scalise, H. Schultink, & A. Spencer (Eds.), *Yearbook of Morphology 2001* (pp. 53–85). Springer Netherlands. https://doi.org/10.1007/978-94-017-3726-5_2

Bobaljik, J. D. (2012). *Universals in comparative morphology: Suppletion, superlatives, and the structure of words*. MIT Press.

Borer, H. (1984). *Parametric syntax, case studies in Semitic and Romance languages*. Foris Publications

OCLC: 12614352.

Brame, M. K. (1970). *Arabic Phonology: Implications for Phonological Theory and Historical Semitic* (Doctoral dissertation). Massachusetts Institute of Technology. Boston, MA.

Brody, M. (2000). Mirror Theory: Syntactic Representation in Perfect Syntax. *Linguistic Inquiry*, *31*(1), 29–56. https://doi.org/10.1162/002438900554280

Bruening, B. (2018). The lexicalist hypothesis: Both wrong and superfluous. *Language*, *94*(1), 1–42. https://doi.org/10.1353/lan.2018.0000

Cable, S. (2012). Pied-Piping: Introducing Two Recent Approaches. *Language and Linguistics Compass*, *6*(12), 816–832. https://doi.org/10.1111/lnc3.12001

Caha, P. (2009). *The Nanosyntax of Case* (Doctoral dissertation). University of Tromsø. Tromsø.

Caha, P. (2019). *Case competition in Nanosyntax. A study of numeral phrases in Ossetic and Russian.* lingbuzz/004875.

Caha, P. (2020). *Modeling declensions without declension features*.

Caha, P. (2021). *Minimalism and Nanosyntax: Reconciling Late Insertion and the Borer- Chomsky Conjecture*.

Caha, P., De Clercq, K., & Vanden Wyngaerd, G. (2019). The Fine Structure of the Comparative. *Studia Linguistica*, *73*(3), 470–521. https://doi.org/10.1111/stul.12107

Caha, P., De Clercq, K., & Vanden Wyngaerd, G. (2021a). *On the difference between a √ and a root.*

Caha, P., De Clercq, K., & Vanden Wyngaerd, G. (2021b). *Zero morphology and change-of-state verbs*.

Caha, P., & Pantcheva, M. (2012). Datives Cross-Linguistically.

Chomsky, N. (1995). *The minimalist program* (20th Anniversary Edition). The MIT Press.

Chomsky, N. (2001). Derivation by Phase. In M. J. Kenstowicz (Ed.), *Ken Hale: A Life in Language* (pp. 1–52). MIT Press.

Cinque, G. (2005). Deriving Greenberg's Universal 20 and Its Exceptions. *Linguistic Inquiry*, *36*(3), 315–332. https://doi.org/10.1162/0024389054396917

Collins, C. (2005). A smuggling approach to the passive in english. *Syntax*, *8*(2), 81–120. https://doi.org/10.1111/j.1467-9612.2005.00076.x

Collins, C., & Kayne, R. (2020). *Towards a Theory of Morphology as Syntax* (Manuscript).

Collins, C., & Stabler, E. (2016). A Formalization of Minimalist Syntax. *Syntax*, *19*(1), 43–78. https://doi.org/10.1111/synt.12117

Comrie, B. (1998). *Aspect: An introduction to the study of verbal aspect and related problems* (reprinted). Cambridge Univ. Press
OCLC: 247564008.

Corbett, G. G. (2000). *Number*. Cambridge University Press. https://doi.org/10.1017/cbo9781139164344

Cysouw, M. (2003). *The paradigmatic structure of person marking*. Oxford University Press.

De Clercq, K. (2020). *The Morphosyntax of Negative Markers: A Nanosyntactic Account*. De Gruyter Mouton. https://doi.org/10.1515/9781501513756

De Clercq, K., & Wyngaerd, G. V. (2019). On the idiomatic nature of unproductive morphology. *Linguistics in the Netherlands*, *36*, 99–114. https://doi.org/10.1075/avt.00026.cle

Demirok, Ö. (2013). *Agree as a Unidirectional Operation: Evidence from Laz* (Thesis). Istanbul.

Embick, D. (2015). *The morpheme: A theoretical introduction* (Vol. 31). De Gruyter Mouton.

Embick, D., & Noyer, R. (2007). Distributed Morphology and the Syntax-Morphology Interface. In G. Ramchand & C. Reiss (Eds.), *The Oxford handbook of linguistic interfaces* (pp. 289–324). Oxford University Press.

Erschler, D. (2022). A dedicated nominal singular morpheme without singulative semantics. *Morphology*. https://doi.org/10.1007/s11525-022-09392-2

Foley, S. (2017). Morphological conspiracies in Georgian and Optimal Vocabulary Insertion. In J. Kantarovich, T. Truong, & O. Xherija (Eds.), *Proceedings of the 52nd Chicago Linguistics Society Meeting*.

Halle, M. (2000). Distributed Morphology: Impoverishment and Fission. In J. Lecarme, J. Lowenstamm, & U. Shlonsky (Eds.), *Research in Afroasiatic grammar: Papers from the Third Conference on Afroasiatic Languages, Sophia Antipolis, France, 1996 / edited by Jacqueline Lecarme, Jean Lowenstamm, Ur Shlonsky* (pp. 125–150). John Benjamins.

Halle, M., & Marantz, A. (1993). Distributed Morphology and the Pieces of Inflection. In S. J. Keyser & K. Hale (Eds.), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger* (pp. 111–176). MIT Press.

Halle, M., & Marantz, A. (1994). Some Key Features of Distributed Morphology. *MIT Working Papers in Linguistics*, *21*, 275–288.

Harley, H. (2014). On the identity of roots. *Theoretical Linguistics*, *40*(3-4). https://doi.org/10.1515/tl-2014-0010

Harley, H., & Noyer, R. (1999). State-of-the-Article: Distributed Morphology. *GLOT*, *4*(4), 3–9.

Harley, H., & Ritter, E. (2002a). Person and Number in Pronouns: A Feature-Geometric Analysis. *Language*, *78*(3), 482–526. https://doi.org/10.1353/lan.2002.0158

Harley, H., & Ritter, E. (2002b). Structuring the Bundle: A Universal Morphosyntactic Feature Geometry. In H. J. Simon & H. Wiese (Eds.), *Pronouns: Grammar and representation*. J. Benjamins.

Harris, A. C. (1982). Georgian and the Unaccusative Hypothesis. *Language*, *58*(2), 290. https://doi.org/10.2307/414100

Heck, F. (2008). *On pied-piping: Wh-movement and beyond / by Fabian Heck* (Vol. 98). Mouton de Gruyter.

Heck, F. (2009). On Certain Properties of Pied-Piping. *Linguistic Inquiry*, *40*(1), 75–111. https://doi.org/10.1162/ling.2009.40.1.75

Heck, F., & Müller, G. (2007). Extremely local optimization. *Proceedings of WECOL*, *26*, 170–183.

Hewitt, B. G. (1995). *Georgian: A structural reference grammar*. John Benjamins publishing Co.

Hiraiwa, K. (2005). *Dimensions of symmetry in syntax: Agreement and clausal architecture* (Doctoral dissertation). Massachusetts Institute of Technology.

Kastner, I. (2018). Templatic morphology as an emergent property: Roots and functional heads in Hebrew. *Natural Language & Linguistic Theory*. https://doi.org/10.1007/s11049-018-9419-y

Kayne, R. S. (1994). *The antisymmetry of syntax* (Vol. 25). MIT Press.

Kayne, R. S. (2017). Antisymmetry and Morphology: Prefixes vs. Suffixes. *Wiener Linguistische Gazette, 82 (Festschrift für Martin Prinzhorn)*, 145–161.

Kayne, R. S. (2020). A note on the silent GO that underlies an instance of apparent suppletion in Spanish. *Isogloss. Open Journal of Romance Linguistics*, *6*, 1. https://doi.org/10/gjts3h

Keine, S. (2010). *Case and Agreement from Fringe to Core: A Minimalist Approach*. de Gruyter. https://doi.org/10.1515/9783110234404

Kiparsky, P. (1973). "Elsewhere" in Phonology. In M. Halle, S. R. Anderson, & P. Kiparsky (Eds.), *A Festschrift for Morris Halle* (pp. 93–106). Holt, Rinehart and Winston.

Kiparsky, P. (2008). Universals Constrain Change; Change Results in Typological Generalizations. In J. Good (Ed.), *Linguistic Universals and Language Change* (pp. 23–53). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199298495.003.0002

Kloudová, V. (2020). *The interaction of functional morphemes inside the nominal phrase*. Lincom GmbH.

Koopman, H. (2017a). A Note on Huave Morpheme Ordering: Local Dislocation or Generalized U20? In G. Sengupta, S. Sircar, M. G. Raman, & R. Balusu (Eds.), *Perspectives on the Architecture and Acquisition of Syntax*. Springer Singapore. https://doi.org/10.1007/978-981-10-4295-9

Koopman, H. (2017b). Topics on the interfaces of syntax with morphology and syntax with semantics.

Kouneli, M. (2019). *The syntax of number and modification: An investigation of the Kipsigis DP* (Doctoral dissertation). New York University. New York.

Kouneli, M. (2020). Number-based noun classification: The view from Kipsigis. *Natural Language & Linguistic Theory*. https://doi.org/10/gh63nn

Kramer, R. (2015). *The Morphosyntax of Gender*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199679935.001.0001

Kramer, R. (2016). The location of gender features in the syntax: The Location of Gender Features in the Syntax. *Language and Linguistics Compass*, *10*(11), 661–677. https://doi.org/10.1111/lnc3.12226

Kratzer, A. (2009). Making a Pronoun: Fake Indexicals as Windows into the Properties of Pronouns. *Linguistic Inquiry*, *40*(2), 187–237. https://doi.org/10.1162/ling.2009.40.2.187

Lacroix, R. (2009). *Description du dialecte laze d'Arhavi: Grammaire et textes* (Thesis).

Lacroix, R. (2011). Ditransitive Constructions in Laz. *Linguistic Discovery*, *9*(2). https://doi.org/10.1349/PS1.1537-0852.A.395

Lebeaux, D. (1991). Relative clauses, licensing, and the nature of the derivation. *Syntax and semantics*, *25*, 209–239.

Lomashvili, L., & Harley, H. (2011). Phases and templates in Georgian agreement. *Studia Linguistica*, *65*(3), 233–267. https://doi.org/10.1111/j.1467-9582.2011.01184.x

Marantz, A. (1996). *"Cat" as a phrasal idiom: Consequences of late insertion in Distributed Morphology* (Manuscript).

Marantz, A. (1997). No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. *University of Pennsylvania working papers in linguistics*, *4.2*, 201–225.

Márkus, A. (2015). *Taming the Hungarian (in)transitivity zoo: Undiagnosed species and a complete derivation of the morphosyntactic patterns.* (Doctoral dissertation). UiT The Arctic University of Norway. Tromsø.

McGinnis, M. (2008). Phi-Feature Competition in Morphology and Syntax. In D. Harbour, D. Adger, & S. Béjar (Eds.), *Phi theory: Phi-features across modules and interfaces* (pp. 155–184). Oxford University Press.

McGinnis, M. (2013). Agree and Fission in Georgian Plurals. In O. Matushansky & A. Marantz (Eds.), *Distributed Morphology Today* (pp. 39–57). The MIT Press. https://doi.org/10.7551/mitpress/9780262019675.003.0003

Middleton, H. J. (2020). **ABA syncretism patterns in pronominal morphology* (Doctoral dissertation). University College London. London, UK
draft Dec 2019.

Middleton, J. (2021). Developing the feature inventory of the inherent cases. *Glossa: a journal of general linguistics*, *6*(1), 68. https://doi.org/10/gj47fg

Müller, G. (2005). Syncretism and iconicity in Icelandic noun declensions: A Distributed Morphology approach. In G. Booij, J. van Marle, S. Anderson, M. Aronoff, M. Baker, L. Bauer, J. Bybee, A. Carstairs-McCarthy, G. Corbett, W. Dressler, M. Haspelmath, J. Hoeksema, R. Lieber, P. Matthews, F. Rainer, S. Scalise, H. Schultink, A. Spencer, & G. Stump (Eds.), *Yearbook of Morphology 2004* (pp. 229–271). Springer Netherlands.

Muysken, P. (1981). Quechua word structure. In F. Heny (Ed.), *Binding and filtering* (pp. 279–327). The MIT Press Cambridge, MA.

Nash, L. (1992). La catégorie AGR et l'accord en Géorgien. *Recherches Linguistiques*, *21*, 65–79.

Nevins, A. (2011). Multiple agree with clitics: Person complementarity vs. omnivorous number. *Natural Language & Linguistic Theory*, *29*(4), 939–971. https://doi.org/10.1007/s11049-011-9150-4

Noyer, R. (1992). *Features, Positions, and Affixes in Autonomous Morphological Structure* (Thesis). Massachusetts.

Öztürk, B., & Erguvanlı Taylan, E. (2017). Omnipresent little v in Pazar Laz. In R. D'Alessandro, I. Franco, & Á. J. Gallego (Eds.), *The verbal domain* (pp. 207–231). Oxford University Press.

Öztürk, B., & Pöchtrager, M. (Eds.). (2011). *Pazar Laz*. LINCOM Europa.

Pantcheva, M. (2011). *The Nanosyntax of directional expressions: Decomposing Path* (Thesis).

Pantcheva, M., & Caha, P. (2012). Contiguity beyond linearity: Modeling cross-dimensional syncretisms.

Preminger, O. (2011). Asymmetries between person and number in syntax: A commentary on Baker's SCOPA. *Natural Language & Linguistic Theory*, *29*(4), 917–937. https://doi.org/10.1007/s11049-011-9155-z

Ritter, E. (1993). Where's gender? *Linguistic Inquiry*, *24*(4), 795–803.

Ritter, E. (1997). Agreement in the Arabic prefix conjugation: Evidence for a non-linear approach to person, number and gender features. *Proceedings of the Canadian Linguistics Association*, 191–202.

Rosenthall, S. (2006). Glide Distribution in Classical Arabic Verb Stems. *Linguistic Inquiry*, *37*(3), 405–440. https://doi.org/10.1162/ling.2006.37.3.405

Ryding, K. C. (2005). *A Reference Grammar of Modern Standard Arabic*. Cambridge University Press. https://doi.org/10.1017/cbo9780511486975

Sportiche, D. (2005). *Division of Labor between Merge and Move: Strict Locality of Selection and Apparent Reconstruction Paradoxes*.

Sportiche, D. (2006). *NP Movement: How to Merge and Move in Tough-Constructions*.

Starke, M. (2009). Nanosyntax: A Short Primer to a New Approach to Language. *Nordlyd*, *36*(1), 1–6.

Starke, M. (2018). Complex Left Branches, Spellout, and Prefixes. In L. Baunaz, K. De Clercq, L. M. V. Haegeman, & E. Lander (Eds.), *Exploring nanosyntax*. Oxford University Press.

Stegovec, A. (2019). Taking case out of the Person-Case Constraint. *Natural Language & Linguistic Theory*. https://doi.org/10.1007/s11049-019-09443-0

Stepanov, A. (2001). Late Adjunction and Minimalist Phrase Structure. *Syntax*, *4*(2), 94–125. https://doi.org/10.1111/1467-9612.00038

Stump, G. T. (2001). *Inflectional morphology: A theory of paradigm structure*. Cambridge University Press.

Svenonius, P. (2016). Spans and words. In D. Siddiqi & H. Harley (Eds.), *Morphological Metatheory* (pp. 201–222). John Benjamins Publishing Company. https://doi.org/10.1075/la.229.07sve

Svenonius, P. (2018). Delimiting the Syntactic Word.

Taraldsen, K. T. (2018). Spanning versus Constituent Lexicalization: The Case of Portmanteau Prefixes. In L. Baunaz, K. De Clercq, L. M. V. Haegeman, & E. Lander (Eds.), *Exploring nanosyntax*. Oxford University Press.

Trommer, J. (1999). Morphology consuming Syntax' Resources: Generation and Parsing in a Minimalist Version of Distributed Morphology. In C. Retoré & E. Stabler (Eds.), *Procceedings of ESSLI Workshop on Resource Logic and Minimalist Grammars*.

van Urk, C., & Sun, Z. (2021). Dinka plural morphology is concatenative and regular.

Vanden Wyngaerd, G. (2018). The Feature Structure of Pronouns: A Probe into Multidimensional Paradigms.

Williams, E. (2003). *Representation theory*. MIT Press

OCLC: 248457959.

Wyngaerd, G. V., Clercq, K. D., & Caha, P. (2021). *Late Insertion and Root Suppletion*.

Wyngaerd, G. V., Starke, M., De Clercq, K., & Caha, P. (2020). How to be positive. *Glossa: a journal of general linguistics*, *5*(1), 23. https://doi.org/10.5334/gjgl.1114

Zwarts, J. (2001). Number in endo-marakwet. *Advances in Nilo-Saharan Linguistics: Proceedings of the 8th International Nilo-Saharan Linguistics Colloquium, University of Hamburg*, 191–206.