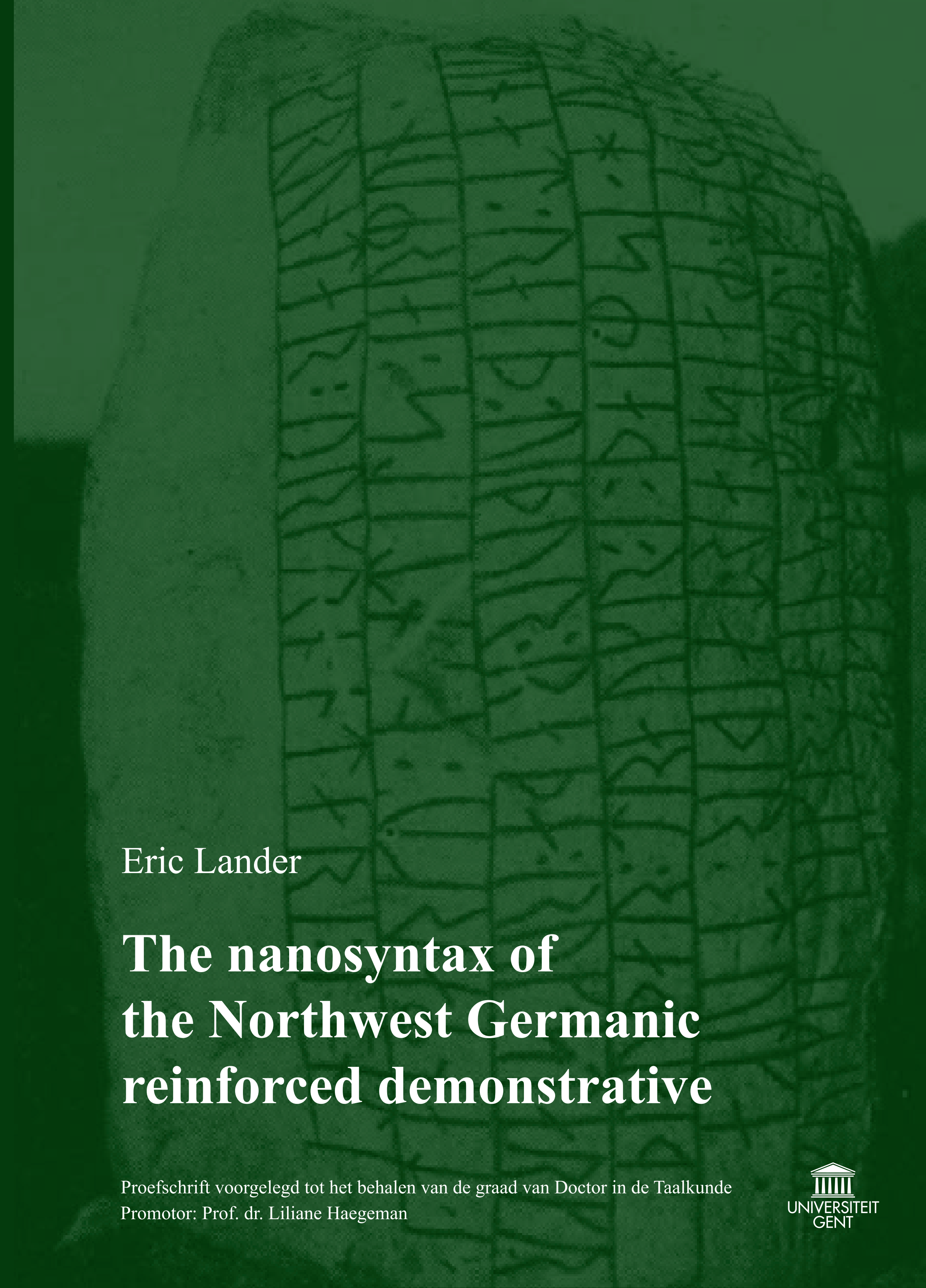


Eric Lander

The nanosyntax of the Northwest Germanic reinforced demonstrative

Proefschrift voorgelegd tot het behalen van de graad van Doctor in de Taalkunde
Promotor: Prof. dr. Liliane Haegeman

UNIVERSITEIT GENT

The nanosyntax of the Northwest Germanic reinforced demonstrative
EricLander

For Britt and Liv

Promotor	Prof. dr. Liliane Haegeman (Universiteit Gent)
Co-promotor	Prof. dr. Michal Starke (Universitetet i Tromsø)
Decaan	Prof. dr. Marc Boone
Rector	Prof. dr. Anne De Paepe

UNIVERSITEIT GENT

Faculteit Letteren & Wijsbegeerte

The Nanosyntax of the Northwest Germanic Reinforced Demonstrative

door

Eric T. Lander

Proefschrift voorgedragen tot het behalen van de graad van
Doctor in de Taalkunde

2015

This research is funded by BOF grant 01D30311
awarded by Ghent University.

GIST
Generative Initiatives in Syntactic Theory

Acknowledgements

First and foremost I would like to thank Liliane Haegeman, a most devoted adviser and steadfast supporter in all aspects of work and life. This dissertation has its beginnings in a suggestion Liliane made to me to investigate the Gothic discourse particle *sai*, which turned out to be an infinitely insightful suggestion. The task required years of work, and it led me down many productive paths. Many of these paths are documented in this dissertation. I cannot express my gratitude enough, Liliane, for everything you have done for me.

I was fortunate enough to participate in a nanosyntax weblab in November and December of 2012 organized by Michal Starke. This experience profoundly inspired the direction this dissertation took. Since then, I am indebted to Michal for countless Skype sessions and meetings in person about my work over the years. My generalizations would be far less neat and clear without Michal's perceptive help and encouragement.

I presented early work on *sai* and the Northwest Germanic demonstrative at the Incontro di Grammatica Generativa 39 (Modena) and the 25th Scandinavian Conference of Linguistics (Reykjavík). I am grateful to the participants for their helpful comments, especially Hilda Koopman, Gertjan Postma, and Fred Weerman at IGG 39, and Jóhanna Barðdal, Joe Salmons, and Ulla Stroh-Wollin at SCL 25. I am also very grateful for the opportunity afforded me by Anne Breitbarth to present my work on *sai* by guest-lecturing in her introductory Gothic class in December 2013 and November 2014. Thanks also to Anne for some very useful insights into the West Germanic data.

Hans Frede Nielsen, Sverre Stausland Johnsen, and Katrín Axelsdóttir were all very helpful when I came to them for help with some of the Old Norse/Icelandic data. Special thanks to Luc de Grauwe for data and literature advice at the beginning of my research quest.

Thanks to two anonymous OUP reviewers for helpful comments on an earlier version of this work in the spring of 2014.

Large parts of Chapters 3 and 4 appear in abridged form as Lander (forthcoming), in a special issue of *The Linguistic Review* for work presented at GLOW 37 in Brussels. I would like to thank Bronwyn Bjorklund, Anne Breitbarth, Jeroen van Craenenbroeck, Marcel Den Dikken, Thomas McFadden, Maria Polinksy, and Coppe van Urk for their questions and comments at GLOW. I would also like to thank two anonymous reviewers for their comments on my *TLR* paper.

I am grateful to members both past and present of GIST (Lobke Aelbrecht, Metin Bagriacik, Lena Baunaz, Liisa Buelens, Jaehoon Choi, Lieven Danckaert, Karen De Clercq, Tijs D'Hulster, Ciro Greco, Will Harwood, Dong-yi Lin, Rachel Nye, Trang Phan, Amélie Rocquet, Aleksandra Vercauteren, and Andrew Weir) for very helpful comments and feedback on various incarnations of this work (especially presentations given on April 2 and 12 of 2013 and March 26, 2014). Special thanks to Jaehoon Choi and Trang Phan for their extensive comments on a prefinal version of this work, and for many interesting discussions about nanosyntax.

To Lena, Ciro, and Andrew – thanks for the many wonderful nights out on the town.

To my parents Kerstin and Tobias I say *Tack för allt ni gjort för mig. Tack vare er har jag ett liv som är både underbart och avundsvärt*. To my brother Max I say thank you for being my best man, my daughter's godfather, and my best friend.

To Leo: you're the best buddy a guy could ask for.

Finally I will attempt to describe in words what my wife Britt means to me. She is a tireless source of love and support, a loving mother and partner, the engine of humor and love and care and all that is good in our life. This thesis is dedicated to her and our beautiful daughter Liv. I am so proud of my girls. I love you both dearly!

All inadequacies are my own.

September 9, 2015
Ghent, Belgium

Submitted: September 17, 2015
Defended: January 22, 2016
Jury: Johan Brandtler, Pavel Caha, Guglielmo Cinque, Øystein Vangsnes

[version 31.03.2016]

List of Abbreviations

#	*number*
1	*first person*
2	*second person*
3	*third person*
A	*(direct modification/descriptive) adjective*
Agr	*agreement*
ACC	*accusative*
ALL	*allative*
C	*complementizer*
COM	*comitative*
DAT	*dative*
D	*determiner/definiteness*
Dem	*(neutral) demonstrative*
DM	*direct modification*
DU	*dual*
EGmc	*East Germanic*
EWAhd	Etymologisches Wörterbuch des Althochdeutschen
foc	*focus*
F	*feminine*
GEN	*genitive*
Gk.	*Greek*
Gm	*geminator with* i-*mutator*
Go.	*Gothic*
I	*inflection*
IE	*Indo-European*
IM	*indirect modification*
K	*(strong adjectival) case*
K_D	*pronominal case*
INS	*instrumental*

Latv.	*Latvian*
M	*masculine*
N	*neuter*
N	*noun*
NGmc	*North Germanic*
NOM	*nominative*
Num	*numeral*
NWGmc	*Northwest Germanic*
OE	*Old English*
OED	Oxford English Dictionary
OF	*Old Frisian*
OFOH	*one feature one head*
OHG	*Old High German*
OLF	*Old Low Franconian*
ON	*Old Norse*
OS	*Old Saxon*
PART	*past participle*
PAST	*past tense*
PIE	*Proto-Indo-European*
PGmc	*Proto-Germanic*
PL / Pl	*plural*
PN	*Proto-Norse*
PRES	*present tense*
Q	*interrogative*
R	*reinforcer*
RDem	*reinforced demonstrative*
RN	*Runic Norse*
σ	*syllable*
SG / Sg	*singular*
Skt.	*Sanskrit*
T	*tense*
V	*verb*
WGmc	*West Germanic*
Φ	*phi (person, number, gender)*

Table of Contents

1

Introduction: The reinforced demonstrative

This dissertation is a detailed study of the Northwest Germanic (NWGmc) reinforced demonstrative pronoun (RDem) and its internal structure. RDem is commonly considered to be a proximal demonstrative and accordingly glossed as 'this'. Its earliest attestations are to be found in the North Germanic (NGmc) and West Germanic (WGmc) languages[1]: Runic Norse (RN) *súsi*, *sási*, *þatsi*; Old Norse (ON) *sjá/þessi*, *sjá/þessi*, *þetta*; Old Frisian (OF) *thius*, *this*, *thit*; Old English (OE) *þēos*, *þe(:)s*, *þis*; Old Saxon (OS) *thius*, **these*, *thit*; Old High German (OHG) *dësiu*, *dësēr*, *diz*. In these languages RDem contrasts both semantically and morphologically with the neutral demonstrative (Dem) 'that': ON *sú*, *sá*, *þat*; OF *thiu*, *thī*, *thit*; OE *sēo*, *se(:)*, *þæt*; OS *thiu*, *the*, *thet*; OHG *diu*, *dër*, *daʒ*.

Though it is considered a crucial defining morphological property of the NWGmc branch, RDem has not received a detailed treatment in the literature. For instance, Robinson (1992: 89-90, my bold) writes:

> Old Norse has developed an 'intensified' demonstrative pronoun…by attaching an intensifying particle *-si* to the regular demonstrative, and then carrying out **a series of analogical changes that render this origin obscure**.

Similarly, Haugen (1982: 100-1, my bold) despairs that we will never be able to tease out the details of the early history of RDem:

> The dem[onstrative] pron[oun] 'this' is formed partly by adding *-se* to the [demonstrative 'that'] (as in WGmc), partly *-a*, a specially Nordic suffix, and is then given a variety of adjectival inflections, which makes it **impossible to set up a C[ommon] Sc[andinavian] paradigm**.

[1] The dictionary forms of RDem and Dem are given in the order F.NOM.SG, M.NOM.SG, N.NOM.SG. Unattested forms are preceded by an asterisk.

Noreen (1923: 314-315), Prokosch (1939: 271-2), and Nielsen (2000: 211) also approach RDem with the philological equivalent of a long stick. In some cases the analysis is especially weak, as when Prokosch (1939: 271-272, my bold) appeals to "emphatic character" to explain some relatively straightforward morphophonological characteristics of RDem:

> Other [Old Norse intensified demonstrative] forms…might also have as their second element the particle *-a* seen in Go. *þata*, etc., **preserved, contrary to phonetic law, by the emphatic character of this pronoun**: acc. sg. masc. *þenna*, nom., acc. sg. neut. *þetta*. **The final consonant is doubled in all of these forms which is at least in part due to the emphatic meaning**.

Clearly much more can and should be said about RDem and its internal structure.

Above all the dissertation treats the RDem paradigm of Old Norse (c.1050 – c.1300). As our understanding of ON grows and an analysis is developed, it will become possible to come to terms with the WGmc facts as well. Thus the paradigms of OE (c.650 – 1066), OF (c.1150 – c.1550), OS (c.830 – c.1200), and OHG (c.750 – c.1050) will be studied in some detail. By then various dialectal and/or more archaic forms from Norse will also fall into place. At times reference will also be made to the runic inscriptions of the Viking Age, written in a language I will term Runic Norse (RN), an older stage of Norse datable to the period c.800 – c.1050.[2] Note that RDem is absent in East Germanic (EGmc), i.e. Gothic (Go.) (4th century but attested in 6th century texts). As we shall see in this chapter, however, Gothic can nevertheless provide us with important insights into the historical development of RDem.

NWGmc is a phylogenetic grouping of the Germanic languages which entails that EGmc (i.e. Gothic) broke away from Proto-Germanic (PGmc) first, leaving the NWGmc dialect group to continue its evolution until it too broke up, into NGmc and WGmc. The NWGmc hypothesis (Kuhn 1952-1953, Lehmann 1966, Nielsen 1989, among others) is overwhelmingly assumed in contemporary Germanic linguistics.[3]

The commonly accepted Germanic 'family tree' looks (more or less) like Figure 1, where boldface indicates that the language will be discussed in this dissertation. Regarding my choice of languages to study, the reader will note that Old Low Franconian (OLF), the ancestor of modern-day Dutch and Flemish dialects, is not included. Even though RDem surely must have existed in this language, it happens to be unattested in our limited sources of OLF (Robinson 1992: 214).

---

[2] For the runic data found throughout this dissertation I make use of the online corpus *Samnordisk runtextdatabas* of Uppsala University. Appendix I lists the specific inscriptions containing the runic forms I cite throughout the dissertation.

[3] Contrast this with the discarded Gotho-Nordic hypothesis (Schwarz 1951), which states instead that EGmc and NGmc underwent a period of common development, separate from WGmc.

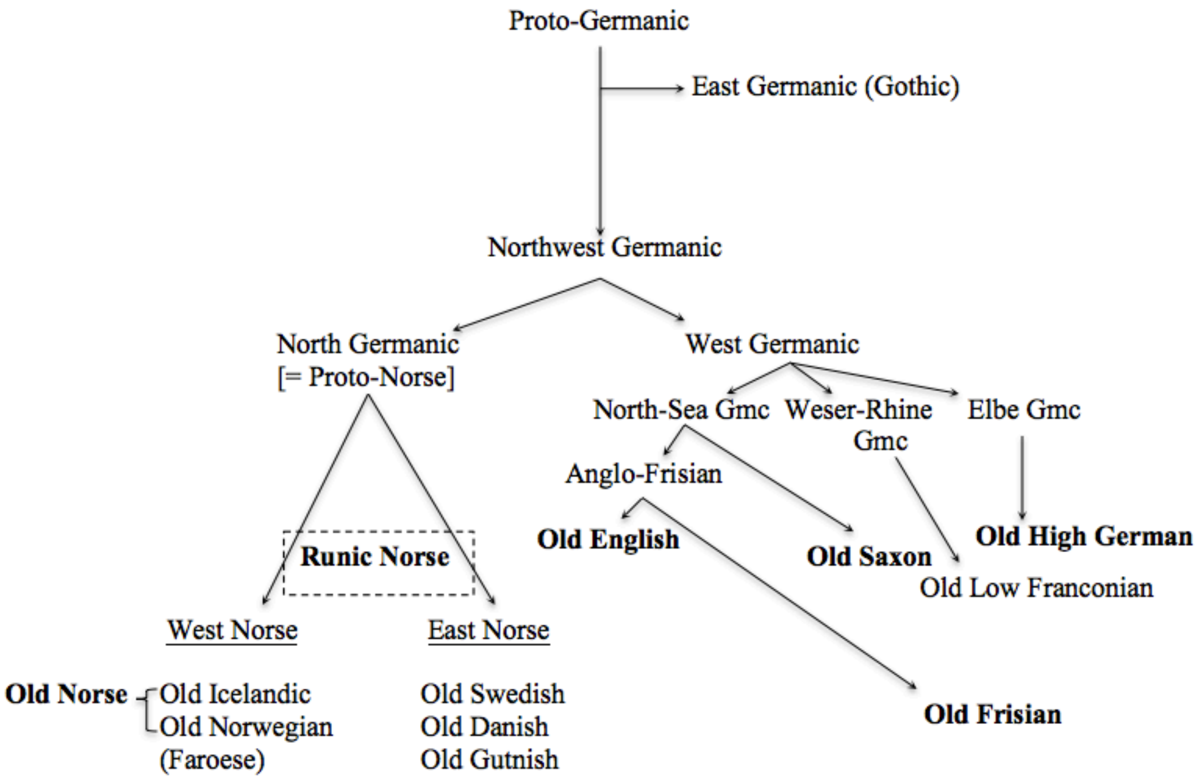


Figure 1 Germanic Stammbaum (cf. Nielsen 1989, 2000)

I have not used the prefix *Proto-* for each and every unattested language in Figure 1. Partly this is due to reasons of space and legibility. Another reason for leaving out *Proto-* comes down to matters of reconstruction. For instance, it is quite difficult to reconstruct the proto-languages of WGmc with much certainty, even though they surely existed in some form. In NGmc, moreover, one could argue that varieties of West and East Norse are actually attested in runic inscriptions of a certain age, and if a language is attested then the *Proto-* prefix is not appropriate (cf. Antonsen 2002: 31-5 on viewing the oldest runic inscriptions as an attestation of NWGmc). On the other hand, I have retained the prefix for PGmc and Proto-Norse (PN). This is because these two languages can be reconstructed with a fair degree of certainty, and for both languages the absence of the prefix would not only be unconventional but also cause confusion, as the terms *Germanic* and *Norse* are too vague if what we intend to refer to are the specific languages PGmc and PN. In this dissertation I use *Germanic* as a cover term for the entire language family in Figure 1.

The usual disclaimers apply about representing closely related languages as a 'family tree' or as a 'dialect continuum'. I have chosen more for the former but have attempted to illustrate in Figure 1, roughly, how speakers of certain languages must have been in close contact with one another. For instance, the continental WGmc languages (as opposed to insular OE) are geographically clustered together and share a great many features, partly due to common ancestry but also due to contact. Thus I have placed OHG, OS, OLF, and

OF in close proximity in Figure 1. At the same time, genetic provenance is important. OE and OF are very closely related, so much so that many Germanicists posit an Anglo-Frisian branch. One level up, we can posit a North-Sea Germanic branch which also encompasses OS. OHG and OLF belong to separate branches.

On the NGmc side, RN is attested in thousands of inscriptions all over Scandinavia which utilize a 16-letter alphabet known as the Younger Futhark. During the Viking Age, the language being written using this runic alphabet can undoubtedly be classified as NGmc, with western and eastern varieties identifiable based on certain dialectal features.[4] Thus in Table 1 I have placed RN below PN and overlapping between the West and East branches of Norse. RN is an especially fascinating linguistic layer of the NGmc branch because it is often quite linguistically conservative, and in this dissertation various interesting properties of its (partially attested) RDem paradigm will come into focus. Finally, it should be noted that while *Old Norse* is commonly used as a broad cover term for the dialects of West and East Norse, the 'classical' ON which is so richly attested in skaldic poetry, the sagas, and the Eddas consists more accurately of the West Norse dialects of Old Icelandic and Old Norwegian, especially the former. Here I have opted for the perhaps misleading but nevertheless traditional name ON, rather than Old Icelandic.

Though I deal with historical material, my primary goals are fundamentally synchronic in nature, in that I consider the RDem paradigms of the Old Germanic languages foremost as objects of study in and of themselves. From there I move into issues of crosslinguistic variation (how the different languages' paradigms compare to and contrast with one another) or of diachrony (how the morphological structures of RDem changed over time), but my starting point will always be synchrony.

The goal of this chapter is to provide some general background information on the demonstrative system of Germanic. It is organized as follows. Section 1.1 discusses Dem in the Germanic languages. Here the reader will become familiar with important morphemes, such as the Dem stem (in its various guises) and inflectional suffixes. Section 1.2 introduces RDem, from its earliest configuration in Proto-NWGmc down to the RDem paradigms of ON, OE, OF, OS, and OHG. The reader will be introduced to a number of distinct reinforcer morphemes in this section. Section 1.3 is a brief discussion of the diachronic and synchronic relations between the adjectival and pronominal inflectional systems. Section 1.4 provides the etymologies of the reinforcers of RDem, which includes

---

[4] There are older runic inscriptions as well. The transitional inscriptions, datable to the period c.550 – c.800, can be considered quite close to PN (also known as *Common Norse*, *Common Scandinavian*, *Proto-North Germanic*, etc.). The runic inscriptions which predate 550 (using the full 24-letter Elder Futhark and called *Early Runic* by Nielsen 2000) are notoriously controversial in Germanic linguistics. Some insist (e.g. Grønvik 1981) that these inscriptions represent a purely NGmc linguistic stage, but there is mounting evidence that these inscriptions are some kind of NWGmc entity. More precisely, Early Runic should be viewed as the ancestor to both NGmc and North-Sea Germanic (= the WGmc languages minus OHG) (Antonsen 1975, 2002; Nielsen 2000).

a discussion of what Gothic can teach us about the history of RDem. In Section 1.5 I put forth the main goals of the dissertation as well as a brief outline of its contents.

1.1 The Germanic neutral demonstrative and its stems

This section introduces Dem, the neutral demonstrative pronoun which is usually glossed as distal 'that'. However, it is very important to note that PIE **seh*$_2$/*sā*, **so*, **tod* could mean both distal 'that' and proximal 'this' (Watkins 2000: 81, Fortson 2004: 129, Beekes 2011: 226). The same was true for PGmc, at least on the basis of the oldest attested Germanic language, namely Gothic. Streitberg (2000 II: 112) writes that Go. *so*, *sa*, *þata* can have either *ich*-deixis 'this' or *der*-deixis 'that'. For Norse, cf. also the Rök Stone's **runaʀ þaʀ** '*these* runes'. For these reasons I designate Dem a *neutral* demonstrative.

As we shall see, Dem served as the basis for the formation of RDem. Though the focus of this dissertation is the RDem paradigm in all its intricacies, it is important to have a general understanding of Dem as well. This section provides an inventory of Dem in various Germanic languages: PGmc (Section 1.1.1), Gothic (Section 1.1.2), ON (Section 1.1.3), OF (Section 1.1.4), OE (Section 1.1.5), and OS and OHG (Section 1.1.6). Section 1.1.7 is an overview. These background sections are especially important in that they identify the morphological subcomponents of Dem, which will be useful when we turn to the internal morphological structure of RDem.

1.1.1 Proto-Germanic

The Dem paradigm of the Germanic mother language, PGmc, is given in Table 1.

Table 1 PGmc Dem paradigm (Nielsen 2000: 217-218, 230-235)[5]

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	sō	sa	þat	þōz	þai	þō
ACC	þō	þan-	þat	þōz	þanz	þō
GEN	þa(i)zōz	þes(a)	þes(a)	þaizō	þaizō	þaizō
DAT	þa(i)zai	þa(i)m, þezmō	þa(i)m, þezmō	þaimiz	þaimiz	þaimiz

Table 2 Decomposition of PGmc Dem

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	sō	sa	**þ-a**-t	**þ**-ōz	**þ**-ai	**þ**-ō
ACC	**þ**-ō	**þ-a**-n-	**þ-a**-t	**þ**-ōz	**þ-a**-nz	**þ**-ō
GEN	**þ-a(i)**-zōz	**þ-e**-s(a)	**þ-e**-s(a)	**þ-ai**-zō	**þ-ai**-zō	**þ-ai**-zō
DAT	**þ-a(i)**-zai	**þ-a(i)**-m	**þ-a(i)**-m	**þ-ai**-miz	**þ-ai**-miz	**þ-ai**-miz

As shown in Table 2, Dem is made up of three main parts. The first component is a demonstrative root **þ-* (< PIE **t-*). This root is present in all forms of the paradigm except two, which instead use an irregular root **s-* (< PIE **s-*). The second ingredient in Dem is a stem vocalism. Where present, the stem vowel is realized either as a monophthong **-a* or as a diphthong **-ai*. A number of forms lack a stem vowel (i.e. F.ACC.SG / N.NOM/ACC.PL **þ-ō*, M.NOM.PL **þ-ai*, F.NOM/ACC.PL **þ-ōz*). The third component of Dem is an inflectional case (K) suffix. In addition to case, the suffix also marks gender, number, and person, or, to use a term current in the generative literature, Φ features. For the purposes of discussion in this dissertation we may think of the components of Dem in terms of just two ingredients, (i) the Dem stem, made up of the root and usually a stem vocalism, and (ii) the inflectional suffix, abbreviated henceforth as K. In Table 2 I have bolded the Dem stem. Again, note that the Dem stem comes in three guises: one with a monophthongal stem vowel (**þa-*), one with a diphthongal stem vowel (**þai-*), and one with no stem

[5] Ringe's (2006: 288-289) reconstruction of Dem differs slightly:

Table (i) Ringe's (2006: 288-289) reconstruction of PGmc Dem

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	sō	sa	þat	þōz	þai	þō
ACC	$þō^n$	$þanō^n$	þat	þōz	þanz	þō
GEN	þaizōz	þas	þas	$þaizō^n$	$þaizō^n$	$þaizō^n$
DAT	þaizōi	þammai	þammai	þaimaz	þaimaz	þaimaz

vowel (*þ-). In this dissertation I will not enter into a detailed discussion of the internal structures of these three stem variants beyond observing that they exist and noting how they are distributed within the paradigm.

As explained in detail by Nielsen (2000: 217-218, 230-235), the PGmc Dem paradigm displayed some amount of variation and was also undergoing a number of complex vowel changes that can be seen in the daughter languages. Based on descendant forms in OE, OF, and ON, it is likely that the stem vocalism **ai* in the singular forms (F.GEN.SG, F.DAT.SG, M/N.DAT.SG) is an intrusion from the plural (GEN.PL **þaizō* and DAT.PL **þaimiz*), meaning that the original vowel in the singular was simply **a*: F.GEN.SG **þazōz*, F.DAT.SG **þazai*, M/N.DAT.SG **þam*.

More relevant for our discussion in the next sections, though, is the fact that these singular forms must have also had variants in **e*, as in F.GEN.SG **þezōz*, F.DAT.SG **þezai*, M/N.DAT.SG **þezmō*, given the forms we see in Gothic, OS, and OHG, to be seen below. Thus it is useful to keep in mind that we need to reconstruct PGmc variants with **e*-vocalism in certain forms. Another form which showed **e*-vocalism was the M/N.GEN.SG form **þes(a)*, and in most daughter forms we end up seeing a continuation of this **e*: Go. *þis*, ON *þes(s)*, OS *thes*, OHG *dës* (in addition to Nielsen 2000: 232-233, see also Prokosch 1939: 269 and Klingenschmitt 1987: 184).

Nielsen (2000: 231, 233) points out that the appearance of **e*-vocalism may very well be bolstered by analogy with the M.NOM.SG third person anaphoric pronoun **ez*, as clearly seen from the vowels of M.NOM.SG Dem forms like OE/OS *se(:)* and OHG *der*. In fact, it is possible that the paradigm of **ez* influenced the Dem paradigm in more slots than just the M.NOM.SG. Indeed, the PGmc third person anaphoric pronoun **sī*, **iz/*ez*, **it* (cf. Go. *si*, *is*, *ita*; OHG *siu*, *ër*, *iȥ*) < PIE **ih_2*, **h_1e*, **id* (Beekes 2011: 227, 229; Fortson 2004: 130 reconstructs **ei-*) displays a significant amount of **e*-vocalism in its paradigm, not just in the M.NOM.SG, as seen in Table 3.

Table 3 PGmc third person anaphoric pronoun (Ringe 2006: 289; see Nielsen 2000: 231, 233 for **ez*)[6]

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	sī	iz ez	it	ijōz	īz	ijō
ACC	ijō^n	inō^n	it	ijōz	inz	ijō
GEN	ezōz	es	es	ezō^n	ezō^n	ezō^n
DAT	ezōi	immai	immai	imaz	imaz	imaz

6 Ringe (2006: 289) also reconstructs a F.INS.SG form with **e*-vocalism: **ezō*. According to Szemerényi's (1996: 207) reconstruction, moreover, M/N.DAT.SG PIE **esmōi* could be inherited into PGmc as **ezmōi* (> Go. *imma*).

The fact that we clearly see a significant amount of **e*-vocalism in the Dem paradigms of both EGmc (Gothic) and WGmc (OS, OHG) suggests that we date this influence from the **ez* paradigm to a relatively late stage of PGmc.

1.1.2 Gothic

Gothic stays quite close to PGmc, as seen in Table 4. In Table 5 I parse the Gothic forms into Dem stem plus K suffix. The Dem stem is bolded, as in Table 2 above, but stem vowels have not been separated out (i.e. instead of M/N.GEN.SG *þ-i-s* I write *þi-s*).

Table 4 Go. Dem (Rauch 2011: 78)

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	sō	sa	þata	þōs	þái	þō
ACC	þō	þana	þata	þōs	þans	þō
GEN	þizōs	þis	þis	þizō	þizē	þizē
DAT	þizái	þamma	þamma	þáim	þáim	þáim

Table 5 Decomposition of Go. Dem

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	sō	sa	**þa**-ta	**þ**-ōs	**þ**-ái	**þ**-ō
ACC	**þ**-ō	**þa**-na	**þa**-ta	**þ**-ōs	**þa**-ns	**þ**-ō
GEN	**þi**-zōs	**þi**-s	**þi**-s	**þi**-zō	**þi**-zē	**þi**-zē
DAT	**þi**-zái	**þa**-mma	**þa**-mma	**þái**-m	**þái**-m	**þái**-m

Gothic continues all three kinds of stems: *þa-* (N.NOM/ACC.SG *þa-ta*, M.ACC.SG *þa-na*, M/N.DAT.PL *þa-mma*, M.ACC.PL *þ-ans*), *þái-* (DAT.PL *þái-m*),[7] and *þ-* (F.ACC.SG / N.NOM/ACC.PL *þ-ō*, F.NOM/ACC.PL *þ-ōs*, M.NOM.PL *þ-ái*). Go. F.NOM.SG *sō* and M.NOM.SG *sa* continue PGmc F.NOM.SG **sō* and M.NOM.SG **sa*. The M/N.DAT.PL form *þamma* goes back to PGmc **þazmō* (< PIE **tosmōi*, Szemerényi 1996: 205) or an old instrumental form **þazmē* (< PIE **tosmē*; Beekes 2011: 226-227, Nielsen 2000: 218).

Gothic also shows a fourth stem, *þi-*, which as just discussed comes from the PGmc stem variant **þe-*: F.GEN.SG *þi-zōs*, F.DAT.SG *þi-zái*, and M/N.GEN.SG *þi-s*; it is also

[7] Grimmean orthographic convention holds that *ái* is a diphthong, which it is historically (in PGmc), but contemporary Germanicists tend to agree that the PGmc diphthong *ái* had become *ɛ:* in Gothic (Rauch 2011: 51-53, 59).

observed in the GEN.PL, i.e. F.GEN.PL *þi-zō* and M/N.GEN.PL *þi-zē*, which can be seen as an influence from the singular (Nielsen 2000: 233).[8]

1.1.3 Old Norse

Consider now the Dem paradigm of ON in Table 6. Again the Dem forms have been divided into the Dem stem (in bold) plus the K suffix.

Table 6 ON Dem (Gordon 1956: 295, Haugen 1982: 101)

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	sú	sá	**þa**-t	**þ**-ær, **þ**-ár	**þei**-r	**þa**-u
ACC	**þ**-á	**þa**-nn	**þa**-t	**þ**-ær, **þ**-ár	**þ**-á	**þa**-u
GEN	**þei**-r(r)ar	**þe**-ss	**þe**-ss	**þei**-r(r)a	**þei**-r(r)a	**þei**-r(r)a
DAT	**þei**-r(r)i	**þei**-m	**þ**-ví, **þ**-ȳ	**þei**-m	**þei**-m	**þei**-m

In ON we see continuations of the three PGmc Dem stems: *þ-* (< **þ-*), *þa-* (< **þa-*) and *þei-* (< **þai-*), as well as the irregular forms with initial *s-* (< PGmc **s-*) for the F.NOM.SG *sú* and M.NOM.SG *sá*. Note also that the M/N.GEN.SG form *þess* shows an irregular stem *þe-*, ultimately reflecting PGmc **þes*.

Observe that M.NOM.PL *þeir* is unexpected. Gothic's M.NOM.PL *þái* points to PGmc **þai* (< PIE **toi*), which should come down as **þei* in ON, contrary to fact. The ON form, then, suggests reconstructing the PN form as **þai-z*, where the appended *-*z* rhotacizes to -*r* in ON. This pleonastic *-*z* is in all likelihood based on the M.NOM.PL endings of the strong noun classes, which all ended in **z*.[9] See Prokosch (1939: 270).

The N.DAT.SG inflection (-*ví*, -*ȳ*) is unexpected given the corresponding PGmc inflection (*-*m*, *-*mmai*, or even *-*smō*). According to Prokosch (1939: 269) the form *því* is an old locative **þī* (< IE **tei*, cf. Go. *þei*), where the insertion of *v* into *þī* is due to analogy with the N.DAT.SG interrogative pronoun *hví* (Prokosch 1939: 269, EWAhd II: 615). As for the N.DAT.SG variant *þȳ*, Haugen (1982: 93) derives this from an old locative plus adjectival dative inflection, i.e. **þī-u*, which straightforwardly gives *þȳ*. This is due to the vowel harmony process known as *u*-umlaut, whereby a non-round vowel is rounded due to a *u* in the next syllable (i.e. *ī* > *ȳ*).

[8] It is also interesting to recognize that these forms display morphological containment of the Go. third person pronoun 'he, she, it': *þ-[izōs]*, *þ-[izái]*, *þ-[is]*, *þ-[izō]*, *þ-[izē]*, and perhaps *þ-á-[im]* (see Rauch 2011: 80 for the third person pronoun paradigm).

[9] From Haugen (1982: 90-1): *a*-stem *-*ōz*, *ia*-stem *-*iōz*, *i*-stem *-*īz*, root-stem *-*iz*, *r*-stem *-*riz*, *nd*-stem *-*iz*.

Finally, note that F.NOM/ACC.PL *þær* (<æ> = /ē/) is unexpected due to its front vowel, since in PN the form was simply **þ-āʀ*. Now, in addition to *u*-umlaut there is another well known vowel harmony process in Scandinavian known as *i*-umlaut, by which a high front vowel or glide (*i* or *j*) will front a back vowel in the preceding syllable. It is well known that PN **ʀ* was also known to trigger *i*-umlaut in western varieties of Norse, so fronting of *ā* to *ē* has taken place in PN **þāʀ* (Prokosch 1939: 270). When **ʀ* changed to *r* in ON, the umlaut environment was erased, resulting in an irregular form, *þær*.

1.1.4 Old Frisian

The OF Dem paradigm is shown in Table 7. The Dem forms have been divided into the Dem stem (in bold) plus the K suffix.

Table 7 OF Dem (Bremmer 2009: 54)

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	**th**-iu	**th**-ī	**the**-t	**th**-ā	**th**-ā	**th**-ā
ACC	**th**-ā	**the**-ne	**the**-t	**th**-ā	**th**-ā	**th**-ā
GEN	**thē**-re	**the**-s	**the**-s	**thē**-ra	**thē**-ra	**thē**-ra
DAT	**thē**-re	**thā**-(m)	**thā**-(m)	**thā**-(m)	**thā**-(m)	**thā**-(m)

As mentioned above, OF and OE share a direct common ancestor. Not surprisingly, then, OF is in many respects very much like OE. The three primary PGmc stems are continued in OF as follows: *th-* (< **þ-*), *the-* (< **þa-*) and *thā-/thē-* (< **þai-*) (Campbell 2003: 52, Robinson 1992: 191-192). Though the various vowel developments in the OF dialects are complex, for the most part OF *ā* corresponds to OE *ā*, and OF *e(:)* corresponds to OE *æ(:)* (see below for OE).

A notable change in OF is that the *s*-initial forms have been leveled and supplanted by *th*-initial forms, giving F.NOM.SG *thiu* (cf. OE *sēo* but also later *þēo*) and M.NOM.SG *thī* (cf. OE *se(:)* but also later *þe(:)*).

In terms of the inflectional system, in OF the F.GEN.SG and F.DAT.SG have been conflated into a single form, *thēre*, as in OE. Also, the NOM/ACC of all genders in the plural is, as in OE, *thā*.

1.1.5 Old English

Consider the OE Dem paradigm in Table 8. The Dem forms have been divided into the Dem stem (in bold) plus the K suffix.

Table 8 OE Dem (Campbell 2003: 290, Lass 1994: 143)

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	sēo	se(:)	**þæ**-t	**þ**-ā	**þ**-ā	**þ**-ā
ACC	**þ**-ā	**þo**-ne	**þæ**-t	**þ**-ā	**þ**-ā	**þ**-ā
GEN	**þǣ**-re	**þæ**-s	**þæ**-s	**þā**-ra, **þǣ**-ra	**þā**-ra, **þǣ**-ra	**þā**-ra, **þǣ**-ra
DAT	**þǣ**-re	**þā**-m, **þǣ**-m	**þā**-m, **þǣ**-m	**þā**-m, **þǣ**-m	**þā**-m, **þǣ**-m	**þā**-m, **þǣ**-m

We again see continuations of the three PGmc stems in OE: *þ-* (< **þ-*), *þæ-* (< **þa-*) and *þā-/þǣ-* (< **þai-*) (Campbell 2003: 52, Robinson 1992: 157-158). OE also preserves *s*-initial F.NOM.SG and M.NOM.SG forms, namely *sēo* and *se(:)*, respectively.

One form which seems to be aberrant is M.ACC.SG *þo-ne*. The *o*-vocalism here is due to PGmc **a* going to OE *o* before nasals (Lass 1994: 41).

Inflectionally speaking, OE corresponds to PGmc, with two exceptions. First, F.GEN.SG and F.DAT.SG have been conflated into a single form, *þǣre*. Second, the NOM/ACC of all genders in the plural have been conflated as well, into *þā*.

1.1.6 Old Saxon and Old High German

Consider now the Dem paradigms of OS in Table 9 and OHG in Table 10. The Dem forms have been divided into the Dem stem (in bold) plus the K suffix.

Table 9 OS Dem (Rauch 1992: 194, Cathey 2000: 37; also Gallée 1910: 238 and Nielsen 2000: 217-218)

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL	
NOM	**th**-iu	**th**-e(:), se(:)	**tha**-t	**th**-ia	**th**-ia	**th**-ia	**th**-iu
ACC	**th**-ia	**the**-na	**tha**-t	**th**-ia	**th**-ia	**th**-ia	**th**-iu
GEN	**the**-ra	**the**-s	**the**-s	**the**-ro	**the**-ro	**the**-ro	
DAT	**the**-ru	**the**-mu	**the**-mu	**thē**-m	**thē**-m	**thē**-m	

Table 10 OHG Dem (Braune & Reiffenstein 2004: 247, Wright 1906: 67)[10]

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	**d-**iu	**dë**-r	**da**-ȥ	**d**-eo	**d**-ē	**d**-iu
ACC	**d-**ea	**dë**-n	**da**-ȥ	**d**-eo	**d**-ē	**d**-iu
GEN	**dë-**ra	**dë**-s	**dë**-s	**dë**-ro	**dë**-ro	**dë**-ro
DAT	**dë**-ru	**dë**-mu	**dë**-mu	**dē**-m	**dē**-m	**dē**-m

The three Dem stems are again continued, but, as we shall see, not as clearly as in OE and OF. The *s-initial forms have, with the exception of OS M.NOM.SG *se(:)*, been leveled out in OS and OHG in favor of *th-* or *d-*, both of which go back to **þ-*. Note also that OHG M.NOM.SG *dër* shows a pleonastic *-r* marker (cf. RN M.NOM.SG **saʀ** and PGmc M.NOM.SG **e-z* < PIE **h_1e*).

In both OS and OHG, the stem vocalism of original PGmc **þa-* has been replaced in many instances by *e*, perhaps mostly due to analogy with the PGmc pronoun **ez* (Nielsen 2000: 222, 231, 233) (cf. Section 1.1.1 above), resulting in forms like OS M.NOM.SG *th-e(:)*, M.ACC.SG *the-na*, M/N.GEN.SG *the-s*; and OHG M.NOM.SG *dë-r*, M.ACC.SG *dë-n*, M/N.GEN.SG *dë-s*. The original **a* has survived in N.NOM/ACC.SG OS *tha-t* and OHG *da-ȥ* (Nielsen 2000: 231).

The long-voweled stems OS *thē-* and OHG *dē-* in the DAT.PL are regular continuations of the PGmc diphthong in **þai-* (Nielsen 2000: 218, citing Dal 1971).

In some cases, however, OS *thē-* and OHG *dē-* are expected but OS *the-* and OHG *dë-* appear instead. Nielsen (2000: 217-218, 233-234) explains that these forms may go back to PGmc variants with **e* (> *e*/*ë*) instead of **ai* (> *ē*), as mentioned above: F.GEN.SG OS *the-ra* and OHG *dë-ra* < PGmc F.GEN.SG **þezōz*, F.DAT.SG OS *the-ru* and OHG *dë-ru* < PGmc F.INS.SG **þezō* (citing Krogh 1996), M/N.DAT.SG OS *the-mu* and OHG *dë-mu* < PGmc M/N.DAT.SG **þezmō* < PIE **tesmō* (citing Krahe 1969; see also Beekes 2011: 228). In addition, the monophthongs in the plural, i.e. GEN.PL OS *the-ro* and OHG *dë-ro* (technically reflecting PGmc **þezō*), are due to analogy with the singular forms just cited, on a par with the Gothic situation (Nielsen 2000: 233).

Another conspicuous feature of the OS and OHG paradigms is the appearance of diphthongs in the following plural forms: OS F.ACC.SG / NOM/ACC.PL *thia* (also *thea*), OHG F.ACC.SG *dea* (also *dia*), F.NOM/ACC.PL *deo* (also *dio*), and N.NOM/ACC.PL OS *thiu* and OHG *diu*. Prokosch (1939: 269-71) suggests that the initial vowel (*i/e*) in these diphthongs comes from a PIE glide **j*. According to him this merits reconstructing doublets in PIE like F.ACC.SG **tām* ~ **tjām* and F.NOM/ACC.PL **tās* ~ **tjās*. By the same reasoning,

[10] NB: <ȥ> = *s*, specifically the *s* which comes from *t*. By the High German consonant shift, *t* in final position goes to *s*.

Klingenschmitt (1987: 182-183) reconstructs **sju* for the WGmc F.NOM.SG forms: OE *sēo*, OF/OS *thiu*, OHG *diu*.

With regard to inflection, OS is similar to OE and OF when it comes to not distinguishing gender in the NOM/ACC.PL, while OHG distinguishes all three genders in the plural while still conflating the nominative and accusative.

1.1.7 Taking stock

For what follows it will be useful to have a general overview of the different stems and inflectional patterns of Dem. Let us therefore summarize the main points of Section 1.1. In Table 1 it was seen that three different stems can be distinguished in the PGmc Dem paradigm. The continuation of these stems into the daughter languages is summarized in Table 11.

Table 11 Dem stems

Proto-Germanic	*þa- (*þe-)	*þai-	*þ-
Gothic	þa- (þi-)	þái-	þ-
Old Norse	þa- (þe-)	þei-	þ-
Old English	þæ-	þā-/þǣ-	þ-
Old Frisian	the-	thā-/thē-	th-
Old Saxon	(tha-) the-	thē-	th-
Old High German	(da-) dë-	dē-	d-

In OS and OHG the stem with *e*-vocalism is more paradigmatically prominent than the stem with *a*-vocalism, which is the reason the stems OS *tha-* and OHG *da-* are in parentheses rather than the stems OS *the-* and OHG *dë-*.

As we shall see next, the languages differ when it comes to which stem appears in the RDem paradigm. ON, OS, and OHG share the fact that they chose their short-voweled Dem stems (ON *þa-*, OS *the-*, OHG *dë-*) when it came to building RDem. OE and OF, on the other hand, did not use OE *þæ-* and OF *the-* but rather OE *þi-* and OF *thi-*, with *i*-vocalism.

1.2 The Northwest Germanic reinforced demonstrative[11]

A morphological innovation shared by the North and West branches of Germanic was the introduction of RDem, formed by the reinforcement of Dem with the suffix **-si*. This combination of Dem with the reinforcer *-*si* is illustrated in Table 12. RDem is frequently considered to be one of the defining pieces of evidence for the development of a Northwest branch of Germanic, distinct from EGmc (see the discussion of phylogenetic grouping in this chapter's introduction). In this section the main diachronic and synchronic characteristics of NWGmc RDem will be presented.

The configuration illustrated in Table 12 represents the earliest stage in the evolution of RDem, and it will often be referred to in the sections and chapters that follow as the 'Dem-*si* stage'.

Table 12 NWGmc Dem plus *-si/-se* (based on Nielsen's 2000: 230-235) reconstruction of Early Runic Dem)

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	sō	sa	þat	þōz	þai	þō
ACC	þō	þan(-)	þat	þōz	þanz	þō
GEN	þaizōz	þas	þas	þaizō	þaizō	þaizō
DAT	þaizē	þaim	þaim	þaim(z)	þaim(z)	þaim(z)

+ -si/-se

Interestingly, the archaic Dem-*si* stage of NWGmc is preserved in some RN inscriptions; see (1).

(1) F.NOM.SG **susi** (*súsi*) < NWGmc **sōsi*

F.ACC.SG **þasi** (*þási*) < NWGmc **þō-si*

M.NOM.SG **saʀ:si** (*saʀsi*)
M.NOM.SG **sasi** (*sási*) < NWGmc **sa-si*

M.ACC.SG **þan:si** (*þansi*)
þansi (*þansi*) < NWGmc **þan-si*

[11] Sections 1.2, 1.3, and 1.4 are closely based on Lander (2013).

N.ACC.SG	**þat:si** (*þatsi*)		
	þatsi (*þatsi*)	<	NWGmc **þat-si*
M.DAT.SG	**þaimsi** (*þœimsi*)	<	NWGmc **þaim-si*
F.ACC.PL	**þaʀsi** (*þársi*)	<	NWGmc * *þōz-si*
M.NOM.PL	**þiʀsi** (*þeirsi*)	<	NWGmc * *þai(z)-si*
N.NOM/ACC.PL	**þausi** (*þausi*)	<	NWGmc **þō-si*

The patterns in Table 12 and (1) typify 'internal inflection' (*Binnenflexion*): the pronominal M.DAT.SG inflectional component (K_D) is to the left of the reinforcer component (R) and thus word-internal, as seen in (2).

(2) Internal inflection in RN

þæi-	-m	-si
D	K_D	R

Roughly speaking, the subsequent overall development of RDem is as follows (language-specific details will be discussed in Sections 1.2.1 through 1.2.5). The most important modification following the Dem-*si* stage is that the sibilant of the *-*si* reinforcer is absorbed into the Dem component **þa*-. This Dem component, moreover, does not retain the original *a*-vocalism, but usually changes from *a* to *e* or *i* by different language-specific mechanisms. Ultimately these developments give rise to a new stem (the RDem stem), which looks something like *þes*- or *þis*-. The RDem stem is usually inflected with strong adjective endings (K) rather than the older pronominal endings (K_D) (Haugen 1982: 100-101, EWAhd II: 611, 613). In the RDem stem *þes*- we can still discern two separate parts, i.e. the Dem component *þe*- (D) and the reinforcer component (R) -*s*.

Here a special note on terminology is in order. According to some definitions, *þe*- in the RDem stem *þes*- is no longer a *stem* but actually a *base*. A base is often said to require some extra stem-forming material in order to become a stem, which then in turn can be inflected. In the case of the RDem stem, *þe*- is a base and -*s* is the stem-forming element. The stem *þes*- is then inflected with K. However, in the context of the Dem paradigm, *þe*- or *þa*- is a stem, since it is directly inflected with K (e.g. ON *þe-s(s)*, *þa-t*, etc.). This terminological distinction is illustrated in (3).

(3) Technical senses of *base* and *stem* in this dissertation

Dem:	þa-	$-K_D$	
	stem	*inflection*	
RDem:	þe-	-s-	-K
	base	*stem-forming element*	*inflection*

Observe that there is no base to speak of in the Dem forms, while in the RDem forms there is both a base and a stem-forming element, namely the reinforcer *-s*. Henceforth the reinforcer *-s* (which as we shall see is present in every single language under discussion in this dissertation) will be referred to as the *sigmatic reinforcer*.

As already alluded to in (3), the coalition between *þe-* and *-s* to form a new RDem stem results in a shift toward external inflection. In the Dem-*si* stage, K appears to the left of R (e.g. **þa-**t**-si*, **þa-**n**-si*). The formation of a new RDem stem, however, puts R immediately adjacent to the D element, giving *þe-s-* or *þi-s-*, and it is to this stem that K attaches. As a result, K ends up to the right of R, a configuration which I refer to as 'external inflection' (*Endflexion*). This is illustrated for ON in (4), where *-um* is the M.DAT.SG (and DAT.PL) strong adjective ending.

(4) External inflection in ON

þe-	-ss-	-um
D	R	K

RDem stem (D + R)

Though the overwhelming trend in the history of RDem is to move towards external, adjectival inflection, there are still forms which show internal positioning of K and forms which show pronominal rather than adjectival inflection (in fact, in Chapter 6 I will argue that these go hand in hand). Such forms can be considered remnants of the older Dem-*si* stage.

As I will show, the developments roughly outlined above can be seen in the RDem paradigms of ON, OE, OF, OS, and OHG. Let us take each language in turn, looking at the following properties:

(i) the nature of the RDem stem,
(ii) the relative position of the inflection,
(iii) the nature of the inflection: strong adjectival (K) vs. pronominal (K_D), and
(iv) additional reinforcers (i.e. non-sigmatic reinforcers).

These properties together make up the empirical core of this dissertation.

1.2.1 Old Norse

The RDem paradigm of ON, the primary language studied in this dissertation, is shown in Table 13. The K endings of RDem are shown in Table 14.

Table 13 ON RDem (stem *þess-*) (Gordon 1956: 294-295)

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	þe-ssi	þe-ssi	þe-tt-a	þe-ss-ar	þe-ss-ir	þe-ssi
ACC	þe-ss-a	þe-nn-a	þe-tt-a	þe-ss-ar	þe-ss-a	þe-ssi
GEN	þe-ss-ar	þe-ss-a	þe-ss-a	þe-ss-a	þe-ss-a	þe-ss-a
DAT	þe-ss-i	þe-ss-um	þe-ss-u	þe-ss-um	þe-ss-um	þe-ss-um

Table 14 ON *n*-type strong adjective endings

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	-Ø	-r	-t	-ar	-ir	-Ø
ACC	-a	-n	-t	-ar	-a	-Ø
GEN	-rar	-s	-s	-ra	-ra	-ra
DAT	-ri	-um	-u	-um	-um	-um

(i) RDem stem

Most of the forms in Table 13 show the stem *þess-*. At first glance, the only forms which do not display a stem *þess-* are N.NOM/ACC.SG *þetta* and M.ACC.SG *þenna* (but see (ii) below). The stem *þess-* is composed of the base *þe-* plus the reinforcer component *-ss-*. This reinforcer component ultimately derives from NWGmc *-*si*. Given that there is only a single *s* in the original NWGmc reinforcer, we have to say something extra about ON *-ss-*. We may think of the ON sequence *-ss-* as containing a (reinforcer) morpheme *-C*, which geminates the sigmatic reinforcer to its immediate left: *þe-s-**C**-* > *þess-*.

(ii) Position of inflection

The forms which are boxed in Table 13 have external inflection, since the K component, provided in Table 14, appears in the rightmost position in these forms: F.ACC.SG / M.ACC.PL *þess-**a***, M.DAT.SG / DAT.PL *þess-**um***, F.NOM/ACC.PL *þess-**ar***, and M.NOM.PL *þess-**ir***. Assimilation (or deletion) of *r* due to a preceding *s* is seen in

F.GEN.SG *þess-**ar*** (< *þess-**rar***), F.DAT.SG *þess-**i*** (< *þess-**ri***), and GEN.PL *þess-**a*** (< *þess-**ra***) (for more discussion see Chapter 3).

The forms M.ACC.SG *þenna*, N.NOM/ACC.SG *þetta*, and M/N.GEN.SG *þessa* show a different kind of structure, with K located word-internally. As we shall see, the *-a* morpheme that appears at the end of these forms is also a kind of reinforcer. Since the K component (*-nn-*, *-tt-*, *-ss-*) appears to the left of *-a* in these forms, inflection should be classified as internal here. Observe here that M/N.GEN.SG *þessa*, though it looks to begin with the RDem *þess-*, does not end with the M/N.GEN.SG ending *-s*. Instead *-s* appears word-internally. This means that M/N.GEN.SG *þessa* does not show the RDem stem *þess-*, rather the initial sequence of morphophonology in this form (i.e. *þe-* plus geminated genitive *-ss-*) just happens to be homophonous with the RDem stem *þess-*.

For the four forms which surface as *þessi* in Table 13, the position of inflection is ambiguous. Since the F.NOM.SG / N.NOM/ACC.PL K ending is null, the forms could have internal inflection (*þe-**Ø**-ssi*) or external inflection (*þe-ssi-**Ø***). And while the M.NOM.SG ending *-r* is not null, it is nowhere to be found in M.NOM.SG *þessi*. These somewhat puzzling issues will be dealt with in Chapter 3.

(iii) Type of inflection

As should already be clear, the ON RDem forms in Table 13 make use of strong adjective endings (K), more specifically the *n*-type endings in Table 14.

(iv) Non-sigmatic reinforcement

In addition to the sigmatic reinforcer *-s* from NWGmc **-si*, ON displays two additional reinforcement strategies. The first is gemination. Throughout the entire paradigm of ON RDem we see gemination of *-s* (the forms with the stem *þess-*) or of K (*þe-nn-a*, *þe-tt-a*, *þe-ss-a*). Thus, as mentioned in (i), ON has introduced into its RDem paradigm a morpheme *-C*, which doubles the consonant appearing to its immediate left: *þe-**s-C**-* > *þe-s-s-*, *þe-**n-C**-* > *þe-n-n-*, or *þe-**t-C**-* > *þe-t-t-*. The second strategy involves the so-called secondary reinforcer *-a*, appearing in *þenn-**a***, *þett-**a***, and *þess-**a***. I will henceforth refer to this reinforcer as the *asigmatic reinforcer*. The asigmatic reinforcer *-a* arose in the NGmc branch only, being totally absent in the RDem paradigms of WGmc. The history of the asigmatic reinforcer is discussed in Section 1.4.3.

1.2.1.1 The history of Norse *þe-*

At this point we may ask why RDem does not simply use as its base one of the Dem stems discussed in Section 1.1.3, such as *þa-* or *þei-*. There are at least two historical processes responsible for why we see *þe-* in the ON RDem paradigm instead of *þa-* or *þei-*. The first process revolves around a phonotactic constraint, and the second has to do with *i*-umlaut. I discuss these in turn.

The NWGmc reinforcer *-*si* has only a single *s*, but the ON RDem stem *þess-* has two. The historical process giving rise to this gemination of *s* involves the sequence **ʀs* in PN becoming assimilated to *ss* (see Gordon 1956: 280, Haugen 1982: 64, 101). (5) gives examples of some reconstructed PN Dem-*si* forms which contain the sequence **ʀs*. By the assimilation rule just mentioned this sequence changes to **ss*.

(5) Assimilation of PN **ʀs* to **ss*

(a)	F.GEN.SG	*þeiʀaʀ-si	>	*þeiʀa-ssi
(b)	DAT.PL	*þeimʀ-si	>	*þeim-ssi
(c)	M.NOM.PL	þeiʀ-si (RN **þiʀsi**)	>	*þei-ssi
(d)	F.NOM/ACC.PL	þēʀ-si (RN **þaʀsi**)	>	*þē-ssi

The outcomes in (5) still have internal inflection, but some time after the Dem-*si* stage, there must have been a switch from internal to external inflection. (6) shows what happens when the inflectional component moves from internal placement to final/external placement in ON.

(6) Inflection going from internal to external[12]

(a)	*þei-~~ʀa~~-ssi	>	*þei-ss(i)-	>	ON þess-ar
(b)	*þei-~~m~~-ssi	>	*þei-ss(i)-	>	ON þess-um
(c)	*þei-/-ssi	>	*þei-ss(i)-	>	ON þess-ir
(d)	*þē-/-ssi	>	*þē-ss(i)-	>	ON þess-ar

---

[12] It is certainly plausible that at some point in the period between the PN and ON stages, external inflection aided in producing *s*-gemination, a process which also resulted in the intermediate stem **þeiss-*.

(i)	F.GEN.SG	*þei-s(i)-rar	>	*þei-ss-ar	>	ON þess-ar
(ii)	F.GEN.SG	*þei-s(i)-ri	>	*þei-ss-i	>	ON þess-i
(iii)	GEN.PL	*þei-s(i)-ra	>	*þei-ss-a	>	ON þess-a

In (i-iii) we see pre-ON forms with K to the right of the reinforcer. The reinforcer is by this time on its way to losing its *i*-vowel, thus it is represented as *-s(i)* here, which puts *s* right next to the *r* of the *r*-initial endings. The sequence *sr* is assimilated to *ss* in ON.

As seen in (6), the process of K becoming repositioned to the right of -*ss*- brings about the emergence of the transitional stems *þeiss*- or *þēss*-, which are surrounded by a dotted box. Note that these forms are not too far from the RDem stem *þess*- found in the ON stage.

In (6) we see that the diphthong and long vowel in the transitional stems **þeiss*- and **þēss*- must have somehow been reduced to *e* on their way to ON, since in ON the stem is *þess*-. There is, in my opinion, a straightforward explanation for this reduction. Observe that without a reduction of *ei*/*ē* to *e*, the RDem stem would contain an 'overlong' syllable, that is to say it would consist of a diphthong/long vowel plus a consonant cluster/geminate. This poses a problem of syllabification, as the transitional stems in (6) would have to be syllabified as **þei.ss*- and **þee.ss*-. Crucially, this kind of syllabification requires the diphthong *ei* or the long vowel *ē* to occupy the entirety of the rime, leaving the geminate *ss* in onset position. The problem, however, is that *ss* is a non-permissible onset in ON. The problem is illustrated in Figure 2.

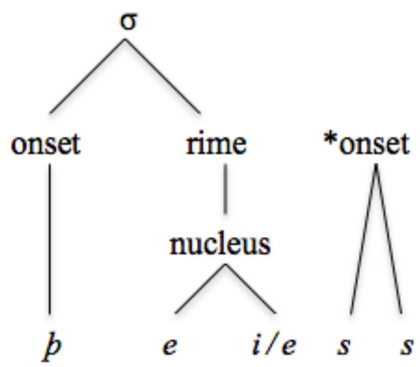


Figure 2 Non-permissible onset

Sandøy (1994) has demonstrated that in the overwhelming majority of cases displaying overlong quantity in ON, a morpheme boundary breaks up the consonant cluster or geminate. This justifies positing, along the lines of Kaye (1990), an empty nucleus for which the first consonant of the cluster/geminate can be the onset. For example, *mǿtti* 'met' is morphologically parsed as *mǿt-ti* 'meet-PAST' and can thus be syllabified *møø.t0.ti*, as shown in Figure 3.

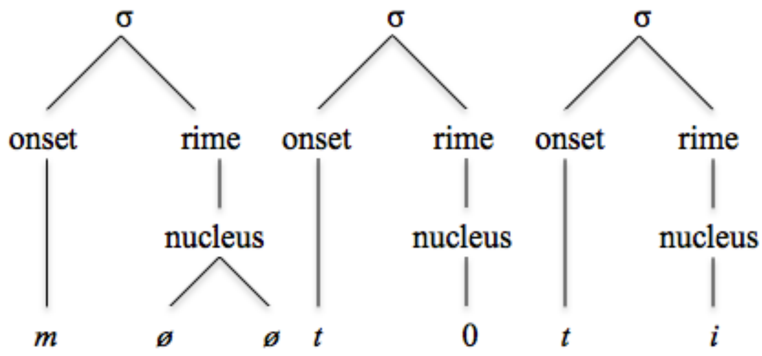


Figure 3 Empty nucleus at morpheme boundary in *mǿt-ti*

Interestingly for RDem, there used to be a morpheme boundary in the PN sequence **R-s* (e.g. F.GEN.SG **þeiRaR-si*), and therefore this sequence could theoretically (according to Sandøy 1994) be syllabified using an empty nucleus. However, a number of things happen during the transitional period between PN and ON to obscure and render meaningless this morpheme boundary, such as the assimilation of **Rs* to *ss* and the restructuring of inflection from internal to external. Due to the morphophonological tumult during this intermediate stage between PN and ON, Sandøy's solution becomes unavailable to the transitional stems **þeiss-* and **þēss-*, which are left with an overlong syllable but no morpheme boundary to justify alleviating it with an empty nucleus. Consequently there would have been phonotactic pressure to eliminate the deviant overlong syllable. The change from long *ei/ē* to short *e* relieved this pressure by allowing for the first *s* in *ss* to be a coda, as seen in Figure 4.

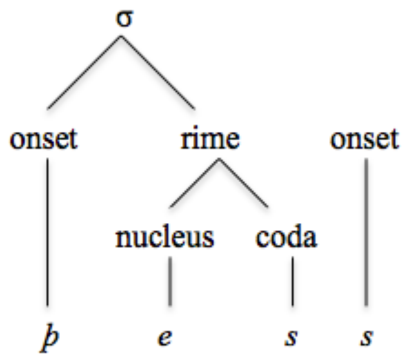


Figure 4 Permissible onset

This phonotactic proposal accounts for the monophthong in the ON RDem stem *þess-*.

A second way the *e*-vocalism in the base *þe-* came about was from *i*-umlaut of the Dem stem *þa-*. It is quite clear that the reinforcer *-si* could trigger *i*-umlaut (Nielsen 2000: 237, n. 3). Evidence for this comes from spelling alternations in various RN Dem-*si* forms; see (7).

(7) (a) M.ACC.SG **þan**(:)**si** (*þansi*) vs. **þensi** or **þinsi** (*þensi*)

(b) N.ACC.SG **þat**(:)**si** (*þatsi*) vs. **þitsi** (*þetsi*)

(c) F.ACC.SG **þa**(:)**si** (*þāsi*) vs. **þesi** or **þisi** (*þēsi*)

Orthographic alternations like these are typical hallmarks of nascent sound change. The cases in (7) are good evidence for *i*-umlaut conditioned by *-si*.[13]

To sum up, at least these two events (elimination of the overlong syllable in **þei.ss-* and *i*-umlaut from *-si*) co-conspired to produce a leveling effect throughout the RDem paradigm which resulted in a uniform root vocalism *e*. We should also not discount a third fact here, which is that the M/N.GEN.SG Dem form was ***þe**-s(s)* (cf. Section 1.1.1 on the PGmc stem **þe-*, as well as Section 1.4.2.2 on genitive stems in OE and OF). Synchronically speaking, then, the ON RDem forms all use the Dem stem *þa-* in its allomorphic guise *þe-*.

1.2.1.2 Reconstructing the development of the reinforced demonstrative in Norse

Looking ahead, in Chapter 6 I will discuss and integrate into my analysis a number of forms which lie outside of the classical ON paradigm given above.[14] For instance, there was a well known variant for M.NOM.SG *þessi* which took the form *þenni*.[15] Another variant which stood in for M.DAT.SG / PL.DAT *þessum* in skaldic Norse was *þeima*, where the final *-a* is the asigmatic reinforcer (cf. also the archaic N.DAT.SG *því-s-a*). Moreover, I have found a number of runic inscriptions which display interesting variants for classical ON M.ACC.SG *þenna* and N.NOM/ACC.SG *þetta*. The RN variants suggest that a noteworthy evolution led up to the classical forms as we know them. There is evidence for at least four stages in this evolution, as seen in (8).

---

[13] We are fortunate to have the clear alternations seen in (7) between <**a**> and <**e**>/<**i**>: technically, in the Younger Futhark <**a**> could represent the back vowels [*a*, *ɔ*] or the front vowel [*æ*/*e*], but <**e**> and <**i**> unambiguously represent front vowels (Haugen 1976: 144).

[14] Most of these variants are quite well known and can be found throughout the literature. I have consulted Haugen (1982), Axelsdóttir (2003), and Faarlund (2004). As far as I know, though, the RN forms in (8) have not been discussed before in any great detail.

[15] Ultimately the M.NOM.SG in Old Swedish, Old Danish, and Old Norwegian syncretized with the M.ACC.SG, giving *þenna* in both slots (> Swedish *denna*, Danish and Norwegian *denne*).

(8) (i) þa-t-si [non-umlauted Dem-*si* stage]
þa-n-si

(ii) þe-t-si [umlauted Dem-*si* stage]
þe-n-si

(iii) þe-t-s-a [asigmatic reinforcement stage]
þe-n-s-a

(iv) þe-t-t-a [classical forms]
þe-n-n-a

Stages (i) and (ii) and their RN forms have been mentioned above. For stage (iii), we have RN **þitsa** (*þetsa*) and **þensa**, **þinsa** (*þensa*) attested.

The final variant that will be discussed in Chapter 6 is the F/M.NOM.SG form *sjá*. This form is often cited instead of *þessi* as the prototypical F/M.NOM.SG form in the ON paradigm (or frequently both *sjá* and *þessi* are provided). The form is derived from **sē-a*, where the special stem **sē-* is ultimately the result of *i*-umlaut due to *-*si* (i.e. *sā-si* > **sē-si*, where the new stem *sē-* takes the asigmatic reinforcer -*a*). It comes as no surprise that the form *sjá* is older than *þessi*, since we clearly observe a leveling effect: *sjá*, being closely related to the *s*-initial Dem forms *sá* and *sú*, began to give way to *þ*-initial *þessi* already by c. 1150 (Axelsdóttir 2003: 46). Eventually *sjá* was ousted from the paradigm completely, leaving only *þessi*.

After this brief discussion of the components involved in the development of RDem, we can return to Haugen's (1982: 100-101) worry, quoted at the beginning of this chapter, that the 'Common Scandinavian' (= PN) RDem paradigm cannot be reconstructed. Fortunately, I do not think the situation is as hopeless as Haugen believed it to be. It is actually quite straightforward to reconstruct the early stages of the PN paradigm, and from there we can infer later stages of PN as well. We can start with the earliest stage, shown in Table 15, when the reinforcer -*si* is simply appended to the regular Dem forms. This stage can be directly inferred from the archaic RN forms which were given in (7).

Table 15 Reconstructed stage (i) of PN RDem (based on Haugen 1982: 101)

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	sū-si	sā(R)-si	þat-si	þāR-si	þeiR-si	þau-si
ACC	þā-si	þan-si	þat-si	þāR-si	þā-si	þau-si
GEN	þeiRaR-si	þes-si	þes-si	þeiRa-si	þeiRa-si	þeiRa-si
DAT	þeiRi-si	þeim-si	þvī-si	þeimR-si	þeimR-si	þeimR-si

Not long after this, *i*-umlaut from *-si* would undoubtedly have begun its work, fronting the root vowels of the forms in Table 15. Furthermore, the asigmatic reinforcer with the shape *-a* was innovated around this time in PN. The new reinforcer *-a* intrudes into a significant portion of the paradigm in Table 15, which we know from looking at forms in both transitional and classical Norse. Based on the available evidence, the asigmatic reinforcer *-a* most likely would have affected the region shaded in Table 16, where I have reconstructed stage (ii)/(iii) of the PN RDem paradigm.

Table 16 Reconstructed stage (ii)/(iii) of PN RDem

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	sē-a	sē-a	þet-s-a	þēʀ-si	þeiʀ-si	þey-si
ACC	þē-si	þen-s-a	þet-s-a	þēʀ-si	þē-si	þey-si
GEN	þeiʀaʀ-si	þes-s-a	þes-s-a	þeiʀa-si	þeiʀa-si	þeiʀa-si
DAT	þeiʀi-si	þeim-s-a	þvī-s-a	þeimʀ-si	þeimʀ-si	þeimʀ-si

The shaded area in Table 16, representing the extent to which the asigmatic reinforcer had intraparadigmatically spread in PN, is based on the following considerations:

- the classical ON forms that display the *a*-reinforcer,
- the F/M.NOM.SG classical form *sjá*, which must come from a PN **sē-a*, and
- the archaic/poetic forms M.DAT.SG / PL.DAT *þeim-a* and N.DAT.SG *því-s-a*.

Note in Table 16 that I have reconstructed **þeim-s-a* instead of (skaldic) *þeim-a* to keep it on a par with the surrounding forms that display cooccurrence of sigmatic *-s* and asigmatic *-a*.[16] It is unlikely that **þeimsa* would have been the DAT.PL form as well, since (as discussed throughout this chapter for various languages) the M.DAT.SG form was typically influenced by the plural rather than the other way around.

Shortly after the stage shown in Table 16, the PN paradigm would have begun to develop its RDem stem *þess-* as sketched above, by assimilation of **ʀs* to *ss* and the strong tendency to move towards external inflection. On the basis of the new stem *þess-*, moreover, gemination – a form of non-sigmatic/secondary reinforcement – would have become generalized throughout the paradigm, turning older forms like *þensa* and *þetsa* into *þenna* and *þetta*. The original domain of the asigmatic reinforcer *-a* eventually shrunk as well, as *þeim(s)a* and *þvísa* were regularized to *þessum* and *þessu*, and *sjá* was

[16] As will become clear in Chapter 6, I take *sē-* to be a portmanteau encompassing the D, K, and sigmatic reinforcer ingredients. This is why I do not reconstruct **sē-s-a*.

regularized to *þessi*. The end result of this development is, of course, the classical, stage (iv) ON paradigm already given in Table 13.

1.2.2 Old Frisian

The OF RDem paradigm is shown in Table 17. The K/K_D endings for RDem are given in Table 18.

Table 17 OF RDem (stem *thiss-*) (Bremmer 2009: 55, Markey 1981: 136; Hewett 1879: 59 gives *a*-final GEN.PL)

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	th-iu-s	th-i-s	thi-t	thiss-e	thiss-e	thiss-e
ACC	thiss-e	thiss-en	thi-t	thiss-e	thiss-e	thiss-e
GEN	thiss-er	thiss-es	thiss-es	thiss-er(a)	thiss-er(a)	thiss-er(a)
DAT	thiss-er	thiss-em	thiss-em	thiss-em	thiss-em	thiss-em

Table 18 OF strong adjective endings, with relevant pronominal inflection (Markey 1981: 125-126, 135; Bremmer 2009: 54 for *thī*)

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	-Ø [Dem th-**iu**]	-Ø [Dem th-**ī**]	-Ø [Dem the-**t**]	-e	-e	-e
ACC	-e	-en	-Ø [Dem the-**t**]	-e	-e	-e
GEN	-er(e)	-es	-es	-era	-era	-era
DAT	-er(e)	-a/-e [*-em* < PL]	-a/-e [*-em* < PL]	-a/-e [*-em*]	-a/-e [*-em*]	-a/-e [*-em*]

(i) RDem stem

All of the boxed forms in Table 17 show the RDem stem *thiss-*. The OF RDem stem shows extensive gemination of the sigmatic reinforcer, i.e. *thi-s-C-* > *thiss-* (reminiscent of ON *þess-*). As we will see below, OF is like OE in that it has adopted an innovatory *i*-vocalism in its RDem base *thi-*.

(ii) Position of inflection

The forms which are boxed in Table 17 have external inflection, since the K morpheme, provided in Table 18, appears in the rightmost position in these forms, to the right of the sigmatic reinforcer component: F.ACC.SG / NOM/ACC.PL *thiss*-***e***, M.ACC.SG *thiss*-***en***, F.GEN/DAT.SG *thiss*-***er***, M/N.GEN.SG *thiss*-***es***, M/N.DAT.SG / DAT.PL *thiss*-***em***, GEN.PL *thiss*-***er(a)***.

The forms F.NOM.SG *th*-***iu***-*s* and M.NOM.SG *th*-***i***-*s* have internal inflection (though see (iii) below for more on the latter form), meaning that the inflectional component is to the left of the (non-geminated) sigmatic reinforcer.

Finally, we may note that the status of N.NOM/ACC.SG *thit* is not perfectly clear at this point. While it appears to have external inflection (*thi*-***t***), it does not take the *thiss*- stem like the other externally inflected forms (which are boxed in Table 17), nor does it show any sign of the sigmatic reinforcer.

(iii) Type of inflection

Most of the inflection in the OF RDem paradigm originates in the strong adjective (K) paradigm, in fact all of the boxed forms in Table 17 take adjectival K from Table 18. However, a portion of the forms display pronominal (K_D) inflection: F.NOM.SG *th*-***iu***-*s*, which does not take K *-Ø* but K_D *-iu*, and N.NOM/ACC.SG *thi*-***t***, which does not take K *-Ø* but K_D *-t*. Again, it is exactly the non-boxed forms which take K_D (see Section 6.1.5).

As for OF M.NOM.SG *this*, it could of course be argued that null K inflection is active here, giving either *thi-s-**Ø*** or *thi-**Ø**-s*. However, considering OF's close affinity to OE (and to OS), it is more likely that OF *th-i-s* is the same as OE *þ-ē-s* (i.e. basically Dem *th-**ī***, with K_D inflection, plus the sigmatic reinforcer *-s*). The only thing standing in the way of this assumption is that the vowel is short in OF RDem *this* but long in OF Dem *thī*. This, however, is not so surprising once we consider that OE *þēs* could also surface in short-voweled form as *þes* (e.g. Campbell 2003: 291). Similarly, the OE M.NOM.SG Dem form *se(:)* also displayed varying length, just like OS M.NOM.SG *the(:)* and *se(:)* (Prokosch 1939: 276, Nielsen 2000: 223).[17] Thus it is quite plausible that OF *this* shows vowel shortening, accounting for the disparity in vowel length between RDem *this* and Dem *thī* in OF. See also Ringe & Taylor (2014: 102) on internal inflection in OF.

Finally, there are two discrepancies between the inflection found in the RDem paradigm in Table 17 and the endings provided in Table 18. First, the OF dative ending *-em*[18] requires some extra explanation. The ending *-em* belongs to an older stage of the language, having eventually been supplanted by *-a*/*-e* (Markey 1981:

[17] Prokosch (1939: 268) claims that sentential stress dictates the vowel length in OE and OS.

[18] Hewett (1879: 56) reports that most dialects display *-em*, but that the Rustringer dialect shows *-on* and the Brockmer dialects shows *-um*.

125-126). The OF RDem forms take this older ending (*thiss-em*) instead of the younger *-a/-e* ending. Moreover, it is likely that the ending *-em* started out in the DAT.PL only but then intruded into the singular as well, resulting in M/N.DAT.SG *thiss-em*.[19] To sum up, *-em* (variants include *-um* and *-on*; Markey 1981: 125-126) is technically adjectival (K), but an older, synchronically opaque kind of K. For this reason the forms inflected with *-em* have been shaded in Table 17. See Section 6.1.1 for more discussion.

The second discrepancy is that the GEN.PL RDem forms in Table 17 show the K ending *-er* even though they are expected to show *-era* according to Table 18. This discrepancy can be explained in terms of an emerging syncretism with the F.GEN/DAT.SG ending *-er(e)*. Indeed, in modern West Frisian, where only remnants of the old case system survive, there is a GEN/DAT syncretism in *-er* (cf. Sipma 1913: 60, 62; Tiersma 1999: 44). The OF RDem form *thisser* apparently anticipates this eventual conflation. To put it simply, these case endings are in flux.

(iv) Non-sigmatic reinforcement

OF displays prevalent use of the consonant geminator *-C*. As mentioned in (i), the RDem stem is *thiss-*, i.e. *thi-s-**C**-*, in all of the boxed forms in Table 17. As will be seen in the next sections, gemination of the sigmatic reinforcer *-s* is much more prevalent in the OF RDem paradigm than in the RDem paradigms of OE, OS, and OHG.

1.2.3 Old English

The OE RDem paradigm is shown in Table 19. The K/K_D endings for RDem are given in Table 20.

---

[19] As touched on above, ON, OE, and OF show plural forms intruding into the singular (ON M.DAT.SG / DAT.PL *þeim*, *þessum*; OE M/N.DAT.SG / DAT.PL *þām*/*þǣm*, *þis(s)um*; OF M/N.DAT.SG / DAT.PL *thā(m)*, *thissem*), as opposed to OS and OHG where these stayed distinct (OS M/N.DAT.SG *themu*, *thesemu* vs. DAT.PL *thēm*, *thesum*, and OHG M/N.DAT.SG *dëmu*, *dësemu* vs. DAT.PL *dēm*, *dësēm*). The DAT.PL of Dem can be reconstructed as PIE **toimis* > PGmc **þaimiz* (> ON *þeim*, OE *þām*/*þǣm*, OF *thā(m)*, OS *thēm*, OHG *dēm*) while the DAT.SG can be reconstructed as PIE **tesmō* > PGmc **þezmō* (> OS *themu*, OHG *dëmu*) (Markey 1981: 136, Prokosch 1939: 269, 271).

Table 19 OE RDem (stem *þis(s)-*) (Lass 1994: 145, Campbell 2003: 291)

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	þ-ēo-s	þ-e(:)-s	þis	þ-ā-s	þ-ā-s	þ-ā-s
ACC	þ-ā-s	þis-ne	þis	þ-ā-s	þ-ā-s	þ-ā-s
GEN	þis-re > þisse	þiss-es	þiss-es	þis-ra > þissa	þis-ra > þissa	þis-ra > þissa
DAT	þis-re > þisse	þiss-um	þiss-um	þiss-um	þiss-um	þiss-um

Table 20 OE strong adjective endings, with relevant pronominal inflection (Campbell 2003: 262, 290)

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	-Ø [Dem þ-**ēo**]	-Ø [Dem þ-**e(:)**]	-Ø	-e, -a [Dem þ-**ā**]	-e [Dem þ-**ā**]	-Ø [Dem þ-**ā**]
ACC	-e [Dem þ-**ā**]	-ne	-Ø	-e, -a [Dem þ-**ā**]	-e [Dem þ-**ā**]	-Ø [Dem þ-**ā**]
GEN	-re	-es	-es	-ra	-ra	-ra
DAT	-re	-um	-um	-um	-um	-um

(i) RDem stem

The RDem stem in this language is *þis(s)-*, seen in the forms which are boxed in the paradigm in Table 19. The RDem stem *þis(s)-* does not show the same vocalism as the OE Dem stems *þæ(:)-* or *þa(:)-*. The *i*-vocalism in the RDem base *þi-* can be thought of as being analogically based on another pronominal stem in OE, namely *hi-* 'he, she, it' (< PIE **ḱi-*, a variant of PIE **ḱo-*; see Watkins 2000: 43, Fortson 2004: 130, Beekes 2011: 226 for PIE), a hypothesis which can be traced back to Kieckers (1917-1920) and which is also mentioned in Ringe & Taylor (2014: 102). Furthermore, the morpheme *-C* only appears to geminate the sigmatic reinforcer in a subset of the OE RDem forms, namely M/N.GEN.SG *þisses* and M/N.DAT.SG / DAT.PL *þissum*, where the stem can be represented as *þi-s-C-* > *þiss-*. Campbell (2003: 183, 291-2) claims that the form without gemination (*þises*, *þisum*) arises if unstressed, meaning that the form with gemination (*þisses*, *þissum*) is the underlying form.

The rest of the boxed forms have the stem *þis-* with no geminator present. Note that sometimes we see geminated *s* in these forms due to assimilation with a neighboring *r*-initial inflectional ending, since *sr* > *ss* in OE: F.GEN/DAT.SG *þis-re* > *þisse* and GEN.PL *þis-ra* > *þissa* (reminiscent of the ON rule *ssr* > *ss* discussed above). To be clear, these forms do not show the RDem stem *þiss-* but rather the stem *þis-*.

(ii) Position of inflection

The forms which are boxed in Table 19 have external inflection, since the K component, provided in Table 20, appears in the rightmost position in these forms: F.GEN/DAT.SG *þis-**re*** (> *þisse*), M.ACC.SG *þis-**ne***, M/N.GEN.SG *þiss-**es***, M/N.DAT.SG / DAT.PL *þiss-**um***, and GEN.PL *þis-**ra*** (> *þissa*).

The forms which are not boxed in Table 19 show the inflectional component to the left of the sigmatic reinforcer: F.NOM.SG *þ-**ēo**-s*, M.NOM.SG *þ-**e(:)**-s*, and F.ACC.SG / NOM/ACC.PL *þ-**ā**-s*. In other words, these forms are internally inflected and can be considered holdovers from the archaic Dem-*si* stage. See also Ringe & Taylor (2014: 102) on this.

The position of inflection in N.NOM/ACC.SG *þis* is ambiguous. Since the N.NOM/ACC.SG K ending is null, the form could have internal inflection with K to the left of the sigmatic reinforcer (*þi-**Ø**-s*) or external inflection with K to the right of the sigmatic reinforcer (*þi-s-**Ø***). For the sake of simplicity, the form is for now assumed to have external inflection and is therefore boxed in Table 19. Nonetheless, Nielsen (2000: 211, 158) considers N.NOM/ACC.SG *þis* to be an internally inflecting form, assuming that *þis* derives from **þi-t-se* (citing Kluge 1920). This issue will be discussed more in Section 6.1.2.

(iii) Type of inflection

Most of the inflection in the OE RDem forms in Table 19 originates in the strong adjective (K) paradigm from Table 20. However, a portion of the forms display pronominal (K_D) inflection: F.NOM.SG *þ-**ēo**-s* does not show K *-Ø* but K_D *-ēo*, F.ACC.SG *þ-**ā**-s* does not show K *-e* but K_D *-ā*, M.NOM.SG *þ-**e(:)**-s* does not show K *-Ø* but K_D *-e(:)*, and in the NOM/ACC.PL of all genders the form *þ-**ā**-s* does not show K *-e/-a/-Ø* but K_D *-ā*. In fact, it is exactly the non-boxed forms which show K_D (see Section 6.1.5). OE therefore displays a mix of K and K_D inflectional types.

(iv) Non-sigmatic reinforcement

In addition to the sigmatic reinforcer *-s* from NWGmc **-si*, OE displays the additional reinforcement strategy of consonant gemination. As mentioned in (i) above, M/N.GEN.SG *þisses* and M/N.DAT.SG / DAT.PL *þissum* show gemination of the sigmatic reinforcer *-s* by means of the morpheme *-C*, i.e. *þi-s-**C**-es* > *þisses* and *þi-s-**C**-es* > *þissum*. It will be noticed that gemination is much more limited in the OE

RDem paradigm than in the ON (or OF) paradigm, where gemination is omnipresent.

1.2.4 Old Saxon

The OS RDem paradigm is shown in Table 21. The K/K_D endings for RDem are given in Table 22.

Table 21 OS RDem (stem *thes-*) (Rauch 1992: 196, Cathey 2000: 37; also Gallée 1910: 240; see Nielsen 2000: 158 for *thitt*)

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL	
NOM	th-iu-s	*the-s-e	thi-t-t (OS_A) thi-t (OS_B)	the-s-e	the-s-e	the-s-e	th-iu-s
ACC	the-s-a	the-s-an	thi-t-t (OS_A) thi-t (OS_B)	the-s-e	the-s-e	the-s-e	th-iu-s
GEN	the-s-ara	the-s-es	the-s-es	the-s-aro	the-s-aro	the-s-aro	
DAT	the-s-aru	the-s-umu	the-s-umu	the-s-um	the-s-um	the-s-um	

Table 22 OS strong adjective endings with relevant pronominal inflection (Rauch 1992: 199, 194; Cathey 2000: 38; also Gallée 1910: 221, 238)

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL	
NOM	-Ø [Dem th-**iu**]	-Ø [Dem th-**e**(:)]	-Ø [Dem tha-**t**]	-e	-e	-e	-Ø [Dem th-**iu**]
ACC	-a	-an	-Ø [Dem tha-**t**]	-e	-e	-e	-Ø [Dem th-**iu**]
GEN	-era	-es	-es	-aro	-aro	-aro	
DAT	-eru	-umu	-umu	-um	-um	-um	

(i) RDem stem

All of the boxed forms in Table 21 show the RDem stem *thes-*. The OS RDem stem, then, does not display any gemination of the sigmatic reinforcer, in contrast with the RDem stems of ON, OE, and OF. Important to note is also that the OS RDem stem *thes-* shows *e*-vocalism in the base component *the-*.

(ii) Position of inflection

The forms which are boxed in Table 21 have external inflection, since the K component, provided in Table 22, appears to the right of the sigmatic reinforcer component: F.ACC.SG *thes*-***a***, F.GEN.SG *thes*-***ara***, F.DAT.SG *thes*-***aru***, M.ACC.SG *thes*-***an***, M/N.GEN.SG *thes*-***es***, M/N.DAT.SG *thes*-***emu***, NOM/ACC.PL *thes*-***a***, GEN.PL *thes*-***aro***, DAT.PL *thes*-***um***.

There are also examples of internal inflection in OS. It is clear that the F.NOM.SG / N.NOM/ACC.PL form *th*-***iu***-*s* is a representative of internal inflection, with the inflectional component *-iu* to the left of the sigmatic reinforcer.[20] See also Ringe & Taylor (2014: 102) on this. Observe, however, that the N.NOM/ACC.SG OS$_A$ variant *thitt* also displays internal inflection, since the inflectional component *-t* is to the left of the geminator reinforcer: *thi*-***t***-*C*.

The M.NOM.SG form **these* is reconstructed with external inflection, i.e. *the*-*s*-***e***, but see Section 6.1.3 for more discussion.

(iii) Type of inflection

Most of the inflection in the OS RDem paradigm originates in the strong adjective (K) paradigm, in fact all of the boxed forms in Table 21 take adjectival K from Table 22. However, a portion of the forms display pronominal (K$_D$) inflection: F.NOM.SG / N.NOM/ACC.SG *th*-***iu***-*s* does not show K *-Ø* but K$_D$ *-iu*, and N.NOM/ACC.SG *thi*-***t****(t)* does not show K *-Ø* but K$_D$ *-t*. Once again, it is exactly the non-boxed forms which take K$_D$ (see Section 6.1.5). Like OE and OF, then, OS shows a mixture of K and K$_D$ in its RDem paradigm.

It may also be noted here that though F.GEN.SG *-era* and F.DAT.SG *-eru* in the adjectival paradigm appear to differ slightly from the RDem endings *-ara* and *-aru*, there is little question that the RDem endings are of adjectival origin.[21] OS orthography displays a massive amount of variation, including writing <e> for <a> before a non-<i> in the next syllable or vice versa, that is, writing <a> for <e> before a non-<i> in the next syllable (Rauch 1992: 134, 131-2). Prokosch (1939: 276), moreover, notes that "interchange between *e* and *a* is common before *r*." See also Twaddell (1963) for discussion.

---

[20] *Thius* is of course directly cognate with OE *þēos*. WGmc **iu* is lowered to *ēo* in OE but not in OS.

[21] Rauch (1992: 196-197): "If the *thius* and *thit* forms are understood as having ø-suffix, then they, as well as the trigender *-e* plural of the intensified deictic pronoun, show this paradigm to be isomorphic with the strong adjective suffix". I will reconsider *thius* and *thit* in Chapter 6.

(iv) Non-sigmatic reinforcement

OS does not show gemination of the sigmatic reinforcer. Nevertheless, the consonant geminator *-C* is present in the OS RDem paradigm, since the N.NOM/ACC.SG OS_A variant *thitt* can be decomposed as *thi-t-**C***. In this case it is not the sigmatic reinforcer *-s* which is being geminated but rather the K_D ending *-t*. Note also the higher than expected stem vocalism in *thit(t)*, namely *i* instead of expected *e* (expected given that the rest of the OS RDem paradigm shows the base *the-*).

1.2.5 Old High German

The OHG RDem paradigm is shown in Table 23. The K/K_D endings for RDem are given in Table 24. An important detail about Table 24 that needs to be mentioned at this point is that all of the nominative slots in both the singular and plural, in addition to the N.ACC.SG, may show the ending *-Ø* (Wright 1906: 56). These null inflections are known as the 'short endings', as opposed to the endings given in Table 24, which are known as the 'long endings'. In fact the short endings correspond to an earlier stage, while the long endings arose at a later stage.[22] In this dissertation I consider the long endings only. The reason for shading in the N.NOM/ACC.SG slots will become clear below.

Table 23 OHG RDem (stem *dës(s)-*) (Braune & Reiffenstein 2004: 250; Wright 1906: 67)

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	dës-iu	dës-ēr	diz < *þitt	dës-o	dës-e	dës-iu
ACC	dës-a	dës-an	diz < *þitt	dës-o	dës-e	dës-iu
GEN	dës-era	dëss-es	dëss-es	dës-ero	dës-ero	dës-ero
DAT	dës-eru	dës-emu	dës-emu	dës-ēm	dës-ēm	dës-ēm

[22] The long endings were modeled on the pronominal endings, but for our purposes they are still instantiations of K, not K_D.

Table 24 OHG strong adjective endings with relevant pronominal inflection (Braune & Reiffenstein 2004: 220; also Wright 1906: 55-6)

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	-iu	-ēr	-aʒ < *-at [Dem da-ʒ < *þa-**t**]	-o	-e	-iu
ACC	-a	-an	-aʒ < *-at [Dem da-ʒ < *þa-**t**]	-o	-e	-iu
GEN	-era	-es	-es	-ero	-ero	-ero
DAT	-eru	-emu	-emu	-ēm	-ēm	-ēm

(i) RDem stem

Most of the forms in Table 23 show the RDem stem *dës-* (PGmc **þ* > *d* by a late stage of the High German Consonant Shift). The OHG RDem stem, for the most part, does not display gemination of the sigmatic reinforcer. The one form which does seem to show a geminated stem *dëss-* (i.e. *dë-s-C-*) is M/N.GEN.SG *dëss-es*.[23] Thus the RDem stem of OHG should be given as *dës(s)-* (or *dis(s)-* in later OHG; Prokosch 1939: 272). Important to note is also that the OHG RDem stem *dës-* shows *e*-vocalism in the base component *dë-*, as was also the case in OS.

(ii) Position of inflection

The forms which are boxed in Table 23 have external inflection, since the K component, provided in Table 24, appears to the right of the sigmatic reinforcer component: F.NOM.SG *dës-**iu***, F.ACC.SG *dës-**a***, F.NOM.SG / N.NOM/ACC.PL *dës-**iu***, F.GEN.SG *dës-**era***, F.DAT.SG *dës-**eru***, M.NOM.SG *dës-**ēr***, M.ACC.SG *dës-**an***, M/N.GEN.SG *dëss-**es***, M/N.DAT.SG *dës-**emu***, F.NOM/ACC.PL *dës-**o***, M.NOM/ACC.PL *dës-**e***, GEN.PL *dës-**ero***, DAT.PL *dës-**ēm***.

While we have found instances of internal inflection in all the languages considered so far, the OHG paradigm does not contain any obvious cases of internal inflection, but see (iii) below for discussion of N.NOM/ACC.SG *diz*.

[23] While the most commonly attested OHG M/N.GEN.SG forms are *dësses* (Muspilli) and *disses* (Notker), with *s*-gemination, the non-geminated form *deses* is attested in Tatian, though it is rare (EWAhd II: 608).

(iii) Type of inflection

Most of the inflection in the OHG RDem paradigm originates directly from the strong adjective (K) paradigm, in fact all of the non-shaded boxed forms in Table 23 take adjectival K from Table 24. There is one form, however, which does not fit this generalization, and for this reason it has been set off from the rest of the paradigm with gray shading.

The form in question is N.NOM/ACC.SG *diz*. It is clear, first of all, that this form does not show the adjectival K ending *-aȥ* (< **-at*), since we do not observe something like **diaȥ* or **dëaȥ*.[24] Notice that *diz* is spelled with <z> and not <ȥ>. The latter spelling, the one seen in K *-aȥ* and K_D *-ȥ* in Table 24, represents *s(s)* which derives from word-final **t* by the High German Consonant Shift. The former spelling, the one observed in the RDem form, represents the affricate *ts*, which in this position must derive from the geminate **tt* (again as a result of the High German Consonant Shift). In other words the RDem form *diz* is pronounced [dits], with an affricate, and this form ultimately derives from the form **þitt* with geminated **tt* (see Braune & Reiffenstein 2004: 250-251). This older layer of the N.NOM/ACC.SG form shows the K_D ending **-t* (cf. Dem *da-ȥ* < **þa-t* in Table 24) which is then geminated by *-C*. All in all, the form is parsed **þi-**t**-C* > **þitt*. In other words, the OHG form *diz* is hiding a history of pronominal inflection. Not only that, the pronominal inflection is found to the left of the reinforcer *-C*, making it an internally inflected form (the only candidate for internal inflection in the OHG RDem paradigm). The proto-form **þitt*, moreover, is a direct counterpart to OS_A *thitt* from Table 21 above.[25]

Synchronically, however, it seems highly unlikely that OHG speakers analyzed the form *diz* as underlyingly /ditt/ (or the like), since the shift from **tt* > *ts* took place hundreds of years before our attested sources of OHG. For this reason *diz* has been shaded in Table 23.

Some variant forms also existed in the N.NOM/ACC.SG slot, most prevalently *dezzi* and *dizzi* (pronounced *detsi* and *ditsi*, respectively), but also spelled *dezi*, *deze*, or *dize* (EWAhd II: 608). According to the EWAhd (II: 613-614) these can be traced back to **þet-þi* > **þetti* > *detsi* and **þit-þi* > **þitti* > *ditsi*, where **-t* is the pronominal ending and **-þi* is a reinforcer (see Section 1.4.1.1). Another option is that *dezzi* and *dizzi* are re-intrusions of the old **-si* reinforcer due to contact or

---

[24] Even though this would be phonologically acceptable: witness the diphthong in F.ACC.SG Dem *dia* or *dea*.

[25] Nielsen (2000: 158) warns that many incorrectly equate OHG *diz* with the single-*t* variants of the N.NOM/ACC.SG, e.g. OS_B/OF *thit*, Dutch *dit*, Early Runic **þit** (DR IK312, 1 $U). We know about the value of <z> in OHG *diz* thanks to Isidor's spelling with <z> (= [ts] in his system) instead of <zs> (= [s] in his system) (Armitage 1911: 207, EWAhd II: 608). Modern editorial convention uses <ȥ> for *s* from word-final **t*.

‘dialect borrowing’ (Campbell 1999: 213-5, 222-4; see also the quote from H.F. Nielsen in Section 1.4.2.1). This would mean that *dezzi* and *dizzi* represent /de-t-si/ and /di-t-si/, with the old *-*si* reinforcer having survived intact. Either way these forms are special in the sense that they would have in all likelihood been morphologically opaque to speakers.

(iv) Non-sigmatic reinforcement

OHG shows gemination of the sigmatic reinforcer in M/N.GEN.SG *dëss-es*. Furthermore, the consonant geminator was also involved in the history of the N.NOM/ACC.SG *diz*, whose proto-form was **þitt*, i.e. **þi-t-**C***, where *-C* geminates the K_D ending *-*t*. Note also the higher than expected stem vocalism in *diz*, namely *i* instead of expected *ë* (expected given that the rest of the OHG RDem paradigm shows the base *dë-*), as was also the case with OS *thit(t)*.

1.2.6 Intermediate summary

Table 25 summarizes the facts discussed in Section 1.2.

Table 25 Basic RDem facts

	ON	OE	OF	OS	OHG
RDem stem (base + sigmatic reinforcer)	þe-s-C- > þess-	þi-s(-C)- > þis(s)-	thi-s-C- > thiss-	the-s-	dë-s(-C)- > dës(s)-
Internal inflection	þe-**n**-C-a þe-**t**-C-a þe-**s**-C-a	þ-**ēo**-s þ-**e(:)**-s þ-**ā**-s	th-**iu**-s th-**i**-s	th-**iu**-s thi-**t**-C	*þi-**t**-C (> *diz*)
External inflection	remaining forms (*þessi*: TBD)	remaining forms	remaining forms	remaining forms	remaining forms
Inflection type	K	K & K_D	K & K_D	K & K_D	K & *K_D
Non-sigmatic reinforcement	*-C* *-a*	*-C*	*-C*	*-C*	*-C*

The following points emerge from the overview:

(i) With regard to the RDem stem, it is fair to say that ON, OE, and OF show a significant tendency towards geminating the sigmatic reinforcer, while OS and OHG do not. In OS there is no evidence for gemination in the stem, while in OHG there is only a single form, M/N.GEN.SG *dësses*, which shows this kind of gemination.

(ii) As for the position of K, all of the languages have a small subset of forms which retain inflection to the left of a reinforcer (internal inflection). Generally speaking, however, RDem forms have inflection to the right of reinforcers (external inflection).

(iii) As for the type of inflection (adjectival K vs. pronominal K_D), there is a fair bit of variation. ON shows purely adjectival (K) inflection, while WGmc shows a mixture of adjectival (K) and pronominal (K_D) inflection, though adjectival inflection is always in the majority. OHG displays a layered system: in addition to making use of the 'long endings' (which are technically K but historically derived from K_D; see fn. 22) in the NOM.SG/PL and N.ACC.SG slots, the paradigm contains the opaque form *diz* in the N.NOM/ACC.SG which derives from **þitt* with pronominal inflection *-*t*. Due to the more diachronic nature of pronominal inflection in the OHG RDem paradigm I have put an asterisk in front of OHG's K_D in Table 25.

(iv) Finally, in addition to the sigmatic reinforcer -*s* which derives from NWGmc *-*si* and is found in every language's RDem stem, there is also a consonant geminator morpheme -*C* found in all of the languages. Its distribution, however, is highly variable crosslinguistically. The asigmatic reinforcer -*a* is found in ON only, not in any of the WGmc languages.

1.3 Notes on adjectival vs. pronominal inflection

As discussed above and as can be seen in Table 25, RDem forms may sometimes use the endings of the demonstrative pronoun system (K_D), but more often the forms take the endings of the strong adjective system (K). This section will briefly discuss the history of the strong adjective endings. Though the strong adjective endings are, historically speaking, very closely related to the pronominal endings, the two systems have become synchronically distinct by the time of ON, OE, OF, OS, and OHG. This is not a controversial claim – hopefully the reader will have reached that conclusion already simply on the basis of the paradigms above – but it is worth making explicit.

In PIE, there were two systems of declension: nominal and pronominal. Adjectives did not have a separate declension class: for the most part they declined according to the nominal pattern. Germanic, however, developed specific adjectival systems of declension. More precisely, Germanic innovated 'strong' (indefinite) and 'weak' (definite) declensions for adjectives. This innovation is datable to the PGmc period since all branches show the strong/weak distinction.

It is widely accepted that the weak declension was based on the endings of the nominal *n*-stems. As for the strong declension, until recently the general concensus was that it was a hybrid, based on a combination of nominal (*a*- and *ō*-stems) and pronominal endings (see, for example, Lass 1994: 146-147). However McFadden (2004) has convincingly demonstrated that the entire strong adjective system originates from the pronominal declension, with no need to assume that nominal *a*- and *ō*-stems were ever involved.

This interaction between the strong adjective endings and the pronominal endings, however, is a diachronic issue and belongs to the prehistory of Germanic. By the time of the daughter languages, the two inflectional systems have become synchronically distinct. There may be syncretic overlaps between the pronominal endings and the strong adjective endings, but there are enough non-syncretic endings to motivate positing synchronically separate inflectional classes, i.e. K_D vs. K.

Take ON as an example. In (9) I have excluded what I take to be irregular forms: F.NOM.SG *sú*, M.NOM.SG *sá*, and N.DAT.SG *því* or *þý*. (9a) lists the syncretisms between the K_D and K systems; (9b) lists the endings that are distinct between the two.

(9) (a) Syncretic

	K_D / K
F.GEN.SG	-rar
F.DAT.SG	-ri
M.ACC.SG	-n(n)
N.NOM/ACC.SG	-t
M/N.GEN.SG	-s(s)
GEN.PL	-ra

(b) Distinct

	K_D	K
F.ACC.SG / M.ACC.PL	-á	-a
M.DAT.SG / DAT.PL	-m	-um
F.NOM/ACC.PL	-ær, -ár	-ar
M.NOM.PL	-r	-ir
N.NOM/ACC.PL	-u	-Ø$^{(u)}$

Details aside, it is undeniable that there are two distinct systems here. A shared prehistory in PGmc and a few syncretisms can do nothing to change that fact.

From the survey of the languages above it should be clear that the inflection of RDem has undergone a shift in classification between the time of the older NWGmc Dem-*si* stage, when pronominal inflection (K_D) reigned, and the time of the daughter languages (ON, OE, OF, OS, and OHG), when adjectival inflection (K) has supplanted most of the pronominal inflection. Nevertheless forms remain in the daughter languages which are holdovers from the Dem-*si* era, in the sense that they show internal inflection of a pronominal nature.

1.4 Etymologies of the reinforcers

In this section I provide some philological discussion of the reinforcers encountered above. Since Section 1.4 is only intended to give a more well rounded view of RDem and its historical development, it is possible to skip this section in its entirety and still understand the rest of the dissertation.

In the discussion above the reinforcer *-*si* was seen to be the crucial catalyst in the historical formation of NWGmc RDem. Two other reinforcers were also identified which arose after the NWGmc Dem-*si* stage, namely the consonant geminator -*C*, found in both NGmc and WGmc, and -*a*, found only in NGmc. All three have distinct etymologies. Here I will briefly survey these etymologies. Section 1.4.1 discusses the sigmatic reinforcer, Section 1.4.2 the consonant geminator, and Section 1.4.3 the asigmatic reinforcer.

1.4.1 The sigmatic reinforcer

1.4.1.1 Two hypotheses for the etymology of *-*si***

One hypothesis for the etymology of *-*si* is that it should be identified with the imperative verb **se*/**si* 'see! look!' (OED, P-Z: 3295; Feist 1939: 403 attributes the verbal etymology ultimately to J. Grimm). Informally speaking, the idea is that a structure like *that-see* 'look at that!' developed into *this*, where the -*s* is actually cognate with the verb *see*.

The detailed development of *-*si* according to the verbal hypothesis is illustrated in Figure 5.

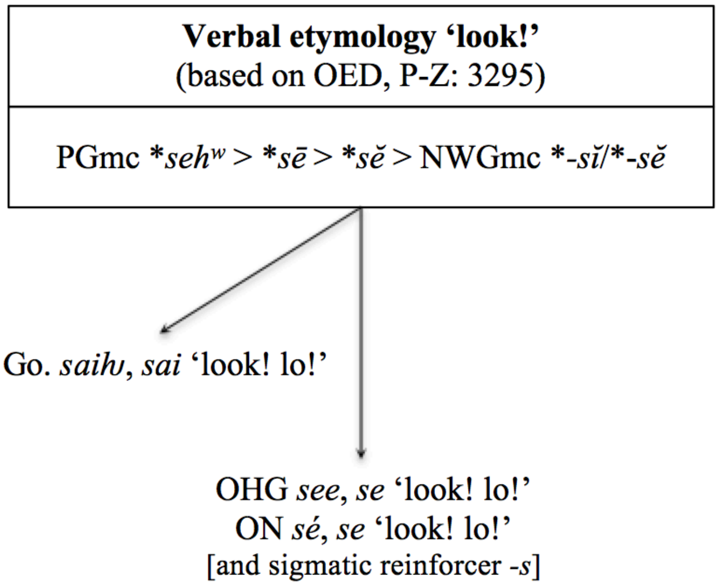


Figure 5 The verbal etymology for *-*si*

The infinitive of the verb 'see' in PGmc was *seh^w-*an*. Taking the stem of this, *seh^w, gives us the imperative singular form 'see!', which is the starting point for the etymology in Figure 5. From here the idea is that the imperative *seh^w grammaticalized into a discourse particle. Part of this grammaticalization involved phonological reduction, whereby the final labiovelar consonant h^w was lost. This deletion led to compensatory lengthening of the vowel, giving *sē*. As the grammaticalization process continued, this vowel was reduced as well, giving *se*. By the time we reach the NWGmc stage, the particle has become a clitic *-*si*/*-*se*. Note that clitics are unstressed. The unstressed vowel system of NGmc was a three-way contrast between **i*, **a*, and **u* (Haugen 1982: 29), so here the unstressed **e* in **se* shifted to **i*, giving PN *-*si*. In WGmc both *-*se* and *-*si* existed.

Evidence for this development – from imperative verb to discourse particle – can be found in all three branches of Germanic. Though Gothic (EGmc) never fully followed through by developing an RDem of its own, it still shows the full imperative form *saiƕ* (<ai> = *ε* before h^w) 'see!', as well as (according to the verbal hypothesis) a reduced discourse particle *sai* [sε]. The same pair of full and reduced versions of 'see!' can be seen in OHG (WGmc) and ON (NGmc), in addition, of course, to the sigmatic reinforcer -*s* (< *-*si*) which is an integral part of the NGmc and WGmc RDem paradigms.

In other words, according to Figure 5 we have a pan-Germanic element *$se(h^w)$: all three branches of Germanic show evidence for an imperative 'see!' and a reduced discourse particle grammaticalized from the full imperative. Indeed, the grammaticalization of imperative verbs into interjections or discourse particles is a very common phenomenon crosslinguistically speaking. See Haegeman & Hill (2013) for contemporary West Flemish and Romanian examples. Consult Derolez & Simon-Vandenbergen (1988) and Haegeman (2014) for more on West Flemish, as well as notes on verbal interjections in OE, Italian, Swiss German, Latin, and Greek. Feist (1939: 403) also provides examples from Finnish, Lithuanian/Latvian, and Icelandic. See also Tanghe & Jansegers (2014) for a study of Spanish and Italian discourse markers derived from verbs of perception.

A second hypothesis is that **-si* has its origins in a locative-pronominal element with a meaning like 'there' or 'here' (see the EWAhd II: 608-17; also Feist 1939: 403, citing proposals by Meyer 1869 and Osthoff 1901). The detailed development of **-si* according to the locative-pronominal hypothesis is illustrated in Figure 6.

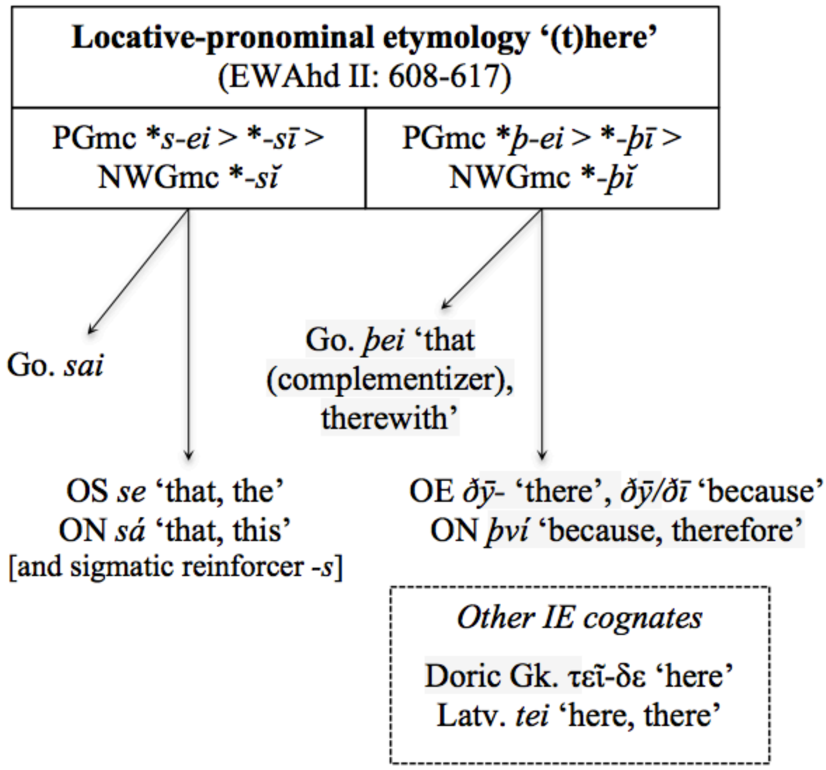


Figure 6 The locative-pronominal etymology for **-si*

According to the locative-pronominal etymology, there was a locatival component **-ei* in PGmc which could combine with the pronominal roots **s-* (M/F.SG) and **þ-* (other/oblique). Thus our starting points here are the items **s-ei* and **þ-ei*. From here the PGmc diphthong **ei* changed to a long **ī* in late PGmc (Antonsen 2002: 28), yielding **sī* and **þī*. Eventually, due to grammaticalization and cliticization of the **sī* and **þī* particles, the vowel is shortened. The result is **-si* and **-þi*.

Evidence for this development appears to be found in all three branches of Germanic. For the item **þei*, the EWAhd provides Go. *þei* 'that, therewith' for EGmc, OE *ðȳ-* 'there' and *ðȳ/ðī* 'because' for WGmc, and ON *því* 'because, therefore' for NGmc, in addition to some relevant non-Germanic cognates from Doric Greek and Latvian. For the item **sei*, however, the evidence is scantier. For this component of the locative-pronominal hypothesis, the EWAhd provides the M.NOM.SG demonstratives *se* from OS and *sá* from ON. While there is no doubt that these derive from a PIE pronominal root **s-*, the claim here is, more precisely, that OS *se* and ON *sá* support the reconstruction of a supposed locatival **sei*. This leap is not very well grounded, since *se* and *sá* do not derive from **sei* but rather from PGmc M.NOM.SG **sa*.[26] This casts doubt on the EWAhd's etymology in Figure 6. As shown in Figure 7, this leaves Go. *sai* as the only item left over to support the reconstruction of a PGmc locative **sei*, which is the item that supposedly gives *-*si* in NWGmc.

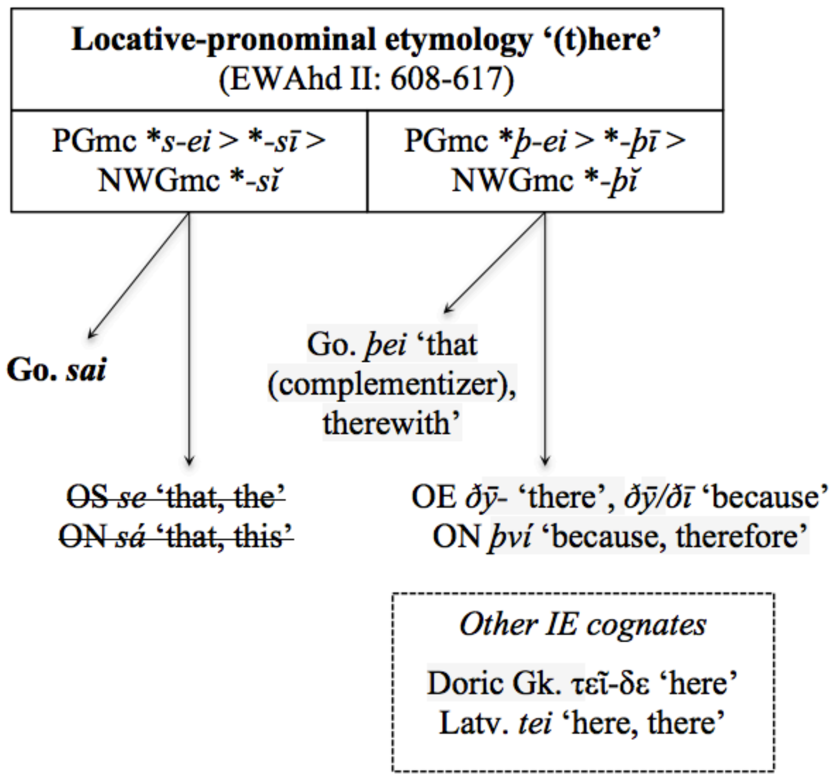


Figure 7 The locative-pronominal etymology without OS *se* and ON *sá*

When all is said and done, the identity of Go. *sai* is the keystone for both the verbal hypothesis and the locative-pronominal hypothesis. The verbal etymology considers *sai* an attestation of a discourse particle derived from the imperative *saiƕ*, while the locative-pronominal etymology presses *sai* into service as the continuation of a PGmc locative with the pronominal root **s*-. Since both hypotheses are vying for the support of *sai*, it is critically important to investigate the identity of Go. *sai* in order to see which hypothesis it actually supports. The 'losing' hypothesis will have to forfeit *sai*, which deals a serious blow to its credibility. For the verbal etymology, the forfeiture of *sai* means the loss of a verb-based discourse particle in EGmc, meaning that only two – rather than all three –

[26] Or **sai* on the basis of the front vowels in WGmc (Klingenschmitt 1987: 182), if we do not accept the hypothesis presented above about influence from the **ez* paradigm.

branches of Germanic support the reconstructed grammaticalization process from imperative to particle. For the locative-pronominal etymology, the forfeiture of *sai* means the loss of the sole evidence for an **s*-based locative in PGmc (since OS *se* and ON *sá* do not make the cut).

1.4.1.2 The identity of Gothic *sai*

Basing this section on Lander (2013), I will present a number of reasons to believe that Go. *sai* is a verbal particle/interjection rather than a continuation of a PGmc locative **sei*. My view is therefore that *sai* supports the verbal hypothesis and not the locative-pronominal hypothesis.

First, it is certainly a weak point in the locative-pronominal hypothesis that the locative was for the most part lost very early on, already in PGmc. One remnant in Gothic which survived was *þei* 'that (complementizer), therewith' (cf. Figures 6 and 7). Importantly, however, this form shows the oblique root *þ-* instead of the root *s-*. This does little to support the proposed etymology of *sai* – in fact the existence of *þei* highlights how unlikely it is that a second locative, based on the highly restricted *s*-root no less, survived into Gothic as well.

Second, the question arises why the proposed Gothic locative would show up as *sai* [sɛ] (or perhaps [saɪ]; see below) and not *sei* [sī], on a par with *þei* [þī]. The '*e*-flavored' locatival morpheme (PGmc *-*ei*) is required for the rest of the Germanic cognates given in Figure 6, and even for the Gothic form *þei*. This would make locative *sai* an odd outlier in both Gothic and Germanic as a whole. In fact, the EWAhd (II: 612) must offer a slight variation on the main hypothesis in order for the vocalism to come out right: M.LOC.SG pre-Gmc ***soī* > PGmc **saī* > Go. *sai*.

Third, it should be noted that the locative-pronominal view of Go. *sai* is similar to an older hypothesis from Meyer (1869), cited in Feist (1939: 403). Meyer's idea is that *sai* is a Gothic-internal innovation whereby the M.NOM.SG pronoun *sa* is suffixed by a deictic particle *-i* (cf. Gk. νυν-ί 'now', οὑτοσ-ί 'this here'), giving *sai*. Meyer's hypothesis implies that *sai* (i.e. *sa-i*) should have a diphthong [aɪ]. Indeed, this is also what the EWAhd (II: 612) predicts with its proposed development of pre-Gmc ***soī* > PGmc **saī* > Go. *sai*. Now, while the Gothic digraph <ai> usually represents the monophthong [ɛ], it could also be used to represent a diphthong [aɪ]. However, in such cases a trema < ¨ > (indicating diaeresis/hiatus) was available in order to avoid ambiguity. Had *sai* had a diphthongal pronunciation, then Wulfila would very likely have written <saï>, but this spelling is *never* found in the original text (Luc de Grauwe, p.c.). This means *sai* should be read [sɛ] and that the Meyer/EWAhd prediction is not borne out.[27, 28]

Finally, and perhaps most convincingly, there is plenty of evidence in favor of the verbal etymology (as opposed to just evidence against the locative-pronominal etymology).[29]

Derolez & Simon-Vandenbergen (1988) point out that *sai*'s most common usage is as a translation for the Greek imperative singular ἰδού or interjection ἴδε, both of which can be translated as 'see! lo!'. I have found that out of a total of 96 instances of *sai* in the Gothic corpus, 82 of them correspond to Greek ἰδού, ἴδε, or the imperative plural ἴδετε (as for the rest of the instances of *sai*, see below and fn. 28). The distribution of these 82 cases is shown in Table 26. Relevant examples are provided in (10), (11), and (12).

Table 26 Translation of Greek forms by *sai*

Greek	***sai***
ἰδού	63/82 = 76.8%
ἴδε	18/82 = 22%
ἴδετε	1/82 = 1.2%

(10) **ἰδού**

(a) jah sai mans bairandans ana ligra mannan
and SAI *men carrying on bed.DAT man.ACC*
'And behold, men brought in a man on a bed'
καὶ ἰδοὺ ἄνδρες φέροντες ἐπὶ κλίνης ἄνθρωπον (Luke 5:18)

---

[27] Meyer's hypothesis also predicts that other forms in the Dem paradigm of Gothic should, like *sa-i*, be able to display deictic appendages, as in F.NOM.SG **so-i*, N.NOM/ACC.SG **þata-i*, M/N.DAT.SG **þamma-i*, M.ACC.PL **þans-i*, etc. Such forms are not attested in the Gothic corpus.

[28] Yet another hypothesis for Go. *sai* cited in Feist (1939: 403) is Osthoff (1901), who proposes that *sai* is the M.NOM.SG Dem pronoun *sa* plus the N.NOM/ACC.SG pronoun *ita*, on a par with Skt. *sḗd* < *sá-íd*. This hypothesis is incoherent, however. In Skt. the *í* of the neuter pronoun is conflated with the preceding *á* according to regular sandhi rules and the dental of the neuter pronoun remains completely intact (*sá-íd* > *sḗd*). In Gothic, on the other hand, one would need to claim that the dental and the following vowel are completely deleted (for no apparent reason), leaving only the initial *i* (*sa-ita* > *sa-i*). These two processes are scarcely on a par.

[29] *Project Wulfila* and Streitberg (2000) are consulted for the Gothic and the Greek. For the *sai* data I have made note of cases which are in some significant way *not* direct word-for-word translations from the Greek, as it is generally agreed upon that such cases give us the best glimpse into native Gothic syntax (to be fair, this should be taken with a grain of salt since no one knows what Wulfila's Greek *Vorlage* might have been). Seventeen cases of *sai* are of interest in this way, by my classification (17/96 = 17.7%), and they are marked by an asterisk at the end (…*). Note that in my figure of 96 I have excluded 13 occurrences of *sai*, two of which are editorial additions by W. Streitberg and 11 of which are duplicates (as Codex Ambrosianus A and B overlap to a great extent).

(b) gaggiþ, sai ik insandja izwis swe lamba in midumai wulfe.
go.2PL SAI *I send.forth you.PL as lambs in middle.DAT wolves.GEN*
'Go your ways: behold, I send you forth as lambs among wolves.'
ὑπάγετε: ἰδοὺ ἀποστέλλω ὑμᾶς ὡς ἄρνας ἐν μέσῳ λύκων. (Luke 10:3)

(11) **ἴδε**

(a) jah sai, andaugiba rodeiþ jah waiht du imma ni qiþand
and SAI *boldly speaks.3SG and anything to him not say.3PL*
'But lo, he speaks boldly and they say nothing to him'
καὶ ἴδε παρρησίᾳ λαλεῖ καὶ οὐδὲν αὐτῷ λέγουσιν (John 7:26)

(b) sai, ƕan filu ana þuk weitwodjand.
SAI *how many.things against thee witness.3PL*
'Behold how many things they witness against thee.'
ἴδε πόσα σου κατηγοροῦσιν. (Mark 15:4)

(12) **ἴδετε**

sai, hvileikaim bokom gamelida izwis meinai handau.*
SAI *what letter wrote.1SG to.you my.DAT hand.DAT*
'See what a letter I wrote to you with my own hand.'
ἴδετε πηλίκοις ὑμῖν γράμμασιν ἔγραψα τῇ ἐμῇ χειρί. (Galatians 6:11)

Let us now compare the distribution of the particle *sai* which was just discussed with the distribution of the morphologically transparent imperative forms *saiƕ* 'see.SG!', *saiƕiþ* 'see.PL!', and *saiƕats* 'see.DU!'. As for *saiƕ* 'see.SG!', there are six attested cases in the Gothic corpus. In half of these cases, the Greek counterpart is ἴδε or ἰδού, and for the other half the counterpart is ὅρα 'see.SG!' or βλέπε 'see.SG!', both imperative singulars of 'see' verbs.[30] The Greek forms which *saiƕ* is used to translate are shown in Table 27.

[30] ὁράω and βλέπω can also mean 'take heed, beware'.

Table 27 Translation of Greek forms by *saihƕ*

Greek	*saihƕ*
ἰδού (Cor. II 7:11)	1/6 = 16.7%
ἴδε (John 7:52, 11:34)	2/6 = 33.3%
ὅρα (Matt. 8:4, Mark 1:44)	2/6 = 33.3%
βλέπε (Col. 4:17)	1/6 = 16.7%

Next, the item *saihƕiþ* 'see.PL!' has 11 attested instances in the Gothic corpus, all of which are translations of Greek ὁρᾶτε 'see.PL!' or βλέπετε 'see.PL!'. This is shown in Table 28.

Table 28 Translation of Greek forms by *saihƕiþ*

Greek	*saihƕiþ*
ὁρᾶτε (Thess.I 5:15)	1/11 = 9.1%
βλέπετε (Luke 8:18, Mark 4:24, 8:15, 12:38, Cor.I 10:18, 16:10, Gal. 5:15, Phil. 3:2 (3x))	10/11 = 90.9%

There is also a single case of *saihƕats* 'see.DU!' (Matt. 9:30), which corresponds to Gk. ὁρᾶτε 'see.PL!'.

These translation data show that there is a direct link between *sai* and the morphologically transparent imperative forms of the verb *saihƕan*. As seen in Table 29, Wulfila used the particle *sai* very much like the imperative *saihƕ* in that both were used to translate ἰδού and ἴδε (though only the morphological imperative *saihƕ*/*saihƕiþ*/*saihƕats* could be used for ὁράω and βλέπω).

Table 29 Wulfila's use of *sai* in relation to imperative verb forms of *saihƕan*

sai	***saihƕ*** 'see.SG!'	***saihƕiþ*** 'see.PL!'
ἰδού	ἰδού	
ἴδε	ἴδε	
	ὅρα	ὁρᾶτε
	βλέπε	βλέπετε

The Greek-to-Gothic translation data support the view that *sai* is simply a derivative of *saihƕ*.

Next let us turn to some particularly interesting cases, including some attestations of *sai* which do not have direct counterparts in the Greek. Consider first the Gothic phrase *sai nu* 'lo now' in the examples in (13). This phrase is clearly cognate with OHG *sē-nu* and ON *sé nu*, both of which are uncontroversially verbal.

(13) (a)
sai nu ju ni sijuþ gasteis *jah* aljakonjai*
SAI *now ye not are strangers and foreigners*
'Behold now, ye are no more strangers and foreigners'
ἄρα οὖν οὐκέτι ἐστὲ ξένοι καὶ πάροικοι (Ephesians 2:19)

(b)
sai nu selein jah ƕassein [garaihta] gudis*
SAI *now goodness and severity* [*righteous*] *god.GEN*
'Behold now the goodness and severity of god'
ἴδε οὖν χρηστότητα καὶ ἀποτομίαν θεοῦ (Romans 11:22)

Note that in the Greek of (13a) there is neither a 'now' nor a 'lo'. In (13b) there is no Greek 'now'. In other words, the Gothic here is not a word-for-word translation of the Greek. This is noteworthy in Wulfila's bible, which usually follows the Greek text extremely closely. The mismatch suggests that Wulfila has inserted something native, and the fact that this insertion is exactly cognate with the WGmc and NGmc items mentioned is evidence for *sai* being an EGmc counterpart of NWGmc *-*si*.[31]

For further evidence, consider the example in (14), also pointed out by Derolez & Simon-Vandenbergen (1988), where *sai* takes an accusative object. This suggests that *sai*

[31] There are a few cases where *sai* instead follows an adverb like *nu* 'now' (iv) or *suns* 'immediately' (v).

(iv) (a)
iþ nu sai, ufkunnandans guþ…*
but now SAI *knowing god*
'But now, after you have known god…'
νῦν δὲ γνόντες θεόν… (Gal. 4:9)

(b)
iþ nu sai, jah taujan ustiuhaiþ…*
but now SAI *and do perform.2.PL.OPT*
'But now, perform the doing of it…'
νυνὶ δὲ καὶ τὸ ποιῆσαι ἐπιτελέσατε (Cor.II 8:11)

(v)
jah suns sai, ahma ina ustauh in auþida.*
and immediately SAI *spirit him drives in wilderness.ACC*
'And immediately the spirit drives him into the wilderness.'
καὶ εὐθὺς τὸ πνεῦμα αὐτὸν ἐκβάλλει εἰς τὴν ἔρημον. (Mark 1:12)

Since the adverbs are clause-initial constituents, *sai* could be taken to occupy a Wackernagel (1892) position in these cases, meaning that it would be unstressed and short-voweled (*sai* = [sɛ]). On this account a case like *þar-uh sai* 'there-and SAI' (Luke 7:12*, 7:37*) would be a 'clitic chain' (Fortson 2004: §15.43). On the other hand, in the *sai nu* cases, where *sai* precedes the adverb, it is more likely that *sai* is stressed and long-voweled (*sai* = [sɛ:]).

is a verb selecting accusative case, especially considering that in Greek the object of the sentence (ὁ τόπος) is in the NOM.SG.

(14) sai þana staþ þarei galagidedun ina*
SAI *the.ACC place where laid.3PL him*
'Behold the place where they laid him'
ἴδε ὁ τόπος ὅπου ἔθηκαν αὐτόν (Mark 16:6)

An additional example I have found of *sai* selecting accusative is (15). Note, though, that the Greek also has accusative objects here (i.e. χρηστότητα and ἀποτομίαν).

(15) sai nu selein jah ƕassein [garaihta] gudis*
SAI *now goodness.ACC and severity.ACC [righteous] god.GEN*
'Behold therefore the goodness and severity of god'
ἴδε οὖν χρηστότητα καὶ ἀποτομίαν θεοῦ (Romans 11:22)

There is also some evidence for *sai* being associated with interjective force. Consider (16), where *sai* is inserted, without a Greek counterpart, to support an exclamative.

(16) sai, ƕaiwa agluba þai faiho gahabandans in þiudangardja gudis galeiþand*
SAI *how hardly they riches having into kingdom god.GEN enter*
'How difficultly they that have riches enter into the kingdom of god!' (Mark 10:23)
πῶς δυσκόλως οἱ τὰ χρήματα ἔχοντες εἰς τὴν βασιλείαν τοῦ θεοῦ εἰσελεύσονται

If we take interrogatives to be like exclamatives in having a focus feature of some kind, a similar thing may be happening in (17). Here *sai* is inserted, again without a Greek counterpart, to support a question (note the interrogative particle *jau*).[32]

(17) sai, jau ainshun þize reike galaubidedi imma aiþþau Fareisaie?*
SAI Q *any these.GEN rulers.GEN believed him or Pharisees*
'Has any one of these rulers or Pharisees believed him?'
μή τις ἐκ τῶν ἀρχόντων ἐπίστευσεν εἰς αὐτὸν ἢ ἐκ τῶν φαρισαίων; (John 7:48)

The sentences in (16) and (17), then, place *sai* even more firmly in the verbal or clausal domain.

With the identity of Go. *sai* as a verbal interjective particle secured, we can be sure that **se(h^{w})* was a truly pan-Germanic element. This lends strong support to the verbal

[32] (16) and (17) happen to be cited together by D&SV (1988: 98-99) as well, but only in order to point out that there are no direct counterparts in the Greek. There is no mention of illocutionary force, even though this would support their claim that *sai* is an interjection.

etymology – and simultaneously undermines the locative-pronominal etymology – of the NWGmc reinforcer *-*si*.

1.4.2 The consonant geminator

In this section I discuss the diachronic origins of consonant gemination in WGmc and ON. Again, the goal in this case is to provide historical background, so the discussion will surround the various processes of sound change and analogy which gave rise to gemination rather than the synchronic status of the morpheme -*C*, which will be discussed in the rest of the dissertation.

1.4.2.1 Gemination of -*t* in the WGmc N.NOM/ACC.SG

Recall that Proto-WGmc had a N.NOM/ACC.SG RDem form **þitt* which is the ancestor of OS$_A$ *thitt* and OHG *diz*. The OHG variants *dezzi*, *dizzi* similarly derive from **þetti* and **þitti*. The main question here is why gemination of the K$_D$ ending *-*t* arose in these forms.

An approach advocated by Kluge ~~(1920)~~ is that WGmc forms like OS *thitt* and OHG *diz* originate in WGmc **þet-ja*, where **ja* is a discourse particle meaning 'verily, yes'. This produces **þitt-j(a)*, because in WGmc the glide *j* geminates an immediately preceding *t* and will also raise *e* to *i*. Kluge is, however, also forced to posit loss of the final vowel *a* in order to make the correct form emerge. The EWAhd (II: 615) rightly points out that this deletion of *a* in **þitt-j(a)* is difficult to justify.

Let us turn, then, to another possibility. Even though Section 1.4.1 was spent refuting the locative-pronominal hypothesis for the etymology of *-*si*, it is important to realize that the EWAhd's locative-pronominal hypothesis may still have a supporting role to play in the history of RDem. For one thing, we know that the NWGmc dialect continuum was very complex.

> When [the introduction of the reinforced demonstrative pronoun] took place and whether it occurred at the same time everywhere, is difficult to say… The raw material for producing new N[orth] G[ermanic]/W[est] G[ermanic] forms was available everywhere, and the innovation could have come about in more than one place. Contact is likely to have accelerated the expansion of the innovation. (Nielsen 2000: 212)

With this in mind, the possibility exists that while the locative-pronominal etymology was not the *original* source of the reinforcer *-*si*, a reanalysis along the lines of the locative-pronominal etymology could very plausibly have taken place at a later stage in the development of RDem. After all, the process of synchronic language acquisition does not have access to historical facts. In other words, even though the reinforcer began as a verbal item, it could have passed through a pronominal stage of development on its way to becoming fully integrated into the internal structure of the Dem pronoun.

Indeed, the locative-pronominal hypothesis has an elegant account for the gemination found in the WGmc forms mentioned at the beginning of this section. The main idea is that during the Dem-*si* stage, the N.NOM/ACC.SG form **þet-**si*** gave way to **þet-**þi*** by analogy (i.e. *þ...s* > *þ...þ*). Importantly, the sequence **tþ* regularly goes to *tt* in WGmc (cf. OE 3SG.PRES *wrītt* 'writes', *unttat* < *unt-þat*, OHG *untazs*; EWAhd II: 613-4, citing Th. v. Grienberger; see also Klingenschmitt 1987: 187). Consequently **þet-**þi*** becomes **þetti*, which becomes OHG *dezzi* [detsi] (for *dizzi* [ditsi] with *i*-vocalism see below). The locative-pronominal approach, then, has the advantage of being able to make use of the pronominal root *þ-*, which straightforwardly explains the gemination of *t*.

Consider now more closely the *i*-vocalism in N.NOM/ACC.SG OHG *diz*, *dizzi* and OS$_{A}$ *thitt* where we would instead expect *e*-vocalism. In Chapter 6 I will propose that this vowel difference is synchronically connected to the geminator morpheme *-C*.

Kluge ~~(1920)~~ attributes the *e*-to-*i* shift to the following **j* in the particle **-ja*, the particle which according to him is also responsible for the gemination of *-t*. Importantly, though, we see *i*-vocalism *without* gemination in Early Runic **þit**, OS$_{B}$/OF *thit*, and Dutch *dit*, so another explanation for the vowel would seem appropriate. More likely is the suggestion by Kieckers (1917-1920) and Ringe & Taylor (2014: 102) that the *i*-vocalism is due to analogy with anaphoric pronouns in **hi-* (cf. OE *hēo*, *hē*, ***hit*** 'she, he, it') < PIE **ḱi-* (see Watkins 2000: 43, Fortson 2004: 130, Beekes 2011: 226). That is, N.NOM/ACC.SG **þat* or **þet* changed to **þit* on the model of N.NOM/ACC.SG **hit*.

Note that this *i*-vocalism analogy has been contained to only the N.NOM/ACC.SG in OHG and OS, while in OE and OF the *i*-vocalism has spread much deeper into the paradigm, considering that their RDem stems are *þis(s)-* and *thiss-*. Thus, we can date the analogical influence of **hi-* on RDem to an early stage of WGmc, since it is most discernible in OE, OF, OS, and OHG.[33]

1.4.2.2 Gemination of the sigmatic reinforcer in WGmc

The gemination of the sigmatic reinforcer in WGmc is seen in the OE and OF RDem stems *þis(s)-* and *thiss-*, as well as the OHG M/N.GEN.SG form *dësses*. For the OHG form, the EWAhd (II: 612) and Klingenschmitt (1987) propose that this started out as an internally inflected form *dë-s-se*, i.e. [[$_{\text{M/N.GEN.SG}}$ *dë-s*]-*se*], where *-se* is a variant of the reinforcer *-si*. The ending *-(e)s* was then analogically added, giving the doubly-inflected

[33] An alternative to Kieckers' account is to reconstruct a PGmc stem **þi-* which is on a par with the regular **þa*-stem (Krause 1971: 159, Klingenschmitt 1987: 183, EWAhd II: 614). This strikes me as a hasty move: it is unnecessary to go all the way back to PGmc for our account of *i*-vocalism. See also Ringe & Taylor (2014: 102) for a similar opinion.

[[[*dë-s*]-*s*]-*es*] but also putting it on a par with the rest of the externally inflected forms in the paradigm. So in the stem *dë**ss**-*, the first sibilant is from the genitive inflection in Dem *dë-s*, while the second sibilant is from the reinforcer **-se*.[34]

The kind of genitive structure seen in OHG *dëss-* must have been available at an early stage of WGmc. According to the EWAhd (II: 613), OE and OF took the genitive-based structure and ran with it, generalizing it more widely than OHG ever did (recall that only M/N.GEN.SG *dësses* shows this gemination in the RDem paradigm). In OE the stem *þiss-* appears not only in M/N.GEN.SG *þiss-es* but also in the M/N.DAT.SG / DAT.PL *þiss-um*. The OE situation most likely reflects the initial, incomplete intraparadigmatic spread of a stem with geminated sigmatic reinforcer in early Anglo-Frisian. In OF, however, the genitive-based structure has spread even further, so that 20 out of the total 24 slots use the stem *thiss-*. This must have happened after the initial, genitive-to-dative expansion which we see preserved in OE. Following this, there would have been enough pressure from forms with gemination for gemination to be generalized throughout the rest of the paradigm in Frisian. Probably relevant here is the fact that OF is attested significantly later than OE, giving Frisian time to show full-fledged intraparadigmatic gemination.

1.4.2.3 Gemination in Norse

Consonant gemination in the RDem of NGmc has origins quite separate from gemination in WGmc. The most striking example of gemination in the ON RDem paradigm is the gemination of the sigmatic reinforcer, as indicated by the shading in Table 30.

Table 30 Gemination of sigmatic reinforcer in ON RDem

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	þe-ssi	þe-ssi	þe-tt-a	þe-ss-ar	þe-ss-ir	þe-ssi
ACC	þe-ss-a	þe-nn-a	þe-tt-a	þe-ss-ar	þe-ss-a	þe-ssi
GEN	þe-ss-ar	þe-ss-a	þe-ss-a	þe-ss-a	þe-ss-a	þe-ss-a
DAT	þe-ss-i	þe-ss-um	þe-ss-u	þe-ss-um	þe-ss-um	þe-ss-um

As already discussed in Section 1.2.1.1, the historical reason for this is that the PN sequence **ʀs* assimilated to *ss* during the development into ON.[35]

There are five slots in Table 30, however, which do not have a sigmatic reinforcer but still have consonant gemination: M.ACC.SG *þenna*, M/N.GEN.SG *þessa*, and N.NOM/ACC.SG

[34] Important to emphasize here is the historical nature of this fact. Synchronically I must insist that the most prudent analysis of *dësses* is that it shows gemination of the sigmatic reinforcer and not gemination of the genitive component *-s*, whatever the diachronic development may have been.

[35] Contrast this with OE, which did not have *rs* > *ss* assimilation, only *sr* > *ss* (Campbell 2003: 195-196).

þetta. In these forms it is not the sigmatic reinforcer which is doubled but the K component (*-n*, *-s*, and *-t*, respectively). Now, considering the fact that RN shows more archaic forms like M.ACC.SG **þinsa**, **þensa** (*þensa*) and N.ACC.SG **þitsa** (*þetsa*), it might be tempting to posit – just as PN **ʀs* changes to ON *ss* – that **ns* changes to *nn* (cf. Haugen 1982: 101) or that **ts* changes to *tt* (cf. Armitage 1911: 207). After all, this would mean there is a smooth development from RN *þensa* to ON *þenna*, and from RN *þetsa* to ON *þetta*. The problem, however, is that we simply do not have any evidence for such sound changes in early Germanic (H.F. Nielsen, p.c.), making this sound-law-based explanation completely *ad hoc*.

Instead we should appeal to analogy. The gemination in M.ACC.SG *þenna* is most likely based on the Dem form *þann*. The same thing could be said for M/N.GEN.SG *þessa*, considering its Dem counterpart *þess*.[36] The doubling of *t* in *þetta* could, in turn, be based on *þenna* and *þessa* (seeing as its Dem counterpart *þat* has only a single *t*). However, the strongest pressure on these five forms may not even have come from the Dem paradigm, but from the RDem paradigm itself. That is, there would have been plenty of intraparadigmatic pressure to have consonant gemination, on the basis of the gemination of *-s* in the RDem stem *þess-*.

1.4.3 The asigmatic reinforcer in Norse

The ON RDem forms with gemination of K display what I have called the asigmatic reinforcer *-a* (< PN **-a*): M.ACC.SG *þenn-**a***, N.NOM/ACC.SG *þett-**a***, and M/N.GEN.SG *þess-**a***. In classical ON this reinforcer is in complementary distribution with the sigmatic reinforcer (< NWGmc **-si*), since *-a* and *-s* do not ever coexist within the same form. In the archaic RN forms M.ACC.SG *þen-**s-a*** and N.ACC.SG *þet-**s-a***, on the other hand, we see that it was possible for the sigmatic and asigmatic reinforcers to coexist within the same form at an earlier stage of Norse.

Noreen (1923: 315), building on Grimm (1831: 27), formulates the hypothesis that PN **-a* is on a par with the Gothic intensifier/conjunction *-(u)h*, which comes from IE **-(u)k^{w}e* (Fortson 2004: 314). However, there are reasons to doubt this hypothesis. Indeed, a final enclitic *-uh* would not survive into NWGmc since unstressed vowels and non-initial *h* were both subject to deletion. Hence Noreen's etymology is not widely accepted.

Instead the etymology of the reinforcer *-a* is usually said to be **-ō$^{(m)}$* (see Krahe 1969 II: §38), which is part of a PGmc N.NOM/ACC.SG doublet **-at ~ -at-ō$^{(m)}$* (cf. Go. *blind* vs.

[36] The gemination in *þann* and *þess* also has to be explained. Prokosch (1939: 269) claims that gemination of *-n* in *þann* is due to analogy with the pronoun *hann* (< **han-m*) 'he'. He then suggests that gemination of *-s* in *þess*, in turn, is due to influence from these forms.

blindata) (McFadden 2004: 130). Note here that the evidence provided is from Gothic (EGmc) rather than from NGmc.

In Lander (2013) I formulate a different hypothesis, namely that PN *-*a* should be identified with the PGmc reinforcer *-*ō* which attached to the 1SG pronoun, as in pre-Gmc **ek-ō* > PGmc **ek-ā* (Feist 1939: 291). The vowel here must have been long, otherwise it would not have survived into NGmc (see Ringe 2006: 137). According to Feist, the following forms are also descendants of PGmc **ek-ā*: Early Runic -**eka** (-*eka*) and **-ika** (-*ika*), as well as RN -**ekA** (-*æka*) and -**kA** (-*ka*).[37]

An advantage of my etymology is that the cognates provided as evidence are from NGmc, instead of EGmc. On the other hand, an advantage of the traditional etymology is that N.NOM/ACC.SG *-*ō*$^{(m)}$ is associated with plausible case and Φ features, since N.NOM/ACC.SG is one of the slots in which the asigmatic reinforcer -*a* took hold in the ON RDem paradigm.

1.5 General goals and outline

The ultimate goal of the present dissertation is synchronic: specifically I want to formulate a hypothesis concerning the morphological makeup of the NWGmc RDem. That is, I am chiefly concerned with how the ingredients discussed above – the demonstrative stem, inflectional endings, the sigmatic reinforcer, the asigmatic reinforcer, and the consonant geminator/reinforcer – fit together in a paradigm. To accomplish this goal, detailed questions about internal structure need to be answered, and we also need a theory which allows us to formulate such questions in a systematic way.

The present dissertation is an attempt to apply formal linguistics (cartography and its offshoot nanosyntax) to the empirical domain of Old Germanic. A foundational assumption is that dead languages can be studied and understood synchronically in the exact same way that we study and understand modern languages (consider in this vein also Danckaert 2012 who takes a cartographic approach to Latin embedded clauses).

The thesis is also a case study in microcomparative syntax, i.e. the study of languages which are genetically very closely related. Studying syntactic variation in this way is especially fruitful because the fact that languages are closely related helps guarantee – on the empirical side – that points of variation are isolated properly and – on the explanatory

[37] Feist also cites the OHG 1SG *ihha* as part of this group, but it is important to note that this form is quite unique: it is used in only two OHG texts in order to translate Latin *egomet* (Braune & Reiffenstein 2004: 242). Thanks to H.F. Nielsen (p.c.) for pointing this out to me.

side – that the real underlying reason for these differences will be pinpointed. While it is still essential to compare unrelated languages, of course, microcomparative syntax is an important research tool in the quest to unravel the interaction between principles and parameters in generative linguistic theory (see also Kayne 2005: Ch.12 for some discussion and references). The languages of this dissertation fit into the microcomparative framework not just because they are all Northwest Germanic languages, but because they are *Old* Germanic languages. ON, OE, OF, OS, and OHG have not diverged from each other nearly as much as today's Germanic languages have, centuries later.

The outline of the dissertation is as follows. In Chapter 2 I introduce the theoretical frameworks which will be used to probe the many RDem structures of ON, OE, OF, OS, and OHG. The first part of Chapter 2 introduces seminal work by Cinque (2005) and the U20 research program in cartography. The second part of Chapter 2 provides an introductioin to the theory of nanosyntax. Nanosyntax is a theory originally developed by researchers at the University of Tromsø. The theory can be seen as a further development of the cartographic framework. It posits a principled morphology with the same rules normally attributed to syntax, making it very useful for the fine-grained morphological decomposition of lexical items. Chapter 3 reintroduces the core empirical data from ON which will be the basis for the analysis and provides various refinements which are not present in this introductory chapter. Chapter 4 provides an analysis of the ON data in terms of the U20 theory of movement developed in Cinque (2005). In Chapter 5 I will reinterpret the analysis in Chapter 4 from a nanosyntactic perspective. Chapter 6 expands the scope of the analysis developed in Chapters 4 and 5 by integrating facts from both the WGmc languages (OE, OF, OS, and OHG) and some additional RDem variants from Norse. Again there will be various refinements made to the WGmc and Norse data in Chapter 6 which are not present in this introductory chapter. Chapter 7 concludes the dissertation.

2
Theoretical background: U20 and nanosyntax

The general theoretical framework adopted in this dissertation is that generally referred to as cartography, whose goal is to identify and map the primitive building blocks of syntactic structure. Specifically, the hypothesis is that structure is built up on the basis of grammatical features. Each feature is merged as a syntactic head, and each head projects according to a rigidly ordered, universal format for syntactic structure-building.

This chapter presents the necessary background on two research areas within the cartographic program which will be of crucial relevance for the thesis. The first is the so-called U20 program, which stems from Cinque's (2005) reinterpretation of Joseph Greenberg's typological work (Greenberg 1963). The second is Michal Starke's theory of nanosyntax (e.g. Starke 2009, 2011abc, 2013).

2.1 Cinque (2005) and Universal 20 (U20)

2.1.1 Updating Greenberg's (1963) original U20

Using a sample of 30 typologically diverse languages, Greenberg (1963) proposes 45 potentially universal generalizations about the order of morphemes in language. This work has given rise to a great deal of subsequent research. The 20th universal ('U20') in Greenberg's list deals with the ordering of the constituents in the the nominal domain:

(18) Greenberg's (1963: 87) Universal 20

> When any or all of the items (demonstrative, numeral, and descriptive adjective) precede the noun, they are always found in that order. If they follow, the order is either the same or its exact opposite.

According to Greenberg, with respect to demonstrative (Dem), numeral (Num), (descriptive or direct modification) adjective (A), and noun (N), only the three orders in (19) are possible: if the modifiers are to the left of N, then Dem precedes Num and Num in turn precedes A (19a); if the modifiers are to the right of N, either the same ordering is found (19b), or its mirror image, in which A precedes Num and Num precedes Dem (19c).

(19) (a) Dem Num A N

(b) N Dem Num A

(c) N A Num Dem

Observe that the discussion here focuses on unmarked orders, i.e. pragmatically neutral orders which do not have special interpretive properties relating to topic or focus, which may have additional operations associated with them, such as movement of APs or PPs (Cinque 2004, Cinque 2010a: 79-85).

Following Hawkins (1983) and much other typological work (I refer to Cinque's original paper for further references), Cinque (2005: 315), points out that while generalization (19a) has held up extremely well (i.e. prenominal modifiers always appear in the order Dem Num A), additional postnominal modifier patterns are attested outside of (19b) and (19c). Thus it has turned out that the generalization in (18)/(19) is too restrictive given the empirical facts.

In light of these findings, Cinque (2005) attempts to update Greenberg's U20 (see also Abels & Neeleman 2009, 2012; Dryer 2009). He starts by looking at all the possible orders of the four items Dem, Num, A, and N. There are 24 possible permutations when four items are involved ($4! = 4 \cdot 3 \cdot 2 \cdot 1 = 24$). According to Cinque's extensive survey of the typological literature, 14 of these orders are attested in the languages of the world, while the other 10 are not. The 24 possible orders and their crosslinguistic attestation are illustrated in (20). Unattested orders are preceded by an asterisk.

(20) 24 possible orders of Dem, Num, A, N and crosslinguistic attestation (Cinque 2005: 319-320)

(a) Dem Num A N
(b) Dem Num N A
(c) Dem N Num A
(d) N Dem Num A
(e) *Num Dem A N
(f) *Num Dem N A
(g) *Num N Dem A
(h) *N Num Dem A
(i) *A Dem Num N
(j) *A Dem N Num
(k) A N Dem Num
(l) N A Dem Num
(m) *Dem A Num N
(n) Dem A N Num

(o) Dem N A Num
(p) ? N Dem A Num [“possibly spurious” (see below)]
(q) * Num A Dem N
(r) Num A N Dem
(s) Num N A Dem
(t) N Num A Dem
(u) * A Num Dem N
(v) * A Num N Dem
(w) A N Num Dem
(x) N A Num Dem

The question arises whether a principled account is available for the pattern found in (20), i.e. whether there is a way to derive the attested strings and exclude the unattested ones. Before discussing Cinque’s proposals for answering this question, I will introduce a few prominent features of the research program in generative grammar known as cartography.

2.1.2 The hallmarks of cartography

The cartographic approach embraces crosslinguistic diversity (see Shlonsky 2010 for some discussion). By studying a wide range of data from many different languages, cartography strives to catalogue the full inventory of grammatical features found in natural language. The cartographic program is undoubtedly part of the broader approach known as Principles and Parameters (Chomsky 1981, 1986), which can be summed up by Chomsky’s (2001: 2) Uniformity Principle:

(21) The Uniformity Principle (Chomsky 2001: 2)

> In the absence of compelling evidence to the contrary, assume languages to be uniform, with variety restricted to easily detectable properties of utterances.

In cartography, the idea of uniformity is most obviously represented by the assumption that both the set of possible grammatical features and their hierarchical organization are universal. This universal ordering of features is referred to as the functional sequence (fseq). Since it is assumed that each feature is associated with a functional head, the building of syntactic structure is theorized to strictly follow the functional sequence, which in effect constitutes the very backbone of syntactic structure. The second part of Chomsky’s principle (2001: 2) has to do with crosslinguistic variation: all crosslingusitic variation is assumed to be reducible to the idiosyncratic properties of the language-specific lexicon. The precise interaction between syntax and the lexicon, and thus principles and parameters, however, is in need of a more concrete formulation.

Cartographers assume a standard Y-model of language (Chomsky 1965, 1981, 1986, 1995, 2000, 2001). This means that there is a (presyntactic) lexicon from which lexical items are selected, as seen in Figure 8.

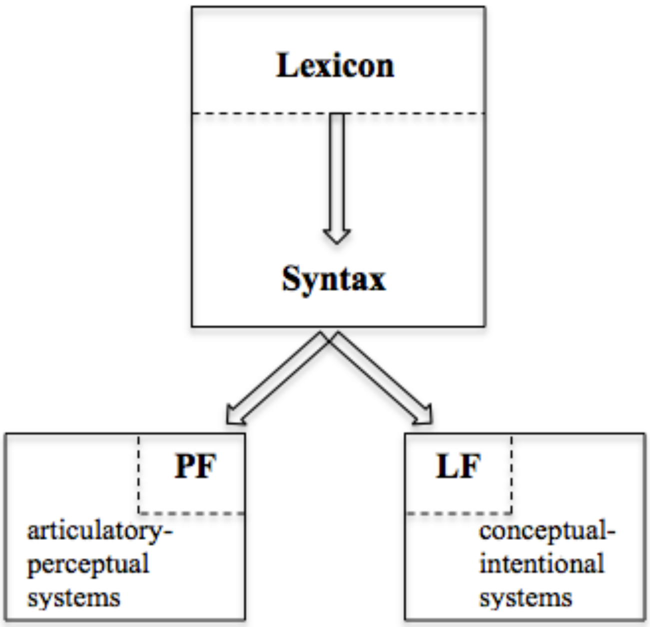


Figure 8 Traditional Y-model

These lexical items are then arranged by the syntax, i.e. the generative component, and the output which is generated by the syntax is interpreted at two interfaces: PF (phonological form) and LF (logical form). In other words, linguistic expressions are made up of sound and meaning, which are related to each other by syntax. Distinguishing cartography from some other generative views is the idea that "scope-discourse semantics, but also prosodic properties are transparently read off from syntactic representations" (Rizzi 2013: 10). That is, there is a general trend in cartography of 'syntacticizing' discourse and semantics as much as possible (e.g. Haegeman & Hill 2013, Ramchand 2008, among others).

A guiding principle of cartographic work is the hypothesis that if there is morphological evidence for a certain syntactico-semantic feature, then this feature deserves to be associated with its own head in the fseq (hence the maxim *one (morphosyntactic) property – one feature – one head*; Cinque & Rizzi 2008: 50). Given the cartographic assumption that each feature correlates with its own syntactic head, the granularity of the resulting syntax is very fine. Over time this kind of methodology has resulted in a steady decomposition of the structure of the clause and of its components, evidenced by a number of milestones in the literature: Abney (1987) on the split NP; Pollock (1989) on the split IP; Hale & Keyser (1993) on the split VP; and Rizzi (1997) on the split CP. Over time this has led to the creation of a number of fine-grained maps of the grammatical domains: adjectives (Cinque 2010a), adverbs and modality (Laenzlinger

1998, Cinque 1999), inflection and agreement (Belletti 1990), negation (Zanuttini 1991, Haegeman & Zanuttini 1991, Haegeman 1995), subject positions (Cardinaletti 1997, 2004), topic and focus (Benincà & Poletto 2004, Aboh 2004a, Rizzi 2004b, Bianchi & Frascarelli 2009), quantifiers (Beghelli & Stowell 1997, Puskás 2000), and determiners (Giusti 1997) and other nominal modifiers (Alexiadou, Haegeman & Stavrou 2007). Seminal papers in cartography are collected in Cinque (2002), Belletti (2004), and Rizzi (2004a).

As pointed out by Van Craenenbroeck (2009: 3), cartographers very often choose to work within the program of antisymmetry (Kayne 1994), according to which syntactic structures are maximally simple. Since in this approach precedence of terminals corresponds to asymmetric c-command, cartographers can use linear order to deduce underlying syntactic hierarchies. In Kayne's system, merge always produces binary-branching structures (Kayne 1984), and multiple specifiers[38] and rightward movement are all ruled out. In other words. This highly restrictive theoretical outlook allows cartographers to build a straightforward map of functional structure, often using a (micro)comparative perspective in their empirical investigations (see Kayne 2005: Ch. 12).[39]

Finally, an important restriction in the Cinquean approach to cartography is the hypothesis that all movement is phrasal (XP) movement. This means that movement is never head (X^0) movement, which results in head adjunction structures (Travis 1984, Baker 1988). I will now briefly discuss some drawbacks of head movement.

Head adjunction occurs when a lower head moves to a higher head and adjoins to the left of this higher head. Further head movement will then involve this complex head adjoining to the left of the next head, and so on.

(22) illustrates this derivation for French verbs. In the endings of French verbs we can distinguish separate tense (T) and subject-verb agreement (AgrS) morphemes, as seen in the verbal structure provided in (4a). Assuming that the verb root *pens-* is a V head, we can derive the structure *pens-i-ons* using head-adjunction. From the base order in (22a), the V *pens-* moves to T and adjoins to its left in (22b). From there the complex head V-T moves up to AgrS and adjoins to the left once again, giving the structure V-T-AgrS (22c).

---

[38] Adjunction is limited to one instance per phrase in the antisymmetric program. Since specifiers are cases of adjunction, consequently there can be only one specifier per phrase. See Kayne (1994: 22, 27-28).

[39] Van Craenenbroeck's (2009) survey of cartography is mostly critical (it is an introduction to a collection of papers offering 'alternatives to cartography'), but we can actually reinterpret properties of cartography which he considers faults as virtues. For instance, *a priori*, there is nothing intrinsically wrong with postulating a highly restrictive theory, since it allows for strong predictions to be made and tested. Cinque's (1995) review of Kayne (1994), for example, makes the case that Kayne's theory of antisymmetry is nothing less than a revolution in generative grammar. Without Kayne's theory, at least, U20-style research could not be conducted.

(22) Head movement in French verbs

pens-i-ons
V-T_{PAST}-$AgrS_{1PL}$
'(we) thought'

(a) [$_{\text{AgrSP}}$ [$_{\text{AgrS}}$ -ons] [$_{\text{TP}}$ [$_{\text{T}}$ -i] [$_{\text{VP}}$ **[$_{\text{V}}$ pens-]**]]]]]

(b) [$_{\text{AgrSP}}$ [$_{\text{AgrS}}$ -ons] [$_{\text{TP}}$ **[$_{\text{V-T}}$ pens-i]** [$_{\text{VP}}$ **[t_{V}]**]]]]]

(c) [$_{\text{AgrSP}}$ **[$_{\text{V-T-AgrS}}$ pens-i-ons]** [$_{\text{TP}}$ **[$t_{\text{V-T}}$]** [$_{\text{VP}}$ **[t_{V}]**]]]]]

The Head Movement Constraint (HMC) states the following.

(23) Head Movement Constraint (Travis 1984: 131)

An X^0 may only move into the Y^0 which properly governs it.

In other words, intermediate heads may not be skipped when head movement takes place. Head movement proceeds from head to adjacent head, meaning there is never 'long' head movement.

Given the HMC in (23), if the underlying order of syntactic heads is A > B > C as in (24), and if the head C needs to move to the head A, C must first move to B, adjoin to B's left (only left adjunction being compatible with Kayne 1994), and then the complex head [C-B] can move to A. Observe that the resulting structure is the mirror image of the underlying structure (i.e. A > B > C), which is 'rolled up', as it were, to give the final inverse order (i.e. C-B-A).

(24)

A	B	C
A	[C-B]	t_{C}
[[C-B]-A]	$t_{\text{C-B}}$	t_{C}

The HMC is an important locality principle in head movement theories. With regard to (24), for instance, the HMC predicts that the head sequence C-A-B should be unattested, since it would require the movement of C to skip B and jump straight to A.[40]

[40] It should be noted that Relativized Minimality (RM) (Rizzi 1990, 2001; Starke 2001) in fact derives the HMC. In short, RM states that movement must not cross an intervening element (bolded in (vi)) which is of the same type as the moving element (italicized in (vi)). This applies not only to instances of A- (via) and A'-movement (vib), but also to cases of head movement (vic).

Assuming that morphemes are heads, head adjunction and the HMC taken jointly predict that whenever movement is involved the resulting morphological structures should be the reverse of the base-generated merge order (i.e. the Mirror Principle; Baker 1985: 375). As Caha (2009: 218-9) rightly points out, however, this prediction is empirically incorrect. There is plenty of crosslinguistic diversity not accounted for by this type of head adjunction and the HMC. The components in the nominal domain provide a telling illustration. If we take the four items Dem, Num, A, and N, and if we assume that the base order is as in (25a), the only orders predicted to exist by head movement theory are the following four:

(25)	(a)	Dem	Num	A	N
	(b)	Dem	Num	[N-A]	t_{N}
	(c)	Dem	[[N-A]-Num]	$t_{N\text{-}A}$	t_{N}
	(d)	[[[N-A]-Num]-Dem]	$t_{N\text{-}A\text{-}Num}$	$t_{N\text{-}A}$	t_{N}

These four patterns are indeed attested, but a quick comparison with (20) above shows that 10 of the attested patterns are still missing in (25). For instance, order (k) (A N Dem Num) is not accounted for in (25), but it can be derived if we assume that phrasal movement is allowed. By moving the phrase [$_{AP}$ A [$_{NP}$ N]] to the left of [$_{DemP}$ Dem [$_{NumP}$ Num]], order (k) (A N Dem Num) is produced. Order (k), an empirically attested order, can simply not be derived using (only) head movement. In other words, the traditional theory of head movement undergenerates.

Many have suggested that the way around this issue of undergeneration is to accept the idea that items may be freely base-generated in different places in the structure. According to this view there is no universal sequence of syntactic projections. Cinque (2010) argues against this view. As a basic example, consider the parallel orders in (26) and (27) (from Cinque 2010a: 38). (26a) illustrates the commonly attested prenominal order of the modifiers Dem, Num, and A. (27a) illustrates the commonly attested prenominal order of different kinds of adjectives: adjectives of size precede adjectives of color, which precede adjectives of nationality (as in English *a* ***big red American*** *car*). (26b) and (27b) are unattested: in these cases the nominal modifiers appear in prenominal position but now in

---

(vi)	(a) **Bill* seems that **it** is possible t_{Bill} to lose.	(A-movement)
	(b) **How* do you think **who** might fix the car t_{how}?	(A'-movement)
	(c) **Have* they **might** t_{have} cheated?	(head movement)

In (vic), for instance, the moved auxiliary *have* crosses over the modal *might*. Since *might* is of the same type as *have*, *might* intervenes and blocks *have*-movement past it.

the reverse order of (26a) and (27a). (26c) and (27c), both crosslinguistically attested, show the same order of modifiers as in (26a) and (27a), but now following the noun. (26d) and (27d), finally, are the mirror image of (26a) and (27a), respectively, and both are again attested.

(26) (a) Dem Num A N
(b) * A Num Dem N
(c) N Dem Num A
(d) N A Num Dem

(27) (a) A_{size} A_{color} $A_{nationality}$ N
(b) * $A_{nationality}$ A_{color} A_{size} N
(c) N A_{size} A_{color} $A_{nationality}$
(d) N $A_{nationality}$ A_{color} A_{size}

The relevant observation here is that there is a single prenominal order of modifiers, namely the ones shown in (26a) and (27a), but there is more than one postnominal order, namely (26c,d) and (27c,d) (among other possibilities). At first glance, given the apparent freedom in the postnominal order of the modifiers, a free base-generation analysis should *ceteris paribus* predict that the same freedom is available prenominally. If the order A > Num > Dem is available in the postnominal domain, then one would expect it to also be available in the prenominal domain. Similarly, if the order $A_{nationality} > A_{color} > A_{size}$ is found in the postnominal domain, then one might expect it to also be available in the prenominal domain. But when we look at the languages of the world, there is a gap: while the postnominal analogues are attested ((26c,d) and (27c,d)) the prenominal orders ((26b) and (27b)) turn out to be unattested. Free base-generation, then, is an inadequate approach to adopt. For an overview of asymmetries observed crosslinguistically and the connection to Kayne's LCA, see Cinque (2009, 2010b, 2012) and Kayne (2010).

2.1.3 Cinque (2005)

Cinque's (2005) paper reinterpreting Greenberg's U20 has become a seminal work in the cartographic tradition.[41] In order to explain the word order pattern in the nominal domain summarized in (20), Cinque proposes a number of principles for structure-building, some with special reference to the projection of the nominal domain (though the principles

[41] See Abels (2011) and Jayaseelan (2010) for an extension of Cinquean reasoning to ordering within verb clusters (cf. Koopman & Szabolcsi 2000, Wurmbrand 2006). See also Muriungi (2008) for a Cinquean analysis of the internal structure of Bantu verbs.

below should apply to all grammatical domains, not just the nominal one). These principles are listed in (28).

(28) (a) There is a universal base-generated order of Dem > Num > A > N.
(b) Only leftward movement is allowed (antisymmetry; Kayne 1994).
(c) No head movement is allowed (i.e. phrasal movement only).
(d) All movements must be movements of structures containing N (i.e. no remnant movement is allowed).

In Cinque's paper Dem, Num, and A are merged as specifiers (DemP, NumP, AP) of abstract heads (Z, Y, X, respectively). For the sake of simplicity and convenience, below I treat these elements instead as simple heads projecting their own phrases (cf. Starke 2004), which has no bearing on the logic of Cinque's argumentation and results. Furthermore, Cinque assumes that the landing sites of movement are specifiers of Agreement Phrases (AgrPs). Departing from certain minimalist assumptions about the nonexistence of the functional category Agr (Chomsky 1995: §4.10.1, but see Belletti 2001), Cinque assumes that there is a dedicated AgrP located above each modifier in the extended projection of the noun: $Agr_{Dem}P$ above DemP, $Agr_{Num}P$ above NumP, and $Agr_{A}P$ above AP. For more on Agr and the extended projection (Grimshaw 1979, 1991, 2000), see Section 2.1.3.3. The basic idea is that each layer of the extended projection must be licensed by a nominal feature [N], and that this licensing can be accomplished either by overt movement of NP or by external merge of the feature [N].[42]

Cinque demonstrates that the principles in (28) suffice to derive the attested patterns in (20) while not generating the unattested ones in (20). I will discuss these two sets of data in turn. In Section 2.1.3.1, I explain how the principles in (28) derive the attested orders. In Section 2.1.3.2, I explain how the principles in (28) rule out the unattested orders. In Section 2.1.3.3, I explain Cinque's reasoning behind (28d).

2.1.3.1 Deriving the attested orders

First consider the attested orders displayed by the nominal modifiers. For the reader's convenience I repeat them in (29):

(29) (a) Dem Num A N
(b) Dem Num N A
(c) Dem N Num A
(d) N Dem Num A

[42] This is also the reason NP does not have an AgrP dominating it: NP is the head of the extended nominal projection and as such does not need to be licensed. It is due to the presence of NP in the first place that the other layers of the extended projection need to be N-licensed.

(k) A N Dem Num
(l) N A Dem Num
(n) Dem A N Num
(o) Dem N A Num
(r) Num A N Dem
(s) Num N A Dem
(t) N Num A Dem
(w) A N Num Dem
(x) N A Num Dem

Let us take each order in turn.

In order (a) (Dem Num A N) there is no movement, so the base-generated order is simply preserved as is, as seen in Figure 9.

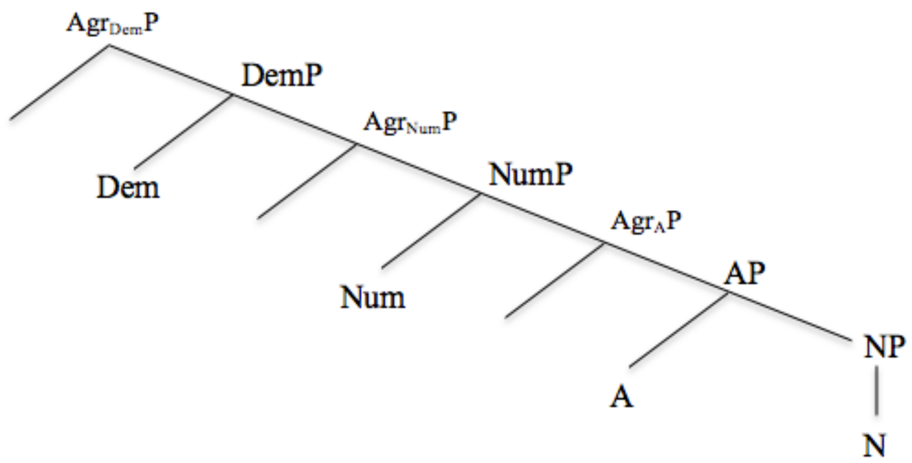


Figure 9 Order (a) $[_{DemP}$ Dem $[_{NumP}$ Num $[_{AP}$ A $[_{NP}$ N]]]]

To derive the closely related order (b) (Dem Num N A), we can postulate NP-raising into Spec-Agr_AP, as seen in Figure 10.

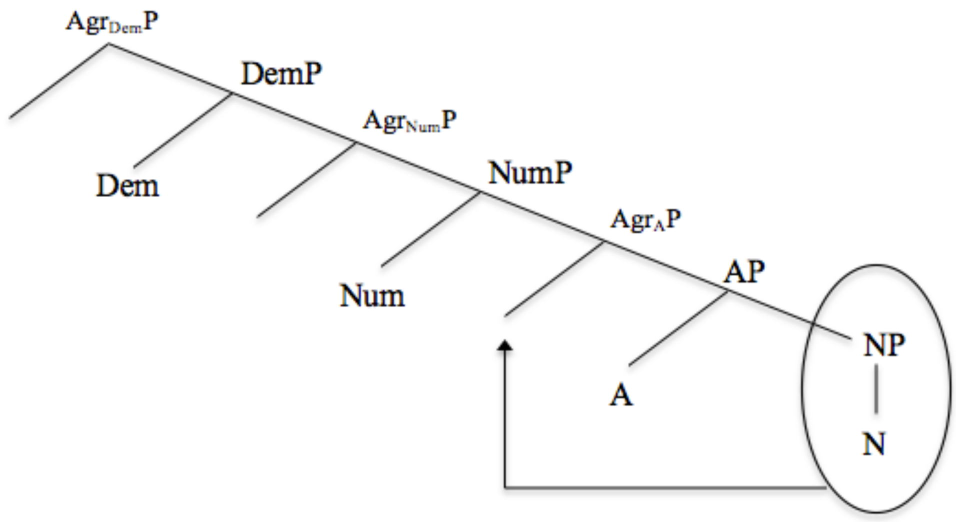


Figure 10 Order (b) [$_{DemP}$ Dem [$_{NumP}$ Num [$_{NP}$ N] [$_{AP}$ A [t_{NP}]]]]

This order, then, is the base-generated structure with NP moved up a step, to the left of AP.

For order (c) (Dem N Num A), NP first moves to the left of AP and then moves again to the left of NumP, where it stops, as seen in Figure 11.

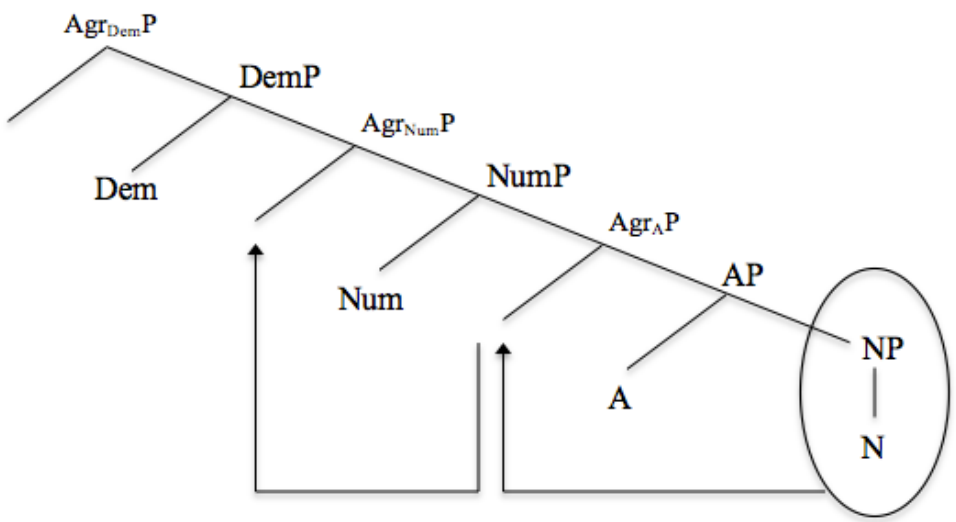


Figure 11 Order (c) [$_{DemP}$ Dem [$_{NP}$ N] [$_{NumP}$ Num [t_{NP}] [$_{AP}$ A [t_{NP}]]]]

That is, NP has moved up two steps, to the left of NumP.

To derive order (d) (N Dem Num A), we postulate cyclic movement of NP all the way up, to the left of DemP, as seen in Figure 12.

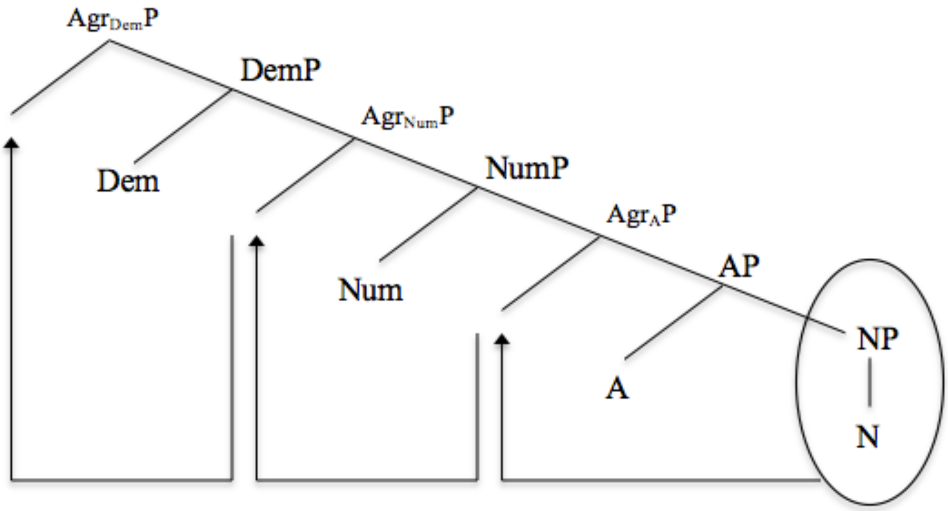


Figure 12 Order (d) [$_{NP}$ N] [$_{DemP}$ Dem [t_{NP}] [$_{NumP}$ Num [t_{NP}] [$_{AP}$ A [t_{NP}]]]]

That is, NP cycles through all of the AgrPs in the structure.

Before moving on to a discussion of the derivations for the rest of the attested word orders, note that the principles set out in (28) allow for pied-piping. In fact, a core requirement is that N(P) must be contained in the moved structure. Pied-piping of additional material is possible, as long as N(P) is contained in the moved structure and the moved structure constitutes a legitimate syntactic object.

In the attested order (n) (Dem A N Num), we see the result of the movement of [AP [NP]], which is to say that NP pied-pipes the layers of AP and Agr$_A$P above it when it moves to the left of NumP. This is illustrated in Figure 13.

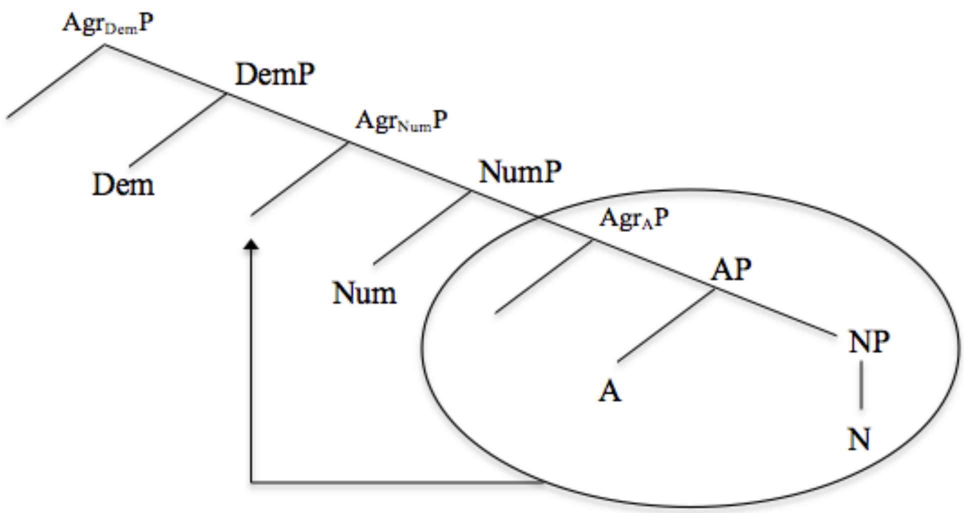


Figure 13 Order (n) [$_{DemP}$ Dem [$_{AP}$ A [$_{NP}$ N]] [$_{NumP}$ Num [$t_{AP\ NP}$]]]

Order (k) (A N Dem Num) follows the same basic procedure as (n), except that the landing site of [AP [NP]] is one step higher, to the left of DemP, as seen in Figure 14.

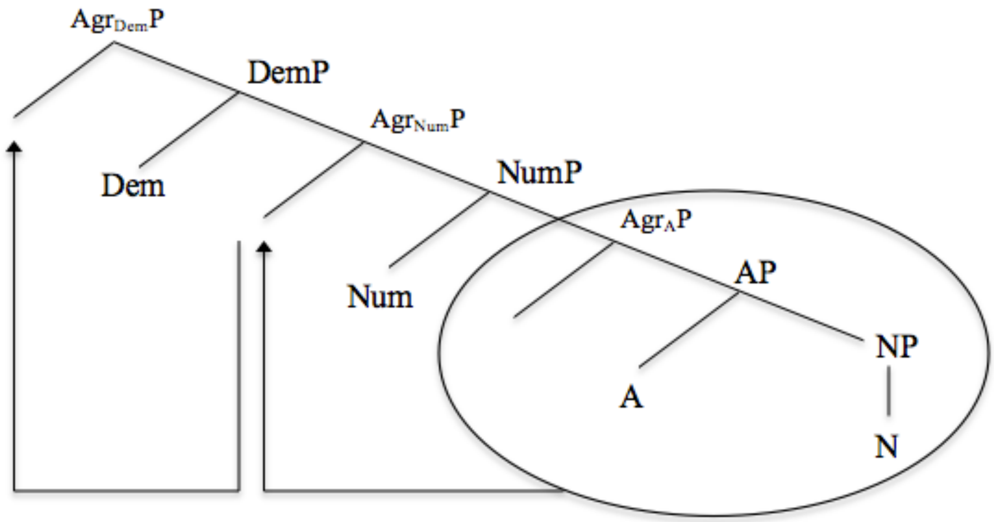


Figure 14 Order (k) $[_{AP}$ A $[_{NP}$ N]] $[_{DemP}$ Dem $[t_{AP\ NP}]$ $[_{NumP}$ Num $[t_{AP\ NP}]]]$

Orders (l) and (o) are also the result of NP pied-piping AP, but they differ from orders (n) and (k) above. In (n) and (k), NP immediately pied-pipes material leading to the movement of a constituent larger than NP at the first step of the derivation. For (l) and (o), however, NP first independently moves to the left of AP, thus reordering NP and AP, and then the constituent containing NP and AP moves up. Thus, in order (o) (Dem N A Num) the landing site of the constituent containing NP and AP is $Agr_{Num}P$, as in Figure 15.

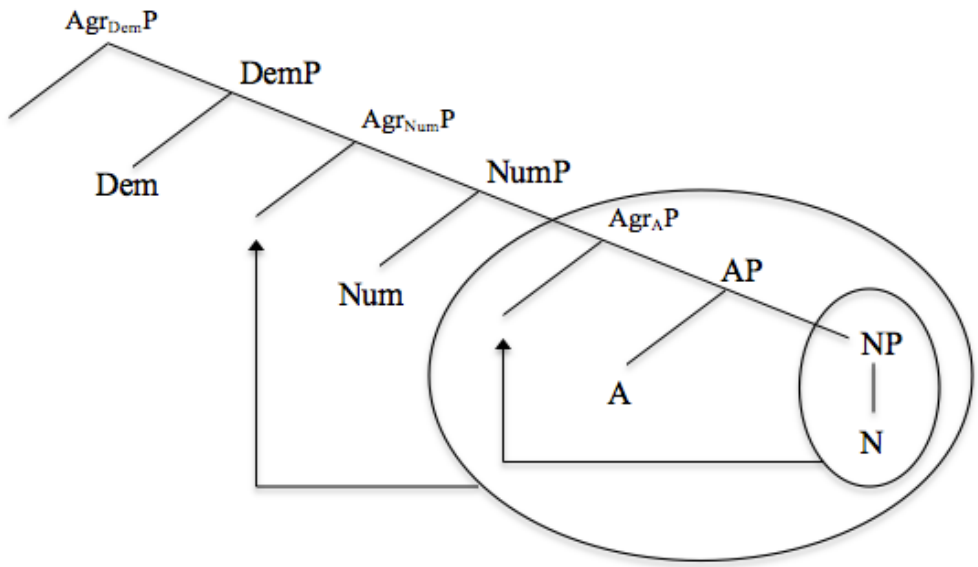


Figure 15 Order (o) $[_{DemP}$ Dem $[[_{NP}$ N] $[_{AP}$ A $[t_{NP}]]]$ $[_{NumP}$ Num $[t_{NP\ AP}]]]$

In order (l) (N A Dem Num), the landing site of the constituent containing NP and AP is one step higher, namely $Agr_{Dem}P$, as in Figure 16.

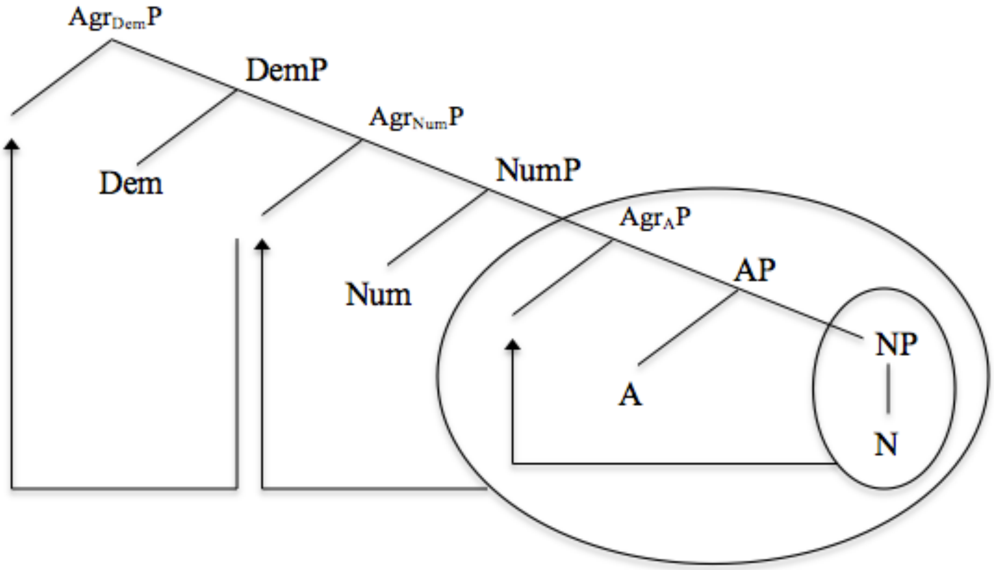


Figure 16 Order (l) [[$_{NP}$ N] [$_{AP}$ A [t_{NP}]]] [$_{DemP}$ Dem [$t_{NP\ AP}$] [$_{NumP}$ Num [$t_{NP\ AP}$]]]

Order (r) (Num A N Dem) is derived by the movement of a constituent reaching all the way up to the Num layer, meaning that NP pied-pipes two layers above it when it raises to the left of DemP, as seen in Figure 17.

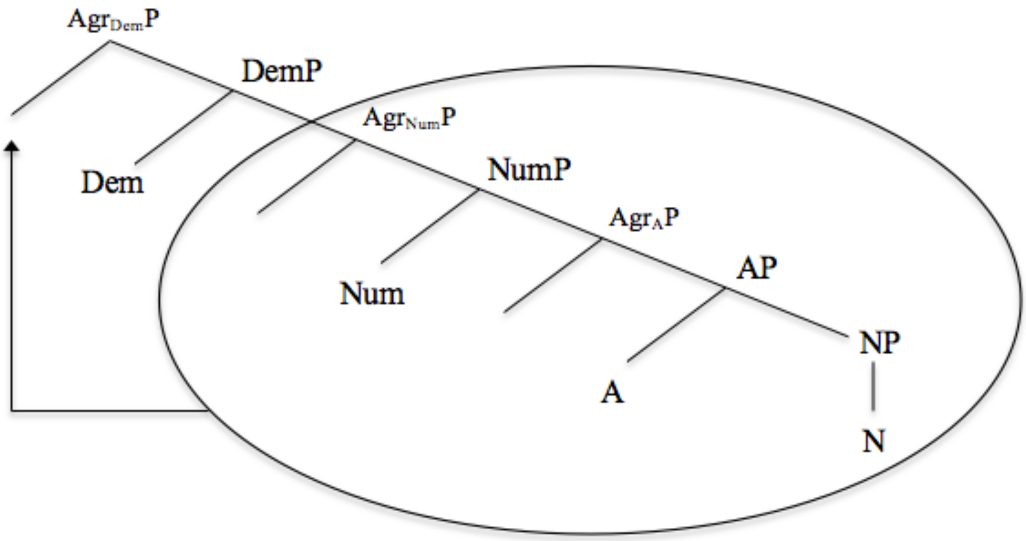


Figure 17 Order (r) [$_{NumP}$ Num [$_{AP}$ A [$_{NP}$ N]]] [$_{DemP}$ Dem [$t_{NumP\ AP\ NP}$]]

Order (s) (Num N A Dem) is derived in the same way as (r) except that there is movement of NP to the left of AP within the constituent that moves to the left of DemP, as seen in Figure 18.

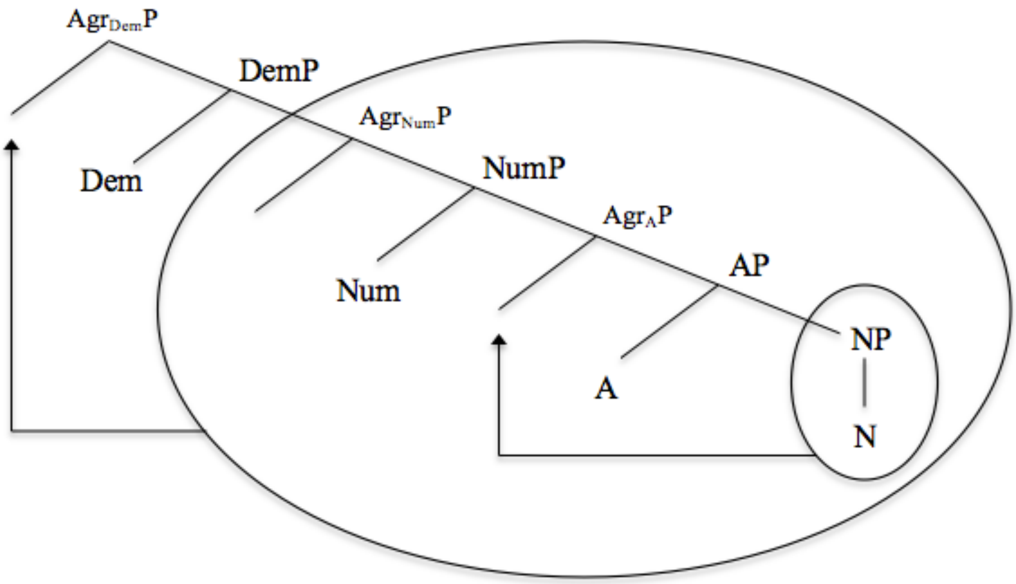


Figure 18 Order (s) $[_{NumP}$ Num $[_{NP}$ N$]$ $[_{AP}$ A $[t_{NP}]]]$ $[_{DemP}$ Dem $[t_{NumP\ NP\ AP}]]$

Order (t) is again similar, with the difference that NP moves all the way up to the left of NumP before the entire constituent containing NP, NumP, and AP moves to the left of DemP, as illustrated in Figure 19.

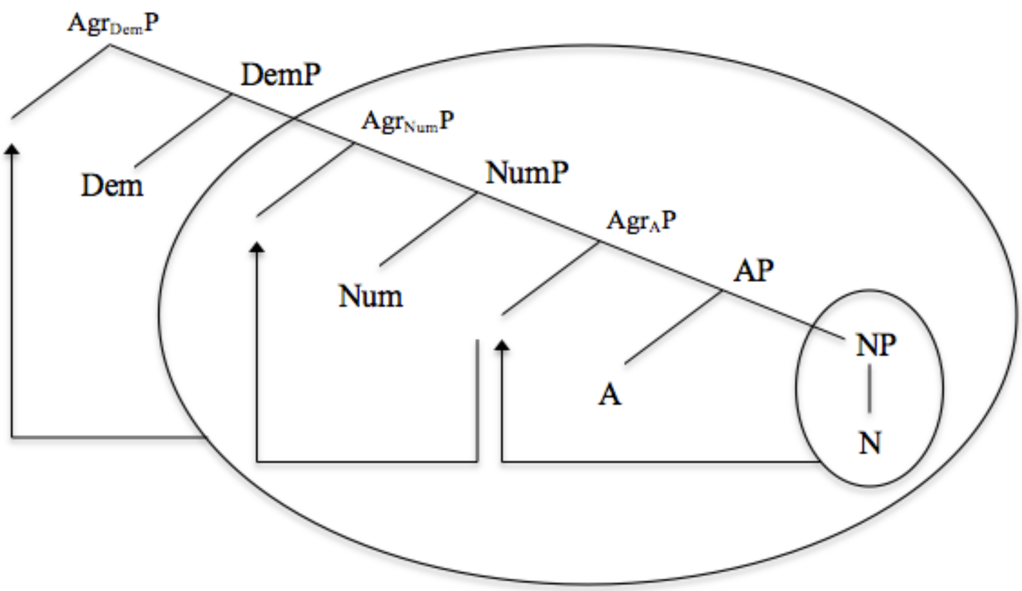


Figure 19 Order (t) $[[_{NP}$ N$]$ $[_{NumP}$ Num $[t_{NP}]$ $[_{AP}$ A $[t_{NP}]]]]$ $[_{DemP}$ Dem $[t_{NP\ NumP\ AP}]]$

Finally, we come to orders (w) and (x). In order (w) (A N Num Dem) the constituent containing AP and NP moves to the left of NumP. The resulting structure then raises to the left of DemP, as seen in Figure 20.

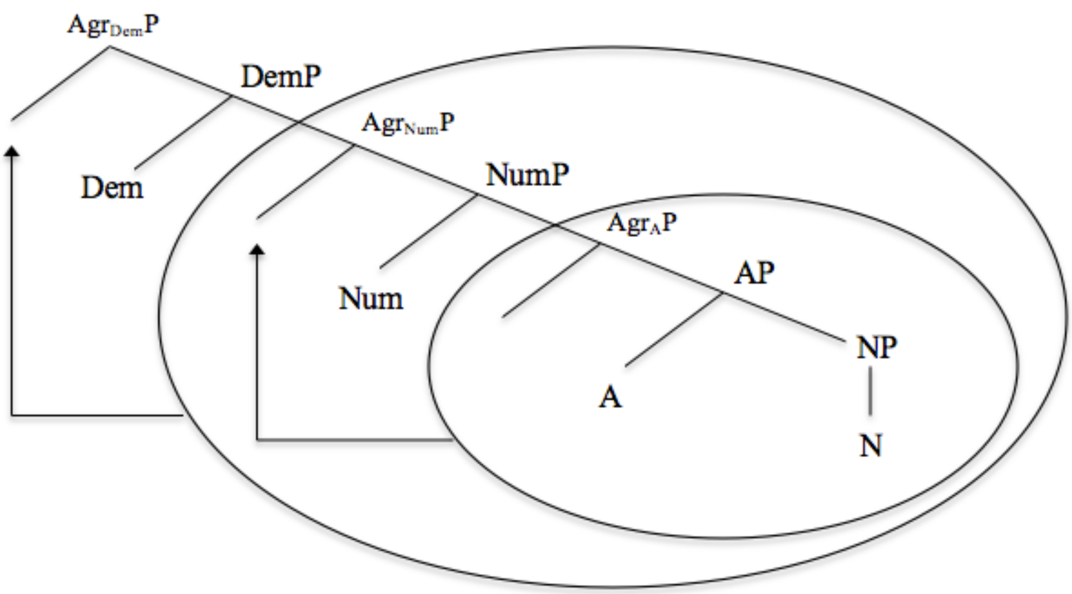


Figure 20 Order (w) [[$_{AP}$ A [$_{NP}$ N]] [$_{NumP}$ Num [$t_{AP\ NP}$]]] [$_{DemP}$ Dem [$t_{AP\ NP\ NumP}$]]

Order (x) (N A Num Dem) is derived by complete 'roll-up', i.e. there is movement and pied-piping at every layer, as seen in Figure 21.

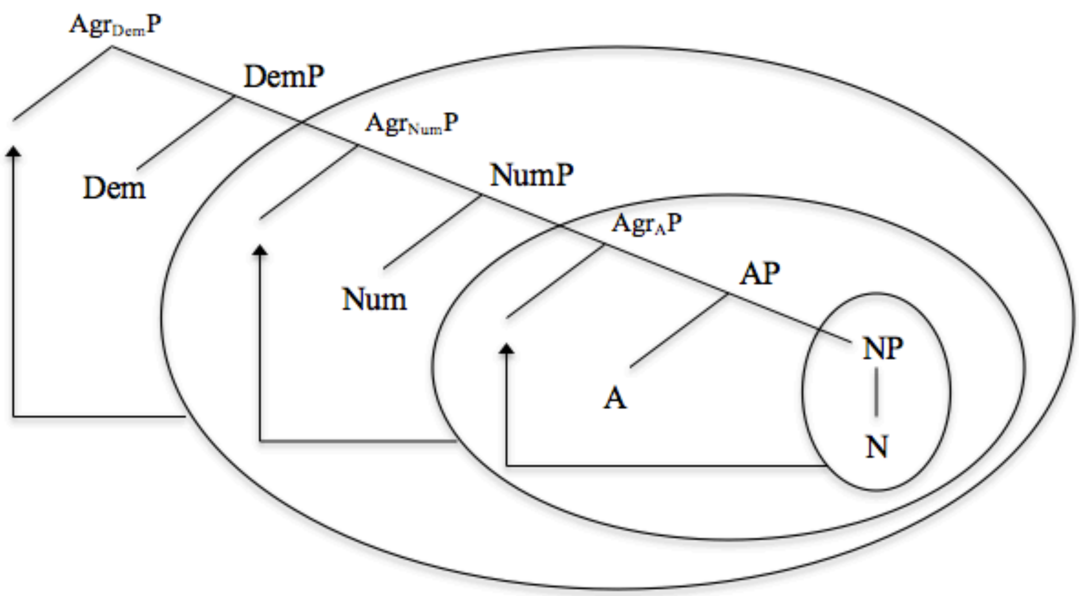


Figure 21 Order (x) [[[$_{NP}$ N] [$_{AP}$ A [t_{NP}]]] [$_{NumP}$ Num [$t_{NP\ AP}$]]] [$_{DemP}$ Dem [$t_{NP\ AP\ NumP}$]]

To summarize the discussion above: the four structure-building principles in (28) allow for 14 possible derivations. Order (a) is simply the base-generated order, and orders (b-d) display raising of NP to various heights in the tree. The remainder of the attested orders are the result of pied-piping chunks of structure larger than (but still including) NP, with or without smaller movements within the larger chunk. In Figures 9-21 above I have illustrated the thirteen derivations which are crosslinguistically attested.

According to the principles in (28) there should be an additional fourteenth derivation which is possible , the so-called 'subextraction' derivation, illustrated in Figure 22, which corresponds to order (p) (N Dem A Num).

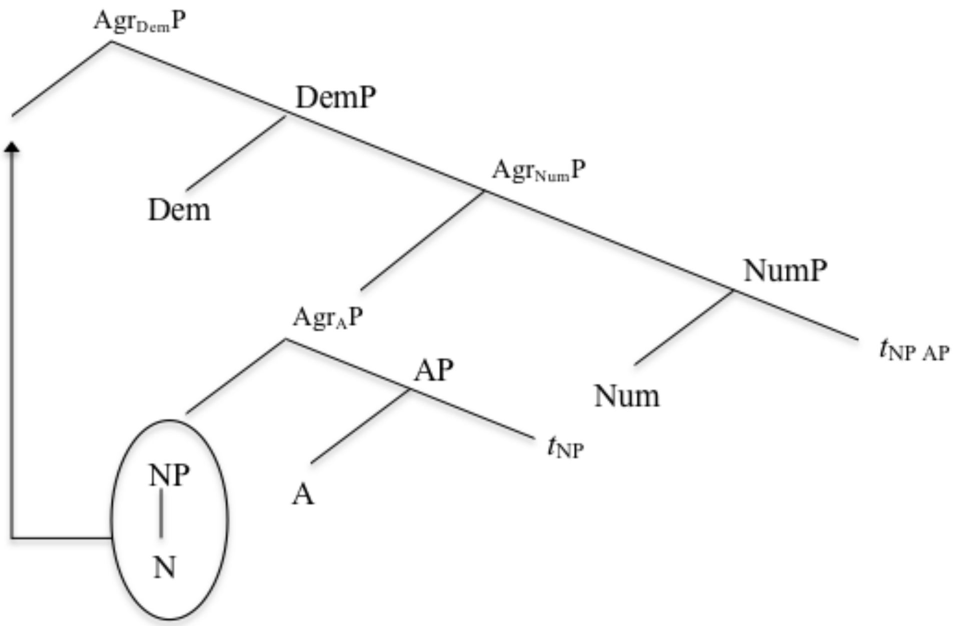


Figure 22 Order (p) $[_{NP}$ N] $[_{DemP}$ Dem $[[t_{NP}]$ $[_{AP}$ A $[t_{NP}]]]$ $[_{NumP}$ Num $[t_{NP\ AP}]]]$

In the derivation in Figure 22, NP first moves to the left of AP, and then the constituent containing NP and AP raises to the left of NumP. From here, however, NP is subextracted from the moved constituent and moves independently to the left of DemP, stranding the constituent containing AP in Spec-Agr_{Num}P. While this is a technical possibility according to the rules in (28), it is perhaps not optimal theoretically speaking. The first two steps look to be the beginning of roll-up movement, with successive pied-piping, i.e. NP raises to the left of AP and then the resulting constituent raises to the left of NumP. From there, however, the roll-up/pied-piping-type of derivation is abandoned, and NP raises independently to the left of DemP, reverting, as it were, to cyclic movement.

It is interesting to note, then, that in addition to requiring a somewhat inelegant derivation, the empirical status of order (p) is fuzzy. The order N Dem A Num is attested in only three languages, and in two of them it alternates with the unproblematic order N Dem Num A. Cinque (2010a) shows that there are two kinds of adjectives: direct modification adjectives (e.g. *former*, *red*, *big*, etc.) and indirect modification adjectives (e.g. participles like *arrived*, *rotten*, etc.). He proposes that direct modification adjectives – the kind of adjective at stake in the U20 generalization (i.e. A) – are merged lower in the extended projection of the noun than indirect modification adjectives (which are analyzed as predicates of reduced relative clauses). As evidence, consider the fact that reduced-relative adjectives like *angekommene* 'arrived' precede direct modification adjectives *ehemalige* 'former' in German (30).

(30) German (Cinque 2010a: 54)

(a) der [kürzlich angekommene] ehemalige Botschafter von Chile
the recently arrived former ambassador of Chile
'the recently arrived former ambassador of Chile'

(b) *?der ehemalige [kürzlich angekommene] Botschafter von Chile
the former recently arrived ambassador of Chile

Cinque (2010a) thus proposes the structure in Figure 23, where reduced relative (or indirect modification) adjectival structures are merged above direct modification adjectives.

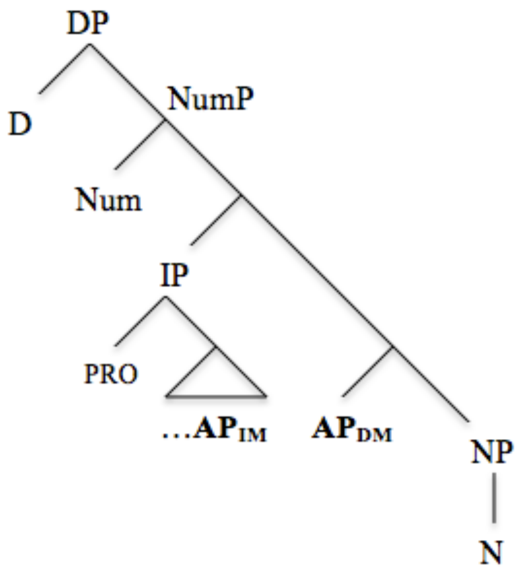


Figure 23 Indirect modification (IM) > direct modification (DM) (based on Cinque 2010a: 55)

The alternation between order (p) (N Dem A Num) and order (d) (N Dem Num A), then, hints at the possibility of A in order (p) being the predicate of a reduced relative clause rather than a direct modifier (Cinque 2005: 323, fn. 27). Thus order (p) is referred to as "possibly spurious" (Cinque 2005: 320). For now I will also put order (p) aside, though the issue of subextraction will resurface later on.

2.1.3.2 Ruling out the unattested orders

Let us now turn to the unattested orders. So far we have seen that Cinque's principles in (28) are able to derive the 13 (again, leaving out order (p)) attested orders of Dem, Num, A, and N. Equally important, however, is that these principles are restrictive enough to be able to rule out the 10 unattested orders, which are repeated in (31).

(31) (e) * Num Dem A N
(f) *Num Dem N A
(g) * Num N Dem A
(h) * N Num Dem A
(i) * A Dem Num N
(j) * A Dem N Num
(m) * Dem A Num N
(q) * Num A Dem N
(u) * A Num Dem N
(v) * A Num N Dem

As seen in (31), the unattested orders (e), (i), (m), (q), and (u) all display the same basic pattern, where the head noun N occurs to the very right of its modifiers. The modifiers to the left of N, however, do not respect the universal base order.

(32) Prenominal ordering of modifers does not respect base order Dem > Num > A

(e) * **Num Dem A** N
(i) * **A Dem Num** N
(m) * **Dem A Num** N
(q) * **Num A Dem** N
(u) * **A Num Dem** N

This kind of ordering is simply not derivable. If N has not moved out of its base position at the very bottom of the structure, then – given that by (28d) all movement must affect at least the projection NP – its modifiers cannot have moved either and should therefore be found only in the order Dem Num A, i.e. order (a) (Dem Num A N).

Observe that it would be possible to derive these orders by remnant movement, i.e. movement that results when part of a projection is extracted and the 'remnant' projection (without the extracted part) is subsequently moved. To demonstrate this, consider the unattested order (e) (*Num Dem A N). Imagine that the projection Agr_AP, containing AP and NP, first moves to the left of NumP, as visible in Figure 24. This would in fact achieve part of order (e), namely the [Dem A N] part. But to achieve the order in which Num appears in the leftmost position, the remnant NumP must then move to the left of DemP, as shown in Figure 24. This movement of the remnant [NumP [$t_{AP\ NP}$]], however, is illicit by (28d) because the remnant does not contain N.

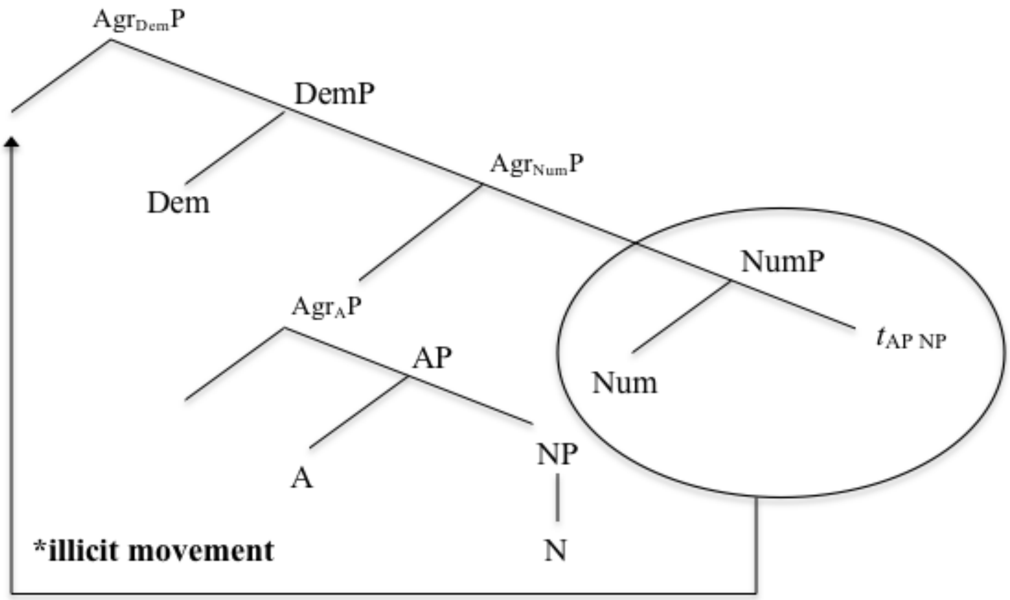


Figure 24 Order (e) with remnant movement

The rule that N must be contained in every moved object thus rules out remnant-movement derivations in the U20 system.

The derivation of a prenominal order of modifiers different from the base-generated order will always require movement of a structure not including N. And since this kind of movement is ruled out by principle (28d), prenominal modifiers can only ever appear in their base-generated order, Dem Num A.

Of course, another option to derive orders like order (e) would be to postulate head movement. Indeed, Num-to-Dem head movement would correctly produce order (e) (Num Dem A N) without any additional operations, as shown in Figure 25.

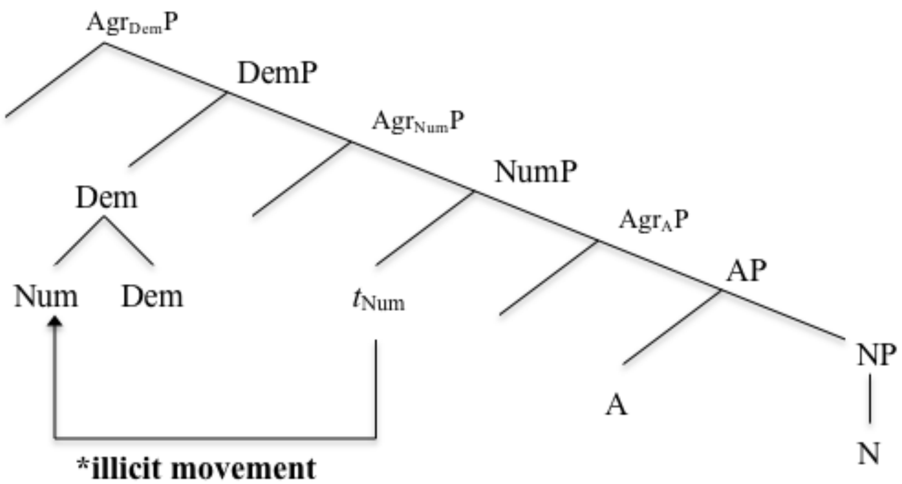


Figure 25 Order (e) with head movement

However, this kind of derivation is ruled out by principle (28c), the requirement that only XPs move.

Orders (f), (j), and (v) are also ruled out for the reasons just cited. They differ from (e), (i), (m), (q), and (u) in (32) in that N is not the rightmost constituent but precedes one modifier. As seen in (33), though, the two prenominal modifiers do not respect the universal merge order.

(33) Prenominal ordering of modifers does not respect base order Dem > Num > A

(f) ***Num Dem** N A
(j) ***A Dem** N Num
(v) ***A Num** N Dem

Once again, there is no way to alter the prenominal order of modifiers unless either head movement or remnant movement of a constituent not containing N are invoked, which are both disallowed in Cinque's system.

Two patterns remain, namely (g) (*Num N Dem A) and (h) (*N Num Dem A). In orders (g) and (h), N and Num – but not A – need to be to the left of Dem. This is impossible, however, since NumP and NP do not form a constituent to the exclusion of AP. Even if NP moves to the left of NumP first, as required in order (h), the only movable constituent containing both NP and NumP will still include AP, as shown in Figure 26.

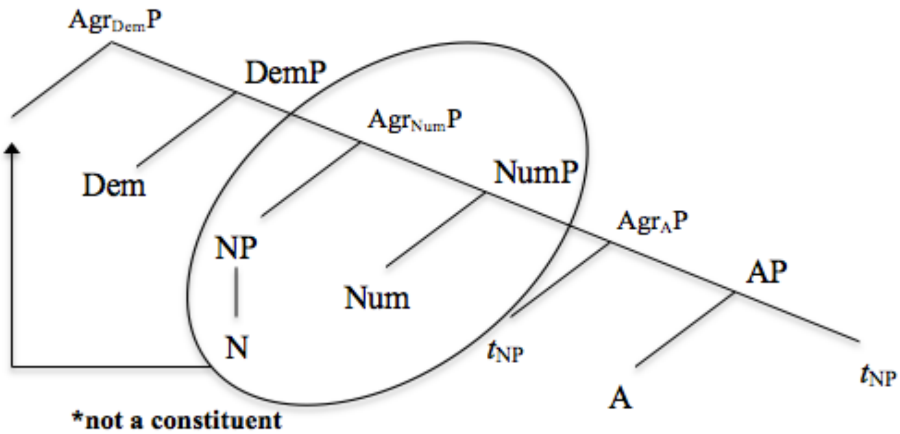


Figure 26 Order (h) needs movement of a non-constituent

Again, head movement is illicit by (28c) and thus cannot be called upon to switch any of the modifiers' positions.

2.1.3.3 Why moved structures must contain the head noun

For the sake of completeness, in this section I explain Cinque's (2005: 325-327) view on why principle (28d), the requirement that moved structures must always contain the head noun, should hold. The discussion will also touch on the role of AgrP.

As members of the noun's extended projection, the elements Dem, Num, and A need to be licensed by a nominal feature [N] (see also Grimshaw 1979, 1991, 2000). In Cinque (2005) this licensing requirement is assumed to take place in the AgrPs interspersed throughout the nominal extension. Some languages move NP itself into AgrP in order to satisfy this licensing requirement. Other languages externally merge an [N] feature in AgrP, and from there the feature [N] establishes an Agree relation with the head noun. Still other languages do a combination of the two (partially movement, partially [N]-merge plus Agree) (Cinque 2005: 325-326).

According to Cinque (2005: 326-327), the question of why the head N must always be contained in structures that move may be reducible to Kayne's (2005: 331-332) condition on movement, given in (34). In fact, (34) can explain not only why movement must always implicate the head noun, but also why remnant movement is ruled out.

(34) Kayne's (2005: 331-332) condition on feature-driven movement

Move to Spec,H the category closest to H (that is not [the complement of H]).

According to (34), movement to the specifier of a head H must always apply to the closest c-commanded category which is not H's complement. Note that this resembles antilocality (Abels 2003).

To take an example, consider first a derivation in which NP moves to the left of AP, landing in Agr$_{A}$P. This is shown in (35).

(35) $[_{AgrNumP}$ **PROBE** $[_{NumP}$ Num $[_{\mathbf{AgrAP}}$ $[_{NP}$ N] $[_{AP}$ A $[t_{NP}]]]]]$

From here, the Agr$_{Num}$ probe will only be able to attract Agr$_{A}$P. This is because Agr$_{A}$P is the closest category which is not the complement of Agr$_{Num}$. The phrase NumP cannot be moved because it is too close, being the complement of Agr$_{Num}$. On the other hand, the remnant [AP t_{NP}] is too far away, since Agr$_{A}$P is a closer non-complement of Agr$_{Num}$. The result, then, is movement of [NP AP t_{NP}] to the left of NumP.

Now, had NP not moved to Agr$_{A}$P and an [N] feature been merged there instead, then Agr$_{Num}$ would still have to target Agr$_{A}$P for movement, by the same logic. This is illustrated in (36).

(36) $[_{AgrNumP}$ **PROBE** $[_{NumP}$ Num $[_{\mathbf{AgrAP}}$ *[N]* $[_{AP}$ A $[_{NP}$ N]]]]]$

This time, though, the movement would be a pied-piping operation of [AP NP] – instead of [NP AP t_{NP}] – to the left of NumP.

Finally, one would like to know why languages can utilize cyclic (Spec-to-Spec) movement of NP alone, without pied-piping. For these types of languages, Cinque (2005: 326, fn. 33) suggests that Kayne's (1994) conception of asymmetric c-command might be at work. According to Kayne (1994), specifiers are adjoined, as indicated in Figure 27.

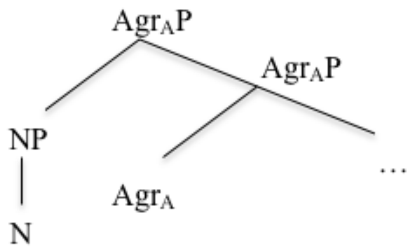


Figure 27 NP in Spec-Agr_AP

By Kayne's (1994: 16) definition of c-command, the NP in Spec-Agr_AP c-commands Agr_AP but Agr_AP does not c-command NP. C-command of Agr_AP over NP is blocked because only a segment of Agr_AP dominates NP, and c-command is restricted to categories by Kayne's assumption. Such an asymmetric c-command relation might entail that NP can qualify as closer to the probe than Agr_AP, accounting for the movement of NP without pied-piped material.

Though these finer details of Cinque's theory are less relevant for this dissertation, it is nevertheless interesting to note the different crosslinguistic possibilities with regard to movement that Cinque lays out. That is, sometimes there is 'no movement', when [N] is externally merged; sometimes there is 'cyclic movement', when just NP moves to check the nominal requirement; and sometimes there is 'movement with pied-piping', when a large structure containing more than just NP moves to the left of the probing layer. These possibilities will appear in a different guise in Section 2.2.

2.2 The theory of nanosyntax

In this section I introduce the theory of nanosyntax (Starke 2009, 2011ab, 2013; Caha 2009, 2010, 2013; Taraldsen 2009; Lundquist 2008; Ramchand 2008; Fábregas 2007, 2009; Muriungi 2008; Pantcheva 2011; Vangsnes 2013, 2014; Rocquet 2013; De Clercq

2014).[43] In Section 2.2.1 I explain the importance of heads being submorphemic. In Section 2.2.2 the architecture of grammar according to nanosyntax is explained. In Section 2.2.3 the structure of lexical entries and their importance to nanosyntactic theory will be discussed. In Section 2.2.4 the basic 'tools' of nanosyntax are introduced, in particular the phenomenon of syncretism, which serves as a central tool for discovering fine-grained structure in nanosyntax. In Section 2.2.5 the principles of spellout (i.e. lexicalization) are introduced: the Superset Principle, the Elsewhere Principle, and the Principle of Cyclic Override. In Section 2.2.6 the Anchor Condition, an important condition on spellout, is explained. In Section 2.2.7 Michal Starke's theory of spellout-driven movement and its ability to derive suffixal positioning is discussed. In Section 2.2.8 the concept of complex heads is introduced and discussed; complex heads, as we will see, are prefix-like constituents. Section 2.2.9 is a brief note on subextraction and its place in the theory and in this dissertation. Section 2.2.10 introduces a shorthand for writing out functional sequences and the way these sequences are divided up into morphemes; it also provides a quick way to compute the effects of spellout-driven movement on a given structure, which will be useful for later chapters. Section 2.2.11 concludes the chapter.

Nanosyntax builds on a foundation of a few basic cartographic assumptions. These are useful to point out right at the beginning of an in-depth discussion of nanosyntactic theory.

(37) Basic cartographic assumptions in nanosyntax

(i) strict syntax-semantics mapping

(ii) simplicity of projection (Kayne 1984, 1994):

(a) binary branching
(b) highly restricted adjunction
(c) only leftward movement

(iii) one feature = one head (OFOH; Cinque & Rizzi 2008: 50)

The OFOH maxim is especially important for the architecture of nanosyntax. Like cartography, nanosyntax has as its aim the identification of the atoms of linguistic structure. As in cartography, these atoms are taken to be syntactico-semantic features. By

[43] There are no textbook introductions to nanosyntax. The theory has been developed by Michal Starke in a series of lectures and seminars at the University of Tromsø from 2005 to the present. The main source for this section is Starke (2011b), a seminar taught at Ghent University. Further source materials come from Starke (2009, 2011a, 2013). The rest of this chapter follows Baunaz & Lander (in prep.), an introductory chapter for an upcoming volume on nanosyntax, quite closely.

OFOH, features are heads, and heads are merged according to the rigid ordering dictated by the functional spine. Thus OFOH is a guiding maxim both for establishing the fine-grained inventory of features shared by all languages and for determining the order in which these features are merged as heads.

Another important guide for nanosyntacticians is Cinque's (2005) U20 program. Various aspects of nanosyntactic theory can be traced to Cinque's system, such as the prominence of phrasal movement. The spellout algorithm of nanosyntax is also parallel to the different kinds of movement options (i.e. different degrees of pied-piping) available in Cinque's system, as we shall see in Section 2.2.7. The idea that any moved constituent must include the head of the extended projection is also underlyingly present in nanosyntactic views on movement.

Taking the assumptions in (37) together gives us a view of linguistic structure which is very restrictive. While many theories of grammar allow for unordered bundles of syntactico-semantic features, nanosyntax rejects this view. In nanosyntax, all syntactico-semantic features are ordered in an asymmetrical sequence as heads, as seen in Figure 28 (see Dékány 2009: 51).

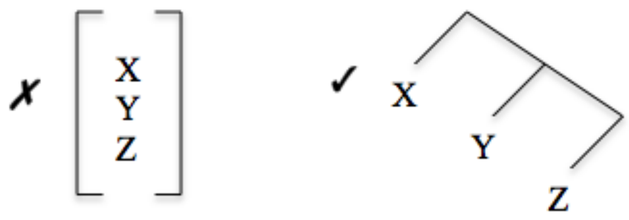


Figure 28 Unordered vs. ordered features

Figure 28 illustrates perhaps the most central concept in nanosyntax, even more so than so-called phrasal spellout, which will be discussed below.

2.2.1 Submorphemic heads

Nanosyntax offers a novel proposal concerning the architecture of grammar. In particular, the radical application of the syntactico-semantic representations as developed in cartographic work has important repercussions for the nature of syntactic structure, the lexicon, and lexicalization (or *spellout*).

One important outcome of the nanosyntactic methodology concerns the relation between syntactic heads and morphemes. According to the traditional view, morphemes are drawn from the lexicon and are inserted at syntactic terminals. In other words, the head-to-morpheme relation is always one-to-one. However, in actual fact, it is common to find morphemes which must consist of more than a single head. Take the difference between agglutinative languages like Finnish and fusional languages like ON. In agglutinative languages, functional categories like case (K) and number (#) are encoded as separate morphemes, while in fusional languages these categories are conflated into a single morpheme. In Finnish (38a), there are separate morphemes for the K (*-lle*) and # (*-i*) ingredients. In ON (38b), one morpheme (*-um*) expresses both K and # at the same time.

(38) (a) Finnish (from Caha 2009: 73)

karhu-i-lle
bear-PL-ALL
'onto the bears'

(b) Old Norse

bjǫrn-um
bear-DAT.PL
'for the bears'

While there is thus a one-to-one correspondence between morphological realization and functional category in Finnish, there is a one-to-many correspondence between morphological realization and functional category in ON (and Indo-European in general). In Finnish, K and # are morphemic, in ON K and # are submorphemic. Put differently, in ON the morpheme is made up of multiple features. By the OFOH maxim this means that one morpheme corresponds to multiple heads.

As just seen, morphemes can be made up of multiple features, and thus multiple heads by the OFOH maxim. By standard cartographic assumptions, moreover, heads are merged in a strict order in the fseq. Since morphemes are made up of heads, then, this means that they have internal hierarchical structure. In Finnish, for instance, we see in (38a) above that the # morpheme *-i* is merged closer to the nominal stem *karhu-* than the K morpheme *-lle*. That # is merged closer to N than K is a crosslinguistically systematic fact, and it is evidence that the underlying merge order at stake is K > # > N. This hierarchy applies to the ON morpheme *-um* in (38b) as well, even though *-um* does not show it overtly as with Finnish *-i* and *-lle*. Thus, *-um* corresponds to the chunk of structure [K [#]] (rather than the unordered bundle [K, #]). In other words, fusional morphemes have internal structure the same way that agglutinative morphemes do.

Traditionally, spellout targets only terminals (X^0). This kind of system, *per se*, cannot lexicalize two separate terminals at the same time, which is exactly what is needed for morphemes like *-um* in ON. This is why nanosyntax adopts *phrasal spellout*, meaning that XPs can be targeted for spellout (outside of nanosyntax, see also McCawley 1968, Weerman & Evers-Vermeul 2002, Neeleman & Szendrői 2007). While in Finnish KP and #P are targeted individually for spelling out *-lle* and *-i*, respectively, as seen in Figure 29, in ON the entire phrase [$_{KP}$ K [$_{\#P}$ #]] can be targeted for spelling out the case ending *-um*, as seen in Figure 30.

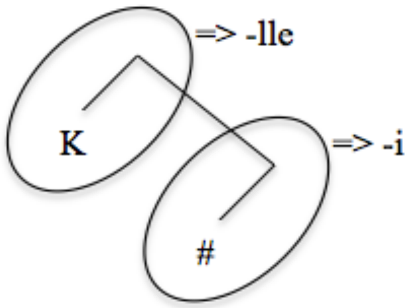


Figure 29 Spellout in Finnish

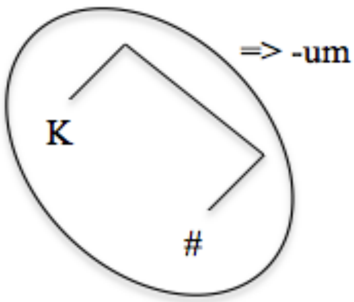


Figure 30 Spellout in ON

The reason *-lle* and *-i* target the phrases KP and #P (rather than just the terminals K and #) will emerge below when spellout-driven movement is discussed.

Fuller derivations for Finnish *karhu-i-lle* and ON *bjǫrn-um* can be seen in Figures 31 and 32.

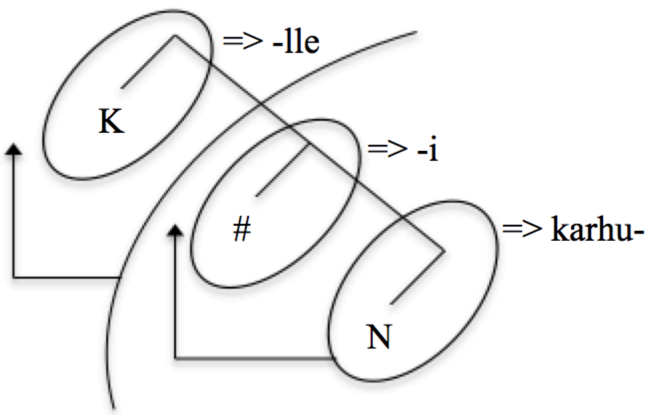


Figure 31 Finnish *karhu-i-lle*

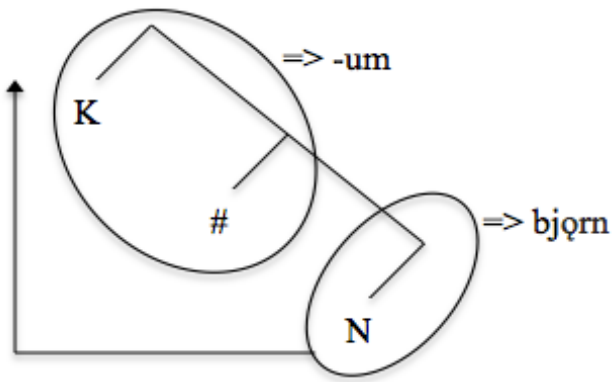


Figure 32 ON *bjǫrn-um*

Roughly speaking, in Figure 31 NP first moves to the left of #P, and then NP pied-pipes #P to the left of KP, giving the rolled-up structure [[[NP] #P] KP]. In Figure 32 NP moves to the left of [KP #P], giving the structure [[NP] KP #P]. Below I will discuss the finer details of these derivations, but for now the crucial observation is that phrasal spellout is needed at least to account for portmanteau morphemes which consist of more than a single layer of functional structure.[44, 45]

Note here that a system with phrasal spellout can still in principle allow for terminal spellout. For some more detailed discussion of this, see Section 5.1.2.

---

44 An alternative that appears to produce results similar to phrasal spellout is the operation Fusion from Distributed Morphology, which causes two adjacent terminals to conflate into a single terminal (Halle & Marantz 1993, Embick & Noyer 2007). There are a number of objections to fusion which make it less desirable than phrasal spellout. First, as Taraldsen (2009: §1.2.2) points out, Fusion is part of neither the syntax nor the spellout process; it is an operation that belongs in the special Morphology module, an expensive addition to the architecture of grammar which nanosyntax avoids by simply positing the existence of phrasal spellout. Taraldsen also points out that the implementation of Fusion will either require a version of look-ahead or a huge amount of 'filtering'. That is, to get Fusion to apply in only the correct environments (e.g. in *mice* but not in *horse-s*) the operation must know in advance which Vocabulary Item is going to be inserted (i.e. before Vocabulary Insertion, which happens late, indeed post-syntactically, in the PF branch according to the Distributed Morphology model), in order to make the decision whether or not to Fuse, or it must produce all possible Fusion products and then filter these results later when the lexicon is accessed. Both of these options are undesirable. The nanosyntactic alternative is the 'spellout loop' between syntax and the lexicon.

45 Another option for mimicking phrasal spellout that should be mentioned is the view that the K/# morpheme *-um* is hosted by, say, K and that # hosts a null morpheme (cf. Kayne 2005 for instance). The problem here is twofold. First we lose the intuition that the *-um* morpheme really consists of both K and #: *-um* is not more of a K morpheme or more of a # morpheme. Second it is very difficult to properly restrict the application of null morphemes, guaranteeing that they appear only when portmanteau morphemes also happen to appear. This type of morphemic coordination is not required if phrasal spellout is adopted.

2.2.2 The architecture of grammar

This section shows how the nanosyntactic perspective leads to a model of grammar that differs from the traditional minimalist Y-model in important ways.

Morphemes and features/heads do not have the same granularity. A typical morpheme will be made up of multiple features. Moreover, since features are merged as heads according to the fseq, morphemes have internal structure corresponding to the fseq. Simply put, syntax is responsible for building morphemes. Morphemes are not inserted prefabricated into syntactic structure; morphological structure *is* syntactic structure.

If syntax is responsible for the morphemes which are found in the lexicon, then we must reconceptualize the architecture of grammar. First, it cannot be that syntax/semantics[46] and morphology are distinct modules. There must instead be a single syntax-morphology-semantics module, call it SMS. SMS is the syntactic engine, the Language Faculty which is central to generative theories of grammar. Second, since SMS (henceforth simply *syntax*) takes features and builds morphemes out of them, it must be the case that syntax feeds the lexicon, that is, that the lexicon is postsyntactic.

The nanosyntactic view of grammar is shown in Figure 33; compare this with the traditional model given in Figure 8 above and repeated for convenience as Figure 34.

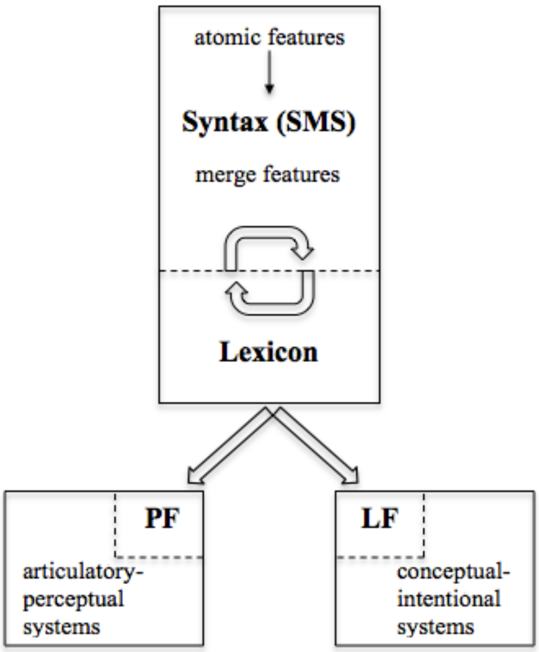


Figure 33 Architecture of grammar according to nanosyntax (based on Caha 2009: 52)

[46] Grammatical or formal semantics, not conceptual semantics (pragmatics).

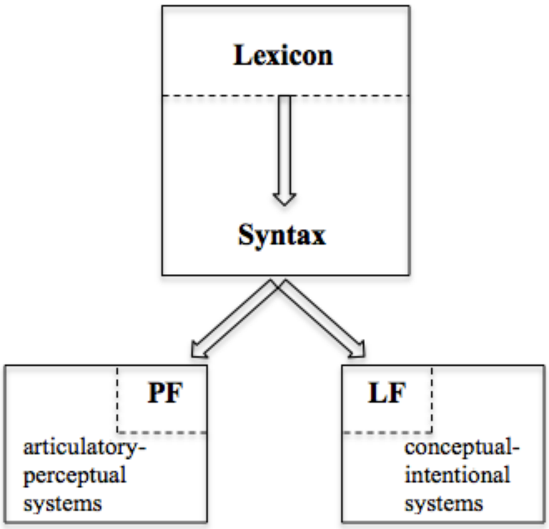


Figure 34 Traditional minimalist architecture of grammar

In the traditional minimalist view of grammar, LIs from the lexicon are inserted into syntax early on. The syntactic derivation is then driven by the featural makeup of the LIs before branching into PF and LF. Structure-building in nanosyntax starts off by syntax assembling atomic features according to the functional sequence.[47] The structures built by the syntax are then checked against and matched by lexical entries in the lexicon. This process of matching L-trees to S-trees is known as *spellout* and is a central component of the theory (see Section 2.2.7). In other words, nanosyntax is a 'late insertion' model (a facet it shares with Distributed Morphology). After lexical insertion the derivation branches into the interfaces with the output systems: the articulatory-perceptual systems at PF and the conceptual-intentional systems at LF.

There are major consequences of postulating a postsyntactic lexicon. Above all, since syntax builds the morphemes which constitute the lexicon, it follows that lexical structure is simply syntactic structure. Lexical items are lexical trees, i.e. syntactic trees located in the lexicon. Though there is no fundamental, substantive difference between syntactic trees and lexical trees, it is nevertheless useful to make a distinction between them. Henceforth, trees built by the syntax will be called *S-trees* and trees stored in the lexicon will be called *L-trees*. In other words we distinguish trees based on *where* they are located (in the syntax or in the lexicon).

With this in mind, observe that the above has consequences for language acquisition and language production. Language acquisition can be formulated in terms of children

[47] We may assume that the fseq is stated in full as a principle of nanosyntax.

storing in their lexicon 'interesting' syntactic structures they come across. As for language production, we can think of S-trees as linguistic expressions of thought. Thoughts are filtered through the syntax, which churns out S-trees ultimately externalized via spellout. Thus language production is equivalent to the production of S-trees by the syntactic engine.

There are also consequences for the entire framework of Principles and Parameters. Starke (2011a) discusses the nanosyntactic perspective on the Principles and Parameters program. According to him, the atoms of syntax and their universal merge order can be viewed as the Principles of language. The way the fseq is divided up from language to language into lexical entries, on the other hand, can be viewed as the Parameters of language. As discussed above, the way L-trees are structured has a direct effect on how the syntactic structure is altered by spellout-driven movement. Thus, the same S-tree made up of the same heads merged in the same order will be realized differently in language A vs. language B because A and B have differently shaped L-trees available to spell out the S-tree. Language variation, in other words, comes down to the language-specific lexicon (cf. Chomsky 2001: 2).

To take a simple example from Pantcheva (2011), consider the item 'to(wards)' in Macedonian as opposed to Dutch. In Macedonian, the item *nakaj* 'to(wards)' can be decomposed into a Place marker *-kaj* 'at' and the Path marker *na-*. In other words, in Macedonian the fseq Path > Place is split up into two separate morphemes. In Dutch, on the other hand, there is a single portmanteau form *naar* meaning 'to(wards)'. Thus both Path and Place are realized together as *naar* in Dutch. See Figure 35.

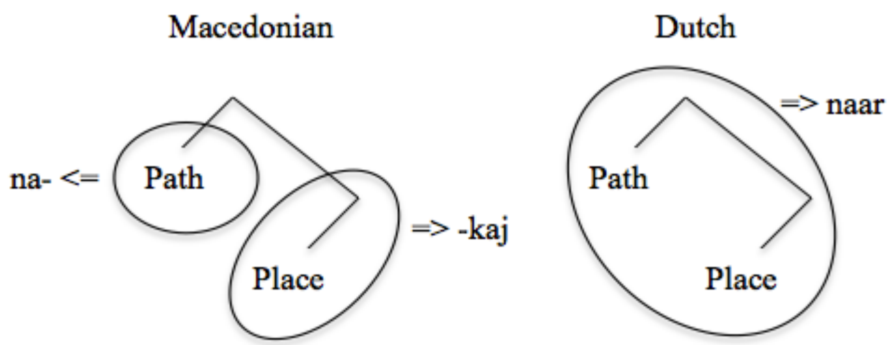


Figure 35 Simple case of crosslinguistic variation

The same fseq is lexically divided up in different ways in Macedonian vs. Dutch. In nanosyntax, the fine-grained fseq offers a precise way to capture crosslinguistic variation. The hope is that fine-grained differences in the way lexical entries are shaped will be able to explain all or almost all of the crosslinguistic variation observed. In Chapter 6 I will provide a case study from Old Germanic which attempts to capture the crosslinguistic variation observed in exactly this way.

2.2.3 Lexical entries

This section introduces the internal structure of lexical entries and explains the importance of lexical entries in understanding idioms.

The lexicon, filled with lexical entries, is a list of unpredictable linguistic information which essentially needs to be, for lack of a better word, memorized by the speaker in the course of acquisition. Lexical entries act as the link between the generative component (syntax) on the one hand and the output systems, PF and LF, on the other hand. Thus a lexical entry can link together three types of information: phonological, syntactic, and conceptual. This is seen in Figure 36.

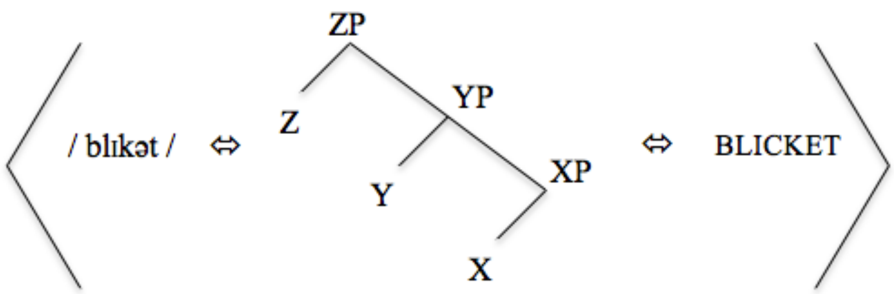


Figure 36 Structure of a lexical entry

The middle slot, for syntax, is the relevant one for the spellout process. The other two slots are needed for the PF and LF branches of the derivation.

2.2.3.1 Idioms

Lexical entries do not always use all three slots, however. For instance, the lexical entry of an irregular form, such as the irregular plural *mice*, does not contain any conceptual information beyond that of the components *mouse* and *-s* (i.e. **mouse-s*, which *mice* replaces). That is, the crucial information about the form *mice* is that it is composed of the NP token *mouse* plus the plural marker *-s* but with a special phonology *mice*. This is shown in (39).

(39) < *mice* ⇔ [[*mouse*]-*s*] ⇔ ___ >

The lexical entry in (39) can be called a *phonological idiom*. Phonological idioms have a special phonological form, like *mice*, that corresponds to a specific syntactic configuration of tokens, like [*mouse*] plus [-*s*].

On the other side of the coin, *conceptual idioms* do not contribute any special phonology but do contribute special conceptual information. For instance, the idiom *hold your horses* meaning ‘be patient’ does not require any special phonology. The particular syntactic combination of [*hold* [*your* [*horses*]]], however, does have a special meaning which needs to be indicated in the lexical entry, as shown in (40).

(40) < ___ ⇔ [*hold* [*your* [*horses*]]] ⇔ BE PATIENT >

In other words, the literal, compositional meaning of *hold your horses* is overridden by the meaning 'be patient'.

Interestingly, the existence of idioms is an argument itself for phrasal spellout, since such a spellout system actually predicts that idioms should exist.

> Idioms are *prima facie* an important source of support for [the nanosyntactic notion of] phrasal spellout. Within the traditional approach, there is no easy way to handle multi-word idiomatic expressions, as witnessed by the clunkiness of existing attempts at handling idioms while at the same time confining spellout to terminals. Under phrasal spellout, idioms are natural: they are cases in which a relatively high-level constituent has been stored. (Starke 2011a: 6)

Starke (2011a: 6-7) goes on to point out that idioms show us that there must be a feedback loop between syntax and the lexicon (see Figure 33), sicne the idiomatic VP *hold your horses* must be constructed word-for-word before the entire VP can then be matched by the idiomatic entry for 'be patient'.

Lexicon

< *hold* ⇔ V ⇔ HOLD >
< *your* ⇔ D ⇔ YOUR >
< *horses* ⇔ N ⇔ HORSES >

< [*hold* [*your* [*horses*]]] ⇔ BE PATIENT >

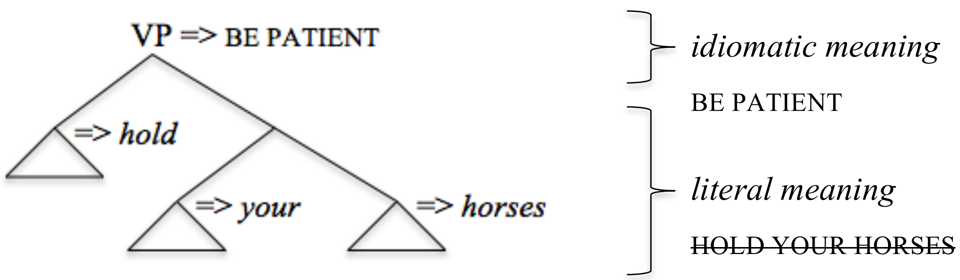


Figure 37 Cyclic spellout of a conceptual idiom

The example in Figure 37 should be interpreted as follows. Imagine that there is a lexical entry each for *horses*, *your*, and *hold* (certainly a simplification considering that *horses* is composed at least of *horse* plus *-s*). Assuming a bottom-up derivation, the structure in Figure 37 will first be spelled out piece by piece, giving [*hold* [*your* [*horses*]]], which corresponds to the literal, compositional meaning of this phrase. However, if there is also a lexical entry like (40) available, which matches the entire VP, then the literal meaning of *hold your horses* will be overridden at VP with the meaning 'be patient'. Without a special lexical entry specifying the exact configuration [*hold* [*your* [*horses*]]], the idiomatic meaning will not arise and will not replace the literal meaning by cyclic override (such as for people who are not familiar with this idiom).

This concept of cyclic override is also relevant for phonological idioms, such as *mice*, as illustrated in Figure 38.

Lexicon

< *mouse* ⇔ N >
< *-s* ⇔ Num >

< *mice* ⇔ [[*mouse*]-*s*] >

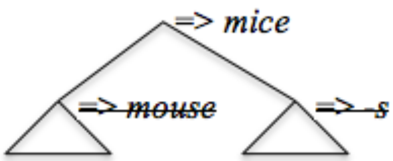


Figure 38 Cyclic spellout of a phonological idiom

As seen in Figure 38, the regular form **mouse-s* is first spelled out. However, since there is a lexical entry structurally specified exactly for the configuration [[*mouse*]-*s*] in the lexicon, the regular phonological configuration **mouse-s* is overridden by the phonology in the idiomatic entry, namely *mice*. Children, for instance, who have not yet acquired the entry < mice ⇔ [[mouse]-s] > will produce the regularly formed *mouse-s*.

2.2.3.2 Pointers

Lexical entries like *mice* contain or refer to other lexical entries. Put differently, we can say that some entries 'point to' other entries. The entry for *mice*, for instance, points to the entries for *mouse* and *-s*.

To illustrate this concept more clearly and in relation to the core topic of this dissertation, consider two irregular Dem forms from ON introduced in Chapter 1: F.NOM.SG *sú* and M.NOM.SG *sá*. The regularly composed forms would be the unattested F.NOM.SG **þa-u* and M.NOM.SG **þa-r*. In order to capture these irregular forms, we can utilize pointer entries in writing the lexical entries for *sú* and *sá*, as shown in (41) and (42).[48]

(41) $<_{210}$ *þa-* ⇔ DP >

$<_{320}$ *-u* ⇔ $K_{F.NOM.SG}P$ >

$<_{325}$ *-r* ⇔ $K_{M.NOM.SG}P$ >

(42) Pointer entries

$<_{2386}$ *sū* ⇔ [[210] 320] >

$<_{2387}$ *sā* ⇔ [[210] 325] >

It has also been argued that pointers can be used to encode complex syncretisms involving multiple featural dimensions (e.g. case, number, gender) (Caha & Pantcheva 2012). The concepts of irregular forms and pointers will be used rather often below.

2.2.4 Tools for uncovering fine-grained structure

In order to identify functional features and map out functional sequences, nanosyntacticians make use of three basic tools: semantic compositionality, syncretism, and morphological containment. In this dissertation only syncretism will play a major role, but for the sake of completeness I will discuss the other two tools here as well.

2.2.4.1 Semantic compositionality

One way to identify functional heads and to deduce their underlying sequence is to investigate how the semantic ingredients fit together compositionally. For instance, Starke

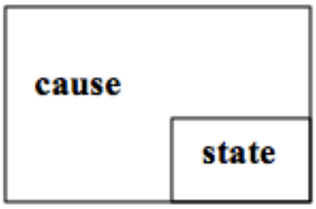
[48] When it comes to the discussion of grammatical items studied in this dissertation it will not be necessary to discuss the interface with the conceptual-intentional module. Nor will I employ the exact, IPA-written phonological form of the relevant morphemes below. Due to these simplifications, the lexical entries used here will have the following basic form: < *form* ⇔ [XP [YP]] >. At times the lexical entries will be numbered with an (arbitrary) numerical index for convenience.

(2012) points out that stative passives (43a) are smaller than eventive passives (43b) in the sense that a state is acted upon by a cause in order to become an event.

(43) (a) The pool is empty *(by the pool guy).

(b) The pool was emptied (by the pool guy).

(43a) describes the state of the pool; it does not invoke any outside force. (43b) invokes a force causing the pool to go from full to empty. The difference between an unacted-upon state and a caused state becomes especially clear in the ability of the eventive in (43b) to take an external argument in the form of a *by*-phrase, whereas this is impossible for the stative in (43a). The compositionality of semantics in this case is roughly sketched in Figure 39.



E.g. *state* = an empty pool
cause = the pool guy
} **event**: the pool guy emptied the pool

Figure 39 cause > state

In other words, eventive structures should be more featurally complex than stative ones because semantically speaking they contain a cause in addition to a state.

Another example of the use of compositionality of semantics comes from Pantcheva's (2011) fseq for Path, namely Route > Source > Goal. The semantics of the highest item, Route '**from** X **to** Y', can be seen as being composed of the Source reading '**from** X' plus the Goal reading '**to** X'. Thus it makes sense from a semantic point of view to say that Route is built on top of Source and Goal. See also Ramchand (2008) and Fábregas (2009) for additional cases.

It should be pointed out that the compositionality of semantics is not the sole diagnostic for determining the order of a functional sequence. It is not always commonsensical to decide how semantic ingredients compose, as researchers may differ in their opinions about what 'makes sense' semantically. Thus semantic compositionality should be used in combination with the other tools (morphological containment and syncretism) and as a tool for verifying results as they arise.[49]

2.2.4.2 Syncretism

When one single morpheme expresses more than a single grammatical function, a syncretism is at stake. A concise definition of syncretism would be “a surface conflation of two distinct morphosyntactic structures” (Caha 2009: 6). Exactly how a single morphological exponent can spell out different syntactic structures will be discussed in Section 2.2.5.1.

As an example of syncretism, consider the expression of Location, Goal, and Source readings in English (44) and French (45) (Pantcheva 2011: 238).

(44) (a) I ran **at** the stadium. Location

(b) I ran **to** the stadium. Goal

(c) I ran **from** the stadium. Source

(45) (a) J’ai couru au stade.
I.have run at.the/to.the stadium
‘I ran **at** the stadium.’ *or* ‘I ran **to** the stadium.’ Location/Goal

(b) J’ai couru du stade.
I.have run from.the stadium
‘I ran **from** the stadium.’ Source

In English (44) there is no syncretism between Location, Goal, and Source. Each reading is expressed by a different preposition: *at*, *to*, and *from*, respectively. In French (45), however, the preposition *à* expresses both the Location and Goal readings (45a), while the preposition *de* expresses the Source reading (45b). Thus there is a Location/Goal syncretism in French, but not in English. Various other syncretism patterns are attested crosslinguistically (see Pantcheva 2011: §§8 and 9, building on work by Svenonius, e.g. Svenonius 2010).

Building on Blake (1994), Caha (2009) investigates nominative-accusative case systems from a crosslinguistic perspective. He finds that case syncretisms target only adjacent or contiguous cells in a paradigm. To demonstrate this concept, consider Table

[49] It is an open question whether or not the fseq can be derived purely from semantic relations (or put differently to what degree syntactic selection needs to be stipulated in the fseq). For some discussion see Nilsen (2003).

31 from Russian, where five cases (nominative, accusative, genitive, dative, and instrumental) are considered.

Table 31 Syncretisms in Russian (from Caha 2009: 12)

	'window.SG'	'teacher.PL'	'one hundred'
NOM	okn-o	učitel-ja	st-o
ACC	okn-o	učitel-ej	st-o
GEN	okn-a	učitel-ej	st-a
DAT	okn-u	učitel-am	st-a
INS	okn-om	učitel-ami	st-a

As indicated by the shading in Table 31, the attested syncretisms are NOM/ACC, ACC/GEN, and GEN/DAT/INS. The order of cases (NOM – ACC – GEN – DAT – INS) is such that only adjacent cells are syncretic. In fact, Caha (2009) shows that even when many additional languages and additional cases are investigated, there is still a single sequence which captures the syncretism patterns in terms of adjacency. His generalization is summarized in (46).

(46) Universal Case Contiguity (Caha 2009: 49)

(a) Non-accidental[50] case syncretism targets contiguous regions in a sequence invariant across languages.

(b) Case sequence: NOM – ACC – GEN – DAT – INS – COM

Assuming (46a), the sequence in (46b) makes predictions about possible syncretisms. Non-contiguous regions that 'skip' a cell in the paradigm should not display syncretism, as illustrated in Table 32. This is often called the *ABA theorem.

Table 32 Unattested syncretisms: *ABA (based on Rocquet 2013: 32)

NOM	A	
ACC	B	
GEN	A	A
DAT		B
INS		A

[50] This is a reference to homophony. Sometimes phonological rules cause two phonological structures which are underlyingly distinct to surface identically, giving the impression that there is a syncretism. However, in these cases the phenomenon is phonological and not morphosyntactic. In other words, homophony belongs to PF and not to narrow syntax (SMS), and so it tells us nothing about the functional sequence.

There are very few exceptions to this generalization. The *ABA theorem has its origins in Distributed Morphology, having been investigated at length by Bobaljik (2007, 2012) in his work on comparative and superlative adjectives and deadjectival verbs, where he finds that the *ABA generalization is crosslinguistically very robust. Thanks to Caha (2009), the *ABA theorem is now central to nanosyntactic theory.

Syncretism patterns are a guide to the relation between heads of a particular domain. By looking at syncretism patterns, one can thus deduce the *linear order* of the functional heads involved. Crucially, the tool is capable of making very fine-grained distinctions, revealing the very atoms of syntactic structure.[51]

2.2.4.3 Morphological containment

Syncretism reveals an underlying order of heads, such as NOM – ACC – GEN – DAT – INS – COM, but it does not reveal the hierarchical order, that is, whether the case sequence is structured as NOM > ACC > GEN > DAT > INS > COM or as COM > INS > DAT > GEN > ACC > NOM. Morphological containment provides a way to handle this problem.

Heads in nanosyntax are usually understood to be unary and additive/cumulative (Caha 2009: §1.2, Starke 2011b). This means, most importantly, that a functional sequence is understood in terms of superset-subset relations, as illustrated in Figure 40 for the case (K) sequence. In Figure 40, the smallest/lowest case, A, corresponds to only a single head, K_1. The second smallest/lowest case in Figure 40, B, corresponds to K_1 plus K_2, and so on. Thus, A is a subset of B (and B is a superset of A).

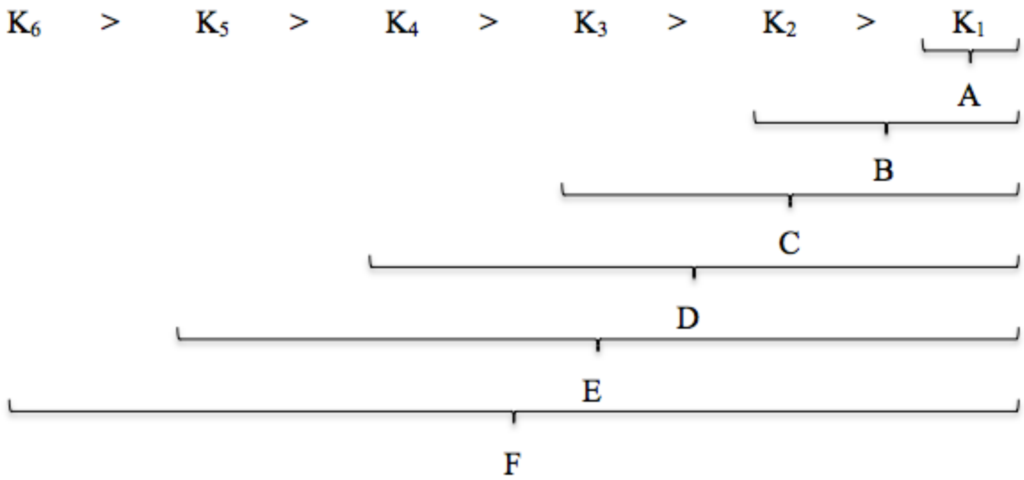


Figure 40 Containment in the K sequence

[51] The nanosyntactic tool of looking at syncretism patterns in terms of structural adjacency resembles the cartographic tool of 'transitive ordering' which is used extensively in work such as Cinque (1999): if X > Y and Y > Z, then X > Y > Z.

The fact that sequences are built up cumulatively in this manner predicts that the superset-subset relations at the abstract level of these formal features might also be realized overtly in the morphology. This is indeed the case, as the following examples illustrate.

In (47) we see two examples of morphological containment from West Tocharian and Russian. In West Tocharian (47a) the accusative ending *-ṃ* is contained within the genitive/dative ending *-ṃts*. In Russian (47b) the dative ending *-am* is contained within the instrumental ending *-ami*.

(47) (a) West Tocharian (Caha 2009: 69)

ACC.PL -ṃ

GEN/DAT.PL -ṃts

[$_{\text{GEN/DAT.PL}}$ [$_{\text{ACC.PL}}$ -ṃ] -ts] **[GEN/DAT [ACC]]**

(b) Russian (Caha 2009: 12)

DAT.PL -am

INS.PL -ami

[$_{\text{INS.PL}}$ [$_{\text{DAT.PL}}$ -am] -i] **[INS [DAT]]**

In other words, (47) shows how ACC is quite literally contained within GEN/DAT, and how DAT is literally contained within INS. Since GEN/DAT is larger than ACC and INS is larger than DAT, we know that the GEN/DAT layers are higher in the fseq than the ACC layer and that the INS layer is higher in the fseq than the DAT layer (given a system that builds structure from the bottom-up).

Caha (2009) extracts a similar generalization from facts about case selection by prepositions. He proposes that both K endings and prepositions consist of K heads. He also suggests that the way a P selects a case-marked nominal complement represents a containment relation. For example, in English (48a) the genitive preposition *of* selects an accusative complement (i.e. GEN contains ACC), in Arabic (48b) the dative preposition *li* selects a genitive complement (i.e. DAT contains GEN), and in German (48c) the instrumental preposition *mit* selects a dative complement (i.e. INS contains DAT).

(48) Case selection by prepositions (Caha 2009: 37)

(a) English *of*$_{\text{GEN}}$ + DP-ACC [GEN [ACC]]

(b) Arabic *li*$_{\text{DAT}}$ + DP-GEN [DAT [GEN]]

(c) German *mit*$_{\text{INS}}$ + DP-DAT [INS [DAT]]

These facts agree with the direction of containment observed in (47), where it was seen that GEN/DAT is larger than ACC (cf. (48a)) and that INS is larger than DAT (cf. (48c)). In (48b), moreover, we also see that DAT is larger than GEN. Taken together, the containment facts lead us to the hierarchy in Figure 41.

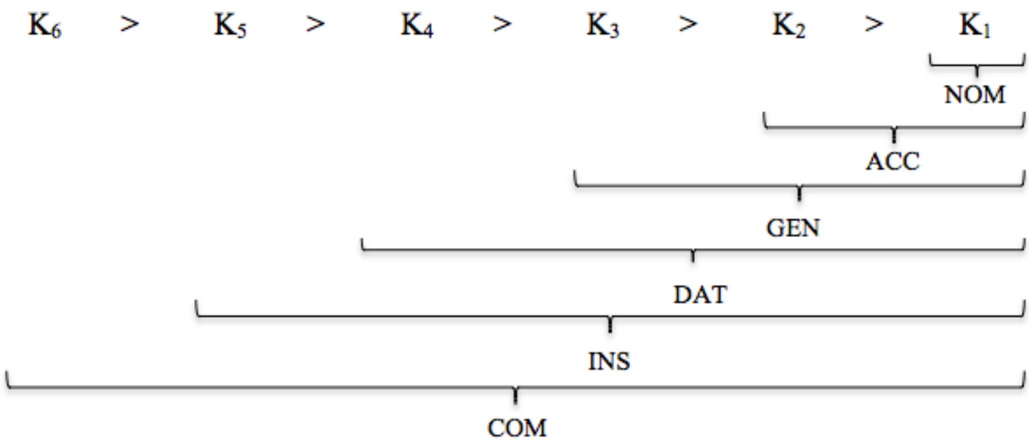


Figure 41 Containment in the K sequence

See also Bobaljik (2007, 2012) for more discussion of the phenomenon of morphological containment from a Distributed Morphology perspective.

2.2.5 The principles of spellout

Recall that syncretism is when a single morpheme can express more than a single grammatical function. In nanosyntactic terms, we can understand a morpheme as a lexical entry, i.e. a piece of structure (an L-tree) linked with some phonology (and perhaps with conceptual content as well). We can understand a grammatical function, moreover, as a morphosyntactic structure built up by the syntax, i.e. an S-tree. In other words, syncretism is the result of a single L-tree being able to match multiple S-trees.

Recall that syntax builds S-trees and that the lexicon is a storage space filled with L-trees. *Spellout* is the name for when S-trees generated by the syntax are matched by L-trees from the lexicon. The way S-trees are lexicalized by L-trees is governed by three principles: the Superset Principle (discussed in Section 2.2.5.1), the Elsewhere Principle, (discussed in Section 2.2.5.2), and the Principle of Cyclic Override (discussed in Section 2.2.5.3).

2.2.5.1 The Superset Principle

The primary principle of nanosyntactic spellout is that an L-tree must be a superset of the S-tree in order for match to succeed. This is called the Superset Principle.

(49) Superset Principle

An L-tree matches an S-tree if and only if the L-tree is a superset, proper or not, of the S-tree.

Put simply, the L-tree must be the same size or larger than the S-tree in order for match to succeed.

To illustrate the principles of spellout, I will use material drawn from Bobaljik's (2007, 2012) work on ablaut in the present, past participial, and simple past forms of verbs, in addition to unpublished work by Michal Starke, who has worked on very similar data from a nanosyntactic point of view. Three attested syncretism patterns of the present (PRES), past participle (PART), and past (PAST) forms of verbs are shown in Table 33.

Table 33 Syncretism patterns across three verb forms

	'come'	'shine'	'sing'
PRES	come	shine	sing
PART	come	shone	sung
PAST	came	shone	sang

In Table 33 we see that there can be a syncretism between the present and the participial form (e.g. *come*) to the exclusion of the past (*came*), a syncretism between the participial and past forms (e.g. *shone*) to the exclusion of the present (*shine*), or no syncretism, i.e. a separate form for the present, participle, and past (e.g. *sing*, *sung*, *sang*, respectively). By Caha's (2009) theory of syncretism, this suggests a hierarchy of verbal (V) heads like the one in Figure 42. In Figure 42, the present corresponds to V_1, the participle to V_1 plus V_2, and the past to V_1 plus V_2 plus V_3. Crucially, the layers responsible for the present and the participle are adjacent, as are the layers responsible for the participle and the past.

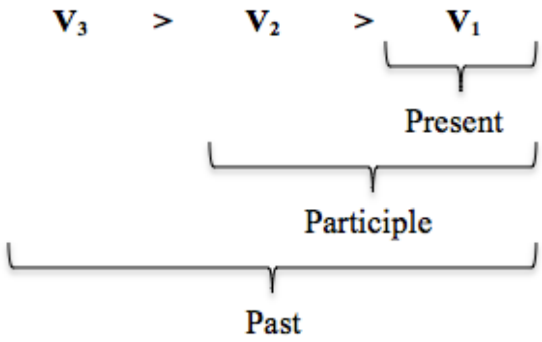


Figure 42 Cumulative fseq of V heads

The direction of containment (i.e. that V_3 corresponds to the past and V_1 to the present) is supported, for instance, by the containment of the present form within the participial form (e.g. English *give* contained within *give-n*, or Dutch *kom* 'come.PRES' contained within *ge-kom-en* 'come.PART'), meaning that the participle is a larger structure than the present.

For the English verb *shine*, there is one lexical entry for both the past tense and the past participle (*shone*). The lexical entry for *shone* is seen in Figure 43.

(L1)

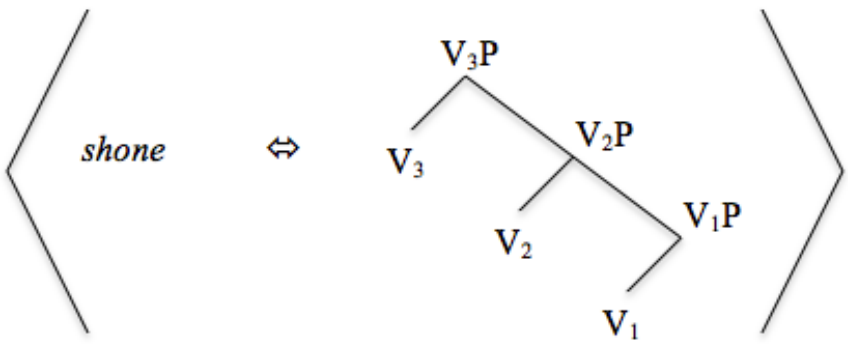


Figure 43 Lexical entry (L1) for *shone*

The L-tree (L1) in Figure 43 contains both the structure [V_2P [V_1P]], corresponding to the participle, and the structure [V_3P [V_2P [V_1P]]], corresponding to the past tense. Thus (L1) contains the appropriate structure for matching both of the S-trees in Figure 44, where (S1) is the S-tree for past tense and (S2) is the S-tree for the past participle.

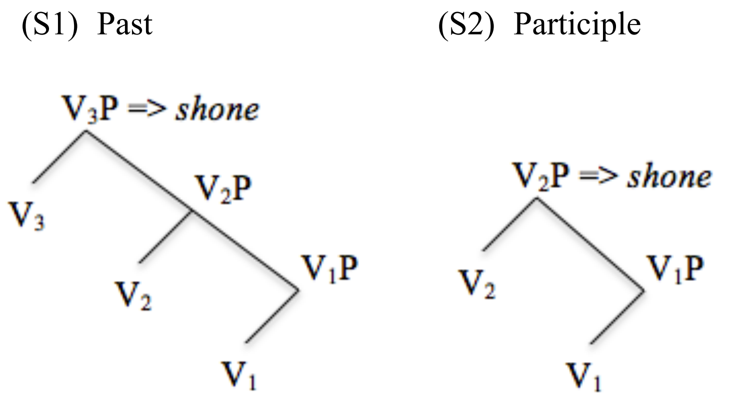


Figure 44 S-trees (S1) and (S2)

(L1) matches the bigger tree (S1) because they share all of their functional structure. (L1) can also match the smaller tree (S2) because the L-tree is a superset of this S-tree. In other words, a single L-tree (L1) matches two separate S-trees, meaning that a single morpheme applies in two separate morphosyntactic environments. This is how the Superset Principle allows us to model syncretism.

2.2.5.2 The Elsewhere Principle

Imagine two L-trees, (L1) and (L2), which can each match a given S-tree, (S2), as sketched in Figure 45.

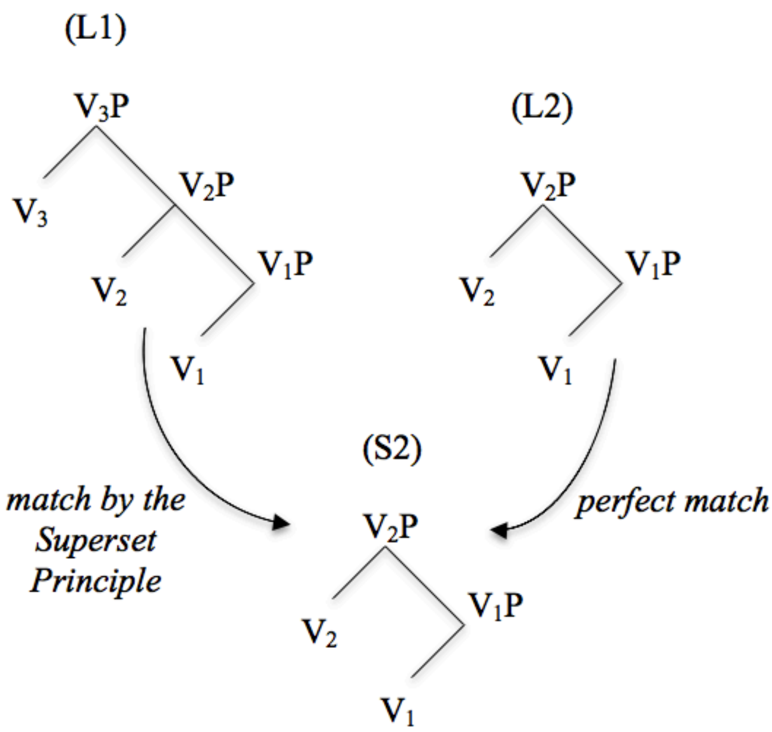


Figure 45 Two L-trees matching a single S-tree

This kind of competition between L-trees to spell out an S-tree is resolved by the Elsewhere Principle. The Elsewhere Principle is defined in (50) (see Kiparsky 1973 for the more general formulation).

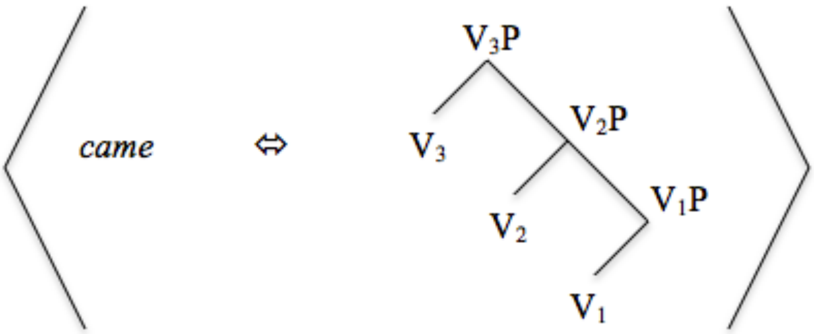
(50) Elsewhere Principle

> When more than a single L-tree can lexicalize an S-tree, the L-tree with the fewest unused features is chosen.

By this principle, a more specific entry (such as (L2) in Figure 45) is always preferred over a less specific one (such as (L1) in Figure 45).

Take the verbal inflection of English *come*. There is a syncretism between the present and past participial forms, both surfacing as *come*, while the simple past of *come* is *came*. This is captured by the two L-trees in Figure 46.

(L1)



(L2)

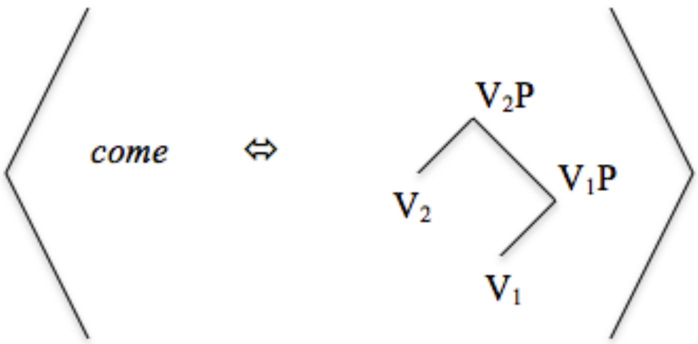


Figure 46 L-trees for *came* and *come*

Three S-trees are provided in Figure 47.

(S1) Past (S2) Participle (S3) Present

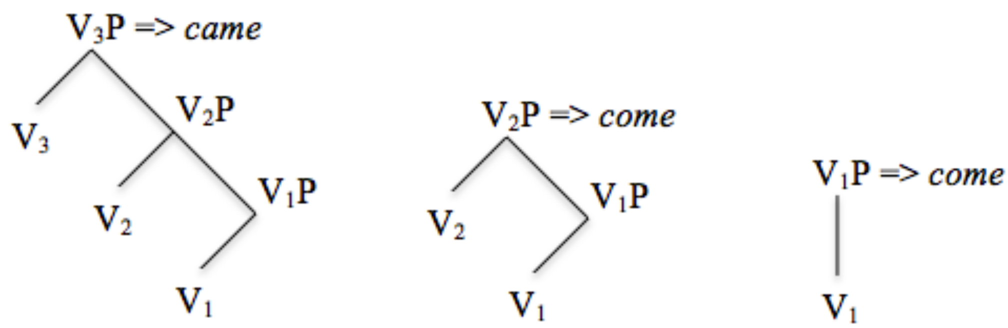


Figure 47 S-trees for past, participle, and present

(S1) [V_3P [V_2P [V_1P]]] can only be matched by (L1) above, since (L2) is not big enough to match this S-tree. Thus (S1) spells out as *came*. (S2), on the other hand, can be matched by either (L1) or (L2) by the Superset Principle: (L1) is a superset of (S2), and (L2) is the same size as (S2). The Elsewhere Principle resolves this competition. Since (L2) has fewer unused features compared to (S2) than (L1) does, (L2) is the winner. That is, (L1) has the extra feature V_3 in comparison to (S2), while (L2) is a perfect match. Thus (S2) spells out as *come*. Finally, (S3) can be matched by either (L1) or (L2), since both of these L-trees are supersets of (S3). The better match, however, is (L2): (L2) has only one extra feature (V_2) compared to (S3), while (L1) has two extra features (V_2 and V_3) compared to (S3). (S3) therefore spells out as *come* by the Elsewhere Principle.

Now that we have seen both the Superset Principle and the Elsewhere Principle and the way they interact, we can derive the *ABA theorem already mentioned in Section 2.2.4, that is, the generalization that only adjacent layers in a sequence can be syncretic. As we shall see, the *ABA theorem derives from the interaction between these two principles (Caha 2009: §2.3).

The *ABA theorem says that the spellout pattern in Figure 48 should be impossible. That is, it should be impossible for both (S1) [ZP [YP [XP]]] and (S3) [XP] to spell out as *a*, while the middle structure, (S2) [YP [XP]], spells out as *b*.

(S1) (S2) (S3)

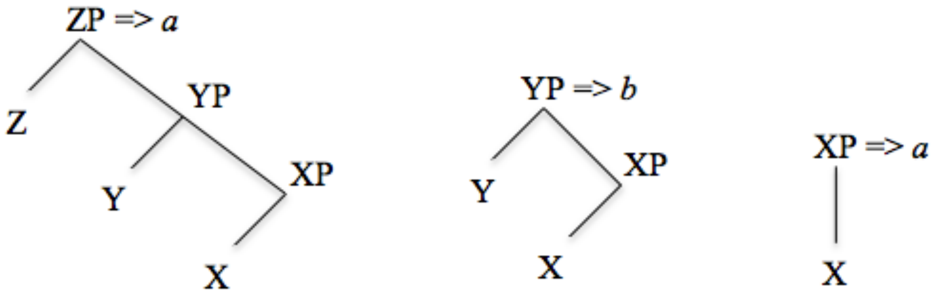


Figure 48 *ABA pattern

In other words, there cannot be a syncretism between the Z and X layers to the exclusion of the intervening Y layer.

The reason that such a spellout pattern is ruled out is as follows. Assume the lexicon to contain the lexical entries in Figure 49.

(L1) (L2)

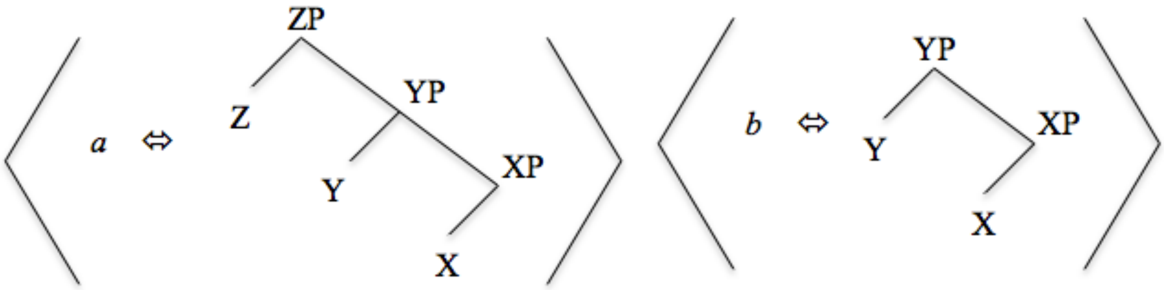


Figure 49 L-trees for *a* and *b*

If we wanted to generate an ABA pattern, these are the only trees that could, at first glance, accomplish this goal: (L1) [ZP [YP [XP]]] is a perfect match for (S1), and it is also a superset of (S3). In other words both (S1) and (S3) could by the Superset Principle spell out as *a*. (L2), moreover, is a perfect match for (S2), spelling out as *b*.

If we only took the Superset Principle into account, the ABA pattern might be possible to generate. But once we consider the effects of the Elsewhere Principle, it becomes clear that the ABA pattern will be blocked. While it is perfectly true that (L1) can, in principle, match (S3), it is not the better match: (L2) is a better match for (S3), and thus the ABA pattern we are trying to generate dissolves into an ABB pattern instead, as shown in (51).

(51) S-tree Spellout

[XP [YP [ZP]]] => *a* (only L1 is a match)
[YP [ZP]] => *b* (L1 and L2 match, but L2 is a perfect fit)
[ZP] => *b* (L1 and L2 match, but L2 is a better fit)

The Elsewhere Principle, then, is crucial for deriving the *ABA theorem. The ABA pattern requires the smallest and the largest S-trees to be spelled out as A, but since there needs to be an intermediate tree spelling out B, this intermediate tree will always be a better fit for the smallest tree by the Elsewhere Principle.[52]

2.2.5.3 The Principle of Cyclic Override

As mentioned above in Sections 2.2.3.1 and 2.2.4.3, nanosyntax assumes that derivations are built bottom-up. This leads us to the third basic principle of nanosyntactic spellout, known as the Principle of Cyclic Override and defined in (52).

(52) Principle of Cyclic Override

Later, higher-level spellouts cancel out previous, lower-level spellouts.

Cyclic Override follows naturally from a system in which derivations are built bottom-up. Because larger constituents arise later in such a system, they can cancel out the lower-level spellouts which have preceded.

For example, take the English verb *drive*. This verb has three separate forms for each of the structures we are interested in: the present form *drive*, the past participle *driven*, and the simple past form *drove*. These facts can be captured using the three lexical entries in Figure 50.

[52] It should be noted that while the *ABA theorem is a very useful diagnostic in studies of syncretism and crosslinguistically very robust, counterexamples do exist. The possibility of gaps in the fseq and attested occurrences of ABA patterns need to be investigated further (see Caha 2009: §9.3, Starke 2013, Vanden Wyngaerd 2014).

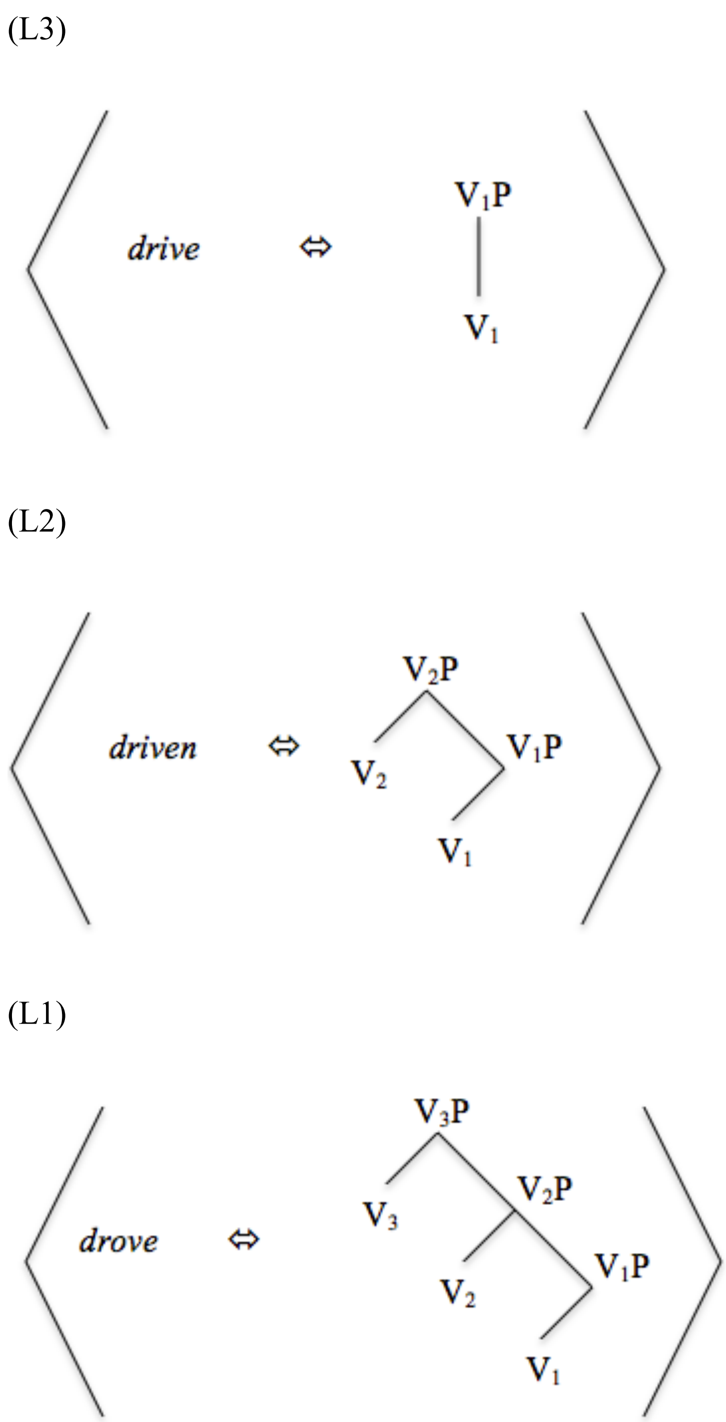


Figure 50 L-trees for *drive*, *driven*, and *drove*

Consider now the derivational history of the past tense form. In a bottom-up derivational system, we assume that the past form is built in three steps, sketched in Figure 51. First the present structure is built (S3), then the participial structure (S2), and finally the past structure (S1).

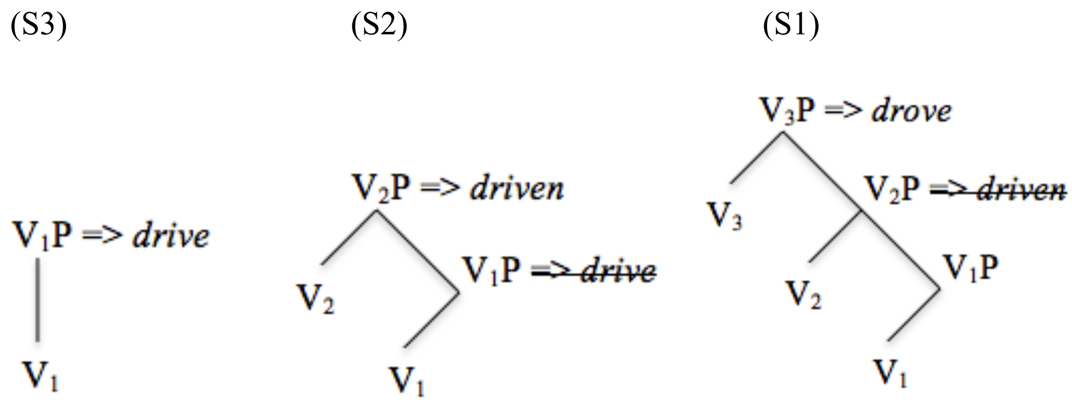


Figure 51 Building the past structure

As we shall see in more detail below, syntactic structures are built one featural head at a time, and a suitable match from the lexicon must be found at each layer. Thus, in order to build the past tense form *drove*, we begin with (S3) V_1P. At this layer the best match is (L3), giving the spellout *drive*. If the derivation continues from here, the next head to be added is V_2. The structure (S2) [V_2P [V_1P]] is then matched by (L2) and spelled out as *driven*, overriding the previous spellout *drive*. Finally, the derivation adds V_3, and the structure (S1) [V_3P [V_2P [V_1P]]] is matched by (L1) and spelled out as *drove*, overriding the previous spellout *driven*.

2.2.6 The Anchor Condition

The cases of spellout considered so far are rather simple, in the sense that we have always seen trees that are subsets or supersets of one another. For the verbal examples, each and every tree has had V_1 as its lowest head. Reality, however, is often more complicated, and trees will not always be in such straightforward subset-superset relations. In fact, there is often some form of ‘overlap’ between L-trees. For example, consider the S-tree (S4) [EP [DP [CP [BP [AP]]]]] in Figure 52.

(S4)

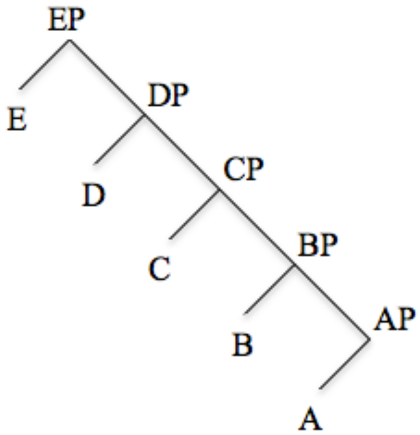
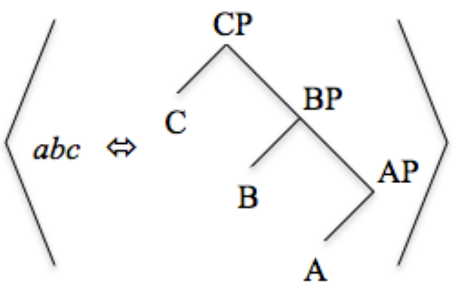


Figure 52 Hypothetical S-tree requiring lexicalization

Imagine now that there are two lexical entries available in the lexicon for spelling out (S4). These lexical entries are seen in Figure 53.

(L4)



(L5)

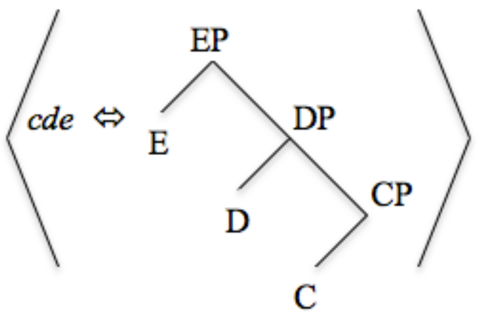


Figure 53 Lexical entries for *abc* and *cde*

Importantly, both (L4) and (L5) contain CP. (L4) has CP as its topmost node, while (L5) has CP as its lowest node. Since (L4) and (L5) share a layer of functional structure, when they go to spell out (S4) there will be an overlap at this shared layer, as shown in Figure 54.

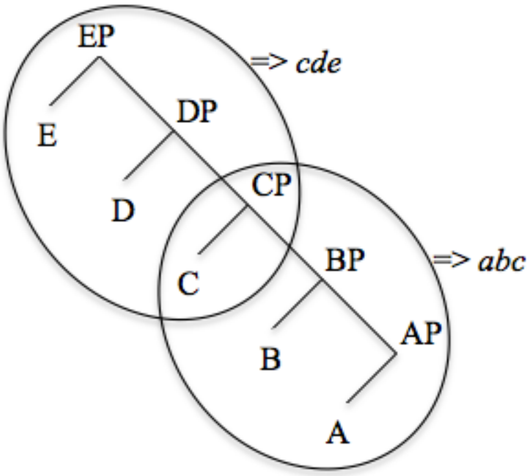


Figure 54 Overlap when (S4) is being spelled out

Thus we can speak of (L4) and (L5) 'overlapping' at CP.

The question posed in Figure 54, then, is which L-tree is ultimately responsible for spelling out CP. To resolve this issue, I adopt what is known as the Anchor Condition, defined in (53).

(53) Anchor Condition (Caha 2009: 89, adopted from Abels & Muriungi 2008)

> In a lexical entry [i.e. L-tree], the feature which is lowest in the functional sequence [i.e. the anchor] must be matched against the syntactic structure [i.e. S-tree].

In the example in Figure 54, two L-trees compete to lexicalize C. (L4), with the phonology *abc*, has C as its topmost feature, while (L5), with the phonology *cde*, has C as its lowest feature. By the definition in (53), then, C is the anchor of (L5). However, C is not the anchor of (L4), since it is not the lowest feature in this L-tree (rather, the anchor of (L4) is A). Since the anchor of an L-tree must be matched by the S-tree according to the Anchor Condition in (53), it is (L5) (*cde*) which gets to spell out C, as shown in Figure 55.

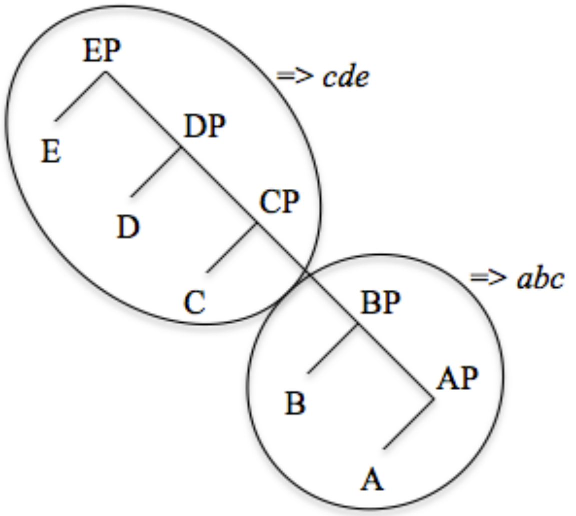


Figure 55 Spelling out the anchor of *cde*

The Anchor Condition is satisfied in Figure 55: *abc*, whose L-structure is [CP [BP [AP]]], gets to spell out its anchor, namely A, in the S-tree, and *cde*, whose L-structure is [EP [DP [CP]]], gets to spell out its anchor, namely C, in the S-tree.

In the case of (L4) (*abc*) and (L5) (*cde*), the Anchor Condition leads to the 'shrinking' of *abc*. Since (L5) (*cde*) must have its anchor matched by the S-tree, the morpheme *abc* no longer gets to spell out C, even though its L-tree contains C.

Overlap at C → *abc* shrinks to accommodate *cde*'s anchor

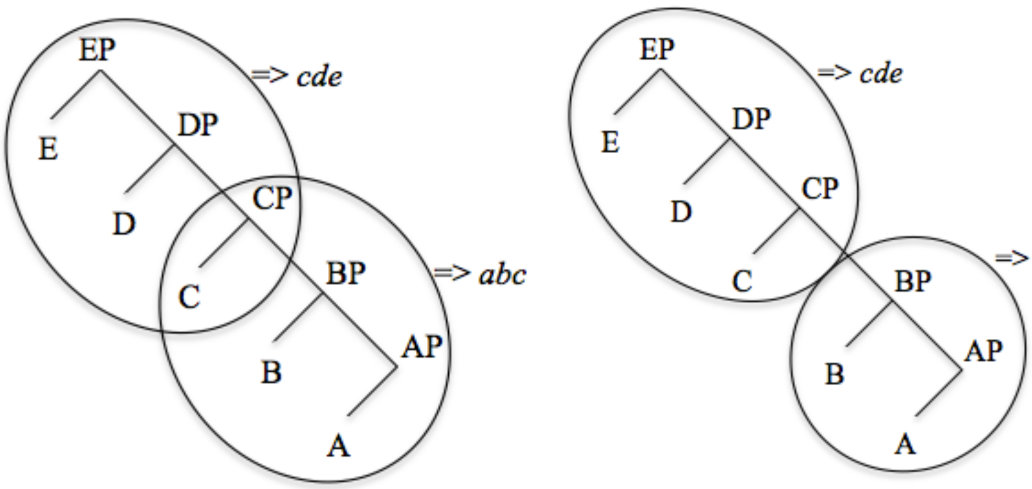
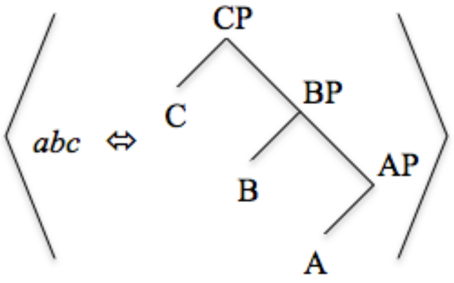


Figure 56 Shrinking of *abc*

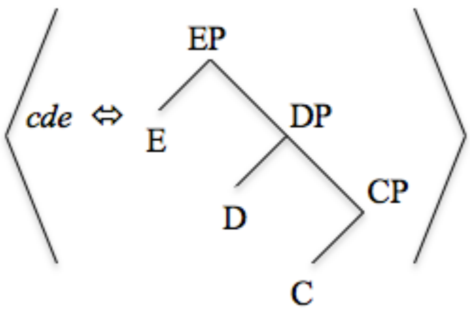
Thus, even though the L-tree of *abc* is [CP [BP [AP]]], in the syntax it can only apply to the subtree [BP [AP]] because of the Anchor Condition.

With this ‘shrinking’ in mind, imagine now that the lexicon contains not only the lexical entries < *cde* ⇔ [EP [DP [CP]]] > and < *abc* ⇔ [CP [BP [AP]]] > but also the entry < *ab* ⇔ [BP [AP]] >. See Figure 57.

(L4)



(L5)



(L6)

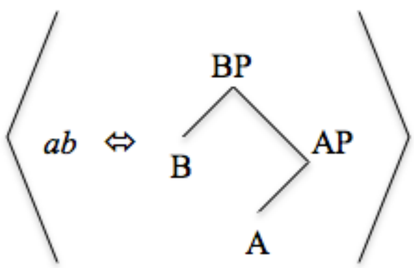


Figure 57 Expanded lexicon

Once again, (L4) and (L5) overlap at C. However, there is also an entry, (L6), which consists simply of [BP [AP]]. (L6), then, overlaps with (L4) at both the A and B layers, since (L4) also contains [BP [AP]]. (L4) and (L6) will therefore be competitors for spelling out A and B in the S-tree.

Return now to (S4) [EP [DP [CP [BP [AP]]]]]. Due to the Anchor Condition, we already know that (L5) must spell out the [EP [DP [CP]]] portion of this S-tree.

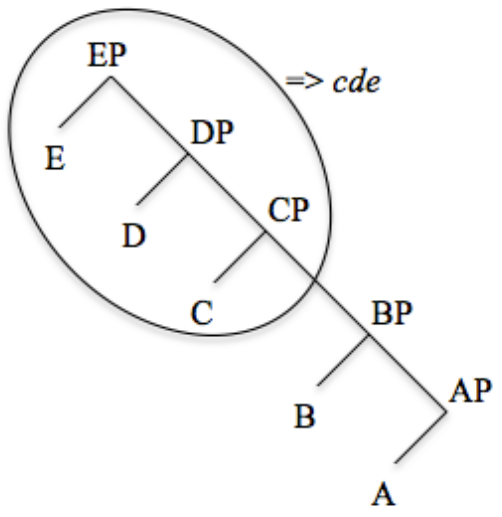


Figure 58 [EP [DP [CP]]] spells out as *cde*

The lower part of the S-tree, namely [BP [AP]], remains to be spelled out. By the Superset Principle there are two L-trees available for spelling out [BP [AP]]: (L4) [CP [BP [AP]]] and (L6) [BP [AP]]. (L4) is a superset of the S-tree [BP [AP]], and (L6) is the same size as the S-tree [BP [AP]]. By the Elsewhere Principle, then, (L6) is a better fit than (L4) for spelling out this part of the S-tree. This results in the spellout shown in Figure 59.[53]

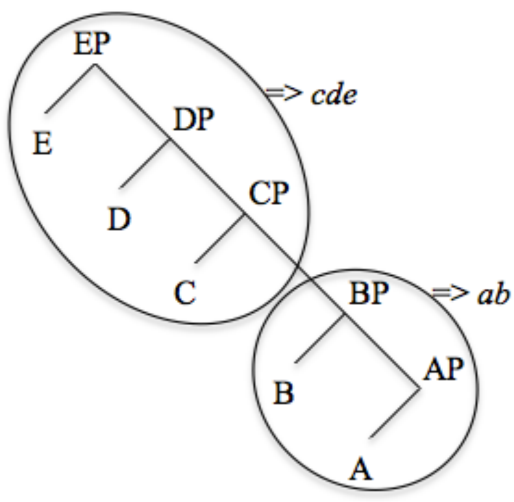


Figure 59 [BP [AP]] spells out as *ab*

Note here that (L4) ends up not being used at all for (S4). (L4) and (L5) overlap at C, but since (L5) has C as its anchor it must spell out C in the S-tree. The truncation of (L4), so to speak, subsequently causes another L-tree, (L6), to be a better match for the remaining

[53] Ignoring, for now, that spellout should target proper constituents, which [EP [DP [CP]]] is not in this case.

S-structure. The result is that the morpheme *abc* is squeezed out of the spellout process completely, in favor of the morpheme *ab*.

To take a concrete example of the Anchor Condition in action, let us consider an example from Caha (2009: §1.8.1 and §2.9). According to Caha, K heads are merged above DP. In order to receive case inflection, DP must move to the left of KP, as sketched in Figure 60 for genitive case.

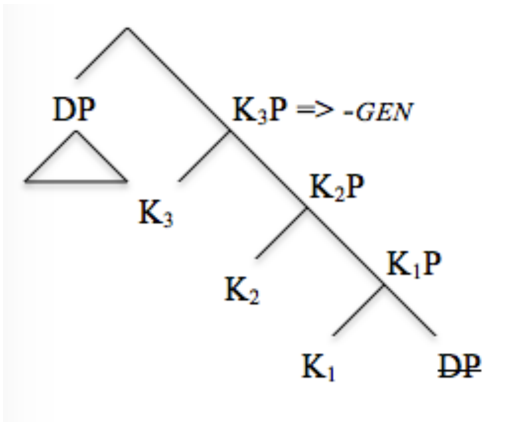


Figure 60 DP movement resulting in genitive case inflection

Recall that, according to Caha, not only case endings but also prepositions consist of K features. He proposes that prepositions are K heads which remain to the left of DP while case endings are K heads which end up to the right of DP after DP movement. For instance, consider the base-generated structure in Figure 61.

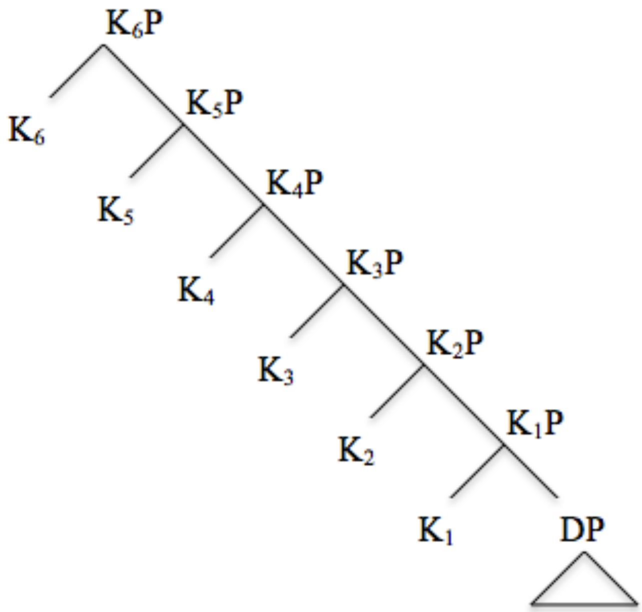


Figure 61 Base-generated [K… [DP]] structure

Imagine that DP now moves to the left of K_3P, as seen in Figure 62. The K heads to the right of DP correspond to genitive case, so the result of such a movement is a genitive-inflected DP (cf. Figure 60). The K heads to the left of DP, moreover, correspond to a preposition, in this case a comitative (see Figure 41 above) preposition with the meaning 'with'.[54] In Arabic, for instance, the preposition *bi* 'with' selects a genitive complement (Caha 2009: 43).

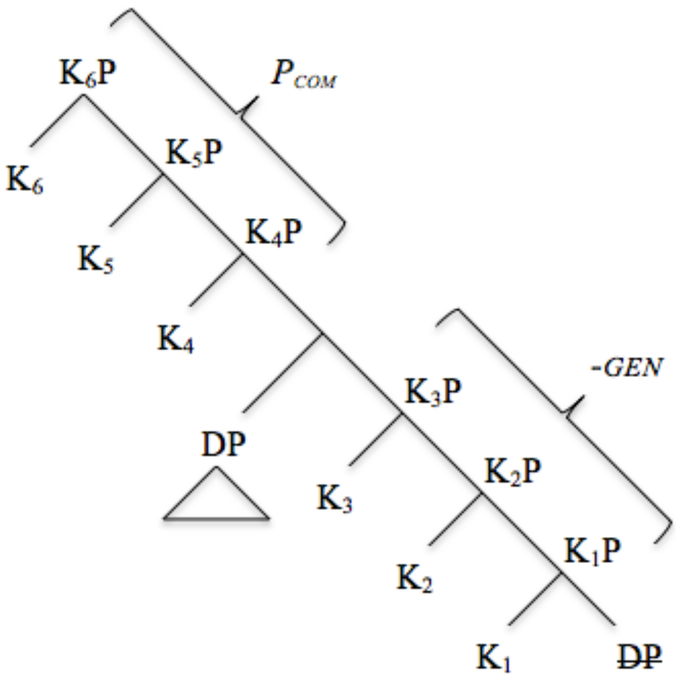


Figure 62 COM preposition + DP + GEN ending

In fact, the position of DP in Figure 62 is as high as DP can move within the K zone in Arabic. Arabic only has the cases NOM (K_1), ACC ($K_1 + K_2$), and GEN ($K_1 + K_2 + K_3$) in its morphological case system, where GEN is the largest/highest case. There are no morphological cases larger or higher than GEN, i.e. DAT ($K_1 + K_2 + K_3 + K_4$), INS ($K_1 + K_2 + K_3 + K_4 + K_5$), or COM ($K_1 + K_2 + K_3 + K_4 + K_5 + K_6$), in this language. Thus the highest position of DP is the one shown in Figure 62.

Now, though Arabic does not have the cases DAT, INS, or COM in the form of inflectional endings, it does have them in the form of functional prepositions, such as COM/INS *bi* 'with' and DAT *li* 'to'. These prepositions, moreover, select the largest morphological case available in Arabic, namely GEN. In Figure 62, then, not only do we see DP movement resulting in case inflection, but also case selection by the preposition.

[54] Again ignoring that spellout should target constituents. See Sections 2.2.7 and 2.2.8.

Interestingly, Caha shows that the fact that functional prepositions select the largest available morphological case extends to other languages as well.

In Greek, as in Arabic, the largest available morphological case is genitive. However, the comitative/instrumental preposition *me* 'with' does not select genitive but the smaller case accusative. Caha (2009: 95) suggests that this problem can be accounted for in terms of the Anchor Condition. He proposes that the L-tree of Greek *me* 'with' does not correspond to the span [K_6 + K_5 + K_4] (as in Figure 62 above for Arabic *bi* 'with') but rather to the slightly larger span [K_6 + K_5 + K_4 + K_3]. This would make the lowest feature (i.e. the anchor) of *me*'s L-tree K_3 instead of K_4. Figure 63 illustrates how this leads to a situation of overlap with genitive case.

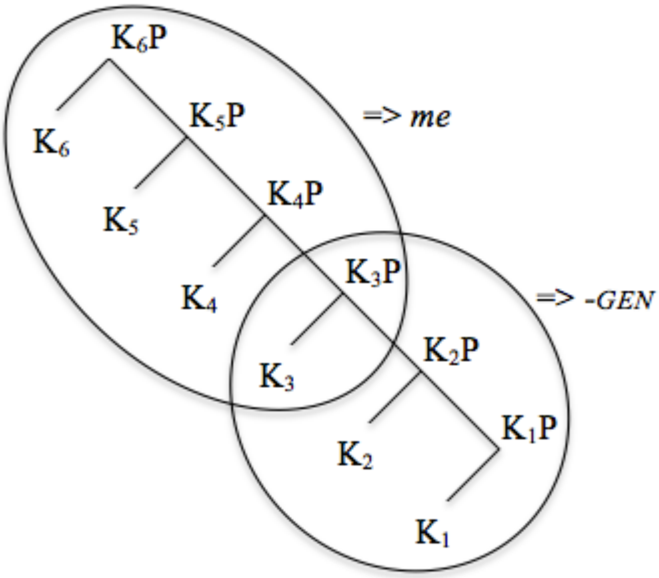


Figure 63 Greek preposition *me* 'with' overlaps with genitive ending

By the Anchor Condition, the lowest feature of an L-tree must have a match in the S-tree. The preposition *me* and the case ending -*GEN* compete to lexicalize K_3 in the S-tree in Figure 63. But since K_3 is the anchor of *me*, the K_3 layer in the S-tree must be lexicalized by *me* instead of -*GEN*. This is shown in Figure 64.

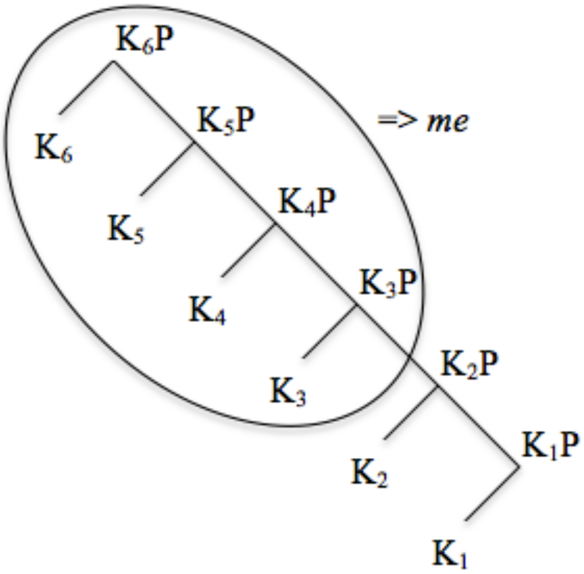


Figure 64 Anchor of *me* (K_3) must be matched by S-tree

Furthermore, this leads to a shrinking effect. The remaining section of the S-tree which still needs to be lexicalized is [K_2P [K_1P]], which is equivalent to the accusative, not the genitive (again, see Figure 41 above). Therefore -*ACC* is now a better match than -*GEN* to spell out the rest of the S-tree.

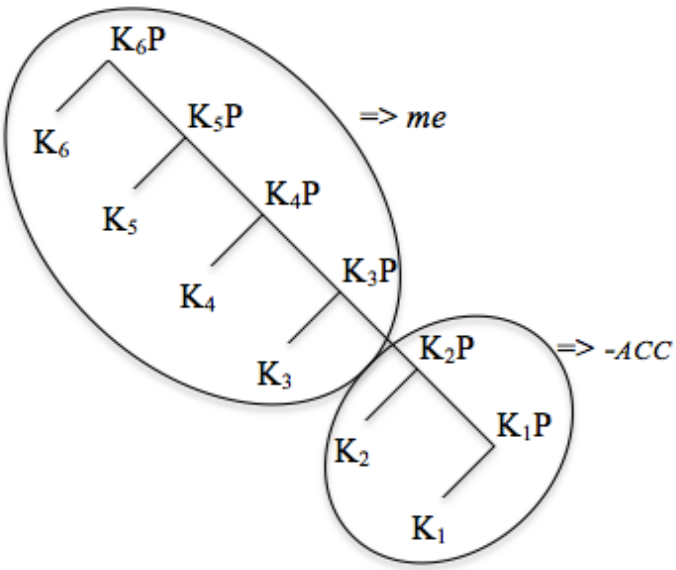


Figure 65 Preposition *me* selects ACC

Thus the Anchor Condition is crucial for understanding why the preposition *me* selects the accusative rather than the genitive in Greek.

To sum up, the Anchor Condition will cause lexical entries that intrude on the anchor of another lexical entry to 'shrink'. This shrinking effect may, in turn, lead to another entry becoming a better fit for that part of the syntactic structure, if such an entry exists in the lexicon. For a more complex case study of shrinking in Spanish, see Fábregas (2009).

2.2.7 Spellout-driven movement

Syntactic structure can be altered in a number of precise ways in order to give rise to constituents that can be lexicalized. This process is known as *spellout-driven movement*. According to the views of Starke (e.g. 2011b, 2013), which I follow here, only constituents can be targeted for spellout. This 'strict constituenthood' view differs from 'spanning' approaches (for which see Abels & Muriungi 2008, Ramchand 2008, Dékány 2009), for which see Section 2.2.8 below.

Spellout matches S-trees with L-trees. In nanosyntax, structures are, moreover, built feature by feature. It is assumed that no feature can be left unlexicalized (Cyclic Exhaustive Lexicalization; Fábregas 2007, Ramchand 2008, Pantcheva 2011). At each layer of the structure, moreover, the lexicon must be consulted to see if there is an appropriate L-tree for spelling out the S-tree thus far built. The system has a set number of attempts and movement strategies to achieve lexicalization at each layer. The algorithm in (54) summarizes these movement strategies, which we call *spellout-driven movement*.

(54) Algorithm for spellout-driven movement

STAY > CYCLIC > SNOWBALL

For example, take the structure in Figure 66.

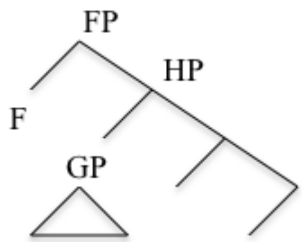


Figure 66 S-tree

Assume that the structure [$_{HP}$ [GP] …] in Figure 66 has already been spelled out, but that F – the most recent feature to have been merged in the structure – has yet to be lexicalized. In order to find an appropriate spellout for the feature F, the first available option to the structure is to *stay as is* (i.e. STAY). The lexicon is consulted to see if there is an L-tree corresponding to [$_{FP}$ F [$_{HP}$ [GP] …], as sketched in Figure 67.

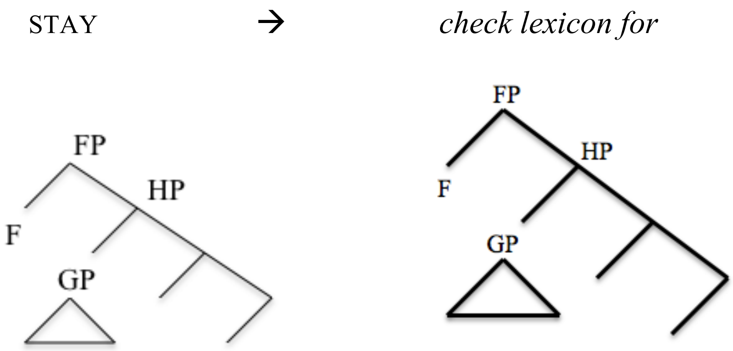


Figure 67 First step in the algorithm

If there is no appropriate lexical entry for the bolded structure in Figure 67, *the daughter of the sister of F is evacuated to the left of F*, in this case meaning that the constituent GP moves to the left of F (i.e. CYCLIC).[55] The lexicon is then checked for a lexical entry that matches what is left over, namely [FP [HP…]], as sketched in Figure 68.

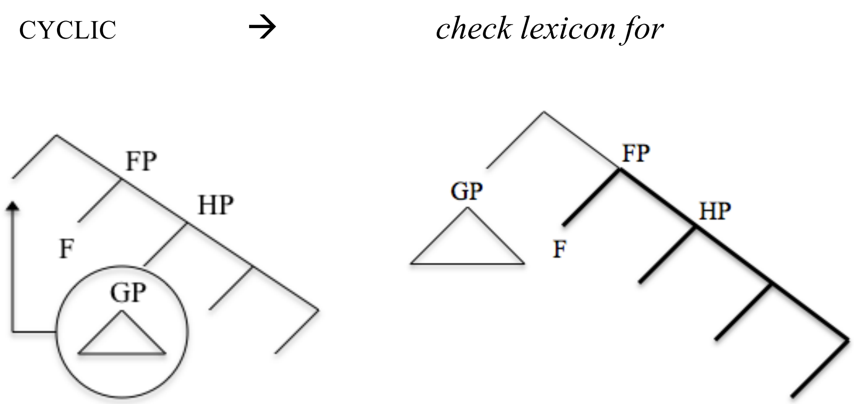


Figure 68 Second step in the algorithm

If there is no appropriate lexical entry for the bolded structure in Figure 68, then the *cyclic movement is undone and the sister of F is evacuated to the left of F* (i.e. SNOWBALL). The lexicon is then checked for a lexical entry that matches what is left over, namely FP, as sketched in Figure 69.

[55] The landing site of GP after cyclic movement is taken to be an unlabeled specifier. Spellout-driven movement, as opposed to syntactic movement on the basis of feature identity (i.e. traditional syntactic movement), is also considered not to leave traces.

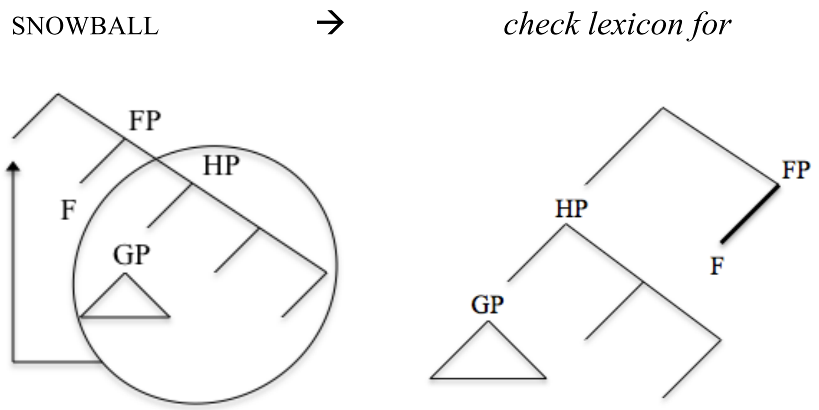


Figure 69 Third step in the algorithm

To sum up, spellout-driven movement is an algorithmic procedure for creating syntactic constituents (S-trees) which can be properly matched by L-trees. As soon as a match is found (whether at STAY, CYCLIC, or SNOWBALL), the next feature is added and the algorithm begins anew. Observe that the system prioritizes no movement over movement, and cyclic movement over snowball movement. Moreover, it is in line with the movement rules of Cinque's (2005) U20 system, in that the head of the extended projection is always part of the moved object, only leftward movement is allowed, and head movement and remnant movement are disallowed.

To see how spellout-driven movement works concretely, we can make use of Pantcheva's (2011: §7.3) nanosyntactic account of the ON-series (a set of morphemes related to the word for 'on') in Karata. In (55) we see how Place 'on' (55a), Goal 'to' (55b), and Route/Source 'from/through' (55c) are expressed in Karata. Note that Route and Source are syncretic in this language, as both readings are expressed by the morpheme *-gal*.

(55) Karata (Pantcheva 2011: 137)

(a) bajdan-ʈ'-a
square-ON-LOC
'on the square'

(b) bajdan-ʈ'-a-r
square-ON-LOC-GOAL
'to the square'

(c) bajdan-ʈ'-a-gal
square-ON-LOC-SOURCE/ROUTE
'from/through the square'

The fseq and relevant lexical entries are provided in (56a) and (56b), respectively.

(56) (a) Route > Source > Goal > Place > AxPart > ... DP

(b) < *bajdan* ⇔ DP ⇔ SQUARE >
< *-ƫ’* ⇔ AxPartP ⇔ ON >
< *-a* ⇔ PlaceP >
< *-r* ⇔ GoalP >
< *-gal* ⇔ [RouteP [SourceP [GoalP]]] >

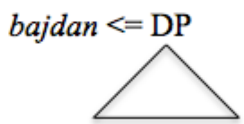
AxPart in (56b) stands for ‘axial part’, having to do with an object’s position with respect to some axis (i.e. ‘front’, ‘back’, and so on; see Svenonius 2006). The Karata AxPart morpheme *-ƫ’* encodes the particular axial position ‘on’. The morpheme *-a* is a locative marker and thus corresponds to the Place layer. The morpheme *-r* is a Goal marker and thus corresponds to the Goal layer. The morpheme *-gal*, which is syncretic between Source and Route readings, corresponds to [RouteP [SourceP [GoalP]]].[56]

To derive the full structure *bajdan-ƫ’-a-gal* ‘through the square’ (i.e. the entire structure from Route down to D), the derivation proceeds as follows. In (57-62) below, there are two additional details to keep in mind. First, spellout-driven movement does not leave traces. Second, spellout-driven movement creates unlabeled specifiers into which displaced constituents move. If movement continues and the unlabeled specifier becomes empty, then the specifier is deleted.

(57) Build up to DP and match with *bajdan* ‘square’

Lexicon

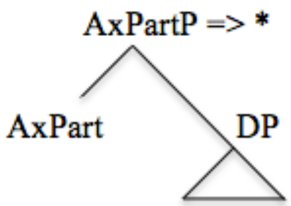
< *-ƫ’* ⇔ AxPartP >
< *-a* ⇔ PlaceP >
< *-r* ⇔ GoalP >
< *-gal* ⇔ RouteP SourceP GoalP >



[56] The reason *-gal* corresponds to [RouteP [SourceP [GoalP]]] rather than just [SourceP [RouteP]] is that *-gal* does not stack on top of the Goal marker *-r*. Since *-gal* and *-r* do not cooccur, *-gal* must spell out whatever structure *-r* is responsible for as well, in order to suppress the spellout *-r*.

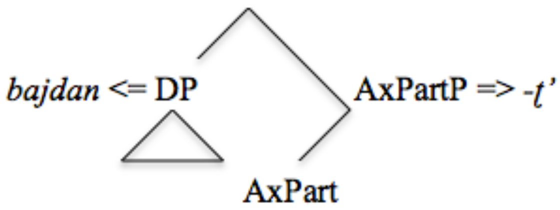
(58) Add AxPart

(a) STAY: No match in the lexicon for [AxPartP [DP]]



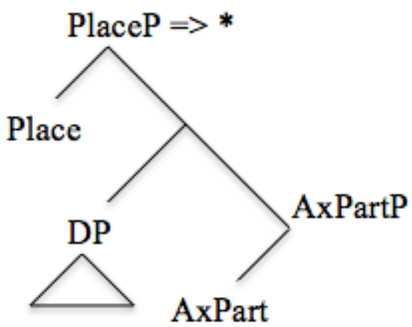
(b) CYCLIC: Not Applicable (NA)

(c) SNOWBALL: Move DP to the left of AxPartP, match AxPartP with *-ʈ'*

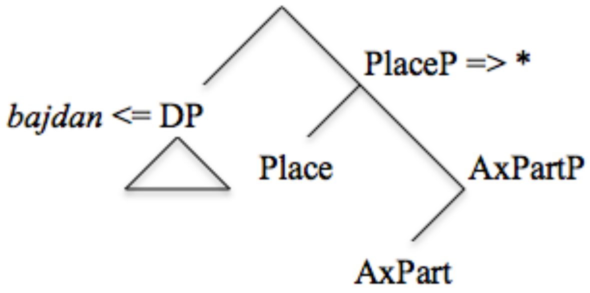


(59) Add Place

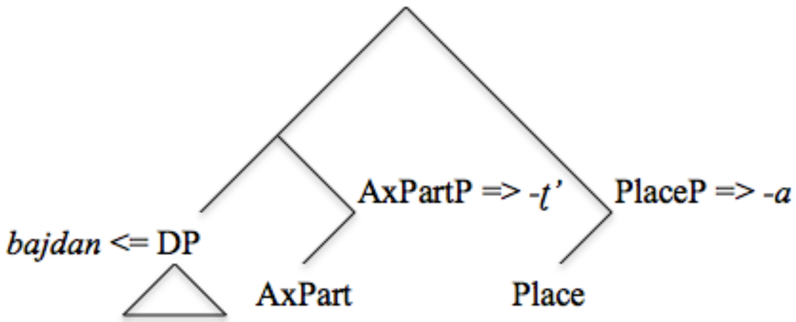
(a) STAY: No match



(b) CYCLIC: Move DP, no match for [PlaceP [AxPartP]]



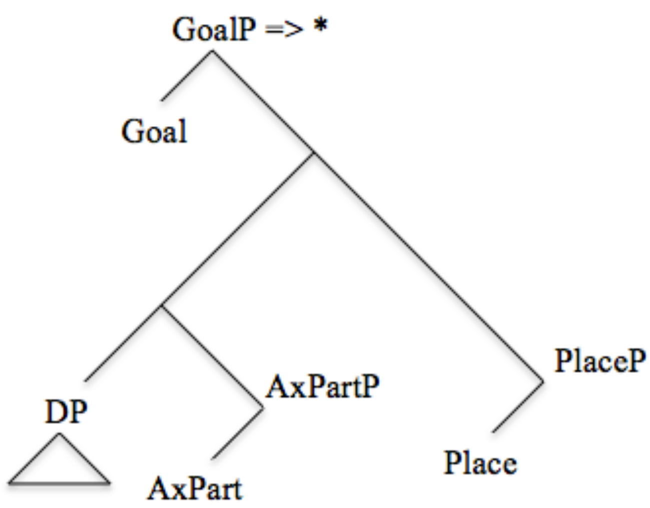
(c) SNOWBALL: Undo CYCLIC and move [[DP] AxPartP]; match PlaceP with *-a*



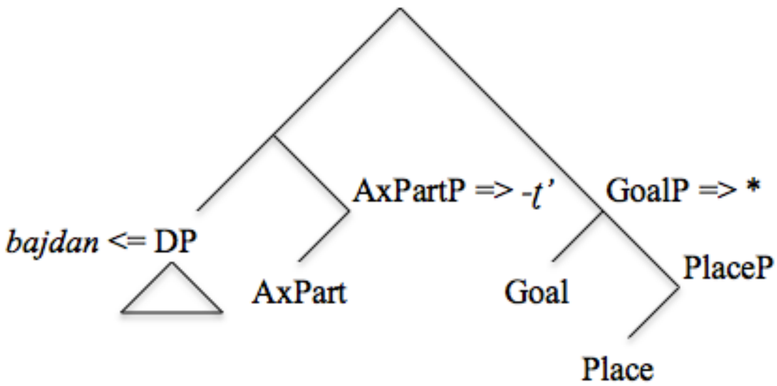
Note that at this point in the derivation we have a structure meaning ‘on the square’, *bajdan-ţ’-a*. However, the aim of syntax in this case is to express ‘through the square’, meaning it has to build up to RouteP. Thus syntax continues building the fseq.

(60) Add Goal

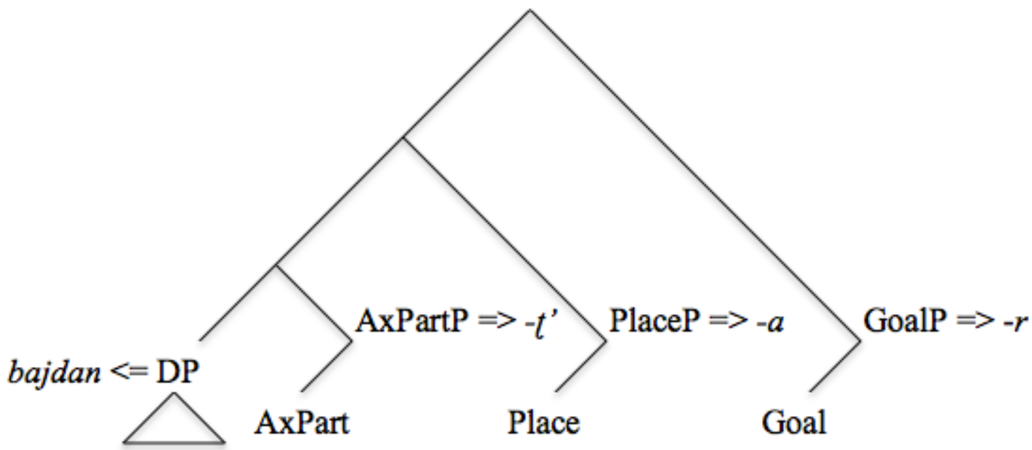
(a) STAY: No match



(b) CYCLIC: Move [[DP] AxPartP]; no match for [GoalP [PlaceP]]

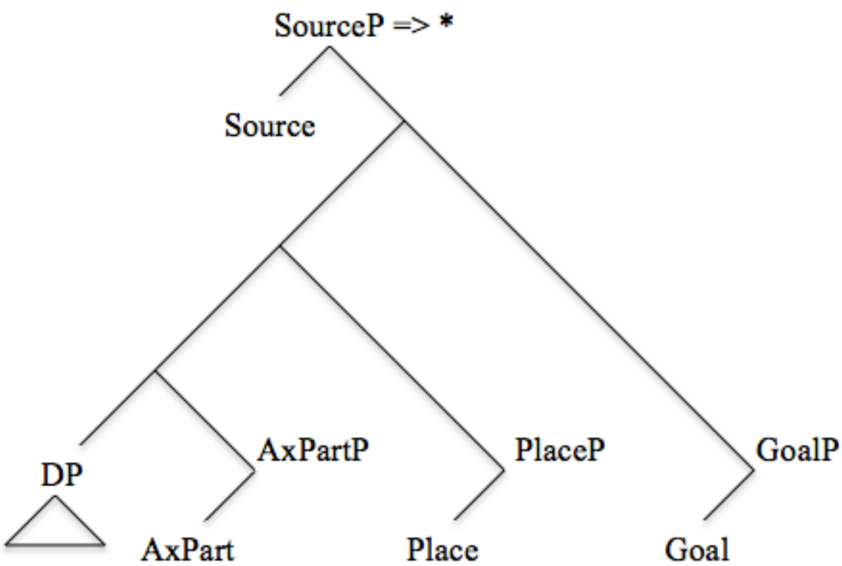


(c) SNOWBALL: Undo CYCLIC and move [[[DP] AxPartP] PlaceP]; spell out GoalP as *-r*

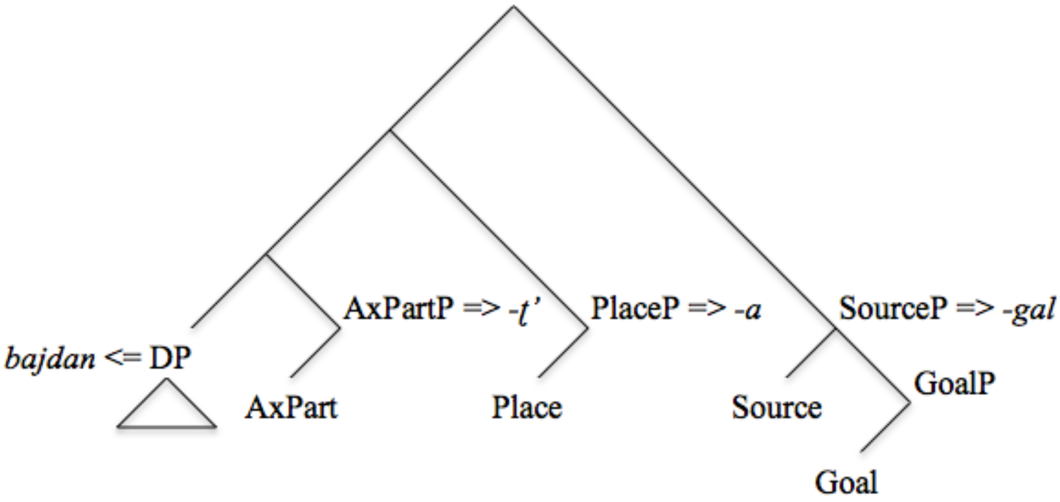


(61) Add Source

(a) STAY: No match

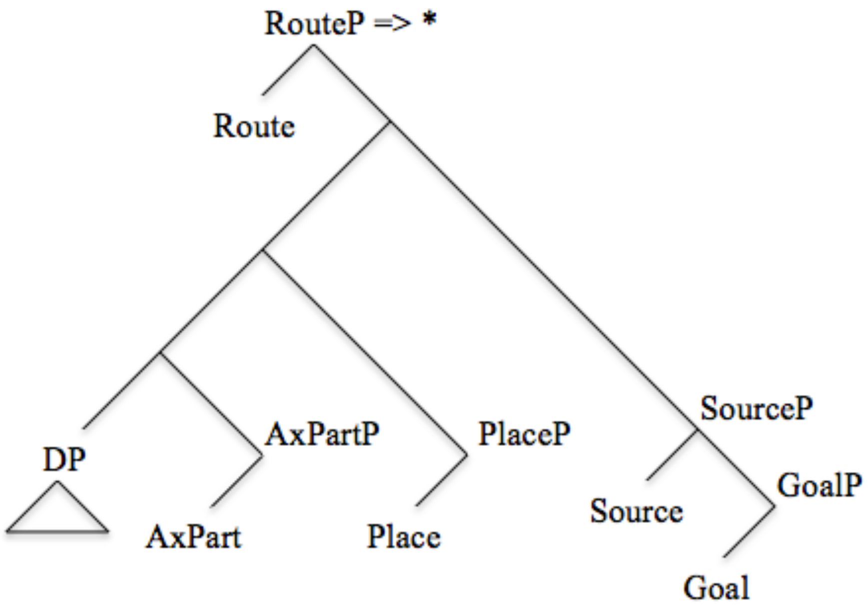


(b) CYCLIC: Move [[[DP] AxPartP] PlaceP]; [SourceP [GoalP]] matches < *-gal* ⇔ RouteP SourceP GoalP > by the Superset Principle[57]



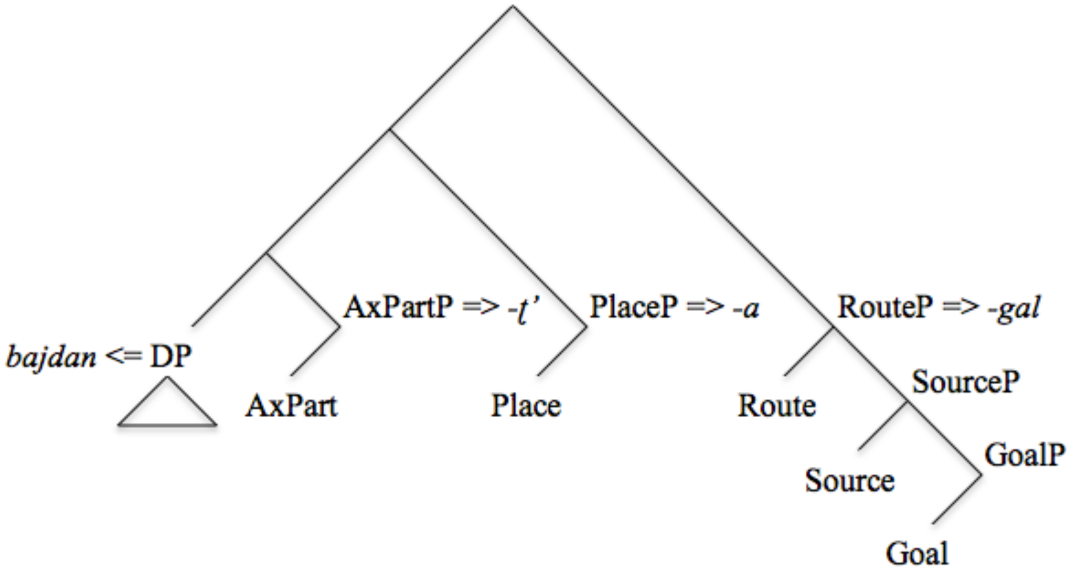
(62) Add Route

(a) STAY: No match



57 Recall from (55c) that *-gal* is syncretic between Source and Route readings, which is accounted for by the Superset Principle.

(b) CYCLIC: Move [[[DP] AxPartP] PlaceP]; match [RouteP [SourceP [GoalP]]] with < -*gal* ⇔ RouteP SourceP GoalP >



The Route structure in (62b) has the final spellout of *bajdan-ţ'-a-gal*. While the road to deriving this structure may seem complicated at first, the way a given structure will be spelled out by spellout-driven movement can actually be predicted rather easily. In Section 2.2.10 I will provide a shorthand for computing the end result of spellout-driven movement.

2.2.8 Complex heads

The spellout algorithm discussed in the previous section succesfully derives suffixation, or in general the right-adjunction of morphemes, because spellout-driven movement evacuates parts of the complement to the left in order to spell out constituents on the right.

However, with the technology presented so far, spelling out prefixes (or morphemes appearing to the left in general) is less straightforward. Let us return to Caha's (2009: 65-67, also Caha 2010) work on prepositions for a concrete example. The German preposition *mit* 'with' is syncretic between instrumental and comitative functions, and it selects dative case. According to the logic of DP movement and prepositional case selection discussed above, this means that DP has moved to the left of K_4P, leaving *mit* to correspond to the features K_6 and K_5 in the K hierarchy, as seen in Figure 70.

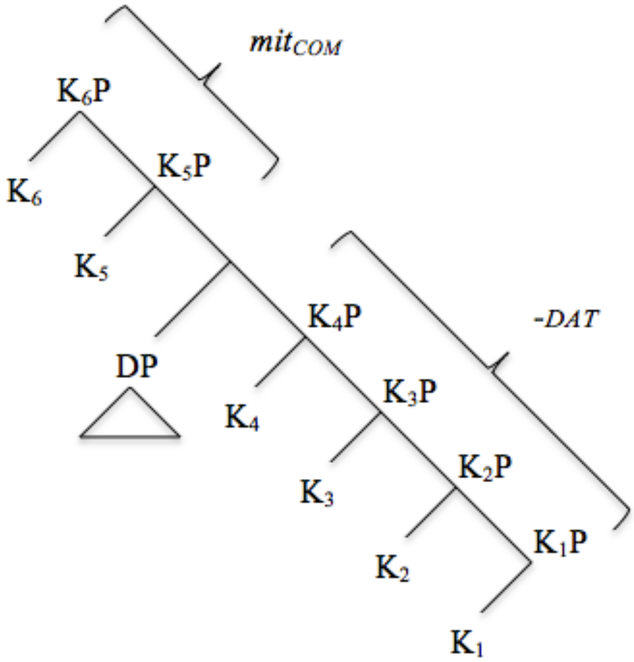
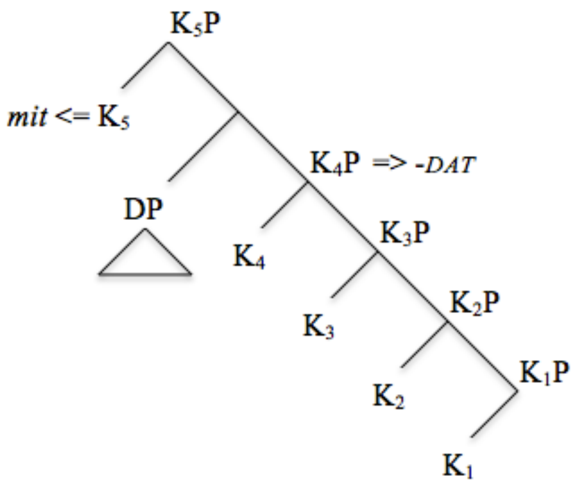


Figure 70 DP movement for DAT inflection

Recall from above, though, that spellout needs to target constituents. As shown in Figure 71, K_6 and K_5 do not form a constituent and therefore cannot be targeted by spellout according to the spellout system adopted in this dissertation.[58]

[58] Recall that *mit* is syncretic between COM ($K_6 + K_5$) and INS (K_5). Had we wanted to spell out *mit*$_{INS}$ rather than *mit*$_{COM}$, then, only K_5 would need to be targeted for spellout. As a single terminal, K_5 could theoretically be targeted for spellout since it is technically a proper constituent. See (vii).

(vii) INS *mit* and terminal spellout



For reasons to be discussed in Chapter 5, however, terminal spellout is problematic. Indeed, Starke (2011b and p.c.) proposes that spellout never targets single terminals but always non-terminals (i.e. phrases or a set of two or

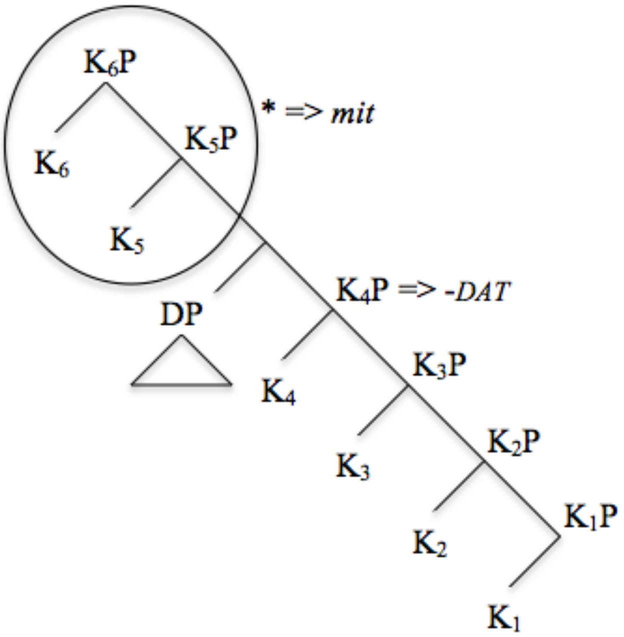


Figure 71 Preposition *mit* is not a constituent

With the technology we have encountered so far, we would expect [[DP] K_4P...] to evacuate to the left in order to make [K_6P [K_5P]] a constituent that can be targeted for spellout. However, this would make *mit* a postposition, contrary to fact.[59]

For some researchers, spellout can target non-constituents. For these 'spanning' approaches see Abels & Muriungi (2008), Ramchand (2008), Dékány (2009). Caha (2009: §2.6.1) also adopts a spanning-type approach for prepositions in positing that elements which have already been spelled out can be ignored for the purposes of later spellouts. In Figure 71, then, [[DP] K_4P…] is essentially invisible for the purposes of spellout, meaning [K_6P [K_5P]] can be treated as a constituent and thereby targeted for spellout.

Instead I adopt a stricter approach to spellout, following unpublished work by M. Starke (and in line with Caha 2010: 28, fn.11, Pantcheva 2011, Starke 2013). According to this view, [[DP] K_4P…] in Figure 71 blocks [K_6P [K_5P]] from being a constituent, so

---

more terminals). As for the particular case of *mit*$_{INS}$, in Starke's approach we would be forced to posit that Caha's K hierarchy is actually more fine-grained than proposed, so that the instrumental preposition actually corresponds to at least two features, K_5 and some feature X, the identity of which is still to be determined. All in all, Starke's approach says that the lexicon will never store a morpheme that consists of only a single feature/terminal. Though I do not want to take a solid stance on whether Starke's approach or an approach which allows for terminal spellout in certain cases is correct, I ultimately do not need to make use of any terminal spellout, thus for all intents and purposes siding with Starke.

[59] Unless we then allow for remnant movement of [K_6P [K_5P]] back to the left of [[DP] K_4P…], which is rejected on the basis that it violates Cinque's U20 rules.

mit must have a different structure. The idea is that *mit* corresponds to a complex head, as shown in Figure 72.

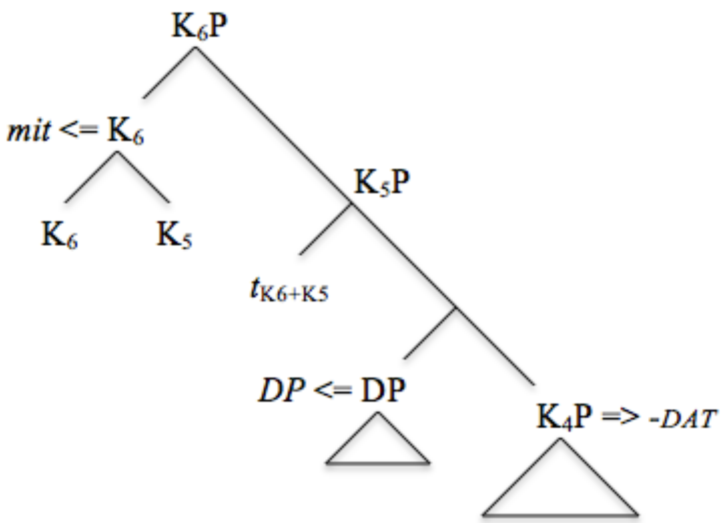


Figure 72 Preposition *mit* as a complex head

The structure in Figure 72 will require some extra explanation, most importantly with regard to how the complex head arises. Leaving this aside for the moment, observe that the complex head involves a binary set of features at the bottom of its structure, while a regular suffix involves a singleton set consisting of a single feature at the bottom of its structure, as sketched in Figures 73 and 74.

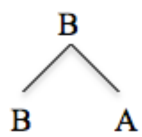


Figure 73 Binary set (prefix)

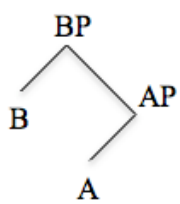


Figure 74 Singleton set (suffix)

According to Starke, this is how the difference between prefixes and suffixes should be lexically encoded. Merge always affects two elements (Kayne 1984). Thus the binary set in Figure 73 implies a complete application of merge. That is, A and B have been merged normally and have not been tampered with any further. The singleton set of $[_{AP}$ A] at the bottom of the structure in Figure 74, on the other hand, does not have a sister and as such is missing its merge partner. According to Starke, then, a singleton set at the bottom of a structure implies that dislocation has occurred, more precisely that the sister of A has been moved out.[60] If the sister of A has been moved out, moreover, and since all movement is leftward (Kayne 1994), this leaves [BP [AP]] to the right of whatever has been moved, making it a suffix. To sum up so far, then, a prefix is a complex head formed by binary merge, and a suffix is a structure which has had the sister of its lowest head evacuated.

An idea which exists in generative grammar but is often not pursued in detail (though see Jayaseelan 2015) is that there is more than a single cognitive workspace in which to build syntactic structures. In fact, for the structure in Figure 72, we need to posit two separate workspaces. In the first workspace, [[DP] K_4P...] is constructed, as sketched in Figure 75.

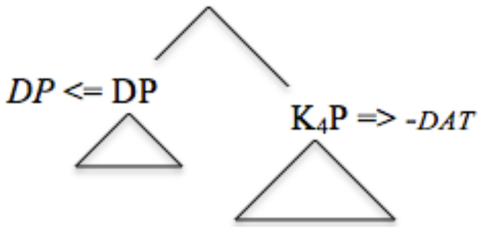
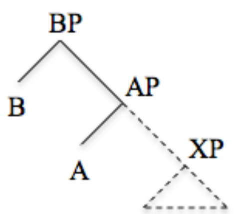


Figure 75 Workspace-1

The next step in the derivation consists of adding K_5. However, at this point, all the steps in the spellout algorithm will fail to produce a match, given that the structure of the preposition *mit* is the complex head [K_6 K_5]. This is illustrated in (63).

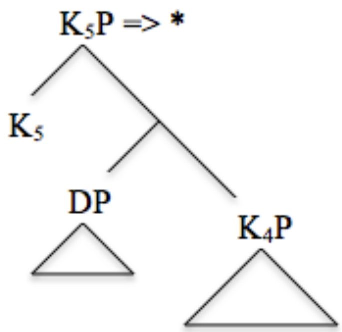
[60] We might illustrate this as follows, where XP has been moved out:

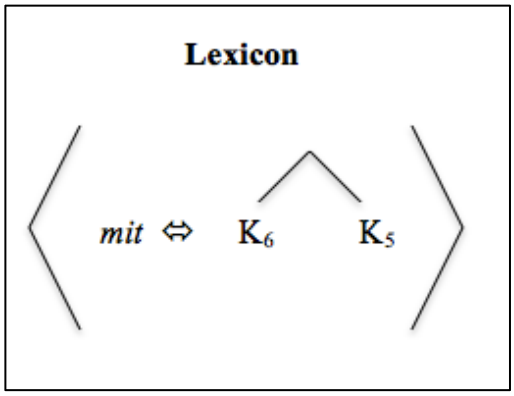
(viii)



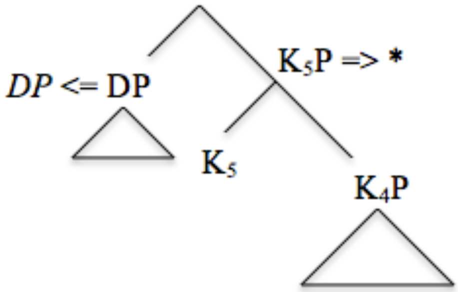
(63) Add K_5

(a) STAY

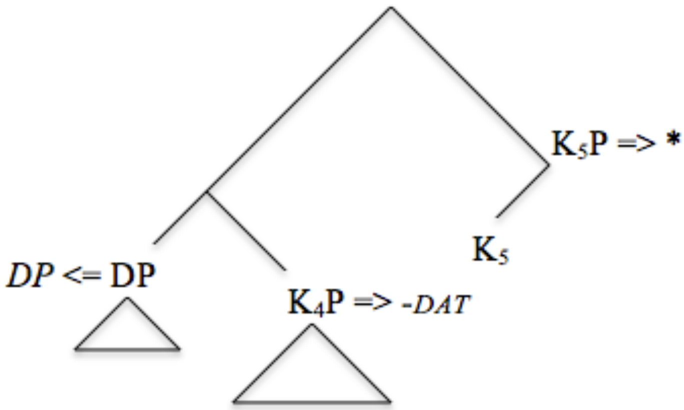




(b) CYCLIC



(c) SNOWBALL



In (63), all three steps in the spellout algorithm fail to produce a constituent that matches the complex head structure [K_6 K_5] found in the lexicon. Rather than crashing, the derivation can at this point activate a second workspace in order to produce a syntactic object that can be lexically matched.

In this second workspace, the syntax takes the feature K_5 (since this is the layer of the structure that we are trying to lexicalize) and merges it with another feature, namely the next feature in the sequence, K_6.[61]

1. K_5 2. Apply (binary) merge

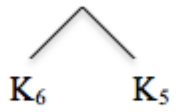


Figure 76 Workspace-2

The complex head from Workspace-2 is then inserted into the structure already sitting in Workspace-1.

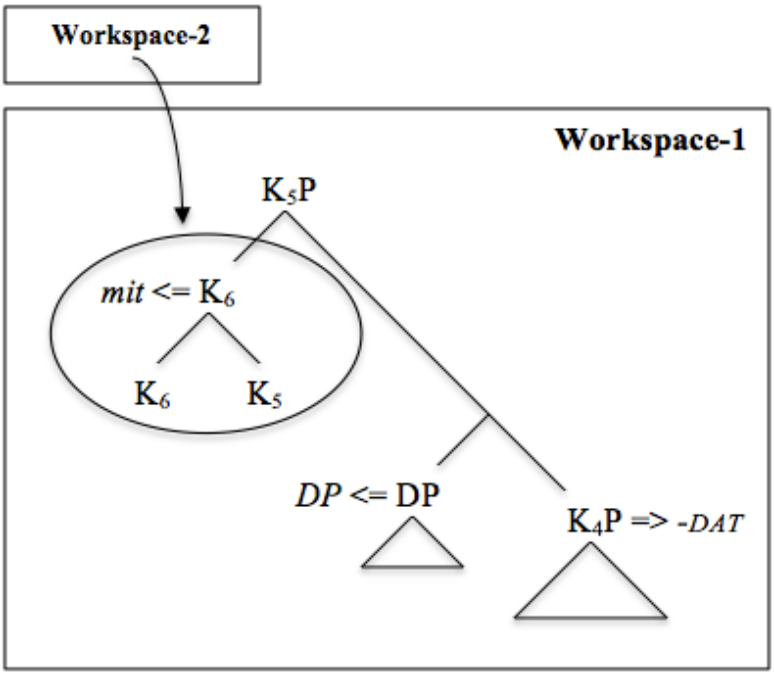


Figure 77 Insert [K_6 K_5] into Workspace-1

The complex head is a constituent that can be targeted for spellout, and there is a suitable lexical entry in the lexicon for this constituent. The result is that *mit* is spelled out, as shown in Figure 77.

The next step in the derivation is the K_6 layer. Recall that K_6 has already been built by the syntax, since the complex head built in the previous stage of the derivation already

[61] In the new workspace, we are more or less starting 'from scratch'. Since merge always applies to two elements, [K_6 K_5] is the smallest possible structure that can be built in the new workspace.

contains K_6. Thus, instead of K_6 being added again, the complex head under K_5P moves up to K_6P to act as its head.[62]

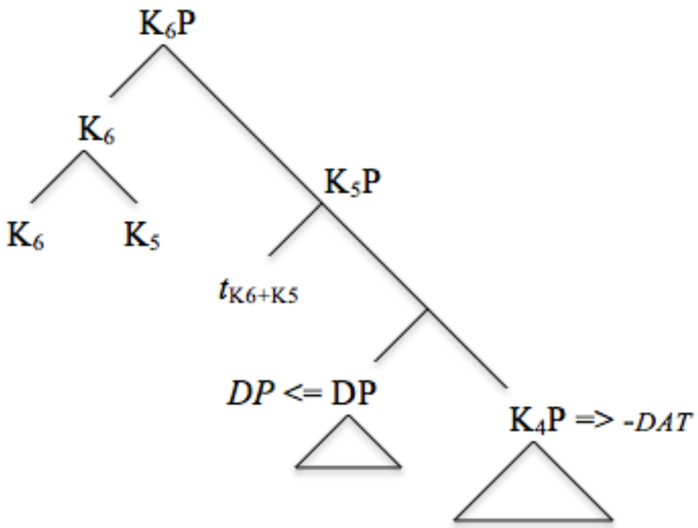


Figure 78 Add K_6 = complex head [K_6 K_5] moves to K_6P to act as its head

The complex head spells out as *mit* at the first step of the algorithm, STAY.

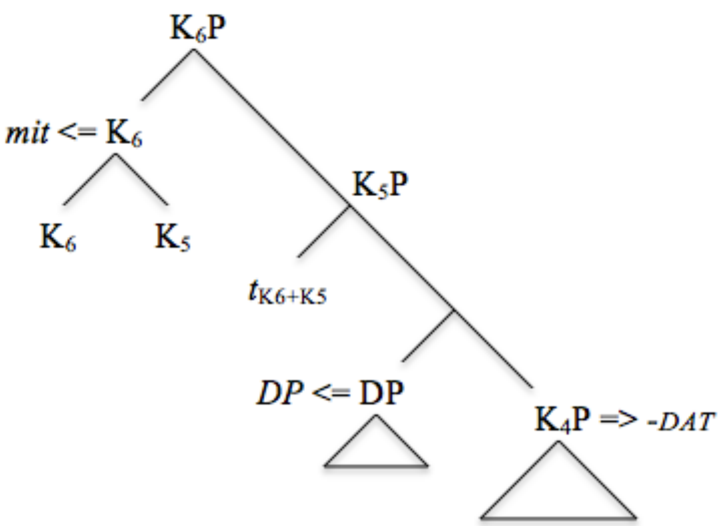


Figure 79 Complex head [K_6 K_5] spells out at STAY

[62] The movement of the complex head to K_6P may be considered a violation of the U20 ban on head movement. The ban on head movement may be reformulated in nanosyntax as a flatout ban on moving *single* features/heads, meaning that only non-singleton sets of features (i.e. complex heads) or projections/phrases may undergo movement.

I adopt this procedure for building complex heads and spelling them out, and in this sense we might say that we are extending the spellout algorithm: STAY > CYCLIC > SNOWBALL > CONSTRUCT.

It can be noted that other ideas are present in the literature for prefix-like items. Caha (2010), for instance, proposes that pre-elements are built by a kind of Travis (1984)-Baker (1988) head movement.

2.2.9 Side note: Subextraction

As a brief side note, recall that Cinque's (2005) problematic order (p) involved a case of subextraction, as illustrated in Figure 80.

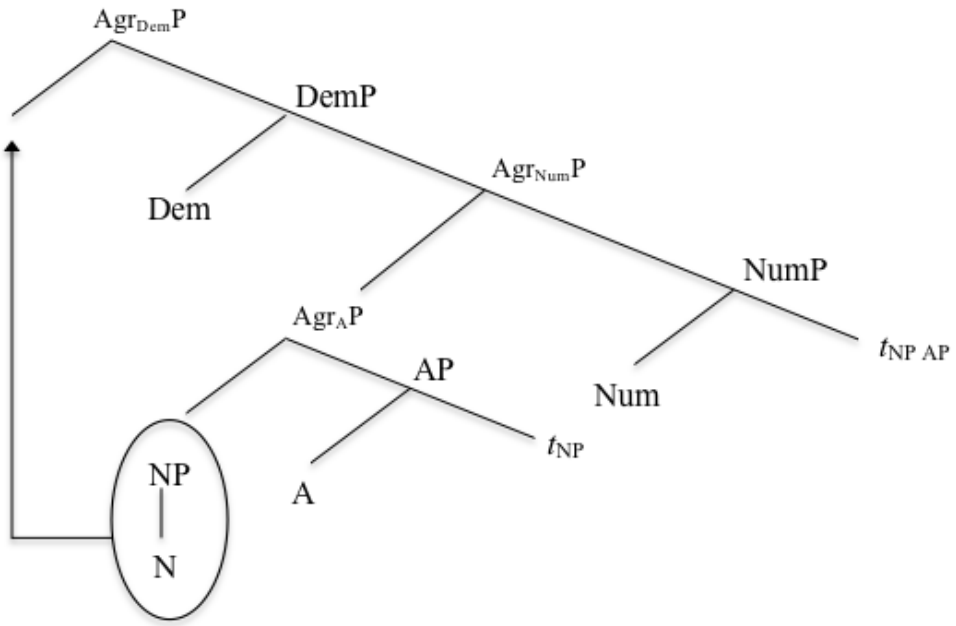


Figure 80 Subextraction in order (p) (N Dem A Num)

In order to derive order (p) (N Dem A Num), NP first moves to the left of AP, after which the entire constituent [[NP] AP] moves to the left of NumP. From there NP subextracts out of the constituent [[NP] AP] to the left of DemP. Also recall from the discussion that data serving as evidence for order (p) were a bit unclear, since only a small number of languages display this order, and even then the order seems to be only one alternative, with the other alternative being a more easily derived order.

Nevertheless, subextraction does not violate any of the U20 rules, and it turns out that there are certain advantages to allowing subextraction in the system. When a constituent is subextracted out of a larger structure, the structure which is left behind is sometimes called a 'peel'. Nanosyntactic analyses involving subextraction are accordingly known as 'peeling' analyses. Following work by M. Starke, Caha (2009: Ch. 4; see also Rocquet

2013), shows that peeling is useful in deriving certain case alternations. In Czech passives, for instance, there is a case alternation between dative and nominative, as shown in (64).

(64) Czech DAT ~ NOM alternation (Caha 2009: 155)

(a) Petr vynadal Karl-**ovi**.
*Peter scolded Charles-**DAT***
'Peter has scolded Charles.

(b) Karel-**Ø** dostal vynadáno (od Petra).
*Charles-**NOM** got scolded from Peter*
'Charles was scolded (by Peter).'

In the active sentence in (64a), the verb 'scold' takes a dative object, which is promoted to a nominative subject in the passive version in (64b). According to Caha (2009: 155-156), this alternation can be captured in the following way. The internal argument of 'scold' is base-generated as a full dative, as shown in Figure 81, where DP is associated with all of the case layers up to K_4.

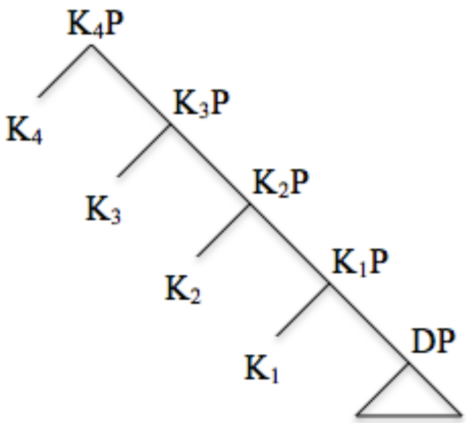


Figure 81 Dative object

In the case of the active pattern in (64a), the structure in Figure 81 is selected by the active verb, and after DP movement to the left of KP, the result is a dative object.

The same structure in Figure 81 is at stake in the passive pattern in (64b), but this time the structure is subject to attraction by a case-selecting head, more precisely a nominative-selector (S-Nom). Thus, the nominative constituent [K_1P [DP]] is subextracted out of the structure in Figure 81, stranding the case layers [K_4P [K_3P [K_2P]]] (the 'peel'). This is illustrated in Figure 82.

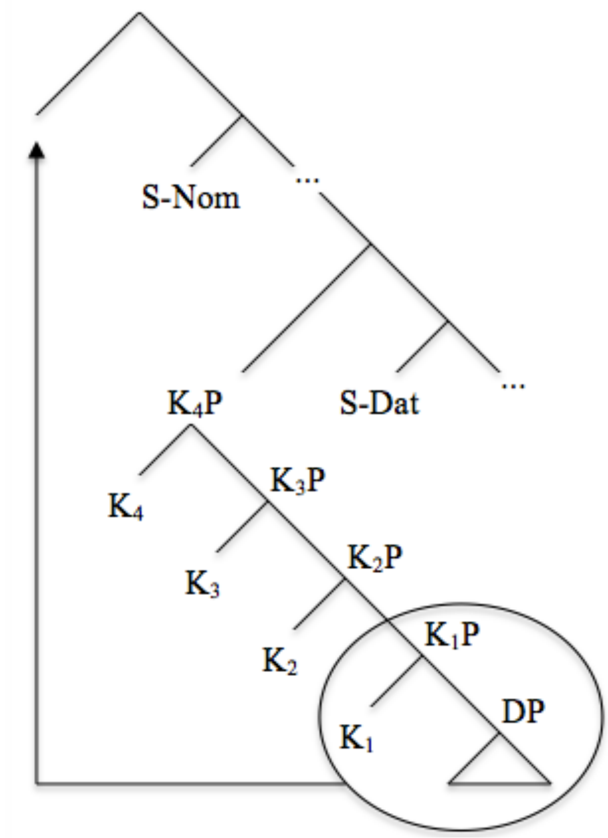


Figure 82 Dative promoted to nominative

The peeling approach can be applied to a wide variety of empirical cases, such as active/passive case alternations, and in some languages the leftover peel can be seen overtly, as discussed by Caha (2009: §4.6) and Rocquet (2013: *passim*). For this reason I do not want to exclude the possibility of subextraction in this dissertation. Indeed, in Section 5.4.3 we shall see that some of my own data can benefit from a subextraction operation.

2.2.10 The Roll-up Shortcut and lexical anchoring

Rather than taking each derivation step by step and seeing how the algorithm creates constituents for matching, it is actually possible to simply look at the way a functional sequence is *packaged*[63] and in doing so predict how spellout-driven movement will treat a given structure.

---

[63] Henceforth I will be using the terms *package(d)* and *packaging* informally, as a way to talk about the way a functional sequence is divided up into lexical entries. *Packaging* is not meant to imply any lack of internal structure.

For example, take the abstract fseq F > E > D > C > B > A, shown in the top line in Figure 83, and let us propose that it is lexically packaged into four separate morphemes, as in the bottom line: the morpheme *na-* corresponds to AP, *no-* to [CP [BP]], *-syn* to DP, and *-tax* to [FP [EP]].

F	E	D	C	B	A
-tax		*-syn*	*-no*		*na-*

Figure 83 Packaging for *na-no-syn-tax*

The tree form of Figure 83 is given in Figure 84.

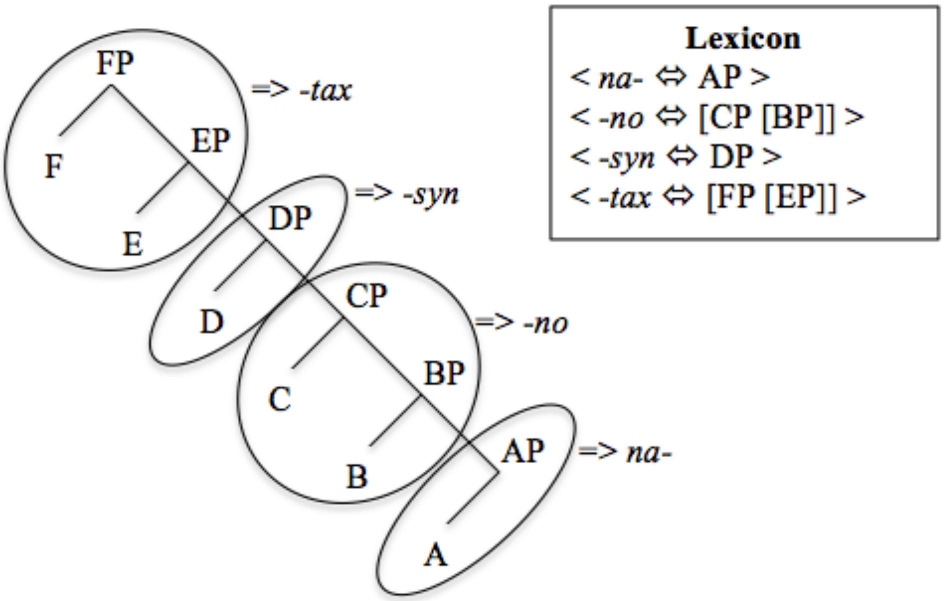


Figure 84 Tree form of Figure 83

Since each morpheme needs to be a constituent on its own in order to spell out, we know that various evacuation movements will be needed in order to create proper constituents in Figures 83 and 84. In other words, spellout-driven movement will be needed.

Again, derivations proceed from the bottom-up. Thus we begin the derivation at the lowest layer of the structure, AP. AP is in fact already a constituent in the structure we have been given in Figure 84 and can spell out as *na-*.

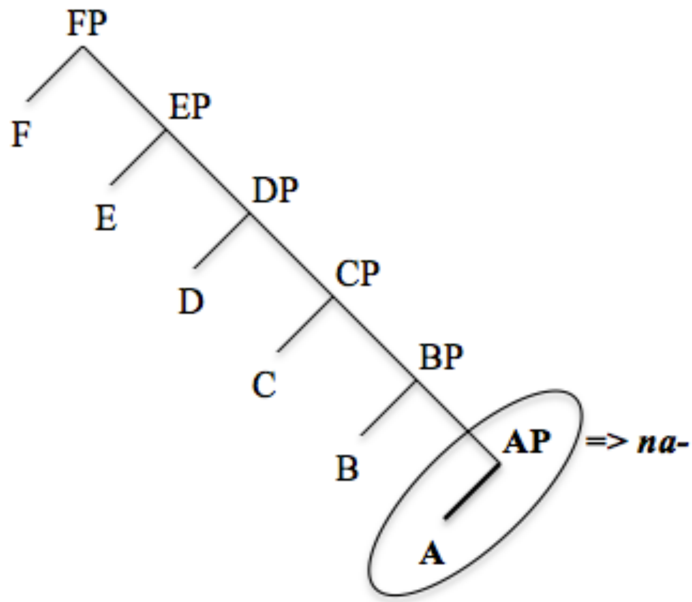


Figure 85 Spelling out AP as *na-*

Moving up the structure, we see that the next feature to be lexicalized is B. However, we already know that the morpheme *-no* corresponds not just to BP but to the entire chunk [CP [BP]]. Since spellout-driven movement operates according to constituency, we also know that [CP [BP]] needs to be a constituent in order to spell out *-no*. This means evacuating AP to the left of [CP [BP]], as in Figure 86.

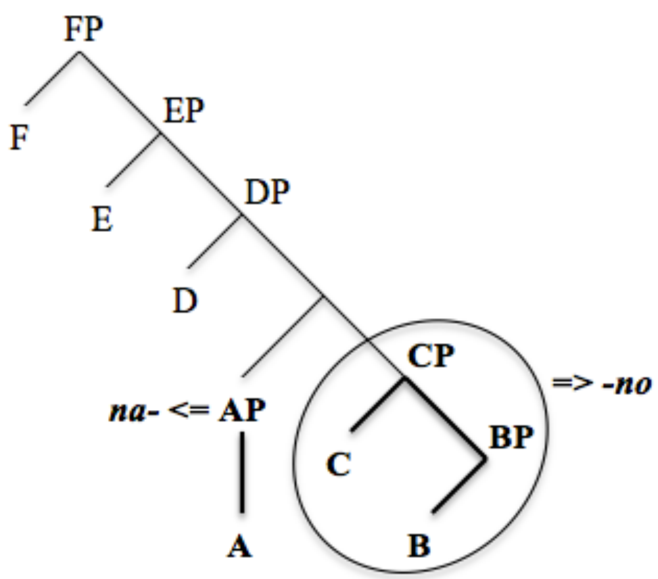


Figure 86 Spelling out [CP [BP]] as *-no*

Next we need DP to be a constituent, so that *-syn* can be spelled out. Thus the entire constituent [[AP] CP BP] will need to move to the left of DP, giving us [[[AP] CP BP] DP] or [[[*na-*] *-no*] *-syn*].

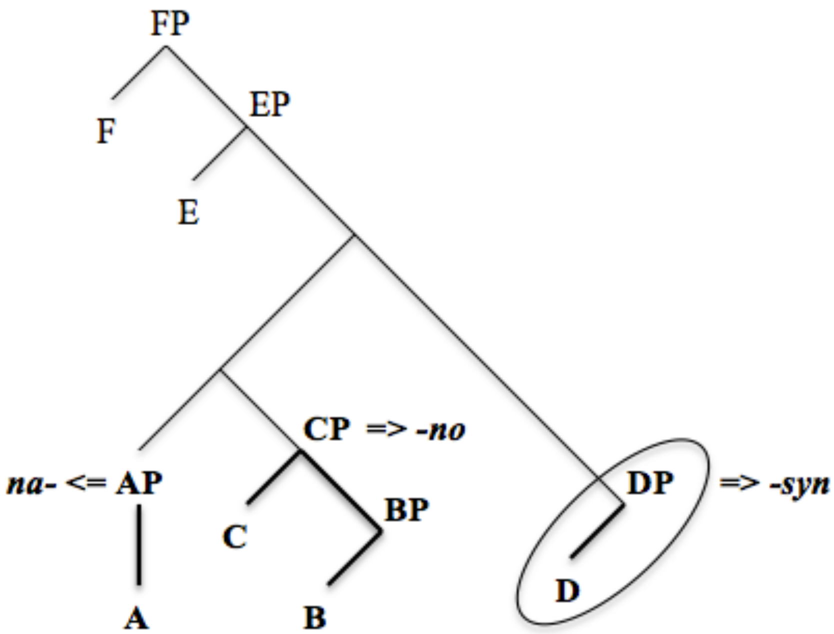


Figure 87 Spelling out DP as *-syn*

Finally, we know that [FP [EP]] needs to be a constituent in order for *-tax* to be realized. Thus the constituent below it – [[[AP] CP BP] DP] – needs to move to the left of FP, giving us [[[[AP] CP BP] DP] FP EP] or [[[[*na-*] *-no*] *-syn*] *-tax*], as seen in Figure 88.

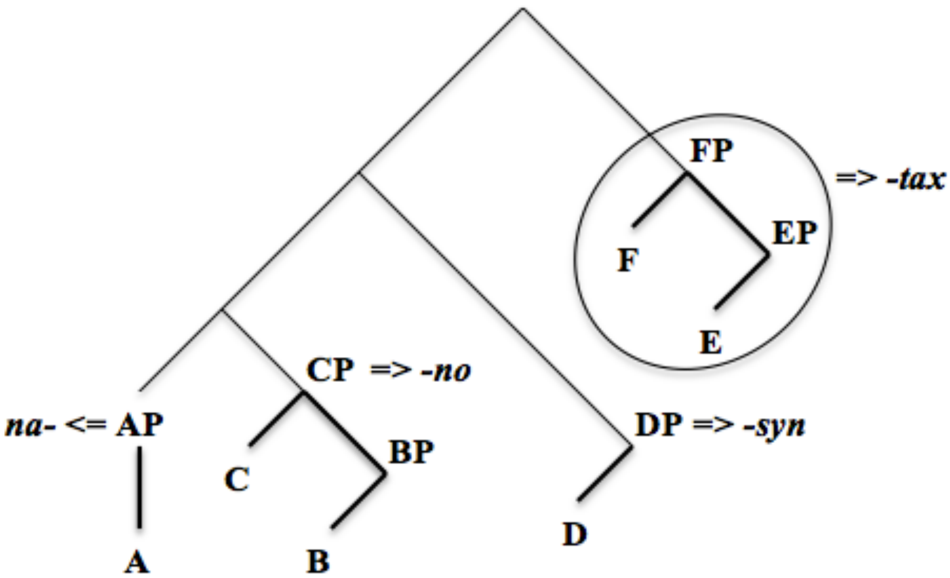


Figure 88 Spelling out [FP [EP]] as *-tax*

In sum, by just looking at the way the fseq is packaged and by making note of which constituents ultimately need to be created in order for spellout to succeed, we can predict how spellout-driven movement will affect a given structure. In short, we effectively skip directly to the SNOWBALL movements which are needed. This spares us the hassle of computing all of the STAY and CYCLIC steps which are interspersed between the SNOWBALL movements. At its core, spellout-driven movement reverses the order of morphemes. At each lexical border in the fseq, then, a SNOWBALL movement is performed which swaps the order of morphemes. Since this results in a final structure which has essentially been 'rolled up', I have called this shortcut the Roll-up Shortcut.

Let us now return to the ON and Finnish portmanteau examples from Section 2.2.1 above. Another way to write these examples is seen in the packaging schemas in Figures 89 and 90.

K	Num	D/N
-um		*bjǫrn*

Figure 89 Packaging for Old Norse *bjǫrn-um* 'bear-DAT.PL'

K	Num	D/N
-lle	*-i*	*karhu*

Figure 90 Finnish *karhu-i-lle* 'bear-PL-ALL'

The derivation of Karata *bajdan-ț'-a-gal* 'through the square' from (57-62) above can also be succinctly rewritten, as in Figure 91.

Route	Source	Goal	Place	AxPart	D
-gal			*-a*	*-ț'*	*bajdan*

Figure 91 Karata *bajdan-ț'-a-gal*

Again, once we know how the fseq is lexically packaged up into morphemes, all one needs to do is perform a SNOWBALL operation at each morpheme border – or in other words, reverse the order of morphemes by roll-up – in order to see the final results of spellout-driven movement.

The exception to the Roll-up Shortcut is when complex heads are involved. In Figure 92, for instance, imagine that the F and E layers lexically correspond to a complex head [F E]. A complex head is marked by a double-edged border in Figure 92 and henceforth.

F	E	D	C	B	A
Proto-		*-ic*	*-an*	*Germ-*	

Figure 92 Packaging involving a complex head

As discussed above, pre-elements (prefixes, prepositions, etc.) can be argued to be complex heads. They spell out *in situ*, without the snowballing/evacuation movements which we need for spelling out suffixes. Thus in Figure 92 the Roll-up Shortcut applies only up to and including the D layer. The complex head [F E] does not partake in this order-swapping, however. It remains *in situ*. In other words, *Germ-* will first move to the left of *-an*, giving [[AP] CP BP] or [[*Germ-*] *-an*]. Next [[AP] CP BP] will move to the left of DP, giving [[[AP] CP BP] DP] or [[[*Germ-*] *-an*] *-ic*]. Next, however, the F and E layers spell out as a complex head and as such put a stop to the roll-up trend. In the end, this results in the order [F E] + [[[AP] CP BP] DP] or [*Proto-*] + [[[*Germ-*] *-an*] *-ic*]. Additional operations transporting constituents to the left of a complex head may be possible, of course, but these will not be spellout-driven movements.

Finally, I would like to point out a situation in which the importance of anchors (cf. Section 2.2.6) becomes quite salient, as this will play a role in the discussion to come in subsequent chapters. Imagine that a certain set of forms share a span of features (D > C > B > A) but are differentiated in their case (K) layers. Consider the difference between a nominative and an accusative form. The NOM form will build only K_1, while the ACC form will build both K_1 and K_2. Both forms, however, need to build the span from A to D. See (65).

(65) NOM form: [DP [CP [BP [AP [K_1P]]]]]
ACC form: [DP [CP [BP [AP [K_2P [K_1P]]]]]]

This difference in the lower K layers between nominative and accusative can in fact trigger a difference in which morphemes are lexicalized later on. For example, if K_2 is packaged together with A, as with the morpheme labeled β in Figure 93, then the next morpheme to be triggered will be morpheme γ.

D	C	B	A	K_2	K_1
γ			β		α

Figure 93 A and K_2 are lexically packaged together in the ACC

Now applying the Roll-up Shortcut to Figure 93, we see that the accusative form will have the spellout α-β-γ.

In the NOM form, on the other hand, there is no K_2 at all. Suppose the packaging for the nominative form is the one seen in Figure 94.

D	C	B	A	K_1
ε		δ		α

Figure 94 No K_2 (and thus no lexical packaging of A with K_2) in the NOM

In the nominative derivation there is no K_2, and with no K_2 in the structure, morpheme β from Figure 93 will not be triggered (since it would be missing its anchor K_2). Instead the feature A is added directly after K_1 in the nominative derivation. This triggers the spellout of morpheme δ, which is anchored at A. The spellout of δ will subsequently trigger morpheme ε to be spelled out. The nominative form, then, will have the spellout α-δ-ε.

To sum up, a difference in one layer of the fseq may trigger a cascade effect that goes all the way up the fseq. In Figure 93, the presence of the K_2 head triggers one lexicalization pattern (β plus γ), while the absence of the K_2 head in Figure 94 triggers another pattern (δ plus ε). This is due to the way the fseq happens to be packaged – more precisely, how this packaging determines the anchoring of lexical entries.

2.2.11 Summary

Nanosyntax is a theory of a principled morphosyntax and a principled lexicon. It proposes to develop a principled theory of morphosyntax by breaking down the barrier between syntax and morphology, positing just a single computational system, SMS (syntax-morphology-semantics). It develops a principled lexicon by explicitly stating what a lexical entry must contain (phonology, syntax, and conceptual information) and how L-trees match S-trees during spellout.

The most important guidelines of the theory have been presented above. Matching lexical trees to syntactic trees during the lexicalization process is governed by three core principles: the Superset Principle, the Elsewhere Principle, and the Principle of Cyclic Override. Furthermore, I adopt a strictly constituent-based spellout system which is regulated by Cyclic Exhaustive Lexicalization and the algorithm STAY > CYCLIC > SNOWBALL (> CONSTRUCT). As discussed, the Anchor Condition also plays an important role in determining lexicalization patterns.

3

A descriptive decomposition of the Old Norse reinforced demonstrative

In this chapter I undertake a fine-grained morphological decomposition of the RDem pronoun in ON. In Section 3.1 I present three primary observations about the RDem paradigm. These observations will reveal that there are in fact three different kinds of RDem structures within the paradigm, and in Section 3.2 the morphological templates of these three structures are presented. Finally in Section 3.3 a more refined take on these templates is discussed. By the end of the chapter, we end up with the empirical results that will serve as input to the formal analysis of the ON RDem forms in Chapters 4 and 5.

3.1 Three primary observations

The focus of this chapter is on the internal structure of the 24 forms in the RDem paradigm in Table 34.

Table 34 ON RDem 'this' (Gordon 1956: 295)

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	þessi	þessi	þetta	þessar	þessir	þessi
ACC	þessa	þenna	þetta	þessar	þessa	þessi
GEN	þessar	þessa	þessa	þessa	þessa	þessa
DAT	þessi	þessum	þessu	þessum	þessum	þessum

Recall from Chapter 1 that an older stage of Norse, recorded in various runic inscriptions, preserves a stage in RDem's development that may be referred to as the Dem-*si* stage. Some Dem-*si* forms are given in (66).

(66) Dem-*si* forms in RN

F.NOM.SG	**susi** (*súsi*)	<	NWGmc **sō-si*
M.NOM.SG	**saʀ:si** (*saʀsi*)		
M.NOM.SG	**sasi** (*sási*)	<	NWGmc **sa-si*
M.ACC.SG	**þan:si** (*þansi*)		
	þansi (*þansi*)	<	NWGmc **þa-n-si*
N.ACC.SG	**þat:si** (*þatsi*)		
	þatsi (*þatsi*)	<	NWGmc **þa-t-si*
M.DAT.SG	**þaimsi** (*þœimsi*)	<	NWGmc **þai-m-si*

The RN forms in (66) are transparently decomposable as the neutral Dem form plus the sigmatic reinforcer *-si*. The ON forms in Table 34, on the other hand, appear to be much more opaque. Clearly there has been a great deal of historical development between the RN and ON stages. In this section we shall see, however, that the ON forms in Table 34 are also perfectly decomposable into separate and discrete morphemes. This has already been hinted at in Chapter 1, but in this chapter we start essentially from scratch with the aim of breaking down the RDem forms in as systematic a way as possible.

I start with three basic observations about the ON RDem paradigm. The first observation is that there is a base *þe-* present in every single RDem form in the paradigm (Section 3.1.1). The second observation is that there is a large subset of forms in the RDem paradigm which shows the basic pattern of the RDem stem *þess-* plus the K component (Section 3.1.2). The third observation is that there are two types of forms which do not fall into the *þess*-K pattern (Section 3.1.3).

3.1.1 The base *þe-*

As for the first observation, every form in Table 34 begins with the sequence *þe-*. This is illustrated in Table 35.

Table 35 ON RDem parsed at *þe-* boundary

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	þe-ssi	þe-ssi	þe-tta	þe-ssar	þe-ssir	þe-ssi
ACC	þe-ssa	þe-nna	þe-tta	þe-ssar	þe-ssa	þe-ssi
GEN	þe-ssar	þe-ssa	þe-ssa	þe-ssa	þe-ssa	þe-ssa
DAT	þe-ssi	þe-ssum	þe-ssu	þe-ssum	þe-ssum	þe-ssum

In order to identify the nature of *þe-*, we must consider the ON Dem paradigm in Table 36.

Table 36 ON Dem 'that' (Gordon 1956: 295, Haugen 1982: 101)

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	sú	sá	þat	þær / þār	þeir	þau
ACC	þá	þann	þat	þær / þār	þá	þau
GEN	þeir(r)ar	þess	þess	þeir(r)a	þeir(r)a	þeir(r)a
DAT	þeir(r)i	þeim	því / þȳ	þeim	þeim	þeim

There are three kinds of Dem stems, a pattern which was inherited from the PGmc paradigm. These Dem stems are inflected with pronominal case endings (K_D), as illustrated in (67). In (68), all of the Dem forms are divided according to their Dem stem (*þ-*, *þei-*, or *þa-*).

(67) Dem stems plus pronominal K

[*þ-* (68a-d)

þei- (68e-j)] + -K_D

þa- (68k-m)]

(68) Dem forms according to stem

(a) F.ACC.SG þ-á
(b) F.NOM/ACC.PL þ-ær, þ-ár
(c) M.ACC.PL þ-á
(d) N.DAT.SG þ-(v)í, þ-ȳ

(e) F.GEN.SG þei-(r)rar
(f) F.DAT.SG þei-(r)ri
(g) M.DAT.SG þei-m
(h) M.NOM.PL þei-r
(i) GEN.PL þei-(r)ra
(j) DAT.PL þei-m

(k) M.ACC.SG þa-n(n)
(l) N.NOM/ACC.SG þa-t
(m) N.NOM/ACC.PL þa-u

The vowel-less stem *þ-* (68a-d) comes from PGmc **þ-*. The diphthongal stem *þei-* (68e-j) comes from PGmc **þai*. The monophthongal stem *þa-* (68k-m) comes from PGmc **þa-*. Note here that PGmc **ai* goes to Norse *ei*, which can historically be considered an early application of the *i*-umlaut rule (*a* + *i* > *ei*). This is not the place for a complete analysis of the Dem paradigm and the internal structures of its three stems, so here I will simply adopt at face value the traditional observation in the philological literature that there are three different Dem stems. For the purposes of this dissertation, I will take these three stems to be more or less on a par, structurally speaking.[64]

The reader will notice that some of the forms in Table 36 are apparently not based on one of the three Dem stems. For instance, the M/N.GEN.SG form *þe-ss* has a stem *þe-*, with *e*-vocalism instead of the expected *a*-vocalism. As mentioned in Section 1.1.2, this form ultimately comes from PGmc **þe-s(a)*, which also had an irregular *e*-vocalism. See Nielsen (2000: 230-235) for the complex history of this form. Two other irregular forms are F.NOM.SG *sú* and M.NOM.SG *sá*. As also mentioned in Section 1.1.2, these go back to irregular forms in PGmc (and even PIE) too: ON *sú* < PGmc **sō* and ON *sá* (< PN **sa-R*) < PGmc **sa*.

There are at least two historical processes responsible for why we see *þe-* in the ON RDem paradigm instead of a Dem stem like *þa-* or *þei-*. These have already been discussed in Section 1.2.1, but for the sake of convenience I briefly repeat them here.

The first process involves a phonotactic constraint. The sequence **Rs* in PN undergoes assimilation to *ss* in ON. Various Dem-*si* forms in PN had the sequence **Rs* and were therefore subject to this sound shift. In addition to this sound change, there was also a

[64] Note that it is very unlikely that we are simply dealing with allomorphs. For instance, the *-a* vs. *-ei* opposition in *þ-a-* vs. *þ-ei-* must be structural and not phonological. An allomorphic hypothesis might claim that *-ei* appears only in the environment right before *m* or *r* (e.g. *þ-ei-m* or *þ-ei-r*). But it is very well known that in ON *r* and *n* often participate together in phonological processes (e.g. assimilation rules), yet there is no M.ACC.SG **þ-ei-n(n)*. The allomorphic hypothesis for *-ei* ~ *-a*, then, would assume that *m* and *r* pattern together against *n*. This *m/r* vs. *n* opposition is a highly unnatural fit for the ON phonological system.

general shift from internal inflection (to the left of the reinforcer -*si*) to external inflection (to the right of the reinforcer -*si*) in the RDem system of Norse. The combined result of the assimilation process and the removal of inflection from internal position resulted in the transitional stems **þeiss-* and **þēss-*, the ancestors of the ON RDem stem *þess-*. These changes are summarized in (69).

(69) Transitional stems *þeiss-*/*þēss-* reduced to *þess-*

(a)	F.GEN.SG	*þei-RaR-si	>	*þei-~~Ra~~-ssi	>	*þei-ss(i)-	>	ON þess-ar
(b)	DAT.PL	*þei-mR-si	>	*þei-~~m~~-ssi	>	*þei-ss(i)-	>	ON þess-um
(c)	M.NOM.PL	þei-R-si	>	*þei-/-ssi	>	*þei-ss(i)-	>	ON þess-ir
(d)	F.NOM/ACC.PL	þ-ēR-si	>	*þē-/-ssi	>	*þē-ss(i)-	>	ON þess-ar

Crucially, the stems **þeiss-* and **þēss-* do not syllabify correctly. Observe that the diphthong/long vowel present in the transitional stems forces the geminate *ss* into onset position: **þei.ss-* and **þē.ss-*. The problem is that *ss* is an impermissible onset in ON. Thus there was phonotactic pressure to reduce the diphthong *ei* and the long vowel *ē* to monophthongal *e*. This results in the observed stem *þess-*, which does syllabify correctly (i.e. *þes.s-*).

The second process responsible for the sequence *þe-* in RDem is *i*-umlaut. More precisely, the vowel *i* in the reinforcer -*si* was known to trigger *i*-umlaut (Nielsen 2000: 237, n. 3), which is evident from spelling alternations robustly attested in RN and given here in (70).

(70) Spelling alternations and *i*-umlaut

M.ACC.SG **þan**(:)**si** (*þansi*) vs. **þensi** or **þinsi** (*þensi*)

N.ACC.SG **þat**(:)**si** (*þatsi*) vs. **þitsi** (*þetsi*)

F.ACC.SG **þa**(:)**si** (*þāsi*) vs. **þesi** (*þēsi*)

In other words, the forms *þansi*, *þatsi*, and *þāsi*, usually spelled with <**a**>, could also be pronounced with a front vowel, something that can be deduced from the alternate spellings with <**e**> and <**i**>.

These processes (and perhaps also the fact that the M/N.GEN.SG Dem form was *þess*, with a stem *þe-*) ensured that the RDem paradigm in ON ended up with a base *þe-* instead of something like *þa-* or *þei-*. Synchronically speaking, I will treat the RDem base *þe-* as the *i*-umlauted allomorph of the Dem stem *þa-*, as seen in (71).

(71) RDem base with *i*-umlaut

/þa-/...^{i} > [þe-]...

That is, the underlying base of RDem is /þa-/, which happens to be mutated to [þe-] in all the forms we see in the paradigm. Synchronic evidence for this *i*-mutation will be provided below.

3.1.2 The main pattern: The K-final forms

The second observation to be made about the RDem paradigm is that a majority of the forms show a stem *þess-* plus strong adjectival endings (K) of the *n*-type class. The forms which show this pattern are surrounded by a border in Table 37 and may informally be called the boxed forms. More formally, we can call these forms the *K-final forms*, since the K component is in final position in these forms. The strong adjectival endings (K) are provided in Table 38.

Table 37 *n*-type strong adjective endings (K)

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	þessi	þessi	þetta	þess-ar	þess-ir	þessi
ACC	þess-a	þenna	þetta	þess-ar	þess-a	þessi
GEN	þess-ar	þessa	þessa	þess-a	þess-a	þess-a
DAT	þess-i	þess-um	þess-u	þess-um	þess-um	þess-um

Table 38 *n*-type strong adjective endings (K)

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	-Ø	-r	-t	-ar	-ir	-Ø
ACC	-a	-n	-t	-ar	-a	-Ø
GEN	-rar	-s	-s	-ra	-ra	-ra
DAT	-ri	-um	-u	-um	-um	-um

It is important to keep in mind the exact formulation of K-final. A K-final form must be composed of *þess-* plus one of the strong adjectival endings (K) in Table 38. Thus, though it begins with the sequence *þess-*, the M/N.GEN.SG form *þessa* is not K-final (and thus not boxed in Table 37) because *-a* is not the M/N.GEN.SG K ending in Table 38 (the relevant ending being *-s*).

The approach taken in this dissertation places RDem in the *n*-type inflectional class, the other members of which are listed in (72).

(72) Members of *n*-type class (Gordon 1956: 290, Faarlund 2004: 38)

- Bisyllabic adjectives ending in -*inn*
- Past participles of strong verbs
- **Various D-like items**: -*inn* 'the' (< *hinn* 'that'); *minn* 'my', *þinn* 'your', *sinn* '3.POSS.REFL'; *einn* 'one'; *nǫkkurr* 'some, any', *hverr* 'who'

Clearly RDem fits well within the *n*-type class considering the other D-like elements which belong to this class. In Table 39 I have provided the paradigm of the *n*-type D item *hinn* 'that, the' in order to show its formal similarities to RDem in terms of inflection.

Table 39 *hin*, *hinn*, *hit* 'that, the' (Faarlund 2004: 38)

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	hin-Ø	hin-n (< hin-r)	hit (< hin-t)	hin-ar	hin-ir	hin-Ø
ACC	hin-a	hin-n	hit (< hin-t)	hin-ar	hin-a	hin-Ø
GEN	hin-nar (< hin-rar)	hin-s	hin-s	hin-na (< hin-ra)	hin-na (< hin-ra)	hin-na (< hin-ra)
DAT	hin-ni (< hin-ri)	hin-um	hin-u	hin-um	hin-um	hin-um

It should be pointed out that there is another class of strong adjective inflection, the *an*-type class, which differs from the *n*-type class in a few endings. The endings of the *an*-type class are given in Table 40. The declension of an *an*-type adjective, *jarp*- 'brown', is given in Table 41.

Table 40 Strong adjective endings (Haugen 1982: 102)

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	-$Ø^{u}$	-r	-t	-ar	-ir	-$Ø^{u}$
ACC	-a	-an	-t	-ar	-a	-$Ø^{u}$
GEN	-rar	-s	-s	-ra	-ra	-ra
DAT	-ri	-um	-u	-um	-um	-um

Table 41 *jarp-* ‘brown’

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	jǫrp (< jarp-Ø^{u})	jarp-r	jarp-t	jarp-ar	jarp-ir	jǫrp (< jarp-Ø^{u})
ACC	jarp-a	jarp-an	jarp-t	jarp-ar	jarp-a	jǫrp (< jarp-Ø^{u})
GEN	jarp-rar	jarp-s	jarp-s	jarp-ra	jarp-ra	jarp-ra
DAT	jarp-ri	jǫrp-um	jǫrp-u	jǫrp-um	jǫrp-um	jǫrp-um

There are two main differences to note between Tables 40 and 41. First, the F.NOM.SG and N.NOM/ACC.PL *n*-type endings are *-Ø*, in contrast to *an*-type *-Ø^{u}*, with a *u*-mutator. The lack of *u*-mutation is evident from the F.NOM.SG and N.NOM/ACC.PL RDem form *þessi*, which would otherwise appear as **þøssi* with *u*-mutation.[65] (As another example, the *n*-type item *hin* ‘that, the’ would appear as **hyn* (< *hin-Ø^{u}*) had it been affected by *u*-mutation.) On the other hand, the *an*-type form *jǫrp* (< *jarp-Ø^{u}*) does display *u*-mutation of the root vowel *a*. The second difference is that the M.ACC.SG *n*-type ending is *-n*, in contrast to *an*-type *-an*. As we will see more clearly below, the M.ACC.SG RDem form *þenna* does not show the inflection *-an* but rather *-n*, while the *an*-type adjective *jarp-an* in Table 41 shows the ending *-an*. We can conclude, then, that the *an*-type declension is not relevant for RDem.

It will also be noticed that my generalization about the boxed forms appears to have an exception: the *r* of the *r*-initial inflectional endings is absent in the corresponding RDem forms. In (73a) I provide the relevant endings. (73b) illustrates that the RDem forms are lacking the ending-initial *r*.

---

[65] Readers familiar with runic may know that many instances of N.NOM/ACC.PL **þausi** and **þusi** (found in Danish and Swedish inscriptions) are transcribed as *þøsi*, where both *i*- and *u*-umlaut have occurred: *þa-^{i}* > *þe-^{u}* > *þø-*. Moreover, this form looks very much like the form **þøssi* which I have claimed does not exist. This might lead us to conclude that RN shows evidence that *u*-umlaut actually was active in the RDem paradigm at some stage. There are two problems with this claim, however. The first problem anticipates my discussion below, where I conclude that the *þessi* forms have internal inflection, meaning that the *u*-umlaut hypothesis would give an underlying *þe-u-ssi*, but according to Haugen (1982: 33), *i*- and *u*-umlaut could both occur only if they were acting from the same syllable (his example being **akuisi* > ON *øx* ‘axe’). The second problem is that *au* goes to *ø:* and *ai* to *i:* by the so-called East Norse monophthongization. The spelling alternation between **þausi** and **þusi** is a typical side effect of the monophthongization (cf. **stain** vs. **stin** ‘stone’, **tauþr** vs. **tuþr** ‘dead’, etc.; see Barnes 2005: 182; Antonsen 2002: 313; see also Haugen 1976, 1982). Thus the forms **þausi** and **þusi** are simply representatives of the older Dem-*si* stage, where the N.ACC.PL Dem pronoun *þau* has undergone a sound change to *þø:*, giving *þø:-si*. Therefore there is in fact no evidence at all from RN for a *u*-umlauted **þøssi*, only for an East Norse Dem-*si* form *þau-si* with monophthongization of *au* to *ø:*.

(73) *r*-initial endings

(a) F.GEN.SG -rar
F.DAT.SG -ri
GEN.PL -ra

(b) F.GEN.SG þess-**ar**
F.DAT.SG þess-**i**
GEN.PL þess-**a**

In F.GEN.SG *þessar* we see *-ar* instead of expected *-rar*, in F.DAT.SG *þessi* we see *-i* instead of expected *-ri*, and in GEN.PL *þessa* we see *-a* instead of expected *-ra*.

In fact, this pattern is not surprising at all in the context of ON phonology. A well known characteristic of the ON morphophonological system is that inflectional *r* is highly vulnerable to assimilation (Gordon 1956: 282, Haugen 1982: 64, among others). Not only does *sr* go to *ss* as illustrated for *laus-r* 'loose-M.NOM.SG' in (74), but *ssr* goes to *ss* as shown for the adjective *hvass-* 'sharp' in (75).

(74) *sr* > *ss*

cf. M.NOM.SG *laus-r* 'loose' > *lauss*

(75) *ssr* > *ss*

cf. *hvass-* 'sharp'

F.GEN.SG hvass-**ar**
F.DAT.SG hvass-**i**
GEN.PL hvass-**a**

The rule in (75) is directly relevant for the RDem forms, as illustrated in (76).

(76) F.GEN.SG þess-rar > þess-ar
F.DAT.SG þess-ri > þess-i
GEN.PL þess-ra > þess-a

The rule in (75) and its application in (76) can be thought of in different ways. On the one hand, it could be an assimilation rule, such that *r* assimilates to *s* first, as in *ssr* > *sss*, after which the triplet *sss* is reduced to a simple geminate *ss*. On the other hand, it could be viewed as a straightforward deletion or cluster reduction instead of assimilation, such that *r* in *ssr* is deleted in order to simplify the consonant cluster to *ss*. The reduction of clusters involving geminates is also a well known ON phenomenon (Noreen 1923: 207-208, Barnes 2004: 101, among others). For my purposes it does not matter which one of these is the correct analysis. The crucial thing for me is that there is a phonological process at work, rather than something structural. See Appendix II for more details on the history of *r*-initial endings and the 'intrusive *r*' of post-classical ON.

3.1.3 The constant forms and the K-internal forms

The third and final main observation to make about the ON RDem paradigm is that the forms which remain – i.e. those forms which are not K-final and hence are not boxed – fall into two categories. As seen in Table 42, one set of forms ends in *-i* (these are lightly shaded), while another set ends in *-a* (these are darkly shaded).

Table 42 Two kinds of non-boxed forms

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	þessi	þessi	þetta	þess-ar	þess-ir	þessi
ACC	þess-a	þenna	þetta	þess-ar	þess-a	þessi
GEN	þess-ar	þessa	þessa	þess-a	þess-a	þess-a
DAT	þess-i	þess-um	þess-u	þess-um	þess-um	þess-um

Unlike the K-final forms, the non-boxed forms do not end with the K endings in Table 38.

The lightly shaded forms in Table 42 (F/M.NOM.SG and N.NOM/ACC.PL) all surface as *þessi* and will thus be referred to as the *constant forms*. It is not immediately obvious where the K component is located in the constant forms: M.NOM.SG *-r* is nowhere to be seen in *þessi*, and for F.NOM.SG / N.NOM/ACC.PL the ending is null anyway (i.e. *-Ø*). Thus it is not clear at this point if K is to the left or to the right of the sigmatic reinforcer, that is, if the constant forms have internal or external inflection.

As for the darkly shaded forms in Table 42 (M.ACC.SG, N.NOM/ACC.SG, and M/N.GEN.SG), all of these end with the asigmatic reinforcer *-a*. Moreover, in contrast with the constant forms, these forms show their K component overtly. K is word-internal and geminated in *þe-**nn**-a* (M.ACC.SG *-n*), *þe-**tt**-a* (N.NOM/ACC.SG *-t*), and *þe-**ss**-a* (M/N.GEN.SG *-s*). Since K is located to the left of the asigmatic reinforcer *-a*, these forms will be referred to as the *K-internal forms* from now on.

My identification of three types of RDem forms – K-final, constant, and K-internal – takes special note of where K is located in each type of structure. This is an improvement on more traditional approaches to RDem, where this is not always a primary concern. Haugen (1976: 156), for instance, claims that three RDem stems are active in the paradigm: *þess-*, *þenn-*, and *þett-*. As should be clear by now, this approach is too superficial since it does not take care to separate out inflectional material. While we can agree with Haugen that *þess-* is a stem, the 'stems' *þenn-* and *þett-* cannot be on a par with *þess-* because they already contain K while *þess-* does not. Thus an approach like Haugen's does not say anything meaningful about the position of K. This oversight is especially pertinent to the non-boxed forms, where K requires extra attention since it is either not overtly visible (as in the constant forms) or found word-internally in geminated form (as in the K-internal forms).

3.2 Morphological templates

In this section I introduce the descriptive device 'morphological template'. A morphological template is intended to show how the morphological ingredients of RDem fit together to make the three structures identified above.

3.2.1 The template of the K-final forms

The K-final forms consist of the stem *þess-* plus K. The stem *þess-* can itself be split into the base *þe-*, the *i*-umlauted allomorph of the Dem base *þa-* (see Section 3.1.1), and the reinforcer component -*ss*^{i}- (which will be explained more below). K is located to the right of the sigmatic reinforcer component, meaning that the K-final forms display what is known as external inflection. The template of the K-final/boxed forms is given in (77).

(77) ***þa-ss*^{i}-K**

þa-	-ss^{i}-	-K
base	*geminated sigmatic reinforcer with floating* i	*case inflection*

As for the reinforcer component -*ss*^{i}-, recall that the ON sequence -*ss*- historically comes from PN *-*R-si*, where the reinforcer *-*si* was known to condition *i*-umlaut. With this history in mind, I have chosen to represent -*ss*- as -*ss*^{i}- (historically, PN **Rsi* > ON *ss*^{i}), in which a 'floating *i*' has been introduced. This floating *i* I take to be a bundle of phonological features, minimally [+front, –low], which make up part of the phonological structure of the reinforcer morpheme -*ss*^{i}-. This phonological feature bundle induces *i*-umlaut and subsequently deletes, as illustrated in (78).

(78) þa-ss^{i}-K > þe-ss-K

The floating *i* introduced here is intended to parallel ideas in Gibson & Ringen (2000), who propose that morphemes inducing ʏ-umlaut in modern Icelandic contain a ‘floating’ bundle of [+ round, – back] features in their phonological structures.[66]

Though for now the morpheme *-ss*^{i}*-* with floating *i* has been based on diachronic evidence, it can be synchronically motivated. As we will see below, floating *i* does not always delete as in (78); in some cases it surfaces overtly (as the full vowel *i*) within the RDem paradigm.

3.2.2 The template of the K-internal forms

The K-internal forms begin with the base *þa-*, plus a geminated K ending, plus the asigmatic reinforcer *-a*. Since K is located to the left of the asigmatic reinforcer, these forms display what is known as internal inflection. This is shown schematically in (79).

(79) ***þa*-KK^{i}-*a***

þa-	-KK^{i}-	-a
base	*geminated K*	*asigmatic*
	with floating i	*reinforcer*

Once again, in order to mutate the base *þa-* to *þe-*, we need to posit a floating *i*-mutator. To stay consistent with the template of the K-final forms postulated above, I postulate that there is a floating *i* after the geminated K component in (79). Floating *i* has the same effect in the K-final forms. That is, it induces *i*-umlaut and is then deleted. This is illustrated in (80).

(80)				
	M.ACC.SG	þa-nn^{i}-a	>	þe-nn-a
	N.NOM/ACC.SG	þa-tt^{i}-a	>	þe-tt-a
	M/N.GEN.SG	þa-ss^{i}-a	>	þe-ss-a

Observe that floating *i* is associated with the geminated sigmatic reinforcer in the K-final template (and also in the constant template; see the next section) but with geminated K in the K-internal template. The exact morphological status of floating *i* will become clearer in Section 3.3.

[66] Cf. also D’Hulst (2006): Latin plural *-s* > Italian *-i*, where the [+coronal] feature of *s* survives all the way into the plural marker of modern Romance.

3.2.3 The constant forms and the phonology of floating *i*

The constant forms have the template shown in (81).

(81) ***þa*-K-*ss*^{i}**

þa-	-K	-ss^{i}
base	*case inflection*	*geminated sigmatic reinforcer with floating* i

There are two important things to notice about (81). The first is that this template invariably results in the form *þessi*, which means that in this template floating *i* is actually phonologically realized as the vowel *i*. Thus these forms are especially important because they provide synchronic support for the existence of the floating *i*-mutator postulated above. The second thing to notice about (81) is that I have placed K to the left of the reinforcer component, meaning that the constant forms have internal inflection. Nothing I have said so far justifies putting K in this position. Interestingly, however, understanding the phonology of floating *i* – i.e. exactly in which environments it deletes and in which environments it surfaces – will give us a test for locating K in the constant forms.

In ON there was a class of weak *i*-stem verbs, defined as those verbs consisting of a verbal root plus a "stem-forming suffix" *-i* (Faarlund 2004: 45). I propose that this "stem-forming suffix" *-i* can be identified with the floating *i* postulated here. An example of a weak *i*-stem verb is *dœm-* 'judge', which has the underlying phonology *dōm*^{i}-, where *o:* goes to *ø:* (orthographically <œ>) by *i*-umlaut. This verb inflects in the present indicative as in Table 44.

Table 43 Verb endings

	SG	PL
1	-Ø	-um
2	-r	-ið
3	-r	-a

Table 44 Present indicative of *dœm-* /dōm^{i}-/ (Faarlund 2004: 49)

	SG	PL
1	dœmi	dœm-um
2	dœm-ir	dœm-ið
3	dœm-ir	dœm-a

Note that postulating an underlying form like /dōm^{i}-/ is synchronically supported by the existence of the noun *dóm-*[67] ‘judgment, opinion’ in ON. Other examples of *i*-stem verbs like *dœm-* include *fœr-*[68] ‘lead’, i.e. /fōr^{i}-/, for which there exists the non-umlauted preterite *fór* for the verb *far-* ‘travel’; *nefn-*[69] ‘mention’, i.e. /nafn^{i}-/, for which the non-umlauted noun *nafn* ‘name’ exists; and *leys-*[70] ‘loosen’, i.e. /laus^{i}-/, for which the non-umlauted adjective *laus-* ‘loose’ exists.

There are three main observations to be made about the phonological behavior of floating *i*, which are presented in the next three subsections. In Section 3.2.3.1 it will be shown that floating *i* is deleted before vowels. In Section 3.2.3.2 it will be seen that floating *i* becomes overt in word-final position. In Section 3.2.3.3. we will see that floating *i* is deleted in medial open syllables.

3.2.3.1 Floating *i* deletes before vowels

The paradigm of *dœm-* in Table 44, though only consisting of six forms, can tell us quite a bit about floating *i*. Perhaps the most obvious observation is that floating *i* is deleted before a vowel, though not before inducing *i*-umlaut, as shown in (82).

(82)	1PL	dōm^{i}-um	>	dœm-um
	2PL	dōm^{i}-ið	>	dœm-ið
	3PL	dōm^{i}-a	>	dœm-a

Transferring this observation over to the RDem paradigm, we see that this characteristic of floating *i* is relevant for a large subset of RDem forms, as indicated in (83) and by the shaded cells in Table 45.

(83)	F.ACC.SG	þass^{i}-a	>	þessa
	M.DAT.SG	þass^{i}-um	>	þessum
	M.ACC.SG	þann^{i}-a	>	þenna
	M.NOM.PL	þass^{i}-ir	>	þessum
		etc.		

[67] From the *o*-grade of the PIE root **d^{h}ē-*, namely **d^{h}ō-* (cf. Eng. *do*), which was used to form the PGmc causative **dōm-jan* ‘to cause doom/judgment’ (Watkins 2000: 17), where the *j* causes *i*-umlaut.
[68] From the *o*-grade of PIE **per-* ‘lead’ > PGmc causative **fōr-jan* (Watkins 2000: 66).
[69] From PIE **no(:)-mn̥* ‘name’ >> PGmc **na-mōn* (Watkins 2000: 59).
[70] From PIE **leu-* ‘loosen’ > PGmc **leu-san*, **lau-san* (Watkins 2000: 48).

Table 45 Floating *i* deletes before vowels

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	þessi	þessi	þett-a	þess-ar	þess-ir	þessi
ACC	þess-a	þenn-a	þett-a	þess-ar	þess-a	þessi
GEN	þess- + -rar	þess-a	þess-a	þess- + -ra	þess- + -ra	þess- + -ra
DAT	þess- + -ri	þess-um	þess-u	þess-um	þess-um	þess-um

As seen in the morphological templates introduced above, the shaded forms in Table 45 all contain (in their underlying forms) a floating *i* to the immediate left of a vowel, which results in the deletion of floating *i*. This applies without exception to all of the K-internal forms, since their asigmatic reinforcer *-a* is a vowel and always positioned to the right of -KK^{i}- (see (79) and (80) above). The rule is also relevant for any K-final form with a vowel-initial K ending. The K-final forms with *r*-initial endings, then, are not accounted for since *r* is a consonant. This is indicated in Table 45 by lack of shading. The rule is irrelevant for the constant forms, since it is exactly in these cases that *i* does not delete.

3.2.3.2 Floating *i* surfaces word-finally

Floating *i* is overtly realized when it happens to be in word-final position, which can be seen in the 1SG verb form in Table 44, repeated here as (84).

(84) 1SG dōm^{i}-Ø > dœmi

This rule is directly relevant for the constant forms, as shown in (85).

(85) þa-ss^{i} > þe-ssi

This means that floating *i* has a concrete, visible reflex in the RDem paradigm.

Now the behavior of floating *i* is accounted for in four more forms in the RDem paradigm, namely the constant forms, which are indicated by dark shading in Table 46.

Table 46 Floating *i* surfaces word-finally in bolded forms

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	þessi	þessi	þett-a	þess-ar	þess-ir	þessi
ACC	þess-a	þenn-a	þett-a	þess-ar	þess-a	þessi
GEN	þess- + -rar	þess-a	þess-a	þess- + -ra	þess- + -ra	þess- + -ra
DAT	þess- + -ri	þess-um	þess-u	þess-um	þess-um	þess-um

At this point the constant forms still raise the problem whether K is internal or external. Consider the F.NOM.SG / N.NOM/ACC.PL ending *-Ø*. Whether this ending is internal (86a) or external (86b), both options would surface as *þessi*.

(86) (a) þa-Ø-ss^{i} > þessi

(b) þa-ss^{i}-Ø > þessi

This is where the M.NOM.SG form *þessi* becomes crucial, because in the M.NOM.SG we expect the K ending *-r* (see Table 38) rather than the null *-Ø*. This is where the third observation comes into play. The relevant facts about floating *i* are introduced next, in Section 3.2.3.3. I then return to the M.NOM.SG and the issue of internal or external inflection in Section 3.2.3.4.

3.2.3.3 Floating *i* deletes in medial open syllables

According to Table 44 it would seem that floating *i* surfaces before *r*, as seen in (87).

(87) 2/3SG dōm^{i}-r > dœmir

It is not so simple, however. Johnsen (2012) has argued that PN **ī* goes to ON zero in medial open syllables (and that it is preserved in closed syllables). He points out that this rule appears to be synchronically relevant also for ON paradigms such as the one for *dróttin-* 'ruler', with short *i*, as seen in (88).[71]

[71] It is unclear why *dróttin-* behaves as if it derives from a form with **ī*, since the proto-form is more likely **druhtina-* with short **i* (see discussion in Johnsen 2012: 42). It should therefore behave differently with regard to syncope. Consider *r*-class nouns, for instance, where short *i* is not preserved in a closed syllable: NOM/ACC.PL *fœtr* (< **fōtiz*) instead of **fœtir*. We may choose to write *fœtr* as underlying /fœt-^{j}r/, with a 'floating *j*' (the descendant of short **i*) that causes umlaut but otherwise behaves differently from floating *i* (the descendant of long **ī*). For RDem it is perhaps less surprising that it has a floating *i* as opposed to a floating *j*, since the phonological history of the reinforcer *-*si* is quite complex. Indeed, it is certainly the case that *-*si* had a long vowel at one point in its history, before vowel reduction took hold during the grammaticalization process.

(88) *i* > *Ø* in medial open syllable

dróttin- 'ruler' (Johnsen 2012: 42)

(a) no *i*-syncope (closed syllable)

NOM.SG	dróttin-r	>	drót.tinn[72]
ACC.SG	dróttin-Ø	>	drót.tin
GEN.SG	dróttin-s	>	drót.tins

(b) *i*-syncope (medial open syllable)

DAT.SG	dróttin-i	>	drót.t~~i~~.ni	>	dróttni
NOM.PL	dróttin-ar	>	drót.t~~i~~.nar	>	dróttnar
GEN.PL	dróttin-a	>	drót.t~~i~~.na	>	dróttna

Thus, what is happening in (87) is not as simple as floating *i* surfacing before *r*. Instead, to put it more accurately, floating *i* is surfacing in a closed syllable, as in *dœ.mir* or *drót.tin*.

Johnsen's rule accounts for the deletion of floating *i* in the RDem forms with *r*-initial endings (darkly shaded in Table 47), since in these cases floating *i* will be in a medial open syllable and will thus be deleted, as summarized in (89).

(89) RDem with *r*-initial K

i-syncope

F.GEN.SG	þess^{i}-rar	>	þes.s~~i~~.rar	>	þess~~r~~ar	>	þessar
F.DAT.SG	þess^{i}-ri	>	þes.s~~i~~.ri	>	þess~~r~~i	>	þessi
GEN.PL	þess^{i}-ra	>	þes.s~~i~~.ra	>	þess~~r~~a	>	þessa

Table 47 Floating *i* deletes in forms with *r*-initial K

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	þessi	þessi	þett-a	þess-ar	þess-ir	þessi
ACC	þess-a	þenn-a	þett-a	þess-ar	þess-a	þessi
GEN	þess-ar	þess-a	þess-a	þess-a	þess-a	þess-a
DAT	þess-i	þess-um	þess-u	þess-um	þess-um	þess-um

[72] The syllabification of this form has been simplified from *droo.tØ.ti…* (see Sandøy 1994: 236-237 on the special status of the sequence *tt* < **ht*, as in **druhtina-*).

As seen in Table 47, the behavior of floating *i* has now been accounted for in all of the RDem forms.

3.2.3.4 Testing for the position of K in M.NOM.SG *þessi*

Now that we have an understanding of the phonology of floating *i*, especially the fact that it surfaces in closed syllables, we can test for the position of K in the M.NOM.SG form *þessi* (and by assumption the rest of the constant forms). The reason it is possible to test for the position of K in this particular form is because this form should inflect with the consonantal ending *-r*, which raises the possibility of a closed syllable being formed.

Consider first the hypothetical scenario in which K is external, meaning that *-r* would be to the right of the reinforcer *-ss^{i}-*.

(90) (a) þe-ss^{i}-r > þes.sir [closed syllable → no syncope] > *þessir

(b) cf. 2/3SG dœm^{i}-r > dœ.mir

With *-r* at the rightmost edge of the form, floating *i* ends up in a closed syllable, which means that it will be preserved (90a), and this is indeed what we see in the 2/3SG present indicative of weak *i*-stem verbs (90b). For RDem, the end result of the external inflection hypothesis would be the incorrect form **þessir* (90a).[73] Since **þessir* is not the M.NOM.SG form we have attested in the RDem paradigm, the external inflection hypothesis makes a bad prediction.

Now consider the scenario in which K is internal, meaning that *-r* is to the left of the reinforcer.[74]

(91) þe-r-ss^{i} > þesssi > þessi

As seen in (91), putting *-r* word-internally results in the correct, expected form. The form *þe-r-ss^{i}* first goes to *þe-s-ssi* due to *rs* assimilating to *ss*, and then the triplet *sss* is reduced to the geminate *ss*. Floating *i*, of course, surfaces word-finally. We can see an almost

[73] Note, however, the existence of the variants M.NOM.SG *þesser* and F.NOM.SG / N.NOM/ACC.PL *þessor* (Axelsdóttir 2003: 68, Katrín Axelsdóttir p.c.), which look similar to **þessir*. These are clear cases of external inflection and thus can be classified as boxed/K-final forms in my system. I return to them in Chapter 5.

[74] There is a M.NOM.SG form *þersi* attested in, for instance, the *Hauksbók* (Rask 1976: 235, Cleasby & Vigfússon 1874: 481). This form obviously looks very similar to (91), and it might be tempting to claim it as overt evidence for internal inflection. In this case, however, <rs> is merely an orthographic choice for spelling *ss* (see Noreen 1923: 316), so it should be ignored.

identical set of phonological operations at work in a different area of ON morphology, namely verbs inflecting in the passive/middle voice; see (92).[75]

(92) cf. verb forms with passive/middle *-sk* (examples from Barnes 2004: 144)

RULES
rs > *ss*
$C_1C_2C_2 > C_1C_2$ (Barnes 2004: 101)

(a) finn-r-sk > finnssk > finnsk 'is found'

(b) get-r-sk > getssk > getsk <gezk> 'is begotten'

In ON the passive/middle is marked by the ending *-sk*. The 3SG ending, moreover, is *-r* (cf. Table 43). When these two are combined the sequence *rs* is produced, which assimilates to *ss*. There is also a rule of consonant cluster reduction in ON, whereby a sequence $C_1C_2C_2$ is simplified to C_1C_2 (Barnes 2004: 101). Thus the cluster *nnss* is reduced to *nns* in (92a), and *tss* is reduced to *ts* (spelled <z>) in (92b).

To sum up, then, since the observed form is *þessi* and not **þessir*, we conclude that M.NOM.SG *þessi* has internal inflection. I will extend this finding to the other constant forms as well, even though the test in this section cannot be applied to them since they take null inflection.

(93) Internal inflection in the constant forms

þa-	-Ø	-ss^{i}
þa-	-r	-ss^{i}
D	K	R

In the next section some supporting evidence for (93) will be provided from outside NGmc.[76]

3.2.3.5 Comparative evidence from West Germanic for internal inflection in the constant forms

OE, OF, and OS overtly show internal inflection in RDem forms that paradigmatically overlap with the constant forms of ON. The forms in (94) show pronominal inflection (K_D) to the left of the sigmatic reinforcer *-s*.

[75] ~~Though see section 5.4.3.~~

[76] Note that even if the F.NOM.SG and N.NOM/ACC.PL constant forms can be shown to have external rather than internal inflection, then they will simply be considered K-final and be accounted for accordingly.

(94) Internal inflection in F.NOM.SG, M.NOM.SG, and (N.)NOM/ACC.PL

(a) F.NOM.SG

OE *þ-**ēo**-s*
OF/OS *th-**iu**-s*

(b) M.NOM.SG

OE *þ-**e(:)**-s*
OF *th-**i**-s*
OS *unattested

(c) NOM/ACC.PL (all genders)

OE *þ-**ā**-s*

N.NOM/ACC.PL

OS *th-**iu**-s*

Since there is significant paradigmatic overlap between the WGmc forms in (94) and the ON constant forms, the position of the K morphemes in the WGmc forms can be taken as evidence for my hypothesis in (93) that all of the constant forms in ON inflect internally. Indeed, in OS the overlap is very conspicuous: there is a F.NOM.SG and N.NOM/ACC.PL syncretism in *th-iu(-s)*, just as in ON *þe-Ø-ssi*; the only form missing to complete the pattern is M.NOM.SG, which happens to be unattested in OS. Observe that the evidence here from WGmc is of a comparative nature, the idea being that internal inflection in these slots was the situation at some late stage in the development of NWGmc.

3.3 Refining the templates

In this chapter I have identified three separate structures within the ON RDem paradigm. The morphological templates for these structures are summarized in (95).

(95) Three descriptive templates

(i)	þa-	$-ss^i$	-K	[K-final]
(ii)	þa-	$-KK^i$	-a	[K-internal]
(iii)	þa-	-K	$-ss^i$	[constant]

This brings us to an important observation, which is that every single form in the RDem paradigm displays consonant gemination of some kind. Taking my methodology of fine-grained morphological decomposition to its logical end, then, means that we should, as a final refinement, separate out consonant gemination in each template in (95). The result is (96).

(96) Consonant gemination in the templates

(i)	þa-	-s	$-C^i$	-K	[K-final]
(ii)	þa-	-K	$-C^i$	-a	[K-internal]
(iii)	þa-	-K	-s	$-C^i$	[constant]

What we are left with is an exceptionally regular RDem paradigm: each and every form in the paradigm has a template with four distinct slots, where each slot must always be filled by appropriate morphological material.

Another important observation that we are now in a position to make is that the sigmatic reinforcer *-s* and the asigmatic reinforcer *-a* never cooccur within the same RDem form. That is, they are in perfect complementary distribution. Accordingly, following standard practice, we can consider *-s* and *-a* to be two realizations of the same syntactic head.

3.4 Summary

In this chapter I have identified five separate RDem morphemes: *þa-*, K, *-s*, $-C^i$, and *-a*. Though there are five ingredients, two of them are in complementary distribution, namely *-s* and *-a*. Connected to this is the fact that the morphological templates discussed above have only four slots available. Thus we are actually concerned with only *four* main ingredients, which I will label as in (97).

(97) *þa-* D <u>d</u>emonstrative or <u>d</u>eterminer

-K K <u>K</u>ase

-C^{i} Gm <u>g</u>eminator with *i*-<u>m</u>utator

-s/-a R <u>r</u>einforcer

The morphemes in (97) combine in three different kinds of structural configurations within the RDem paradigm. The morphological templates of the K-final (boxed in Table 48), K-internal (darkly shaded in Table 48), and constant (lightly shaded in Table 48) forms are shown below, and the strong adjective endings (K) are repeated for convenience in Table 49.

Table 48 Four-slot templates with K highlighted

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	þa-Ø-s-C^{i} D-K-R-Gm þessi	þa-r-s-C^{i} D-K-R-Gm þessi	þa-t-C^{i}-a D-K-Gm-R þetta	þa-s-C^{i}-ar D-R-Gm-K þessar	þa-s-C^{i}-ir D-R-Gm-K þessir	þa-Ø-s-C^{i} D-K-R-Gm þessi
ACC	þa-s-C^{i}-a D-R-Gm-K þessa	þa-n-C^{i}-a D-K-Gm-R þenna	þa-t-C^{i}-a D-K-Gm-R þetta	þa-s-C^{i}-ar D-R-Gm-K þessar	þa-s-C^{i}-a D-R-Gm-K þessa	þa-Ø-s-C^{i} D-K-R-Gm þessi
GEN	þa-s-C^{i}-rar D-R-Gm-K þessar	þa-s-C^{i}-a D-K-Gm-R þessa	þa-s-C^{i}-a D-K-Gm-R þessa	þa-s-C^{i}-ra D-R-Gm-K þessa	þa-s-C^{i}-ra D-R-Gm-K þessa	þa-s-C^{i}-ra D-R-Gm-K þessa
DAT	þa-s-C^{i}-ri D-R-Gm-K þessi	þa-s-C^{i}-um D-R-Gm-K þessum	þa-s-C^{i}-u D-R-Gm-K þessu	þa-s-C^{i}-um D-R-Gm-K þessum	þa-s-C^{i}-um D-R-Gm-K þessum	þa-s-C^{i}-um D-R-Gm-K þessum

Table 49 Strong adjective endings of *n*-type class

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	-Ø	-r	-t	-ar	-ir	-Ø
ACC	-a	-n	-t	-ar	-a	-Ø
GEN	-rar	-s	-s	-ra	-ra	-ra
DAT	-ri	-um	-u	-um	-um	-um

In the next chapters a more formal approach will be taken to explain these intraparadigmatic patterns.

4
A Cinquean analysis of the Old Norse reinforced demonstrative

In examining and refining Greenberg's (1963) 20th universal, Cinque (2005) takes a very broad crosslinguistic approach, with data drawn from a wide array of descriptive and typological sources. In contrast, in this chapter I will take a very narrow approach in my own U20-style study. Not only will I look at only one single language, but I will look at a single paradigm within that language. Broadly speaking, I hope to demonstrate in this chapter that intraparadigmatic morphological variation should be studied in the same way that crosslinguistic syntactic variation is, and that the two kinds of variation are actually governed by the same basic principles.

The chapter has three main parts. In Section 4.1 I aim to deduce, within the context of Cinque's (2005) theory of syntactic movement, the functional sequence of RDem. In Section 4.2 I discuss the derivations of the three RDem structures identified in Chapter 3 from the single underlying functional sequence. In Section 4.3 I explain why only these three structures/derivations are available. The discussion focuses primarily on ON, but aspects of OE and modern Icelandic will also be relevant to the discussion.

4.1 Deducing the functional sequence of RDem

In the previous chapter it was seen that five morphological ingredients can be isolated in the ON RDem paradigm. For each instantiation of RDem, however, there are really only four morphemes, since two of the morphemes (the reinforcers *-s* and *-a*) are in complementary distribution. In this chapter I will, along conventional lines, take each of the four morphemes to correspond to a particular syntactic head. The four syntactic heads and the labels I have chosen for them are given in (98), where D stands for determiner or definiteness,[77] K for Kase, Gm for geminator with *i*-mutator, and R for reinforcer (though Gm should also be considered a reinforcer of sorts).

(98) Four syntactic heads

(a)	Dem stem/RDem base *þa*-[78]	D
(b)	strong adjective endings[79]	K
(c)	geminator with *i*-mutator -C^i	Gm
(d)	complementary *-s*/*-a*	R

Gm and R should be considered first and foremost easy-to-remember labels, with very little theoretical importance attached to these names. With regard to this, it is important to point out that the head Gm has been identified on the basis of phonology only. Though I do not investigate this issue here, surely Gm also has its own semantic import in the domain of reinforcement. But whatever the semantics of Gm (and R) may be, it can reliably be shown to be a head purely on the basis of the morphophonological decomposition I have performed.

Returning to (98), the first and most important step in my analysis is to determine the underlying functional sequence (fseq) of the syntactic heads postulated in (98). There are 24 possible combinations of D, K, Gm, and R (4! = 4 • 3 • 2 • 1 = 24). Consequently, there are also in principle 24 possible orders in which D, K, Gm, and R can be base-generated, or, put differently, there are 24 possible functional sequences for RDem. These are listed in (99).

---

[77] It is reasonable to assume that demonstratives and definite articles in modern Germanic (usually characterized by a *th*- or a *d*- morpheme; see Leu 2007, 2008, 2015; Déchaine & Wiltschko 2008; Kayne & Pollock 2010) are instantiations of D in the sense of Abney (1987). For Old Germanic this assumption is not as straightforward, as it is well known that Germanic gradually grammaticalized a genuine DP system over time. It is in fact very likely that ON had not yet developed a true definite article. See Lander & Haegeman (2014) for evidence that ON fits in typologically with NP (article-less) languages, rather than DP (article) languages, in the sense of Bošković (2005, 2008, 2009, 2010).

[78] In Chapter 3 we also came across the stem *þei*-, which is not relevant in this chapter since RDem only makes use of the stem *þa*-. A reasonable hypothesis might be that *þa*- and *þei*-should ultimately be decomposed into *þ*-*a*- and *þ*-*ei*-, where *þ*- marks the specificity of D and *-a-*/*-ei-* the pronominal substance, or NP. I will not go into the differences in NP structure that would, on this hypothesis, underlie *-a-* vs. *-ei-*.

[79] Again, though these case (K) endings also contain information about person, number, and gender (Φ), I will abstract away from Φ features here.

(99) 24 possible RDem fseqs

(a)	R Gm K D	=	-s/-a + -C^i + -K + þa-
(b)	R Gm D K	=	-s/-a + -C^i + þa- + -K
(c)	R D Gm K	=	-s/-a + þa- + -C^i + -K
(d)	D R Gm K	=	þa- + -s/-a + -C^i + -K
(e)	Gm R K D	=	-C^i + -s/-a + -K + þa-
(f)	Gm R D K	=	-C^i + -s/-a + þa- + -K
(g)	Gm D R K	=	-C^i + þa- + -s/-a + -K
(h)	D Gm R K	=	þa- + -C^i + -s/-a + -K
(i)	K R Gm D	=	-K + -s/-a + -C^i + þa-
(j)	K R D Gm	=	-K + -s/-a + þa- + -C^i
(k)	K D R Gm	=	-K + þa- + -s/-a + -C^i
(l)	D K R Gm	=	þa- + -K + -s/-a + -C^i
(m)	R K Gm D	=	-s/-a + -K + -C^i + þa-
(n)	R K D Gm	=	-s/-a + -K + þa- + -C^i
(o)	R D K Gm	=	-s/-a + þa- + -K + -C^i
(p)	D R K Gm	=	þa- + -s/-a + -K + -C^i
(q)	Gm K R D	=	-C^i + -K + -s/-a + þa-
(r)	Gm K D R	=	-C^i + -K + þa- + -s/-a
(s)	Gm D K R	=	-C^i + þa- + -K + -s/-a
(t)	D Gm K R	=	þa- + -C^i + -K + -s/-a
(u)	K Gm R D	=	-K + -C^i + -s/-a + þa-
(v)	K Gm D R	=	-K + -C^i + þa- + -s/-a
(w)	K D Gm R	=	-K + þa- + -C^i + -s/-a
(x)	D K Gm R	=	þa- + -K + -C^i + -s/-a

Obviously, assuming that there is one (universal) functional sequence, the combinatorial possibilities in (99) need to be narrowed down, until only one functional sequence remains. This section will introduce various tests to do so. The reader should keep in mind that since the overall theoretical framework adopted in this dissertation is nanosyntax, some of the diagnostics used below are nanosyntactically flavored, even though this chapter is primarily set in a Cinquean framework. Specifically this means that the diagnostics are offered with an eye toward eventually integrating the facts into a system which allows for phrasal spellout and lexical packaging.

4.1.1 The three templates

As discussed in Chapter 3, the five morphemes in (98) combine with each other in three distinct patterns or templates. Each template has four distinct slots, each of which must be filled with morphological material.

(100) Three templates

(i) D-R-Gm-K = þa-s-C^i-K => *þessum*, *þessu*, *þessir*, etc.

(ii) D-K-Gm-R = þa-K-C^i-a => *þenna*, *þetta*, *þessa*

(iii) D-K-R-Gm = þa-K-s-C^i => *þessi*

The fseq must be able to derive all the templates in (100) by U20 rules. The U20 rules from Section 2.1.3 are summarized in (101).

(101) U20 rules (Cinque 2005)

(a) There is a universal base-generated order.
(b) Only leftward movement is allowed (antisymmetry; Kayne 1994).
(c) Only phrasal movement is allowed (i.e. no head movement).
(d) No remnant movement is allowed.

Given these rules, the fseq candidate in Figure 95, for instance, fails to derive template (i), given in (102).

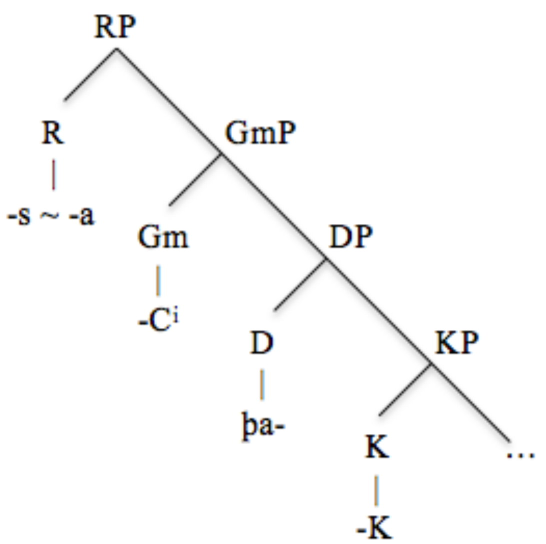


Figure 95 Fseq (b) R Gm D K

(102) D-R-Gm-K
þa-s-C^i-K

To derive the order in (102) from the fseq in Figure 95, we have to raise a constituent consisting of D only all the way to the top of the structure without also implicating KP in this movement. Such an operation, however, is impossible in the U20 system. As seen in Figure 96, the movement of DP without KP would be movement of a non-constituent, which is disallowed.

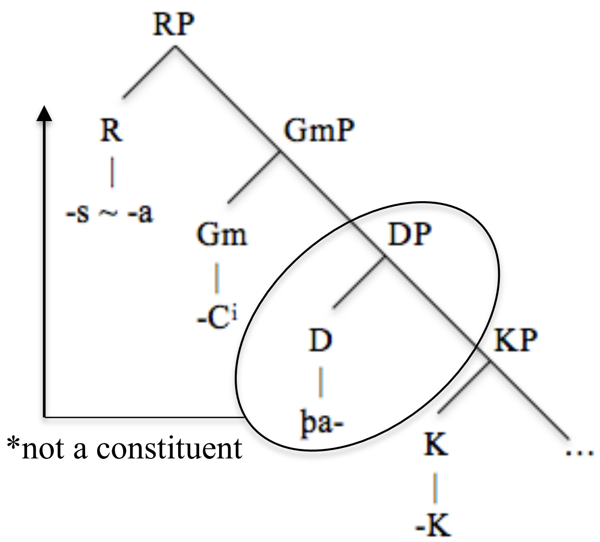


Figure 96 Movement of a non-constituent

Recall that neither head movement nor remnant movement can be invoked to save this type of derivation. In sum, we can conclude that the fseq in Figure 95 (i.e. fseq (b)) is unable to derive template (i), which is a good reason in our search for the correct fseq to eliminate fseq (b) from the running.

4.1.2 Adjacency of D and K

Recall from Chapter 1 and Chapter 3 that in its oldest attested incarnation RDem was formed by adding the suffix *-si* to Dem, as in RN N.ACC.SG **þatsi** (*þat-si*) or M.DAT.SG **þaimsi** (*þœim-si*). In other words, Dem is a standalone form which exists independently of *-si*.

(103) D K *-si*
Dem

As sketched in (103), Dem itself is composed of D plus an inflectional ending K. The reinforcer *-si*, moreover, must in some way correspond to the remaining heads, Gm and R. For the time being, let us assume along more conventional lines (e.g. Kayne 2005) that Gm hosts a null morpheme, since gemination is not observed at the Dem-*si* stage, and that the R head hosts the morpheme *-si*. The order of R and Gm cannot be determined from the Dem-*si* forms alone. For the sake of argument let us place R to the left of Gm. All in all, this yields the template in (104) for the Dem-*si* forms.

(104) D-K-R-Gm
þa-t-si-Ø
þæi-m-si-Ø
etc.

The correct fseq must allow for this basic configuration to be derived. More importantly, it needs to allow for the existence of an independent constituent consisting of D and K only, since this is what a Dem form corresponds to.

Consider an fseq in which D and K are not base-generated next to each other, for example fseq (d), given in Figure 97.

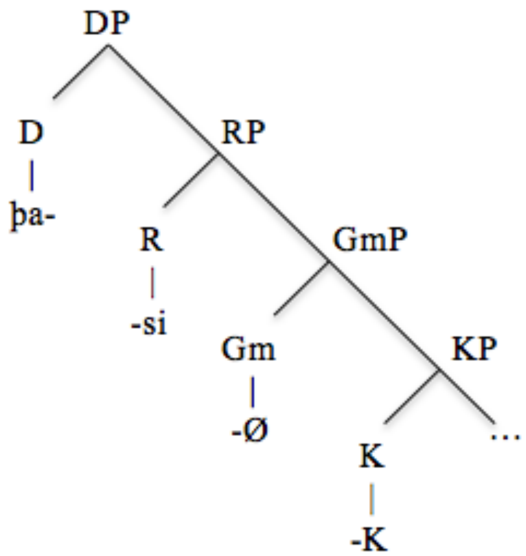


Figure 97 Fseq (d) D R Gm K

Consider now what happens when KP moves to a position right below DP, as illustrated in Figure 98.

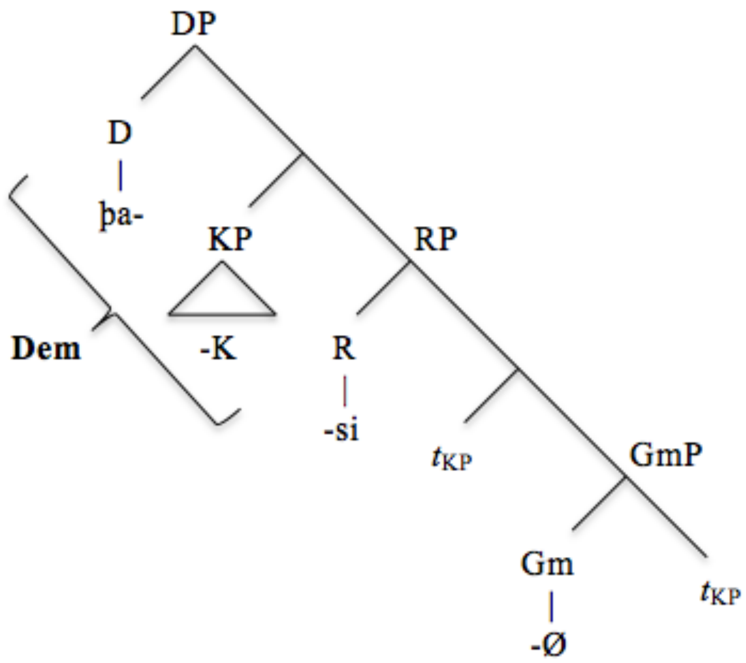


Figure 98 Movement of KP to position immediately below D

The cyclic movement in Figure 98 is perfectly allowed in Cinque's system. Thus we are able to derive the surface order of a Dem-*si* form with the fseq in Figure 97 as our starting point. However, while fseqs like fseq (d) in Figure 97 can derive D-K adjacency via the kind of movement in Figure 98, it is important to note that the resulting adjacency is only superficial, in that D and K never constitute a syntactic object to the exclusion of R and Gm. The only node that contains both D and K in Figures 97 and 98 (i.e. the node labeled DP) contains R and Gm as well. It is simply a structural fact about fseqs which do not base-generate D and K adjacently that there will never be a node which contains D and K only, no matter how many (licit) movements are performed. This, in turn, means that these fseqs cannot treat Dem, which consists of D and K, as a constituent, but must treat it simply in terms of linear adjacency between D and K in the course of a derivation (e.g. as in Figure 98).

Since Dem is a standalone form which exists independently of -*si*, we want to be able to construct a syntactic constituent which is composed of D and K to the exclusion of R and Gm. This requirement is a problem for any sequence in (99) which does not base-generate D and K next to each other, because such fseqs, simply by their nature, cannot accommodate a constituent composed of D and K only.

Integrating an argument derived from nanosyntax, further evidence that a constituent [[DP] KP] is a necessary ingredient comes from non-compositional Dem forms like F.NOM.SG *sú* and M.NOM.SG *sá*. These *s*-initial forms are irregular: if we knew nothing about the history of pronouns in Indo-European and Germanic, we would expect ON to have F.NOM.SG **þa-u* and M.NOM.SG **þa-r* in these slots. In other words, *sú* and *sá* are phonological idioms, and as such, in nanosyntactic terms, they are prime candidates for pointer entries (see Section 2.2.3 for discussion and illustration). See (105).

(105) (a) Regular entry

<$_{210}$ *þa-* ⇔ DP >

<$_{320}$ *-u* ⇔ $\mathrm{K_{F.NOM.SG}P}$ >

<$_{325}$ *-r* ⇔ $\mathrm{K_{M.NOM.SG}P}$ >

(b) Pointer entries

<$_{2386}$ *sū* ⇔ [[210] 320] >

<$_{2387}$ *sā* ⇔ [[210] 325] >

The entries in (105b) are a way of saying that whenever the structures *þa-u* and *þa-r* are built in the syntax, they receive the idiomatic phonologies *sú* and *sá*, respectively, in the same way that **give-d* is replaced by *gave*, **bring-ed* by *brought*, or **we's* by *our* in English.

The availability of the irregular forms *sú* and *sá* and their entries in (105b) show the need for assuming that there is a constituent [[DP] KP], but they also tell us something interesting about Figures 97 and 98. As already mentioned, forms like *þa-t-si* and *þœi-m-si*, which are obviously compositional with regard to D (*þa-*/*þœi-*) and K (*-t*/*-m*), can be derived by movement using an fseq like fseq (d) in Figure 97. However, observe that it is not possible to derive *sú-si* and *sá-si* using fseq (d), even in the 'superficial' movement-based way. This is because there needs to be an appropriate node in the structure at which the pointer entries in (105b) can spell out *sú* and *sá*, but the structure in Figure 98 contains no such node, for the same reason that the fseq in Figure 97 can never have a constituent made up only of D and K. Thus the irregular forms in (105b) are another reason that the correct fseq should base-generate D and K next to each other, since otherwise a constituent [[DP] KP] for the pointer entries cannot arise.

Consider another way in which to maintain an fseq like the one in Figure 97, in which D and K are not merged adjacently. One last-ditch option with regard to the *sú-si/sá-si* problem would be to postulate an idiomatic entry for the entire structure [[[DP] KP] RP GmP]. In the M.NOM.SG, for instance, this would mean that whenever **þa-r-si* is built it will be overridden by the phonology *sási*. This kind of solution is undesirable, to say the least, because it would imply that *sási* is not decomposable, even though it is transparently the case that *sási* is made up of *sá* plus *-si*. As far as I can see, given the empirical facts there is no way to avoid the need for the lexical packaging (see Section 2.2.10) of D and K.

A good (nano)syntactic analysis of RDem should be able to smoothly account for the diachronic transition between the earliest, archaic patterns of RDem, preserved in RN (103), and the later patterns found in ON, OE, OF, OS, and OHG. This is one of the major goals of Chapter 6.

4.1.3 Adjacency of K and Gm

In this section I take a brief excursion to West Germanic, specifically Old English, which provides us with another way of testing the 24 possible base-generated orders of D, K, Gm, and R. The OE RDem paradigm is given in Table 50.

Table 50 OE RDem (Campbell 2003: 291)

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	þ-ēo-s	þ-e(:)-s	þi-s	þ-ā-s	þ-ā-s	þ-ā-s
ACC	þ-ā-s	þi-s-ne	þi-s	þ-ā-s	þ-ā-s	þ-ā-s
GEN	þi-s-re > þisse	þi-s-s-es	þi-s-s-es	þi-s-ra > þissa	þi-s-ra > þissa	þi-s-ra > þissa
DAT	þi-s-re > þisse	þi-s-s-um	þi-s-s-um	þi-s-s-um	þi-s-s-um	þi-s-s-um

In the OE RDem, we observe an alternation between forms with non-gemination of the sigmatic reinforcer (-*s*) (non-shaded) and forms with gemination of the sigmatic reinforcer (-*ss*-) (shaded). The geminated forms can be thought of as having the template in (106).

(106) D-R-Gm-K
þi-s-C-K

Note that the OE template in (106) is equivalent to the ON K-final forms (template (i)).[80]

[80] It may be wondered whether gemination in ON and OE are in fact realizations of the same head. Here I must insist that the burden of proof is not on me to show that Gm in these two closely related languages is the same, but rather on anyone who wants to argue that it is not. The simplest starting point – on the basis of genetic relatedness and clearly observable similarities in the two paradigms – is that Gm is responsible for gemination in both ON and OE, simply by Occam's Razor. Indeed, a major methodological advantage of studying closely related languages is that one can be (more) assured that similar-looking phenomena in two languages can in fact be compared. Now, I will happily admit that it is highly unlikely that RDem in non-Germanic languages *must* lexicalize Gm in the form of gemination, or put the other way around that gemination (if present) *must* be located at Gm. The label Gm, in other words, is simply a relic of my having restricted my focus to Germanic in this dissertation.

In the non-shaded forms with single *-s*, we observe the suppression of gemination. We may wonder why some forms in OE have gemination while others do not. Thinking nanosyntactically, I think the answer to this question should be phrased in terms of lexical packaging (see Section 2.2.10). The idea would be that in those cases where we do not see gemination, the geminator ingredient Gm is lexically stored along with some other ingredient. The only ingredient, moreover, that can track (non-)gemination in the paradigm is K, since D (*þi-*) and R (*-s*) do not vary intraparadigmatically, while K does. Because of this, it is reasonable to suggest that Gm is packaged along with K in the OE non-shaded (non-geminated) forms, while in the shaded (geminated) forms Gm is not packaged with K and thus independently expressed as *-C*.

The packaging hypothesis is illustrated in (107). In (107a), the lexical entry for the M.GEN.SG ending *-es*, for instance, does not package Gm and K together, so Gm is free to be expressed as a separate element, namely the geminator morpheme *-C*, resulting in gemination in a form like M.GEN.SG *þisses*. In (107b), the lexical entry for the M.ACC.SG ending *-ne* does package Gm and K together, so Gm is not free to be expressed as *-C*, resulting in non-gemination in a form like M.ACC.SG *þisne*.

(107) (a) No K/Gm packaging → gemination

K Gm

e.g. *-es* *-C*

=> þi-s-C-es > M.GEN.SG *þisses*

(b) K/Gm packaging → no gemination

K Gm

e.g. *-ne*

=> þi-s-ne > M.ACC.SG *þisne*

An important consequence of this packaging analysis of the distribution of gemination in OE is that our fseq must merge K and Gm adjacently in the underlying sequence. If a given sequence does not merge K and Gm adjacently, then the lexical packaging sketched in (107b), which is needed for OE, cannot be accommodated by the fseq and it should be discarded.

4.1.4 KP constituency (Caha 2009)

Inspired by seminal work by Blake (1994), Caha (2009) has proposed that the case domain (K) be decomposed into multiple additive heads. As discussed already in Section 2.2.4.3, dative case is built on top of genitive case, which in turn is built on accusative, which in turn is built on nominative. This is illustrated in (108).

(108)	DAT =	K_4	K_3	K_2	K_1
	GEN =		K_3	K_2	K_1
	ACC =			K_2	K_1
	NOM =				K_1

Caha's approach captures a number of empirical phenomena: case syncretisms, case-stacking, case-selection of prepositions and in argument structure, among others. If we want to endorse Caha's (2009) nanosyntactic theory of case, then the correct fseq needs to be compatible with a decomposed K domain. Any fseq which turns out to be incompatible with constituents made up of numerous K features will therefore be dismissed. In other words, in addition to deriving all of the templates in Section 4.1.1, the fseq should be compatible with representing numerous K heads as a syntactic object to the exclusion of other, non-K heads.

As an example, consider fseq (u) in Figure 99.

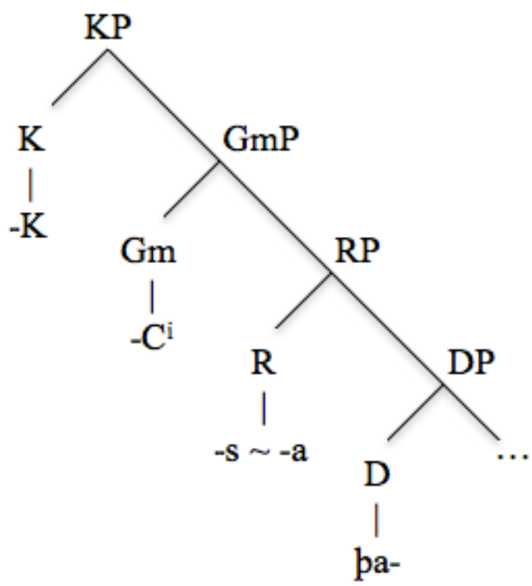


Figure 99 Fseq (u) K Gm R D

Deriving the ON template (ii) (D-K-Gm-R) from the fseq in Figure 99 will require moving DP cyclically all the way up, as shown in Figure 100.

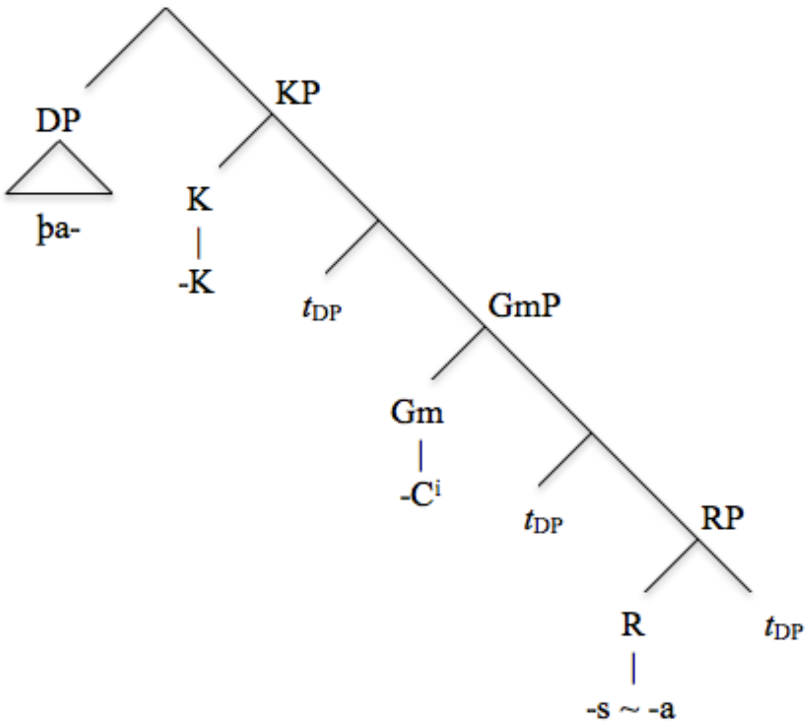


Figure 100 Deriving template (ii) from fseq (u)

While the basic order of template (ii) has been derived in Figure 100, there is a constituency problem for K once we allow this component to represent more than a single head. For example, the M/N.GEN.SG form *þessa* requires three K heads ($K_3 + K_2 + K_1$) for the genitive morpheme *-s*. As seen in Figure 101, once this representation is implemented there is no constituent made up only of K heads to the exclusion of the heads Gm and R.

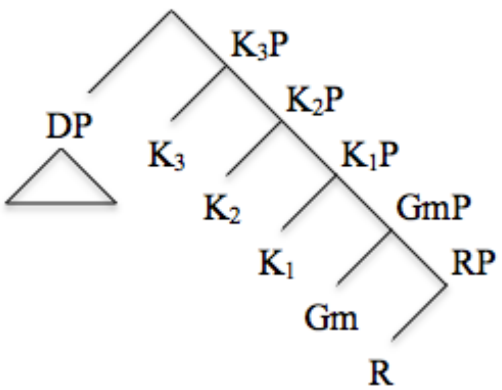


Figure 101 No phrase containing only K heads

In other words, a system with constituent-based spellout cannot target the case ending [K_3P [K_2P [K_1P]]] alone, and thus the fseq candidate in Figure 99 can be said to be incompatible with a decomposed K domain.[81]

If we want to preserve Caha's (2009) findings within a framework of constituent-based spellout, functional sequences like the fseq (u) in Figure 99 will need to be discarded.

4.1.5 Testing the 24 possibilities

There are 24 possible orders of D, K, Gm, and R, and we now have four tests to narrow down this pool of competitors. In (109) I show the results of applying these tests to each of our 24 fseq candidates.

(109) Testing the 24 fseqs

TESTS

The correct fseq must:

- Be able to derive all three ON templates by Cinque's (2005) U20 rules.
  - (i) D-R-Gm-K = **þa-s-C^i-K**
  - (ii) D-K-Gm-R = **þa-K-C^i-a**
  - (iii) D-K-R-Gm = **þa-K-s-C^i**
- Have D and K adjacent so that Dem can be a constituent. (**DK**)
- Have K and Gm adjacent so that gemination can be suppressed in OE. (**KGm**)
- Allow multiple K heads to pattern as a single constituent (Caha 2009). (**K**)

[81] A possible way out of this predicament would be to treat [K_3 K_2 K_1] as a complex head in cases such as these. Recall from Section 2.2.8 that complex heads esesentially spell out *in situ*, that is, without requiring spellout-driven movement from below. In other words, a complex head is not a suffixal structure but rather a prefixal kind of structure, and thus an approach in which [K_3 K_2 K_1] is the case ending would lose us the intuition that K is a suffix in ON and would instead imply that K is a postposition (Michal Starke, p.c.). I will take the traditional stance and assume that we want to think of K in ON as a suffix.

KEY

* = failed test

(subex.) = requires subextraction of DP (possible in Cinque's system but not ideal)

KP constituency test is given after the template.

(a) R Gm K D = -s/-a + -C^{i} + -K + þa-

þa-s-C^{i}-K / K
þa-K-C^{i}-a / K
þa-K-s-C^{i} / K
DK
KGm

(b) R Gm D K = -s/-a + -C^{i} + þa- + -K

*þa-s-C^{i}-K
þa-K-C^{i}-a / K
þa-K-s-C^{i} / K
DK
*KGm

(c) R D Gm K = -s/-a + þa- + -C^{i} + -K

*þa-s-C^{i}-K
þa-K-C^{i}-a / K
*þa-K-s-C^{i}
*DK
KGm

(d) D R Gm K = þa- + -s/-a + -C^{i} + -K

þa-s-C^{i}-K / K
þa-K-C^{i}-a / K
þa-K-s-C^{i} / K
*DK
KGm

(e) Gm R K D = $-C^i$ + -s/-a + -K + þa-

*þa-s-C^i-K
þa-K-C^i-a / K
þa-K-s-C^i / K
DK
*KGm

(f) Gm R D K = $-C^i$ + -s/-a + þa- + -K

*þa-s-C^i-K
þa-K-C^i-a / K
þa-K-s-C^i / K
DK
*KGm

(g) Gm D R K = $-C^i$ + þa- + -s/-a + -K

*þa-s-C^i-K
*þa-K-C^i-a
þa-K-s-C^i / K
*DK
*KGm

(h) D Gm R K = þa- + $-C^i$ + -s/-a + -K

*þa-s-C^i-K
þa-K-C^i-a / K
þa-K-s-C^i / K
*DK
*KGm

(i) K R Gm D = -K + -s/-a + $-C^i$ + þa-

þa-s-C^i-K / K
?? þa-K-C^i-a (subex.) / *K
þa-K-s-C^i / *K
*DK
*KGm

(j) K R D Gm = -K + -s/-a + þa- + -C^i

*þa-s-C^i-K
*þa-K-C^i-a
*þa-K-s-C^i
*DK
*KGm

(k) K D R Gm = -K + þa- + -s/-a + -C^i

þa-s-C^i-K / K
*þa-K-C^i-a
*þa-K-s-C^i
DK
*KGm

(l) D K R Gm = þa- + -K + -s/-a + -C^i

þa-s-C^i-K / K
þa-K-C^i-a / *K
þa-K-s-C^i / *K
DK
*KGm

(m) R K Gm D = -s/-a + -K + -C^i + þa-

?? þa-s-C^i-K (subex.) / K
þa-K-C^i-a / *K
*þa-K-s-C^i
*DK
KGm

(n) R K D Gm = -s/-a + -K + þa- + -C^i

*þa-s-C^i-K
*þa-K-C^i-a
*þa-K-s-C^i
DK
*KGm

(o) R D K Gm = -s/-a + þa- + -K + -C^i

*þa-s-C^i-K

þa-K-C^i-a / *K

*þa-K-s-C^i

DK

KGm

(p) D R K Gm = þa- + -s/-a + -K + -C^i

þa-s-C^i-K / K

þa-K-C^i-a / *K

*þa-K-s-C^i

*DK

KGm

(q) Gm K R D = -C^i + -K + -s/-a + þa-

þa-s-C^i-K / K

*þa-K-C^i-a

þa-K-s-C^i / *K

*DK

KGm

(r) Gm K D R = -C^i + -K + þa- + -s/-a

þa-s-C^i-K / K

*þa-K-C^i-a

*þa-K-s-C^i

DK

KGm

(s) Gm D K R = -C^i + þa- + -K + -s/-a

þa-s-C^i-K / *K

*þa-K-C^i-a

þa-K-s-C^i / *K

DK

*KGm

(t) D Gm K R = þa- + -C^i + -K + -s/-a

þa-s-C^i-K / K
*þa-K-C^i-a
þa-K-s-C^i / *K
*DK
KGm

(u) K Gm R D = -K + -C^i + -s/-a + þa-

þa-s-C^i-K / K
þa-K-C^i-a / *K
?? þa-K-s-C^i (subex.) / *K
*DK
KGm

(v) K Gm D R = -K + -C^i + þa- + -s/-a

þa-s-C^i-K / K
*þa-K-C^i-a
*þa-K-s-C^i
*DK
KGm

(w) K D Gm R = -K + þa- + -C^i + -s/-a

þa-s-C^i-K / K
*þa-K-C^i-a
*þa-K-s-C^i
DK
*KGm

(x) D K Gm R = þa- + -K + -C^i + -s/-a

þa-s-C^i-K / K
þa-K-C^i-a / *K
þa-K-s-C^i / *K
DK
KGm

Failure to package (i.e. the tests DK and KGm) or failure to derive a template are considered fatal flaws. Twenty-two of the orders, (109b-w), possess such shortcomings and should be disposed of completely.

Among the 24 options in (109), only a single fseq passes all of the tests, namely fseq (a) R Gm K D. Interestingly, the reverse order of fseq (a), fseq (x) D K Gm R, fares next best out of all the orders. Fseq (x) runs into problems with the KP constituency test in two of the templates, namely template (ii) and template (iii). Observe that this problem, however, could be solved by allowing K to be a complex head (see fn. 81). This would make fseq (x) an (extended) survivor of the battery of tests in (109). However, as also mentioned in fn. 81, this option would imply a postposition-style analysis of ON case in at least these two templates. This is a complication which is not encountered with fseq (a), which remains the simplest and most efficient solution.

4.1.6 Discussion

Regarding these results, it is important to note that the R/Gm domain here is word-internal and therefore perhaps not quite equivalent to the adverbial locative reinforcers of Bernstein (1997), since Bernstein's reinforcers are 'external' in the sense that they appear in phrases composed of an already fully formed demonstrative and its modified noun (e.g. Italian *questo libro* ***qui*** 'this book here'). Instead, my R/Gm domain resembles, broadly, proposals by Kayne (2005) and Leu (2007, 2008, 2015). They propose that a silent locative HERE combines with a determiner to give a proximal demonstrative. Their HERE, crucially, is part of the internal structure of the demonstrative and explicitly said to be distinct from the external locative reinforcer *here*. Note here, however, that RDem is a reinforced, not necessarily a proximal, demonstrative. More work is needed on the connection between reinforcers and the various functions of demonstratives, especially the spatial-deictic/exophoric function (Lyons 1977; Fillmore 1971, 1982, 1997; Levinson 1983; Himmelmann 1997, 2001; Diessel 1999; Lander & Haegeman 2015). See Section 7.2.2 and Appendix V for some more discussion.

It may be noted that the R/Gm domain is located above the K domain, perhaps surprisingly since K is often considered the highest functional layer in the nominal extension.[82] Overt evidence for the idea that there is a reinforcer domain above K can be seen in Latin and Romance. The Italian demonstratives *questo* and *quello* both derive from a structure with a prefixed reinforcer in Latin (*eccu* + *istum* and *eccu* + *illum*, respectively). Similarly for Spanish, the distal *aquel* derives from Latin *accu* (a variant of *eccu*) + *ille* (Adams 2013: 465-466, 469). These items can be considered to display the

[82] Thanks to Thomas McFadden and Marcel Den Dikken for bringing this issue up.

order R/Gm + D-K, that is, the reinforcer component *eccu*/*accu* plus the inflected demonstrative pronoun D-K. In other words, they represent the crosslinguistic possibility of the reinforcer surfacing in the highest position (where it is base-generated according to my results), above K and D. The reinforcer surfaces in the highest position even in some Germanic languages: consider for instance Afrikaans *hier-die* or Yiddish *ot(-o) der* (Roehrs 2010: 226-227, 243).

Finally we may consider what the winning fseq tells us about the semantics of reinforcement. Observe first that RDem seems to have developed something akin to a proximal (spatial-deictic) meaning in many of the modern Germanic languages (e.g. Swedish *den*, *det* 'that' [< Dem] vs. *denna*, *detta* 'this' [< RDem]). This, crucially, was not the case in (Proto-)NWGmc, where Dem was a neutral demonstrative with the reading 'this' *or* 'that', which could then be reinforced with **-si*. Interestingly, modern Swiss German appears to retain this older semantics of reinforcement. In Swiss German, the defnite Dem item F.SG *dɛ*, M.SG *diä*, N.SG *das* can mean either 'this' or 'that', while the RDem item F.SG *disi*, M.SG *disä*, N.SG *dises* has a more complex reading, i.e. discourse-salient, contrastive 'the other' (Leu 2015: 24–25). With this in mind, it seems reasonable to posit that R and Gm together contribute contrastiveness of some kind when they are added, quite literally, on top of the Dem ingredients D and K.

4.2 Derivations

We are now equipped to more fully understand the internal syntax of RDem. Cinque (2005) and the U20 program will be my foundation in this section. The analysis presented here constitutes a crucial first step towards the full nanosyntactic analysis provided in the next chapter.

4.2.1 Cyclic vs. non-cyclic

Using fseq (a), which is argued for in Section 4.1 and repeated for convenience in (110), we can derive the three ON templates using U20 rules, as shown in (111). I will refer to the derivation of the K-final forms as *fully cyclic*, to the derivation of the constant forms as *partially cyclic*, and to the derivation of the K-internal forms as *roll-up*. These names will become clearer when we look more closely at the derivations.

(110) Fseq: R > Gm > K > D

[$_{RP}$ *-s* / *-a* [$_{GmP}$ *-C*^{i} [$_{KP}$ -K [$_{DP}$ *þa-*]]]]

(111) *U20 derivation* *Template*

DP RP t_{DP} GmP t_{DP} KP t_{DP}	fully cyclic	(i)	þa-s-C^{i}-K	[K-final]
[DP KP t_{DP}]$_{i}$ RP t_{i} GmP t_{i}	partially cyclic	(iii)	þa-K-s-C^{i}	[constant]
[[DP KP t_{DP}]$_{i}$ GmP t_{i}]$_{j}$ RP t_{j}	roll-up	(ii)	þa-K-C^{i}-a	[K-internal]

In the fully cyclic derivation in Figure 102, DP moves first to the left of KP. Next DP raises to the left of GmP, without pied-piping anything. Finally DP raises to the left of RP, again without any pied-piping.[83]

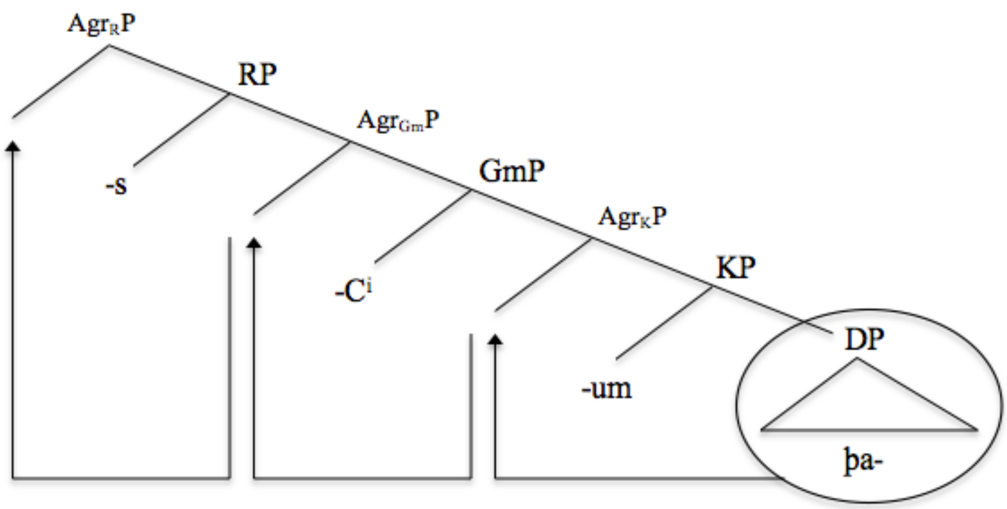


Figure 102 Fully cyclic derivation: M.DAT.SG þa-s-C^{i}-um > *þessum*

Figure 102 shows the fully cyclic derivation: it involves no pied-piping of any material outside of the base [$_{DP}$ *þa-*], which moves by itself all the way up to Agr$_{R}$P.

In the partially cyclic derivation in Figure 103, DP raises first to the left of KP. In the next step, DP and KP (i.e. [DP KP t_{DP}]) raise together to the left of GmP, in a roll-up step. Finally [DP KP t_{DP}] moves to the left of RP, without pied-piping GmP.

[83] In Figures 102-104 I follow Cinque in labeling landing sites as AgrPs, but nothing hinges on this. In the next chapter I will discuss these movements and landing sites in more detail.

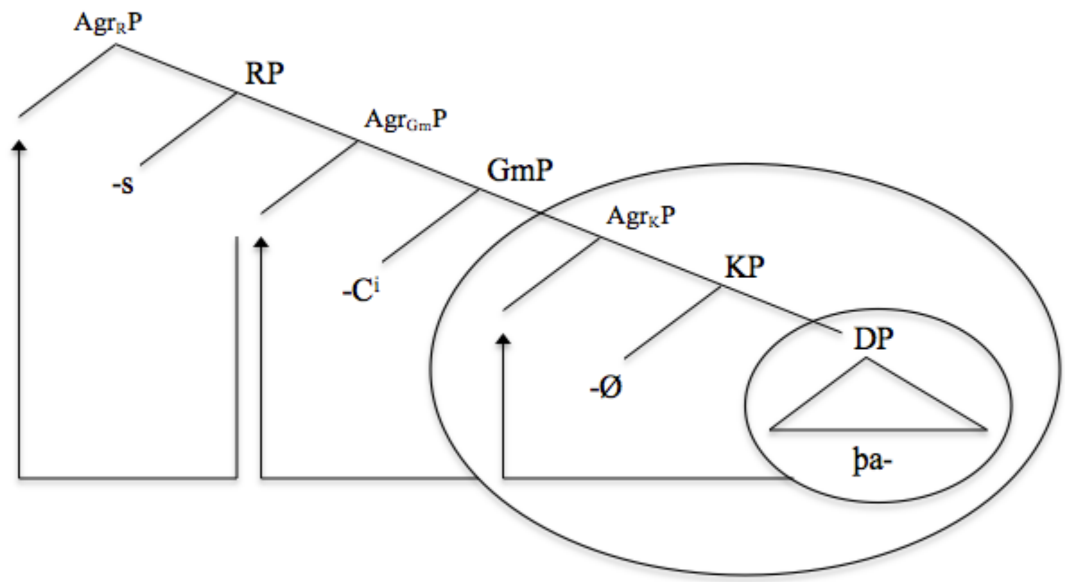


Figure 103 Partially cyclic derivation: F.NOM.SG / N.NOM/ACC.PL þa-Ø-s-C[i] > *þessi*

Figure 103 shows the partially cyclic derivation: the first part of the derivation involves pied-piping of KP along with DP, but after that step there is no pied-piping of any additional material. Figure 103 is therefore a mix of roll-up and cyclic.

In the roll-up derivation in Figure 104, DP first raises around KP. In the next step, the constituent [DP KP t_{DP}] moves to the left of GmP. Finally the constituent [[DP KP t_{DP}]$_i$ GmP t_i] raises around RP.

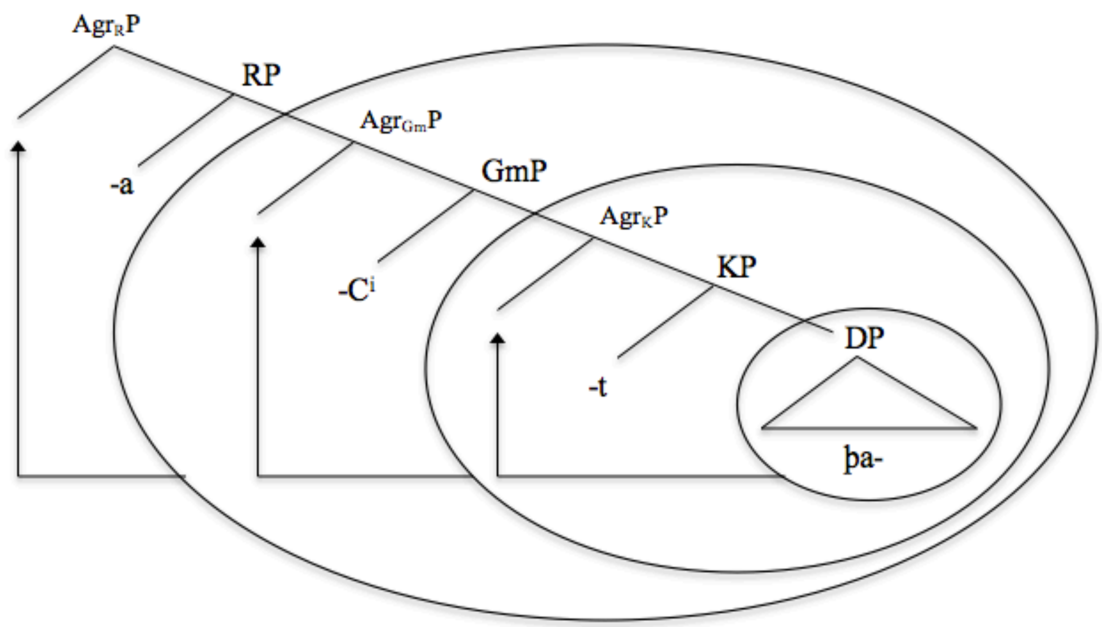


Figure 104 Roll-up derivation: N.NOM/ACC.SG þa-t-C[i]-a > *þetta*

Figure 104 is roll-up derivation because there is pied-piping at every step: DP first pied-pipes KP, then DP (along with KP) pied-pipes GmP. The base-generated order ends up being completely reversed by the end of the derivation.

Cinque (2005: 321) claims that roll-up and cyclic movements are both relatively unmarked.[84] Furthermore, total movement of NP to the very top of the structure is unmarked compared to partial movement of NP, where NP stops in some intermediate position. In all three ON templates, DP (whose internal structure may very well include NP) undergoes total movement. Thus the ON RDem paradigm seems to be highly unmarked according to Cinque's (2005) findings.

Let us return at this point to the realization of the syntactic head R in the ON RDem paradigm. Recall that two instantiations of R have been identified: the sigmatic reinforcer *-s* and the asigmatic reinforcer *-a*. In the fully cyclic derivation in Figure 102 and in the partially cyclic derivation in Figure 103, R is lexicalized as *-s*, while in the roll-up derivation in Figure 104 it is lexicalized as *-a*. Based on the *-s* ~ *-a* allomorphy, I would like to propose that we distinguish only two main *types* of derivations in the RDem paradigm: cyclic and non-cylic/roll-up. A cyclic-type derivation means that R is realized as *-s*, while a roll-up derivation means that R is realized as *-a*. This is summarized in Table 51.

Table 51 Different derivations in RDem

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	þessi	þessi	þetta	þessar	þessir	þessi
ACC	þessa	þenna	þetta	þessar	þessa	þessi
GEN	þessar	þessa	þessa	þessa	þessa	þessa
DAT	þessi	þessum	þessu	þessum	þessum	þessum

KEY

shaded = cyclic-type derivation R => *-a*
non-shaded = roll-up derivation R => *-s*

[84] Furthermore according to Cinque (2005: 321), cyclic movement of NP alone is slightly more marked than roll-up movement of the [[NP] XP] type, but still less marked than movement of the [XP [NP]] type, which we do not observe in ON. We also do not observe the 'no movement' option, which Cinque claims is just as unmarked as roll-up. Presumably we do not observe this unmarked option because this would not put suffixes in the right position (see Section 4.3).

Interestingly, we have evidence here for derivation-sensitive allomorphy, which can be taken as clear morphological support for Cinque's (2005) derivational theory. This is an important result of the investigation so far.

4.2.2 Converging evidence from modern Icelandic for the two derivation types

In this section I will provide some supporting evidence from modern Icelandic for the proposed split in derivational type within the RDem paradigm. The Icelandic RDem paradigm is given in Table 52.

Table 52 Modern Icelandic RDem

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	þessi	þessi	þetta(-ð)	þessar	þessir	þessi
ACC	þessa	þenna-n	þetta(-ð)	þessar	þessa	þessi
GEN	þessarar	þessa(%-s)	þessa(%-s)	þessara	þessara	þessara
DAT	þessari	þessum	þessu	þessum	þessum	þessum

Most of the Icelandic forms in Table 52 are identical to the ON forms. There are, however, two main differences between the Icelandic and the ON paradigms, only one of which is interesting for our purposes. The first, less interesting difference is that the Icelandic K-final forms with *r*-initial endings (underlined in Table 52) display a new stem, *þessa-*, which allows *r* to surface (see Appendix II). Tentatively I propose that this stem, being the result of an analogical change in post-classical Norse, should be analyzed as an irregularity, and I do not have much more to say about it at this point.

The second, more interesting difference is that the Icelandic roll-up forms (shaded in Table 52) – and only the roll-up forms – show (the possibility of) an extra K marker in modern Icelandic. As seen in (112), the M.ACC.SG K ending *-n* is found both word-internally and at the end of *þennan*, the N.NOM/ACC.SG K ending *-t* is doubled by word-final *-ð* in *þettað*,[85] and the M/N.GEN.SG K ending *-s* is found both word-internally and at the end of *þessas* (a form produced by children and adolescents).

[85] It will be noticed that the external marker *-ð* is different from internal *-t*. This is because the marker *-ð* is taken from the final *-ð* of the N.NOM/ACC.SG Dem form/expletive *það* (cf. also the determiner *hið*), which historically comes from ON *þat* (*t* > *d* > *ð* in weakly stressed words in medieval Scandinavian; Haugen 1982: 64). The N.NOM/ACC.SG adjective ending remains *-t* in Icelandic, not *-ð*.

(112) (a) M.ACC.SG þa-**n**-C^i-a-**n** > *þennan*

(b) N.NOM/ACC.SG þa-**t**-C^i-a-**ð** > *þettað*

(c) M/N.GEN.SG þa-**s**-C^i-a-**s** > *þessas*

This means that the non-cyclic/roll-up forms are, as a class, undergoing a change in modern Icelandic which is not available to the cyclic forms, which remain more or less the same as in ON. The two derivational types, then, are shown to have further morphological ramifications in modern Icelandic.[86]

4.3 Why there are only three templates

Recall that in (99) we listed the 24 mathematically possible base-generated orders for RDem. Now that the appropriate underlying functional sequence has been determined, there are still 24 mathematically possible structures that can potentially be derived from this sequence. These are given in (113). Note that these are the same 24 orders as in (99); this time, however, I mean them to be possible structures arising from the single underlying functional sequence R Gm K D.

[86] As pointed out to me by Maria Polinsky, it is perhaps not the case that the roll-up forms are changing, but rather that the cyclic forms are not. In nanosyntax (see also Jayaseelan 2010), roll-up is a hallmark of morphological inflection, while cyclicity is a more syntactic operation. Therefore it is conceivable that the roll-up forms are morphologically developing along a natural route that is unavailable to the cyclic forms precisely because they are cyclic.

(113) 24 possible RDem structures (where base order is R > Gm > K > D)

(a)	R-Gm-K-D	=	-s/-a + -C^i + -K + þa-	
(b)	R-Gm-D-K	=	-s/-a + -C^i + þa- + -K	
(c)	R-D-Gm-K	=	-s/-a + þa- + -C^i + -K	
(d)	**D-R-Gm-K**	**=**	**þa- + -s/-a + -C^i + -K**	**[K-final]**
(e)	Gm-R-K-D	=	-C^i + -s/-a + -K + þa-	
(f)	Gm-R-D-K	=	-C^i + -s/-a + þa- + -K	
(g)	Gm-D-R-K	=	-C^i + þa- + -s/-a + -K	
(h)	D-Gm-R-K	=	þa- + -C^i + -s/-a + -K	
(i)	K-R-Gm-D	=	-K + -s/-a + -C^i + þa-	
(j)	K-R-D-Gm	=	-K + -s/-a + þa- + -C^i	
(k)	K-D-R-Gm	=	-K + þa- + -s/-a + -C^i	
(l)	**D-K-R-Gm**	**=**	**þa- + -K + -s/-a + -C^i**	**[constant]**
(m)	R-K-Gm-D	=	-s/-a + -K + -C^i + þa-	
(n)	R-K-D-Gm	=	-s/-a + -K + þa- + -C^i	
(o)	R-D-K-Gm	=	-s/-a + þa- + -K + -C^i	
(p)	D-R-K-Gm	=	þa- + -s/-a + -K + -C^i	
(q)	Gm-K-R-D	=	-C^i + -K + -s/-a + þa-	
(r)	Gm-K-D-R	=	-C^i + -K + þa- + -s/-a	
(s)	Gm-D-K-R	=	-C^i + þa- + -K + -s/-a	
(t)	D-Gm-K-R	=	þa- + -C^i + -K + -s/-a	
(u)	K-Gm-R-D	=	-K + -C^i + -s/-a + þa-	
(v)	K-Gm-D-R	=	-K + -C^i + þa- + -s/-a	
(w)	K-D-Gm-R	=	-K + þa- + -C^i + -s/-a	
(x)	**D-K-Gm-R**	**=**	**þa- + -K + -C^i + -s/-a**	**[K-internal]**

At first, in Chapter 3, it might have been surprising to discover that there were as many as three structures within a single paradigm, but from the perspective of (113), it now becomes surprising that there are not more than three structures in the paradigm. The question to ask now is why only three of these structures, namely (d), (l), and (x), are attested in the paradigm. In other words, why are 21 of the 24 structures in (113) ruled out?

We can dramatically narrow down the possible structures in (113) by introducing three constraints. The first constraint is syntactic, and it states that a structure must be derivable by the U20 rules given in (101) above. Ten structures are derivationally impossible given U20 rules. That is, illicit operations such as head movement or remnant movement would be required in order to derive these orders from the base order R Gm K D. Structure (p), furthermore, requires subextraction, which as discussed has an ambiguous status. For now let us place structure (p) with the U20-incompatible structures. In total, then, we can rule out 11 structures due to incompatibility with Cinque's (2005) derivational rules. These are listed in (114).

(114) Syntactic constraint: Must be U20-derivable

(e) *Gm-R-K-D = -C^i + -s/-a + -K + þa-
(f) *Gm-R-D-K = -C^i + -s/-a + þa- + -K
(g) *Gm-D-R-K = -C^i + þa- + -s/-a + -K
(h) *D-Gm-R-K = þa- + -C^i + -s/-a + -K
(i) *K-R-Gm-D = -K + -s/-a + -C^i + þa-
(j) *K-R-D-Gm = -K + -s/-a + þa- + -C^i
(m) *R-K-Gm-D = -s/-a + -K + -C^i + þa-
(p) ?? D-R-K-Gm = þa- + -s/-a + -K + -C^i
(q) *Gm-K-R-D = -C^i + -K + -s/-a + þa-
(u) *K-Gm-R-D = -K + -C^i + -s/-a + þa-
(v) *K-Gm-D-R = -K + -C^i + þa- + -s/-a

For instance, if one were to try to derive structure (h) (D-Gm-R-K) from the fseq R Gm K D, one would run into various issues. First of all, DP would need to raise to the left of GmP, since structure (h) involves the ordering D-Gm. This is roughly sketched in Figure 105.

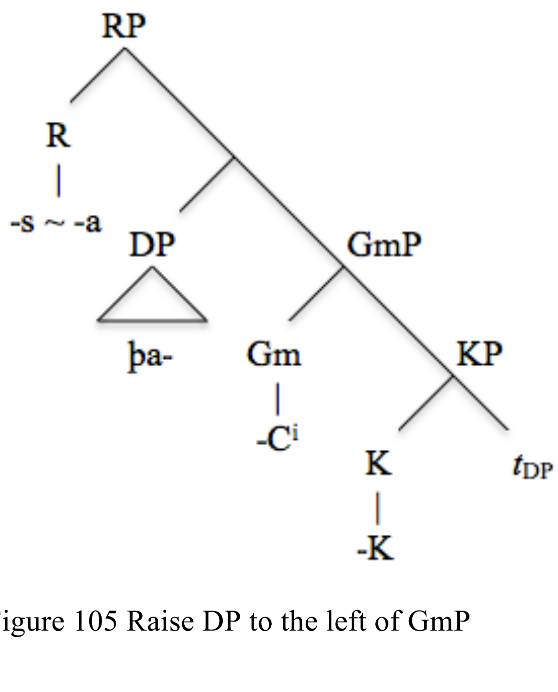


Figure 105 Raise DP to the left of GmP

Next both D and Gm need to reach a position to the left of R, considering the ordering D-Gm-R in structure (h). Crucially, however, D and Gm need to get to the left of R *without* K, but as seen in Figure 106 there is simply no constituent consisting of D and Gm to the exclusion of K.

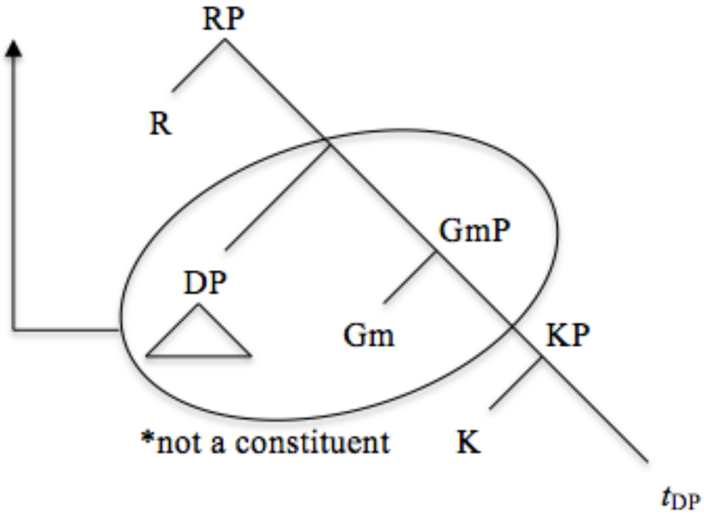


Figure 106 No constituent containing D and Gm only

Thus we can say that structure (h) is not derivable from our fseq.

There are now 13 structures left over, all of which are derivable by U20 rules. These are listed in (115).

(115)	(a)	R-Gm-K-D	=	-s/-a + -C^i + -K + þa-
	(b)	R-Gm-D-K	=	-s/-a + -C^i + þa- + -K
	(c)	R-D-Gm-K	=	-s/-a + þa- + -C^i + -K
	(d)	D-R-Gm-K	=	þa- + -s/-a + -C^i + -K
	(k)	K-D-R-Gm	=	-K + þa- + -s/-a + -C^i
	(l)	D-K-R-Gm	=	þa- + -K + -s/-a + -C^i
	(n)	R-K-D-Gm	=	-s/-a + -K + þa- + -C^i
	(o)	R-D-K-Gm	=	-s/-a + þa- + -K + -C^i
	(r)	Gm-K-D-R	=	-C^i + -K + þa- + -s/-a
	(s)	Gm-D-K-R	=	-C^i + þa- + -K + -s/-a
	(t)	D-Gm-K-R	=	þa- + -C^i + -K + -s/-a
	(w)	K-D-Gm-R	=	-K + þa- + -C^i + -s/-a
	(x)	D-K-Gm-R	=	þa- + -K + -C^i + -s/-a

The second constraint stems from the fact that we are applying Cinque's theory, which is basically developed for syntax, to a smaller domain, usually referred to as morphology. In morphology a distinction has to be made between different kinds of affixes, based on where in a word these inflectional pieces attach: prefixes attach to the left, suffixes to the right, and infixes in the middle. For RDem, we can make the basic observation that, except for the base *þa-*, we are dealing with morphological ingredients which are suffixes. Suffixes, moreover, need to occur to the right of something, which is very often not the case in (115). For instance, structure (r) (Gm-K-D-R) places the suffix -*C*^{i} word-initially, meaning it is not right-adjoined to anything. In total, nine structures fail to fulfill this requirement, and thus we can eliminate these morphologically improper structures, as seen in (116). In all of the potential structures in (116), there is a word-initial suffix.[87]

(116) Morphological constraint: Word-initial suffixes

(a) *R-Gm-K-D = -s/-a + -C^{i} + -K + þa-
(b) *R-Gm-D-K = -s/-a + -C^{i} + þa- + -K
(c) *R-D-Gm-K = -s/-a + þa- + -C^{i} + -K
(k) *K-D-R-Gm = -K + þa- + -s/-a + -C^{i}
(n) *R-K-D-Gm = -s/-a + -K + þa- + -C^{i}
(o) *R-D-K-Gm = -s/-a + þa- + -K + -C^{i}
(r) *Gm-K-D-R = -C^{i} + -K + þa- + -s/-a
(s) *Gm-D-K-R = -C^{i} + þa- + -K + -s/-a
(w) *K-D-Gm-R = -K + þa- + -C^{i} + -s/-a

Put quite simply, RDem structures must begin with the (non-suffix) base *þa-*.

At this point, the potential structures in (113) have been narrowed down to four, namely structures (d), (l), (t), and (x). These remaining structures, shown in (117), are not ruled out by an impossible U20-style derivation or by the morphological ban on word-initial suffixes.

(117) Four remaining structures

(d) D-R-Gm-K = þa- + -s/-a + -C^{i} + -K
(l) D-K-R-Gm = þa- + -K + -s/-a + -C^{i}
(t) D-Gm-K-R = þa- + -C^{i} + -K + -s/-a
(x) D-K-Gm-R = þa- + -K + -C^{i} + -s/-a

At this point we can observe that structure (t) is phonologically problematic, since a vowel cannot be geminated. In other words, the third constraint is a phonological one requiring that the consonant geminator be immediately preceded by a consonant. If this is not the case, the structure will crash at PF. This is indicated in (118).

---

[87] I assume that the property of being a suffix is encoded in the structure of a lexical entry and that postsyntactic "readjustment rules" (see Halle and Marantz 1993) are not available.

(118) Phonological constraint: Consonant must be input to consonant gemination

(t) *D-Gm-K-R = þa- + -C^i + -K + -s/-a > **crash at PF**

This phonological constraint is the last piece of the puzzle. Once structure (t) is filtered out, we are left with only structures (d), (l), and (x), as seen in (119).

(119) (d) D-R-Gm-K = þa- + -s/-a + -C^i + -K > *þe-s-s-ar, þe-s-s-um, þe-s-s-u*, etc.
(l) D-K-R-Gm = þa- + -K + -s/-a + -C^i > *þe-r-s-*s^i*, þe-Ø-s-*s^i > *þessi*
(x) D-K-Gm-R = þa- + -K + -C^i + -s/-a > *þe-t-t-a*, *þe-n-n-a*, *þe-s-s-a*

To summarize, in (120) I have shown how each potential structure fares on each constraint.

(120) Testing the 24 possible RDem structures (where base order is R > Gm > K > D)

<u>CONSTRAINTS</u>

An admissible structure must:

- Be derivable by Cinque's (2005) U20 rules. (**U20**)
- Have right-adjoined suffixes. (**suffixes**)
- Have a consonant to the immediate left of the consonant geminator. (**<u>C</u>-C^i**)

<u>KEY</u>

* = failed test

(subex.) = requires subextraction of DP (possible in Cinque's system but not ideal)

(a) R-Gm-K-D = -s/-a + -C^i + -K + þa-

U20
*suffixes
<u>C</u>-C^i

(b) R-Gm-D-K = -s/-a + -C^i + þa- + -K

U20
*suffixes
<u>C</u>-C^i

(c) R-D-Gm-K = -s/-a + þa- + -C^i + -K

U20

*suffixes

*C-C^i

(d) D-R-Gm-K = þa- + -s/-a + -C^i + -K

U20

suffixes

C-C^i

(e) Gm-R-K-D = -C^i + -s/-a + -K + þa-

*U20

*suffixes

*C-C^i

(f) Gm-R-D-K = -C^i + -s/-a + þa- + -K

*U20

*suffixes

*C-C^i

(g) Gm-D-R-K = -C^i + þa- + -s/-a + -K

*U20

*suffixes

*C-C^i

(h) D-Gm-R-K = þa- + -C^i + -s/-a + -K

*U20

suffixes

*C-C^i

(i) K-R-Gm-D = -K + -s/-a + -C^i + þa-

*U20
*suffixes
*$\underline{C}$-C^i

(j) K-R-D-Gm = -K + -s/-a + þa- + -C^i

*U20
*suffixes
*$\underline{C}$-C^i

(k) K-D-R-Gm = -K + þa- + -s/-a + -C^i

U20
*suffixes
$\underline{C}$-C^i

(l) D-K-R-Gm = þa- + -K + -s/-a + -C^i

U20
suffixes
$\underline{C}$-C^i

(m) R-K-Gm-D = -s/-a + -K + -C^i + þa-

*U20
*suffixes
$\underline{C}$-C^i

(n) R-K-D-Gm = -s/-a + -K + þa- + -C^i

U20
*suffixes
*$\underline{C}$-C^i

(o) R-D-K-Gm = -s/-a + þa- + -K + -C^i

U20
*suffixes
$\underline{C}$-C^i

(p) D-R-K-Gm = þa- + -s/-a + -K + -C^i

?? U20 (subex.)
suffixes
C-C^i

(q) Gm-K-R-D = -C^i + -K + -s/-a + þa-

*U20
*suffixes
*C-C^i

(r) Gm-K-D-R = -C^i + -K + þa- + -s/-a

U20
*suffixes
*C-C^i

(s) Gm-D-K-R = -C^i + þa- + -K + -s/-a

U20
*suffixes
*C-C^i

(t) D-Gm-K-R = þa- + -C^i + -K + -s/-a

U20
suffixes
*C-C^i

(u) K-Gm-R-D = -K + -C^i + -s/-a + þa-

*U20
*suffixes
C-C^i

(v) K-Gm-D-R = -K + -C^i + þa- + -s/-a

*U20
*suffixes
$\underline{C}$-C^i

(w) K-D-Gm-R = -K + þa- + -C^i + -s/-a

U20
*suffixes
*$\underline{C}$-C^i

(x) D-K-Gm-R = þa- + -K + -C^i + -s/-a

U20
suffixes
$\underline{C}$-C^i

As seen in (120), only three structures can be derived while complying with the three constraints. The surviving structures are (d), (l), and (x), which correspond to the K-final (template (i)), constant (template (iii)), and K-internal (template (ii)) forms, respectively. See (121).

(121) (d) D-R-Gm-K = þa- + -s/-a + -C^i + -K => *þessum*, etc. [K-final/(i)]
(l) D-K-R-Gm = þa- + -K + -s/-a + -C^i => *þessi* [constant/(iii)]
(x) D-K-Gm-R = þa- + -K + -C^i + -s/-a => *þetta*, etc. [K-internal/(ii)]

The only other structure that fares reasonably well in (120) is structure (p), as shown in (122).

(122) D-R-K-Gm = þa- + -s/-a + -K + -C^i

?? U20 (subex.)
suffixes
$\underline{C}$-C^i

The only problem for (p) is that it requires subextraction. Note that structure (p) forms are certainly conceivable, as illustrated in (123).

(123) Conceivable but unattested (p) structures

(a) M.DAT.SG *þa-s-um-C^i^* > **þasummi*

(b) F.GEN.SG *þa-s-rar-C^i^* > **þassarri*

As mentioned, such forms are unattested. The fact that (p)-type structures with subextraction are not attested may be seen as problematic or at least surprising, since (as discussed in Section 2.2.9) so-called 'peeling' accounts have had plenty of success in other domains, and these accounts systematically require subextraction. As we shall see in Section 5.3.2, however, in some dialects of Old Norwegian we do observe structures that can be understood in terms of subextraction. A kind of subextraction will even make an appearance in certain ON forms, as part of the analysis presented in Chapter 5.

4.4 Summary

In this chapter an analysis of the ON data discussed in Chapter 3 was presented in line with Cinque's (2005) U20 theory of movement.

In Section 4.1 I identified a number of criteria that the correct functional sequence of RDem must meet. The 24 possible base-generated order of the four ingredients D, K, Gm, and R were then tested. Only a single fseq, (a) R Gm K D, survived all the tests and can fully account for the RDem facts.

In Section 4.2 the three RDem structures were derived using the fseq R Gm K D. The three derivations are fully cyclic, partially cyclic, and roll-up. It was seen, moreover, that the allomorphy between *-s* and *-a* can be captured by identifying two main types of derivations: cyclic and non-cyclic/roll-up. When a derivation is cyclic, R is lexicalized as *-s*, but when a derivation is non-cyclic, R is lexicalized as *-a*. Supporting evidence for this derivational split was seen in modern Icelandic.

A major result of my study is that we have a morphological reflex of Cinque's derivational system. However, it is important to observe also that with regard to the *-s* ~ *-a* allomorphy this approach entails that the derivation controls the lexicalization: "If cyclic, spell out *-s*; if not cyclic, spell out *-a*."[88] From the perspective of the Principles and

[88] A reviewer for *The Linguistic Review* brings up the interesting possibility that the realization of *-s* vs. *-a* is governed by whether or not the complement of R has been moved in the syntax (cf. ellipsis phenomena). Even if this is so, however, the syntactic structure would still be driving the lexicalization: "If complement of R has been moved *in the syntax*, spell out *-a*; if complement of R has not been moved *in the syntax*, spell out *-s*."

Parameters framework, however, the opposite is expected. That is, one would like to say that all variation is lexical and that it is the content of the lexicon which dictates how the syntactic derivations proceed (see Starke 2011a). This is an indication that our analysis of RDem should go even deeper. As we will see in the next chapter, it is here that nanosyntax can be of particular help.

In Section 4.3, it was observed that the fseq R Gm K D has, in principle, 24 mathematically possible structures which can be derived from it, even though it was already established in Chapter 3 that ON displays only three distinct structures. In other words, 21 structures needed to be ruled out. This was accomplished by applying a combination of syntactic, morphological, and phonological constraints to the 24 potential structures. In the end only three structures remained, which correspond exactly to the three structures identified in Chapter 3.

5
A nanosyntactic analysis of the Old Norse reinforced demonstrative

In this chapter I will pursue a nanosyntactic approach to the ON RDem paradigm established in the previous chapters, which builds on the Cinquean analysis developed in Chapter 4. The main goal will be to model the *-s* ~ *-a* 'allomorphy' of R that has been uncovered. In Chapter 4 this allomorphy was understood as a derivation-sensitive alternation, whereby derivations involving cyclic movement lexicalize the feature R as the sigmatic reinforcer *-s*, while derivations involving roll-up movement lexicalize the feature R as the asigmatic reinforcer *-a*. The precise goal in this chapter is to make the spellout of morphemes dependent not on the derivation as such but rather on the shape of lexical entries, which in turn determine how the derivation proceeds. The intuition I follow in this chapter is that we have reached a certain level of detail or granularity at which more traditional tools and analyses are not appropriately equipped to help us make any more progress (specifically in the area of the *-s* ~ *-a* alternation). We are now at the point where a nanosyntactic approach becomes essential.

5.1 Preliminary nanosyntactic structures

In this section I take the U20 structures for RDem elaborated in the previous chapters and reinterpret them in light of nanosyntactic spellout mechanisms. Various important issues about terminal vs. phrasal spellout will arise in the course of this endeavor, some of them problematic.

5.1.1 Translating U20 structures into nanosyntactic structures

In Chapter 4 we distinguished three types of RDem in ON: K-final, constant, and K-internal. The following structures were proposed for the K-final (Figures 107 and 108), constant (Figures 109 and 110), and K-internal (Figures 111 and 112) forms.

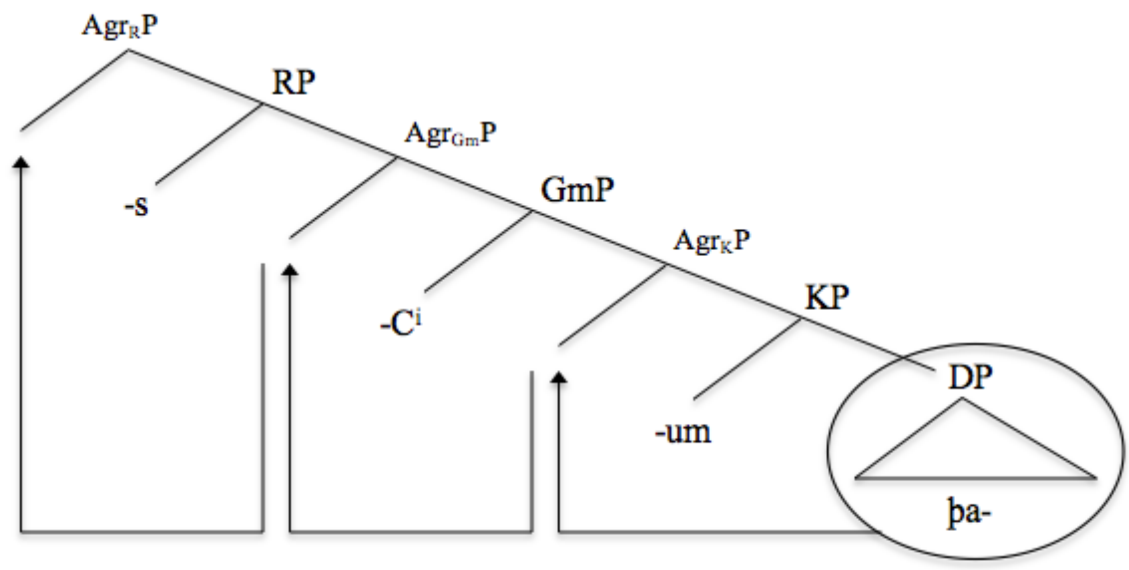


Figure 107 K-final derivation: M.DAT.SG þa-s-C^i-um > *þessum*

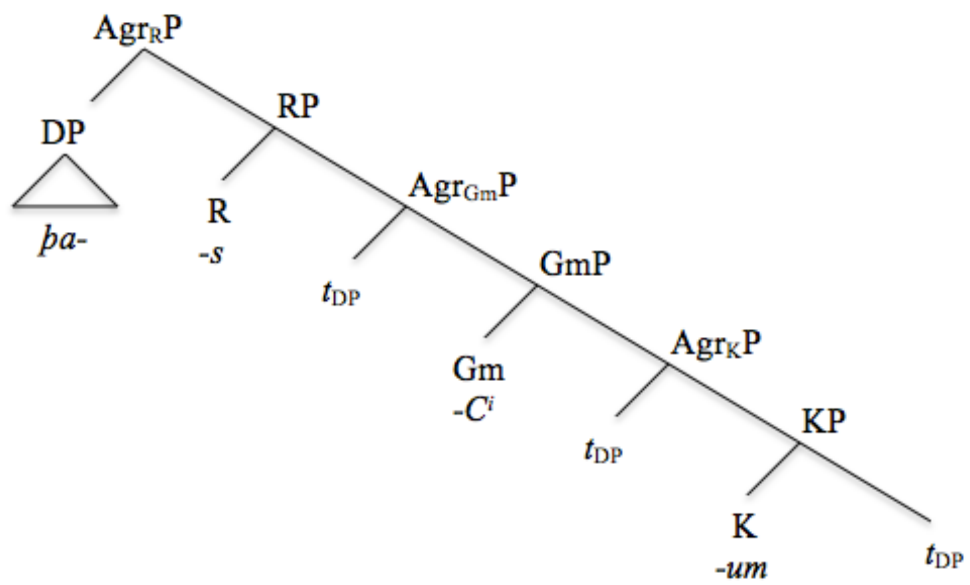


Figure 108 Final structure of K-final form

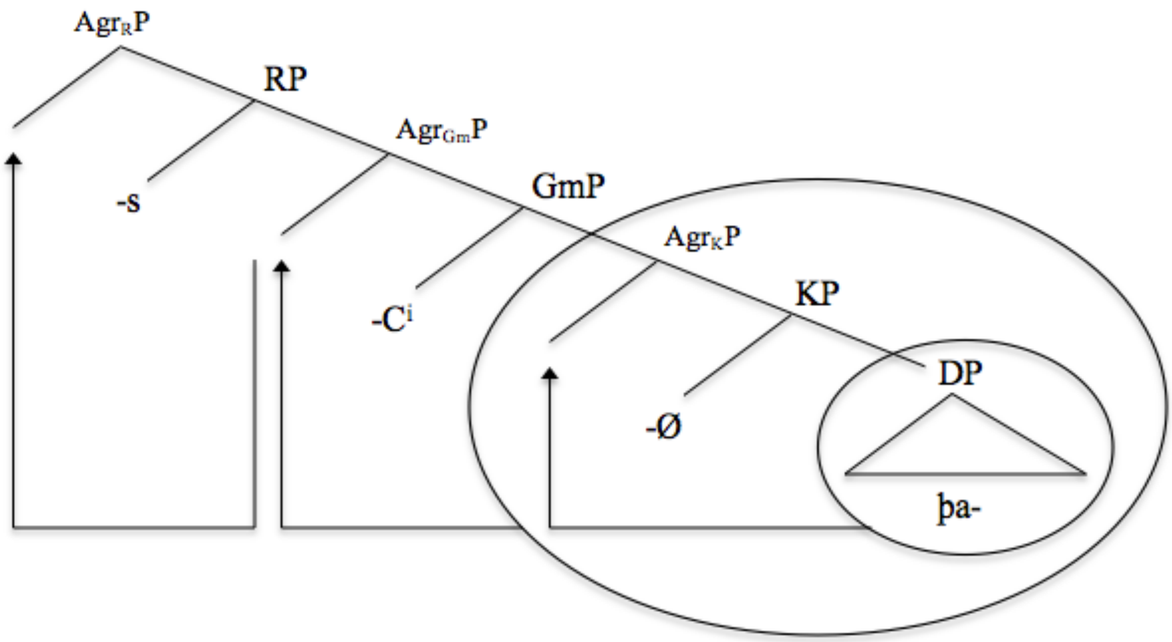


Figure 109 Constant derivation: F.NOM.SG / N.NOM/ACC.PL þa-Ø-s-C^i > *þessi*

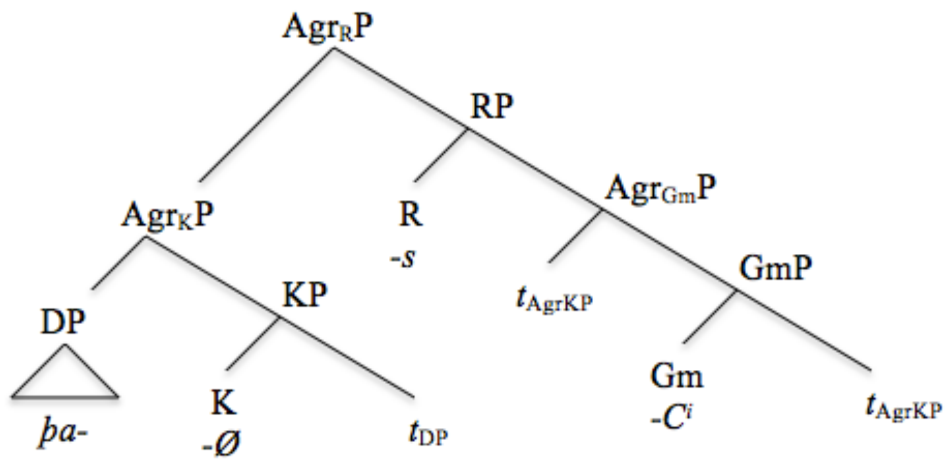


Figure 110 Final structure of constant form

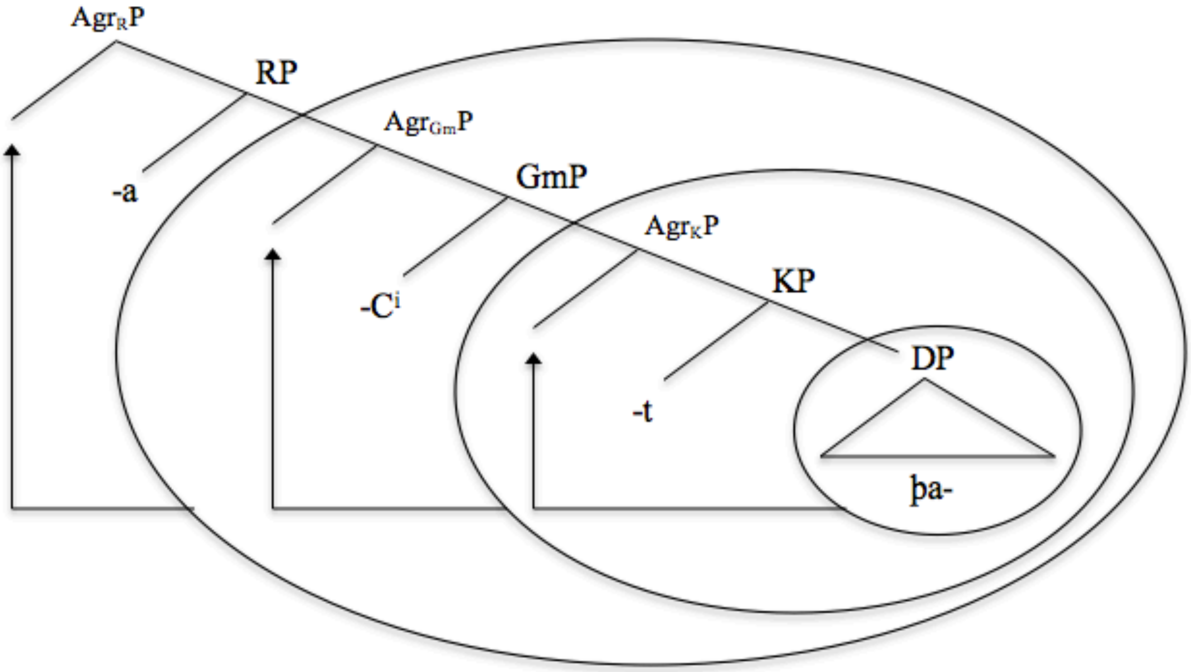


Figure 111 K-internal derivation: N.NOM/ACC.SG þa-t-C^i-a > *þetta*

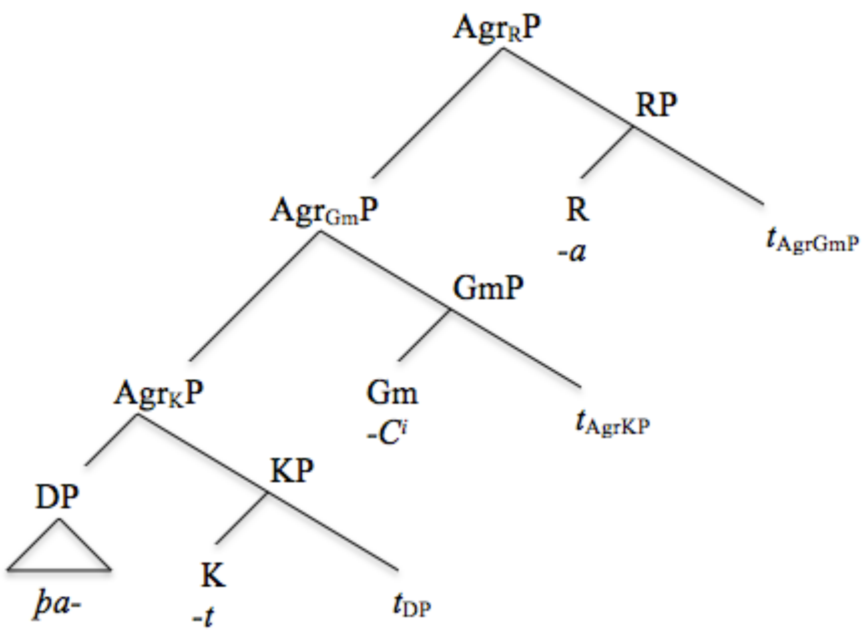


Figure 112 Final structure of K-internal form

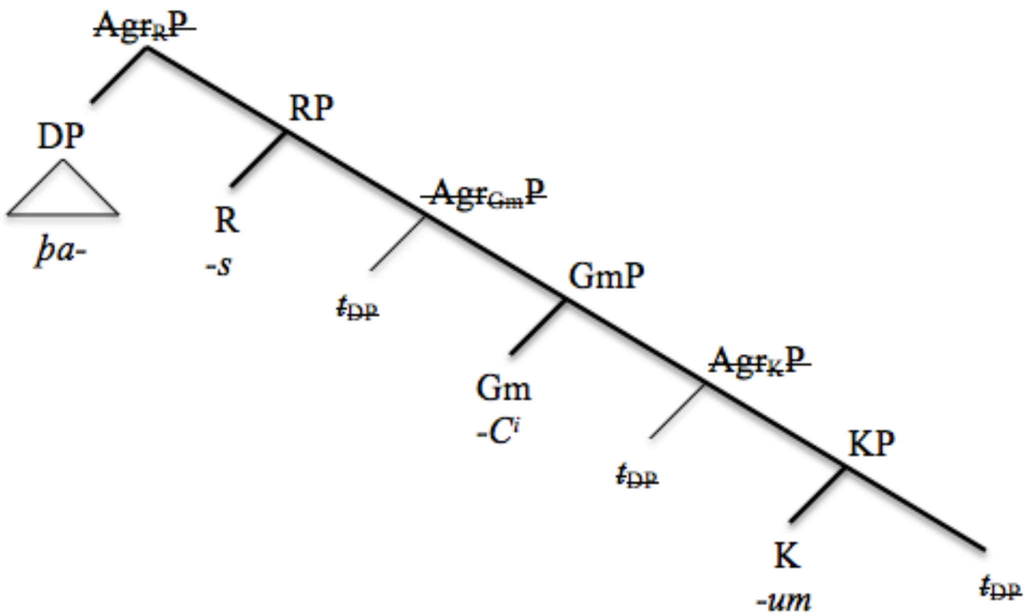
These derivations and structures have been discussed at length in Chapters 3 and 4.

Let us now make an attempt to 'translate' these Cinquean structures into nanosyntactic structures and derivations. First of all, I will propose that the AgrPs in Cinque's (2005) system can be dispensed with. In the U20 structures, AgrPs serve as landing sites for NP-containing constituents moving up the structure. In other words, AgrPs 'open up' the structure, providing available slots through which movement can proceed. The existence of AgrPs is further justified with a theory of licensing within the extended projection of the noun (Cinque 2005: 325-327). In Chapter 4, the purpose served by the AgrPs in the RDem structures was also to provide landing sites. Now, in Chapter 5, I would like to propose that what we observe in the RDem structures in Figures 107-112 above is actually (for the most part) spellout-driven movement, not the presence of AgrPs. Recall from Chapter 2 (cf. examples (57-62)) that spellout-driven movement:

(i) does not leave traces,
(ii) creates unlabeled specifiers for moved constituents, and
(iii) that these unlabeled specifiers are deleted if they are empty (i.e. once the constituent undergoing spellout-driven movement has moved further).

In sum, then, I propose that the AgrPs in Figures 107-112 should be reinterpreted as specifiers created by spellout-driven movement. If this is so, then the 'AgrPs' above do not actually contain traces of movement (cf. (i)), nor do they have a label (cf. (ii)); furthermore, any 'AgrP' above that is not hosting a moved constituent can be expunged from the structure completely (cf. (iii)). Thus we can rewrite the structures above as in Figures 113-115 (with additional refinements to come later).

K-final structure



becomes:

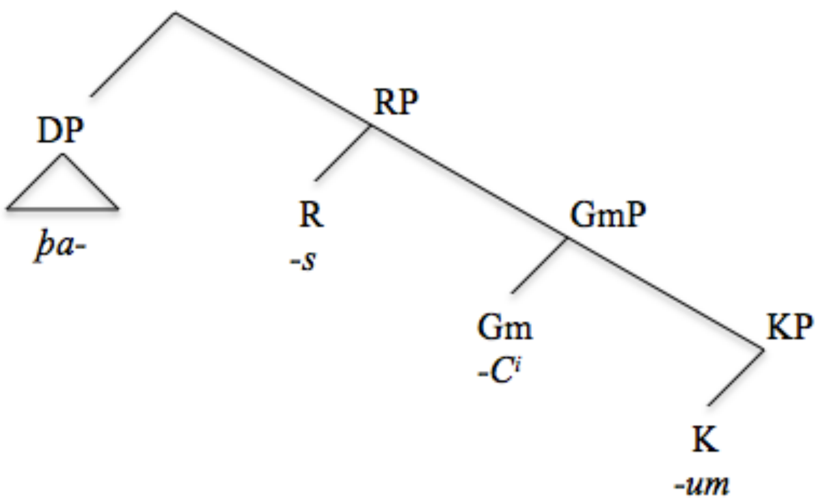
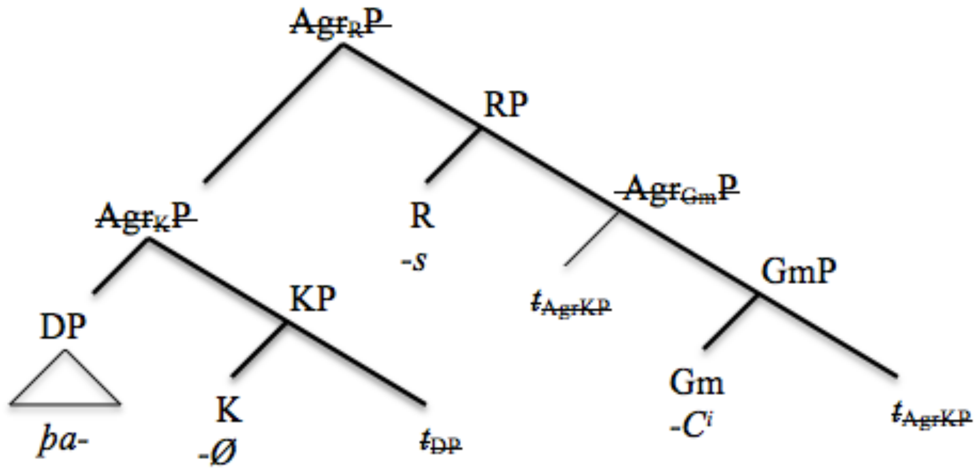


Figure 113 Rewriting the K-final structure

Constant structure



becomes:

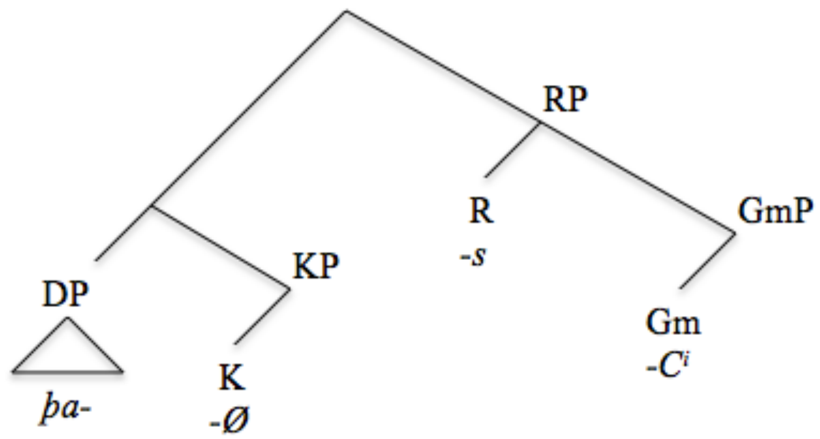
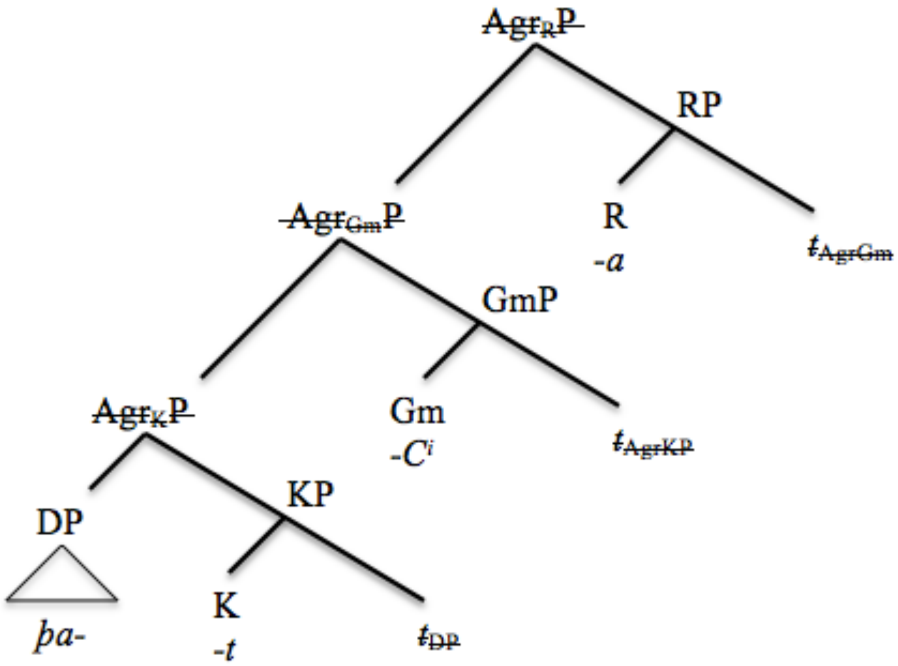


Figure 114 Rewriting the constant structure

K-internal structure



becomes:

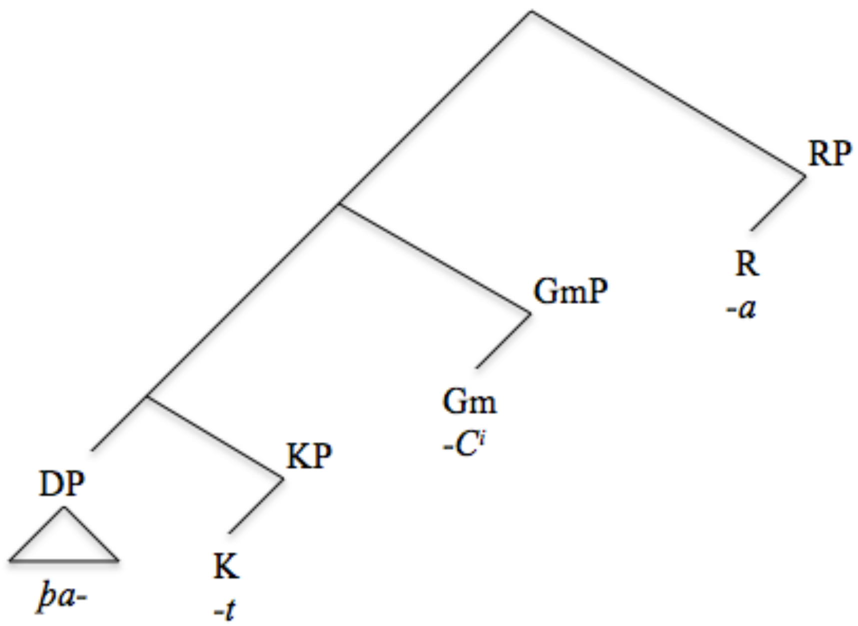


Figure 115 Rewriting the K-internal structure

The next modification that I will propose to the representations above is that most of the morphemes we see in RDem are not to be analyzed as heads, as in the more traditional Cinquean structures, but as XPs. Recall from Chapter 2 that suffixes have the basic shape seen in Figure 116, i.e. an XP with a singleton set at the bottom of the structure.

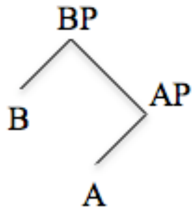


Figure 116 Basic shape of a suffix

The singleton set at the bottom of the structure, [$_{AP}$ A], indicates that 'traceless' movement has taken place, as illustrated in Figure 117, where XP has moved to the left (without leaving a trace), putting [BP [AP]] in a suffixal position in relation to XP.

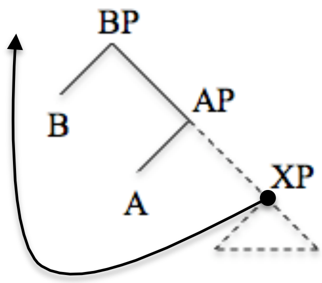


Figure 117 Basic shape of a suffix

Recall from Chapter 4 that most of the morphemes making up RDem are suffixes. Thus we want these morphemes to have a lexical shape like [BP [AP]]. Recall also that all RDem forms begin with the base *þa-*. I will propose that *þa-*, then, corresponds to a constituent like XP in Figure 117. Since *þa-* makes up the bottom chunk of the structure (including presumably the head of the entire nominal projection, N), it is the constituent which moves first during spellout-driven movement and thus the constituent that goes to the left of suffixes. Indeed, we have already seen in Chapter 4 that D is at the bottom of the functional sequence, making the base *þa-* equivalent to the phrase DP.

If we turn back to the structures in Figures 113-115, we will notice that many of the morphemes that are taken to be heads there can be reinterpreted as phrases instead. In the structure of the K-final form (Figure 113), both the base *þa-* and the K ending *-um* can easily be reinterpreted as phrasal constituents (DP and KP, respectively) which can be targeted for spellout, as indicated by boldface in Figure 118.

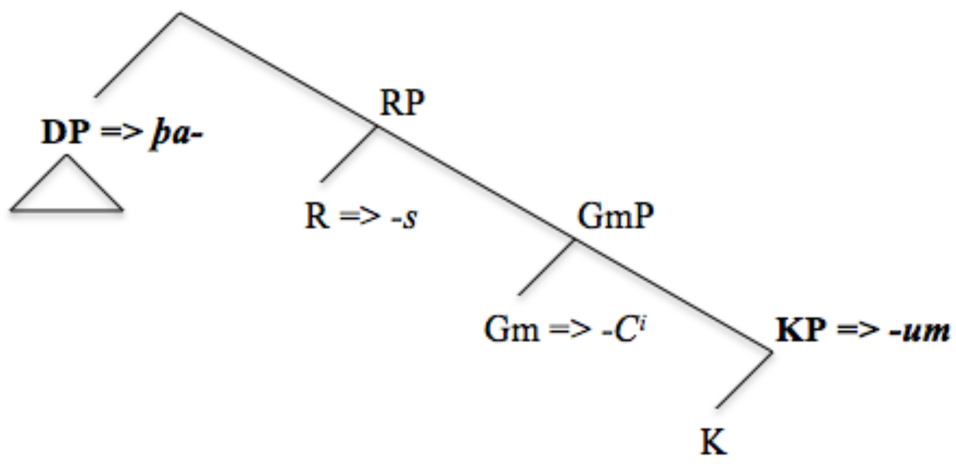


Figure 118 K-final form with XP morphemes

Note that the sigmatic reinforcer -*s* and the geminator -C^i, on the other hand, would have to spell out as terminals in this structure: RP is not an independent constituent to the exclusion of GmP and KP, and GmP is not an independent constituent to the exclusion of KP.

In the structure of the constant form, the base *þa*-, the K ending -*Ø*, and the geminator -C^i can all be considered phrasal constituents (DP, KP, and GmP, respectively) which at some stage of the derivation can be targeted for spellout, as indicated in Figure 119 by boldface.

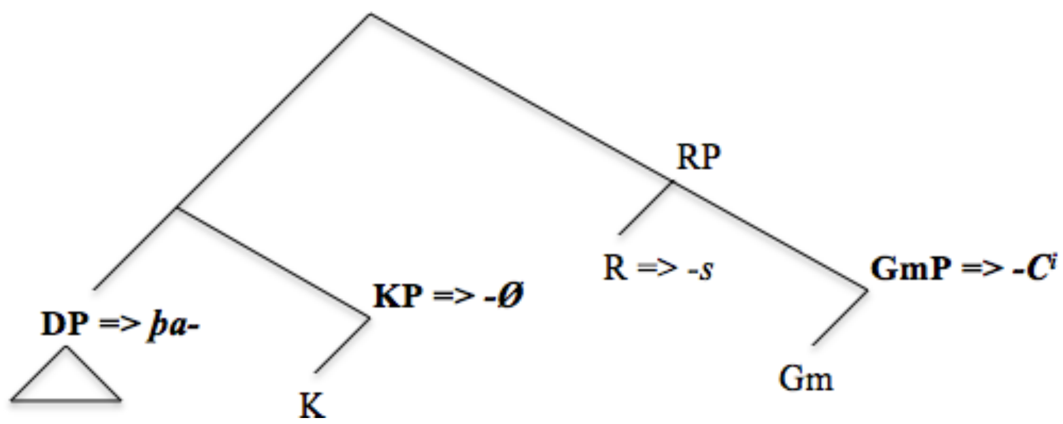


Figure 119 Constant form with XP morphemes

The sigmatic reinforcer -*s*, on the other hand, would again have to spell out as a terminal in this structure: RP is not a constituent to the exclusion of GmP.

Finally, in the structure of the K-internal form, all four of the morphemes, including the asigmatic reinforcer -*a*, can be considered phrasal constituents which at some stage of the derivation can be targeted for spellout, as marked in Figure 120 by boldface.

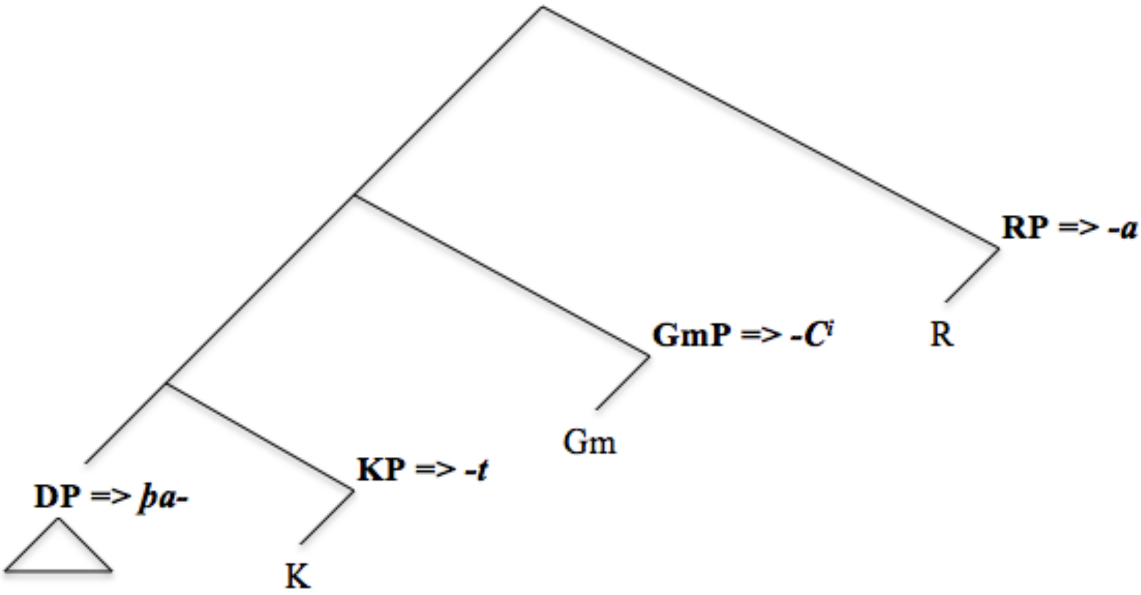


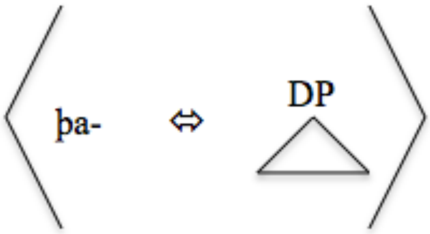
Figure 120 K-internal form with XP morphemes

Note that this derivation does not require any morpheme in the structure to spell out as a terminal. Each morpheme corresponds to its own unique phrasal constituent.

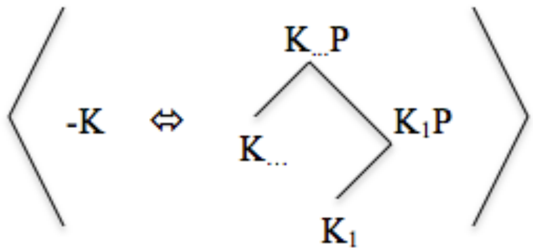
Based on the considerations so far, let us now build a preliminary lexicon of nanosyntactic lexical entries for the ON RDem. In the structures in Figures 118-120 above, the base *þa-* consistently corresponds to DP, hence (124a) (compare also the shape of DP in (124a) to XP in Figure 116 above). The K endings also always correspond to a phrasal constituent, KP, hence the lexical entry in (124b). Recall from Chapter 2 that I follow Caha (2009) in assuming K endings to have a fine-grained internal structure with numerous features/heads like K_1, K_2, K_3, etc. Consider next the lexical structures of the reinforcers in Figures 118-120. Observe first that the sigmatic reinforcer *-s* always corresponds to the *terminal* R, as shown in the K-final form in Figure 118 and in the constant form in Figure 119, while the asigmatic reinforcer *-a* always corresponds to the *phrase* RP, as shown in the K-internal form in Figure 120. Hence we might lexically encode *-s* as the head R in (124c) and *-a* as the phrase RP in (124d). Finally, observe that the geminator *-C^i* shifts between phrasal and terminal status in the structures above. In the K-final form in Figure 118, the geminator must spell out as a terminal. In the constant form in Figure 119 and the K-internal form in Figure 120, however, the geminator can spell out at the phrasal level. According to the Superset Principle, the terminal Gm is a subset of the phrase GmP. Thus, if we say that the lexical structure of the geminator is GmP, then by the Superset Principle the terminal Gm can also be matched if necessary (i.e. in the K-final form). Hence we encode *-C^i* as in (124e), where the lexical structure is the phrase GmP.

(124) Preliminary lexical entries

(a) Base

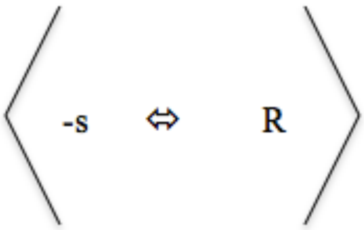


(b) K endings

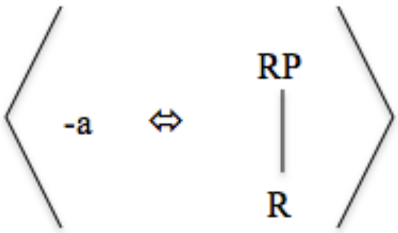


...etc.

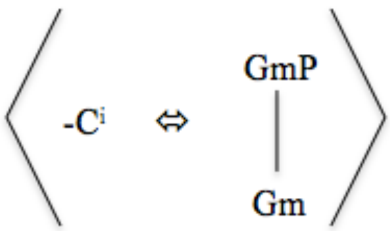
(c) Sigmatic reinforcer



(d) Asigmatic reinforcer



(e) Geminator



This is a first attempt at 'translating' the Cinquean RDem structures of Chapter 4 into a framework like nanosyntax which allows for phrasal spellout.

Importantly, the lexical entries in (124) which introduce a structural distinction between the sigmatic and the asigmatic reinforcers appear to present a potential solution to the *-s* ~ *-a* allomorphy problem. Recall from Chapter 4 that both *-s* and *-a* seem to be two realizations of the same head, R. That is, R is spelled out as a terminal and realized as *-s* in the cyclic-type derivations but is phrasal and spelled out as *-a* in the roll-up derivation. This derivation-specific allomorphy, however, is conceptually problematic in that the type of syntactic derivation should not be the driving force behind lexical alternations, rather it should be the lexical items in the lexicon that decide which syntactic derivation takes place, at least according to the view of linguistic variation underlying the Principles and Parameters framework. Thus, the idea presented in (124) – that the sigmatic and asigmatic reinforcers are actually structurally distinct rather than both corresponding to the terminal R, as in our Cinquean analysis in Chapter 4 – would account for the alternation in a more principled way. However, one crucial component of this approach is that we need to invoke terminal spellout, specifically in the derivations of the K-final and constant structures. This is more complicated than it may seem, as the next section will show.

5.1.2 Exploring the issue of terminal spellout

The entries in (124) make sense from the point of view of the structures after they have already been derived/spelled out, as in Figures 118-120. However, when we consider how the forms are built up and derived step by step, then the entries in (124) present some problems and inconsistencies. These problems are mostly related to the concept of terminal spellout, which – as we have just seen – plays a crucial role in the structures above.

Consider the derivation of a K-final form, the (intended) final derivational structure of which is repeated in Figure 121.

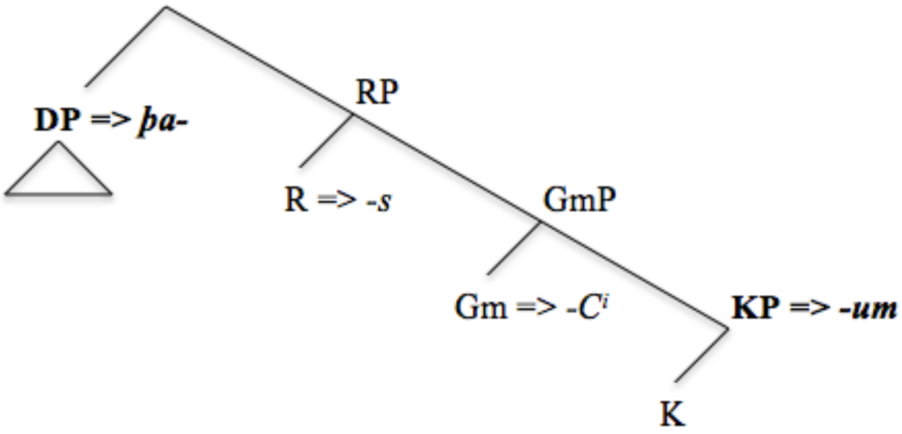


Figure 121 Final structure of K-final form: M.DAT.SG þa-s-C^i-um > *þessum*

Recall that in this structure DP and KP are constituents containing only D and K material, respectively. This means that the base *þa-* and the K ending can both be taken to have phrasal L-structures. The morphemes *-s* and *-C^i*, on the other hand, must correspond to terminals in this structure, since the constituent at RP contains more than just R material and the constituent at GmP contains more than just Gm material.

Now let us begin at the very beginning of the K-final derivation and try to build the structure in Figure 121. The fseq, lexicon, and spellout algorithm are repeated in (125).

(125) Fseq: R > Gm > K > D

Lexicon: < *þa-* ⇔ DP >

< *-um* ⇔ $K_{M.DAT.SG}$P >[89]

< *-s* ⇔ R >

< *-a* ⇔ RP >

< *-C^i* ⇔ GmP >

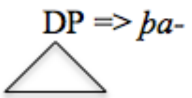
Spellout algorithm: STAY > CYCLIC > SNOWBALL

The first step in the derivation is to build DP, which unproblematically spells out as *þa-* at STAY.

[89] This entry has been simplified for the sake of brevity. This has no bearing on the main points of the discussion here.

(126) Build up to DP

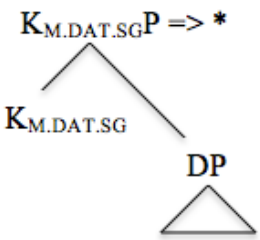
STAY



Next, K features are added. Here, for the sake of simplicity, let us simply say that the K feature $K_{M.DAT.SG}$ is added, giving the structure [$K_{M.DAT.SG}$P [DP]]. This structure does not spell out at STAY, since there is no L-tree with KP on top of a DP. This is seen in (127a). The next step in the algorithm, CYCLIC, is not applicable since there is no daughter of DP which can be moved cyclically. The final step, SNOWBALL, though, targets DP and results in $K_{M.DAT.SG}$P being matched by the entry < -um ⇔ $K_{M.DAT.SG}$P >, as seen in (127c).

(127) Add $K_{M.DAT.SG}$

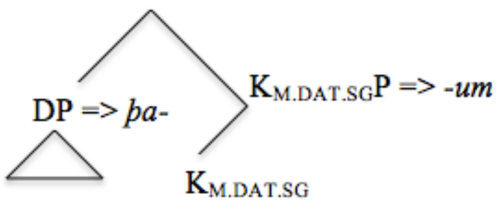
(a) *STAY



(b) *CYCLIC

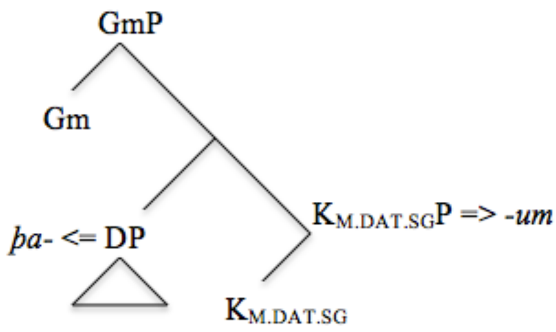
N.A.

(c) SNOWBALL



The next step in the derivation is to add Gm. This gives the structure in (128).

(128) Add Gm



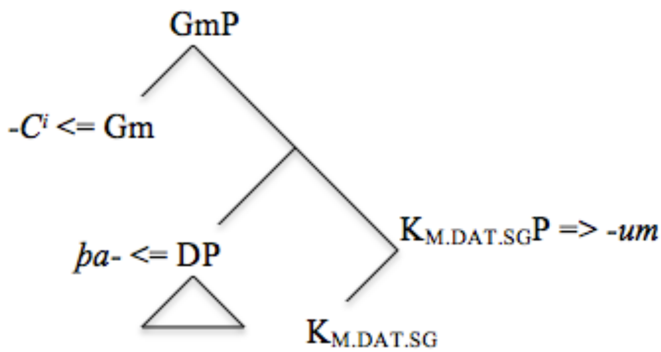
At this juncture in the derivation, an important issue arises. Observe that in the intended final structure in Figure 121, the geminator morpheme -*C*ⁱ corresponds to the terminal Gm. It is unclear, however, how the spellout algorithm would accommodate for terminal spellout. In Chapter 2 there was no explanation as to why or when terminals might spell out.

To explore the concept of terminal spellout, I will at this point branch off into an 'alternate-universe' derivation. In this alternate universe, spellout will not proceed exactly as in Chapter 2. Instead the spellout process will be tinkered with in various ways, in order to see if the structures in Figures 118-120 can be derived in a coherent way by allowing for terminal spellout.

Since the entire point of CYCLIC and SNOWBALL is to create new phrasal constituents, the most likely stage at which terminal spellout would apply is STAY. In (128), for instance, the phrase GmP is not a constituent that can be targeted for spellout without also implicating DP and KP; however, the terminal Gm is in fact a (trivial) constituent that can be targeted for spellout on its own, as shown in (129). Recall that the lexical entry for the geminator is < -*C*ⁱ ⇔ GmP >, meaning that in (129) the Superset Principle can be invoked: the S-tree Gm is a subset of the L-tree GmP and thus the two can be successfully matched.

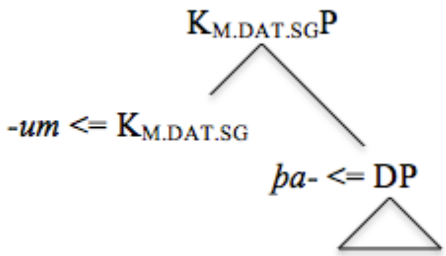
(129) Add Gm

STAY



Indeed, this is basically what is needed for the intended final structure in Figure 121, in which both Gm and R spell out at the terminal level. However, once we allow for terminal spellout at STAY in (129), the question immediately arises why we did not, in the previous stage of the derivation in (127), allow for $K_{M.DAT.SG}$ to spell out as a terminal. See (130).

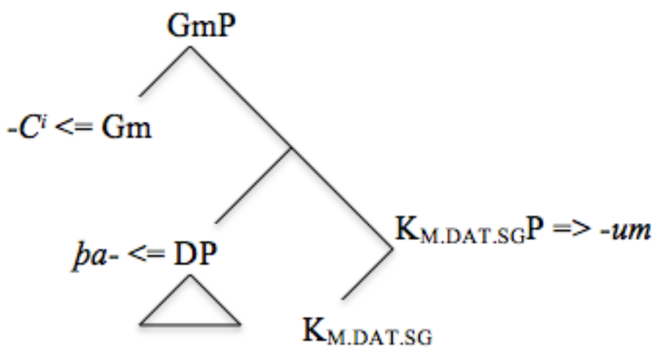
(130) Terminal spellout of $K_{M.DAT.SG}$ at STAY?



Once again, the Superset Principle could theoretically be at work in (130): the S-tree $K_{M.DAT.SG}$ is a subset of the L-tree $K_{M.DAT.SG}P$ from the lexical entry < *-um* ⇔ $K_{M.DAT.SG}P$ >. In other words, the reasoning that leads to the terminal spellout of Gm in (129) is exactly the same reasoning that would allow for $K_{M.DAT.SG}$ to spell out as a terminal in (130). The problem with this is the following. If the system is able to target the terminal $K_{M.DAT.SG}$ for spellout, this would result in *-um* being in a position to the left of *þa-*, as seen in (130). That is, no spellout-driven movement of DP to the left of KP would be triggered. Recall, though, that the ending *-um* is a suffix and as such belongs to the right of *þa-*.

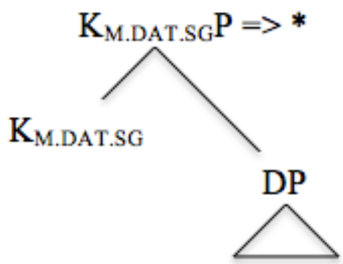
In other words, our 'alternate-universe' derivation shows that allowing for terminal spellout makes it unclear how suffixes should be derived. Moreover, it shows that we would be putting inconsistent demands on the spellout system: in some cases it seems we want STAY to consider terminals for spellout, as shown in (131) for Gm. But in other cases we want STAY not to target terminals for spellout so that the algorithm will continue and trigger spellout-driven movement, as shown in (132) for $K_{M.DAT.SG}P$.

(131) STAY targets terminal Gm for spellout → **match**



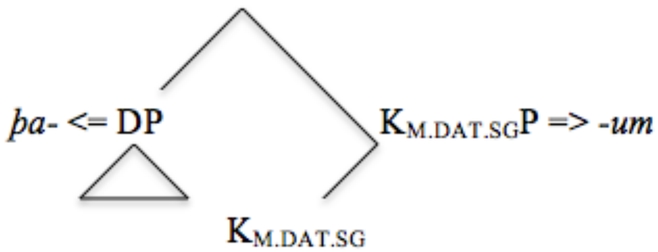
(132) Add Gm

(a) STAY targets phrase $K_{M.DAT.SG}P$ for spellout → **no match**



(b) CYCLIC is not applicable

(c) SNOWBALL moves DP to the left of $K_{M.DAT.SG}$ → match



In sum, there is no obvious way for the spellout system to decide if it should target a terminal at STAY (as for Gm) or continue on to CYCLIC and SNOWBALL (as for $K_{M.DAT.SG}$). This is a problem for our imaginary system that allows for terminal spellout.

Recall, furthermore, that the geminator $-C^i$ is a suffix. However, in the configuration in (129) and (131) it surfaces as a prefix instead. More generally, if we allow for terminal spellout the question arises why we cannot spell out at the terminal level at every single layer of the structure. After all, the minumum amount of structure in an L-tree will be a terminal, meaning there will always be a match between a given terminal in the S-tree and an L-tree. In Figure 122, terminal spellout is possible at every layer thanks to the Superset Principle.

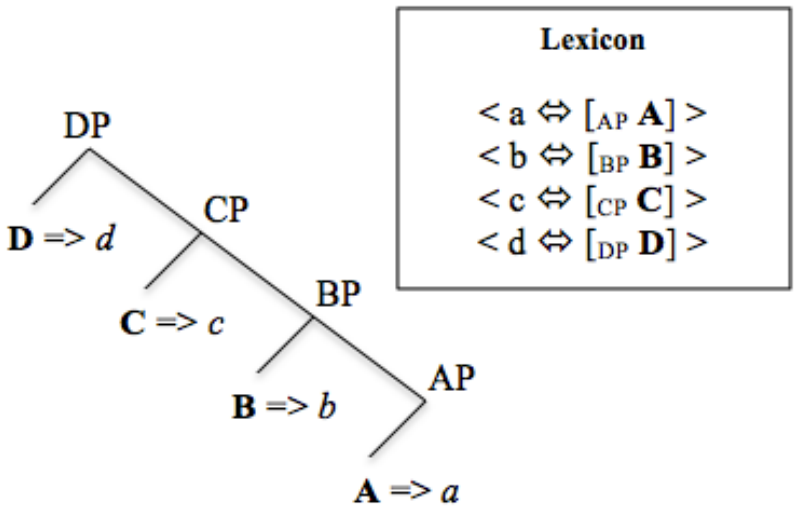


Figure 122 Terminal spellout at every layer

As seen in Figure 122, allowing for terminal spellout at STAY seems to get rid of (or at least weaken) the system's need to do spellout-driven movement: terminals are always (trivial) constituents and thus there will be no need to create new constituents by evacuation movements. In other words, derivations will essentially be stalled in their base-generated form, and the entire idea of spellout-driven movement as an account of suffix formation is lost. This is a serious problem that a system allowing for terminal spellout needs to deal with.

One attempt at solving this dilemma might be to say that it is the L-tree which determines whether or not terminal spellout takes place. This idea is actually developed by Pantcheva (2011: §6.3.2). Pantcheva's system is to my knowledge the most well developed theory of nanosyntactic spellout which allows for both phrasal and terminal spellout. In order to further illustrate the problems with the sorts of terminal spellout needed for the RDem structures as they appear in Section 5.1, I will temporarily adopt Pantcheva's spellout system.

In Pantcheva's system, spellout proceeds from right to left and bottom to top. Thus in the structure $[_{BP}$ B A] in Figure 123, the order of lexicalization will be A first, then B, and finally BP (because the feature A was picked out first and then merged with B, resulting in the phrase BP).

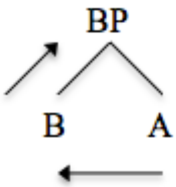


Figure 123 Basic order of lexicalization of nodes

With the lexical entries provided in (133), spellout of [$_{BP}$ B A] will proceed as in Figure 124.

(133) Lexicon

< *a* ⇔ A >

< *b* ⇔ B >

< *c* ⇔ BP >

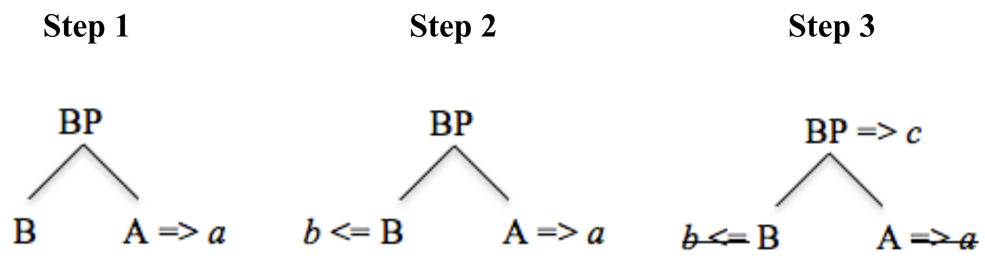


Figure 124 Spellout of [$_{BP}$ B A]: Phrasal spellout overrides terminal spellout (Pantcheva 2011: 116)

In Figure 124, spellout begins at the rightmost and bottommost feature, A. The terminal A is matched by the lexical entry < *a* ⇔ A > and thus spells out as *a*. The leftmost, bottomost feature, B, is then matched by the lexical entry < *b* ⇔ B > and spelled out as *b*. Finally, spellout moves up to the phrase BP, which is matched by the lexical entry < *c* ⇔ BP >. In spelling out BP as *c*, furthermore, the lower spellouts are overridden by Cyclic Override. In other words, Cyclic Override means that phrasal spellout takes priority over terminal spellout: spelling out a phrase composed of two features like A and B is preferred over spelling out the features A and B separately. This idea has already been discussed with regard to idioms in Chapter 2: for instance, spelling out the higher node *mice* is preferred to spelling out the individual components **mouse-s*.

Had there not been a lexical entry like < *c* ⇔ BP > in the lexicon, however, then terminal spellout will survive in Pantcheva's system.

(134) Lexicon

< *a* ⇔ A >

< *b* ⇔ B >

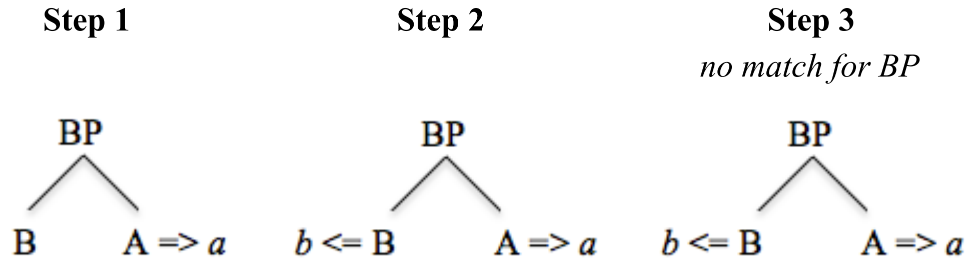


Figure 125 Spellout of [$_{BP}$ B A]: Terminal spellout survives (Pantcheva 2011: 118)

Note that even though there is no match for BP at Step 3 in Figure 125, the node BP is "lexicalized by inheritance" (Pantcheva 2011: 119-120) in the sense that its component ingredients A and B have been properly lexicalized. Thus Cyclic Exhaustive Lexicalization, according to which no features may go unlexicalized, is still satisfied.

However, it is not the case that direct lexicalization (i.e. spelling out an entire phrase like BP in Figure 124) is on equal footing with lexicalization by inheritance (i.e. spelling out only the component parts of BP as in Figure 125). Pantcheva (2011: 140) proposes that direct lexicalization of an XP is always to be preferred over lexicalization by inheritance. Importantly, this preference holds *even if structure needs to be moved* in order to make the XP a matchable constituent. In other words, even if a terminal X^0 has been properly matched, the system will still go so far as to perform spellout-driven movement if it means that XP can find a match this way (thereby overriding the spellout of X^0).[90] Put simply, phrasal spellout is strongly prioritized over terminal spellout.

Pantcheva's system, then, treats terminal spellout in a highly consistent way: terminal spellout is always available, but it will be overridden by later phrasal spellouts if the lexicon contains such lexical entries. What determines whether or not terminal spellout survives is the content of the lexicon. In Figure 125, for instance, since only lexical entries with terminals are available, A and B will spell out as terminals and not be deleted by Cyclic Override.

Let us now return to the RDem forms. Even with Pantcheva's system, in which terminal spellout and phrasal spellout coexist, the problem with the geminator morpheme

---

[90] For Pantcheva's (2011) take on spellout-driven movement, which differs in various ways from my own system of spellout-driven movement, see her Chapter 7.

-C^i discussed above remains: how does the spellout algorithm know to spell out -C^i as a terminal in the K-final forms but as a phrase in the constant forms and in the K-internal forms? Indeed, since the L-tree for -C^i is the phrase GmP (see (124e)), the phrase GmP will override the terminal spellout of Gm by the logic of Pantcheva's system.

The same problem rears its head at the R layer as well. In (124) we posited that the asigmatic reinforcer *-a* corresponds to RP while the sigmatic reinforcer *-s* corresponds to the terminal R. Now, according to Pantcheva's system, once the derivation of an RDem form reaches the R layer, the terminal R is first targeted for spellout, as shown in Figure 126.[91]

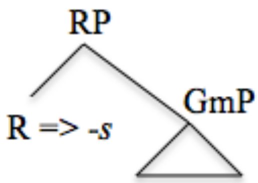


Figure 126 Terminal spellout: R spells out as *-s*

R can be matched by the lexical entry < *-s* ⇔ R > at this stage in the derivation. Next, Pantcheva's system will attempt to spell out the node RP. As mentioned above, the spellout of phrases is preferred to the spellout of terminals, even if spellout-driven movement must occur. Thus the system will eventually evacuate GmP to the left of RP in order to make RP accessible for matching the entry < *-a* ⇔ RP >. This, in turn, will override the spellout of *-s*, as seen in Figure 127.

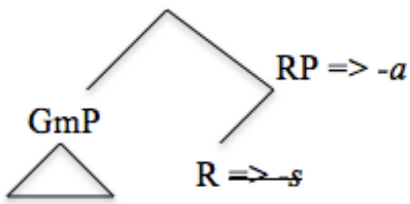


Figure 127 Phrasal spellout of RP overrides spellout of R

The only way the terminal spellout of R as *-s* would survive is if there had not been a lexical entry like < *-a* ⇔ RP > in the lexicon.

We are left with something of a paradox. If the lexical entries < *-a* ⇔ RP > and < *-s* ⇔ R > coexist in the lexicon, then we predict that the sigmatic reinforcer will never surface

[91] Technically the node GmP is targeted first, then R. However, there will not be a singular spellout for the entire phrase GmP, but since its component parts will have been properly spelled out, the node GmP is "lexicalized by inheritance" in Pantcheva's (2011: 119-120) sense.

overtly, since the spellout of RP, namely *-a*, will always override the spellout of R, namely *-s*. So if we want the terminal R to spell out as *-s*, then we cannot tolerate the competing entry < *-a* ⇔ RP > in the lexicon. However, not having the entry < *-a* ⇔ RP > in the lexicon would mean that the asigmatic reinforcer does not exist, which is obviously not the case.

So the main hypothesis from this section – that *-s* corresponds to R and *-a* corresponds to RP – gets us nowhere. We need to construct appropriate L-trees for *-s* and *-a* that do not interfere so aggressively with each other. These L-trees need to compete in just the right way, allowing *-s* to appear in the appropriate subset of RDem forms and allowing *-a* to appear in its appropriate subset of RDem forms. The next section will offer a way of accomplishing this.

In response to some of the difficulties with terminal spellout discussed above, Starke (2011b, p.c.) has proposed that only *non*-terminals are allowed to spell out (i.e. phrases or complex heads consisting of two or more terminals). In other words, Starke proposes a flatout ban on terminal spellout in his nanosyntactic spellout system. I, however, will take a more agnostic stance and leave open the issue of terminal spellout and whether or not it can coexist with phrasal spellout. For RDem, at least, I think the issues encountered in this section suggest that a more fine-grained structure is needed to account for the lexical structures of $-C^i$, *-s*, and *-a*, and in the next section I will show that it is better to think of $-C^i$ and *-s* as complex heads consisting of two features each. This way these entities are subject to different types of lexicalization depending on their precise structural composition. Given the complex head analysis, we will not need to invoke terminal spellout for any of the RDem forms. For all practical purposes, then, the system developed here does not require terminal spellout. Whether a wholesale ban on terminal spellout must be maintained (along the lines of Starke's work) or whether terminal spellout is required elsewhere and thus must be present in the system in general is a matter which I cannot pursue here.

5.2 A more fine-grained functional sequence

In the previous discussion we encountered problems with terminal spellout in the derivations of the K-final and constant forms. Informally, the problem can be summed up in that *-s* and $-C^i$, which apparently need to lexically correspond to terminal nodes in the structure, cannot be spelled out without modifying the entire spellout system, and even then the lexical competition between *-s* and *-a* does not fall out correctly. This section proposes that one way to deal with the problems arising in the previous section is to refine the fseq. More specifically, I propose that R and Gm actually consist of two features each.

In other words, the four-layered sequence R > Gm > K > D becomes the six-layered sequence $R_2 > R_1 > G > m > K > D$.

5.2.1 Packaging in the K-internal forms

So far the RDem fseq has basically been packaged as in Figure 128.

R	**Gm**	**K**	**D**
-s -a	-C^i	-K_{final} -$K_{constant}$ -$K_{internal}$	þa-

Figure 128 Fseq packaging so far

The packaging in Figure 128, however, is too coarse and too crowded: if we have to squeeze both the sigmatic and the asigmatic reinforcer into the R layer, then in order to get a structural difference between the two we need to posit that one is the terminal R (-*s*) and the other is the phrase RP (-*a*), leading to the host of problems encountered in Section 5.1.2. More features in the sequence would provide more material for making the distinction between -*s* and -*a*.

Consider now the three different types of K endings which appear under the K layer in Figure 129. We know from Caha (2009) that the K layer in Figure 129 actually decomposes into numerous K features. Indeed, features for number, gender, and person (Φ features) must also be involved at or around this layer. Thus the K endings are not nearly as 'crowded' in Figure 128 as they seem to be, since we assume a fine-grained sequence of K features. I will propose, then, that the problem of the -*s* ~ -*a* alternation can be resolved if we treat R similarly to K, that is, if we assume that there is some internal structure to the R domain. I propose to divide R into at least two features, call them R_2 and R_1.

Suppose now that the sigmatic reinforcer -*s* corresponds to R_2 and R_1 together, as shown in Figure 129. Since the sigmatic reinforcer surfaces only in the K-final and constant forms, only these K endings are shown under K in Figure 129.

<table>
<tr><th>R_2</th><th>R_1</th><th>Gm</th><th>K</th><th>D</th></tr>
<tr><td colspan="2">-s</td><td>-C^i</td><td>-K_{final}
-$K_{constant}$</td><td>þa-</td></tr>
</table>

Figure 129 Sigmatic reinforcer = $R_2 + R_1$

If *-s* is the result of packaging R_2 and R_1 together, then obviously we do not want to say that *-a* also corresponds to R_2 and R_1. Let us suppose instead that *-a* corresponds to the R_2 layer, as shown in Figure 130.[92] Since the asigmatic reinforcer surfaces only in the K-internal forms, only these K endings are mentioned under K in Figure 130.

R_2	R_1	Gm	K	D
-a	?	-C^i	-$K_{internal}$	þa-

Figure 130 Asigmatic reinforcer = R_2 only

As shown in Figure 130, the question arises how R_1 fits in with the K-internal forms. This question is especially puzzling since all of the morphemes we need for the K-internal forms seem to be accounted for already, so it is not clear what R_1 could correspond to at all.

Before solving this issue, let us press ahead with the decomposition strategy as a way of moving away from the terminal spellout approach. Suppose that not only R but also Gm should be divided into two features. In fact, it is suggestive of internal structure that the geminator *-C^i* could easily be analyzed as bimorphemic, i.e. the consonant geminator plus the *i*-mutator (*-C* + - ^{i}). Thus we divide Gm into the component features G (the geminator *-C*) and m (the *i*-mutator). The packaging schema for the K-final and constant forms in Figure 129 above now looks like Figure 131. Crucially, both *-s* and *-C^i* now correspond to two features each rather than one, as was the case in our initial decomposition.

R_2	R_1	G	m	K	D
-s		-C^i		-K_{final} -$K_{constant}$	þa-

Figure 131 Both *-s* and *-C^i* correspond to two features

The packaging schema for the K-internal forms in Figure 130 now looks like Figure 132. Recall that there was no obvious candidate for lexicalizing the element R_1.

[92] The alternative that *-a* corresponds to R_1 only is not workable, since this suggests that *-s* will always override *-a* on account of [R_2 + R_1] being a superset of R_1.

R_2	R_1	G	m	K	D
-a	?	$-C^{i}$		$-K_{internal}$	þa-

Figure 132 No obvious candidate for R_1

Let us now turn our attention to the K-internal packaging in Figure 132. The core properties of Figure 132 are:

(i) R has been split into R_1 and R_2,
(ii) Gm has been split into G and m, and
(iii) R_1 does not have a specific instantiation yet.

The packaging in Figures 131 and 132 suggests that R_1 and R_2 are separate but that G and m are always packaged together as *-C^{i}*. In fact, I propose that the packaging is slightly different in the K-internal forms in Figure 132. Specifically my hypothesis is that R_1 is packaged with the neighboring feature G. Put differently, the K-internal morpheme *-C* does not correspond simply to G but rather to the combination of R_1 and G. Furthermore, I will posit that the mutator feature m is not packaged with G but rather that it is packaged with K. This means that m helps to spell out the K endings, and moreover it predicts that the K endings contain an *i*-mutator, i.e. $-{}^{i}K_{internal}$. As shown in Figure 133, once we repackage the morphemes in this way for the K-internal forms, the Roll-up Shortcut (see Section 2.2.10) gives us exactly the right results.

R_2	R_1	G	m	K	D
-a	-C		$-{}^{i}K_{internal}$		þa-

=> *þa-*^{i}K-*C-a*, e.g. þa-^{i}t-C-a > *þetta*

Figure 133 Roll-up Shortcut gives correct result

The derivation summarized in Figure 133 makes the following claims about the K-internal forms:

(i) The asigmatic reinforcer *-a* corresponds to R_2 only.
(ii) The geminator morpheme is not *-C^{i}* but rather *-C*, and R_1 happens to be part of the lexical structure of *-C*.
(iii) The *i*-mutator is packaged with K (rather than with *-C*).

Claim (i) is half of the solution to the *-s* ~ *-a* allomorphy problem, the other half being that *-s* corresponds to R_1 plus R_2. The solution to the allomorphy problem, however, has consequences for the packaging of the other morphemes in the K-internal forms, which is evident from claims (ii) and (iii). Though I have no direct morphological evidence for claim (ii) (that R_1 is part of the structure of *-C*), straightforward evidence can actually be given for claim (iii), namely that the *i*-mutator is packaged with the case endings of the K-internal forms. I will discuss this evidence in the next section.

5.2.2 Supporting evidence for packaging m with K in the K-internal forms

So far the K endings of the K-internal forms have been glossed as M.ACC.SG *-n*, N.NOM/ACC.SG *-t*, and M/N.GEN.SG *-s*. I have now made a proposal that these K endings are actually packaged with the *i*-mutator component, entailing that the phonological shapes of these endings are more like $-^{i}n$, $-^{i}t$, and $-^{i}s$. There is independent empirical support for this hypothesis, which I present here.

Consider the Dem paradigm given in Noreen's (1923) classic grammar of ON.

Table 53 Dem paradigm (Noreen 1923: 314; my italics): Unexpected front vowels

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	sú	sá	þat ***þæt*** ***þet***	þær	þeir	þau
ACC	þá	þan(n) ***þæn(n)*** ***þen(n)***	þat ***þæt*** ***þet***	þær	þá	þau
GEN	þeir(r)ar	***þes(s)***	***þes(s)***	þeir(r)a	þeir(r)a	þeir(r)a
DAT	þeir(r)e þæir(r)i	þeim	þuí þí	þeim	þeim	þeim

As seen in the shaded cells in Table 53, Noreen provides some alternate forms which are not typically cited in the ON Dem paradigm, most notably M.ACC.SG *þæn(n)*/*þen(n)* (cf. RN **þin**, **þen**) and N.NOM/ACC.SG *þæt*/*þet* (cf. RN **þit**, **þet**). And while M/N.GEN.SG *þes(s)* does not have more than one variant, recall from Section 1.1.3 that its front vowel *e* is itself already unexpected (or irregular) as well. Note that the shaded forms in Table 53 overlap precisely with the K-internal forms in the RDem paradigm. I have distinguished the forms with unexpected front vowels by putting them in italics and boldface.

What is interesting about the italicized/bolded forms in Table 53 is that there is no immediately obvious reason for them to have the vowels *æ* or *e*. The other forms containing front vowels in the paradigm are expected: *æ* in F.NOM/ACC.PL *þær* is due to *R*-umlaut and the forms with *ei* (F.GEN.SG *þeir(r)ar*, F.DAT.SG *þeir(r)e* and *þæir(r)i*, M.DAT.SG / DAT.PL *þeim*, F.NOM.PL *þeir*, and GEN.PL *þeir(r)a*) are due ultimately to the even older sound change **ai* > *ei*. It should also be noticed that while an alternate spelling with <æ> is given for the F.DAT.SG (i.e. *þæir(r)i*), the normal spelling of this form is also with a front vowel, namely <e> (i.e. *þeir(r)e*). What makes the italicized/bolded forms even more interesting is that other forms in the paradigm with the back vowel *a* (*sá*, *þá*, *þau*) do *not* have front-vowel alternates.

The interesting forms in Table 53 overlap exactly with the K-internal forms in the RDem paradigm, for which we postulated that m packages with K as seen in Figure 133. Instead of labeling the Dem variants 'irregular' (i.e. phonological idioms), hence unexplained, in order to account for their fronted vowels we might invoke the following special case endings for these forms: -^{i}*n*, -^{i}*t*, -^{i}*s*, yielding *þa-*^{i}*n* > *þen*, *þa-*^{i}*t* > *þet*, *þa-*^{i}*s* > *þes*.[93]

In sum, the packaging approach elaborated above to account for the *-s* ~ *-a* alternation makes an interesting prediction about K in the K-internal forms, namely that precisely these K endings (and only these) should show some trace of an *i*-mutator. The prediction is borne out by variants in the Dem paradigm with fronted root vowels. I take this to be support for my packaging hypothesis in the previous section.

5.2.3 The K-final forms and the constant forms: Complex heads and focus movement

In Section 5.1.1, I 'translated' U20 structures into preliminary nanosyntactic structures. I repeat the preliminary structures for the K-final and the constant forms in Figures 134 and 135, respectively.

---

[93] To be more precise, we have a situation in which a single lexical entry can apply to both RDem and the front-vowel Dem variants. By the Superset Principle, the entries for these endings will match both [mP KP] for RDem or just KP for Dem – either way, the *i*-umlauted phonology will be spelled out. Those varieties of ON with non-fronted Dem variants, on the other hand, would have two separate entries: < -^{i}K ⇔ mP KP > for RDem and < -K ⇔ KP > for Dem.

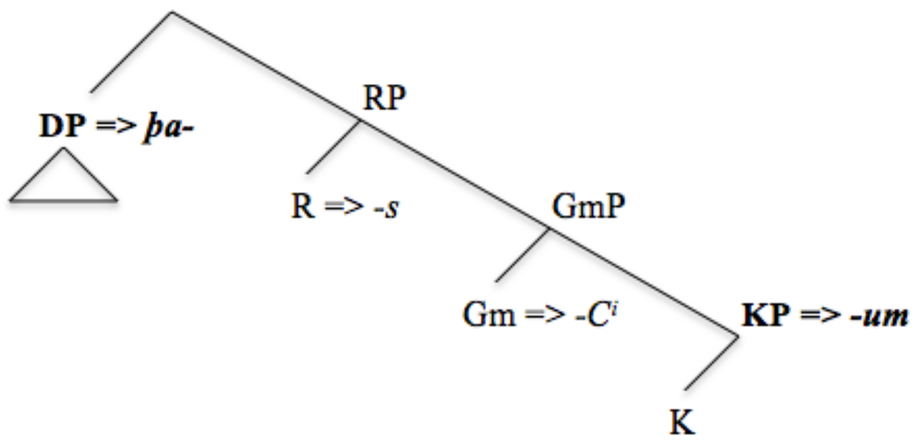


Figure 134 Preliminary nanosyntactic structure of K-final form:
M.DAT.SG þa-s-C^i-um > *þessum*

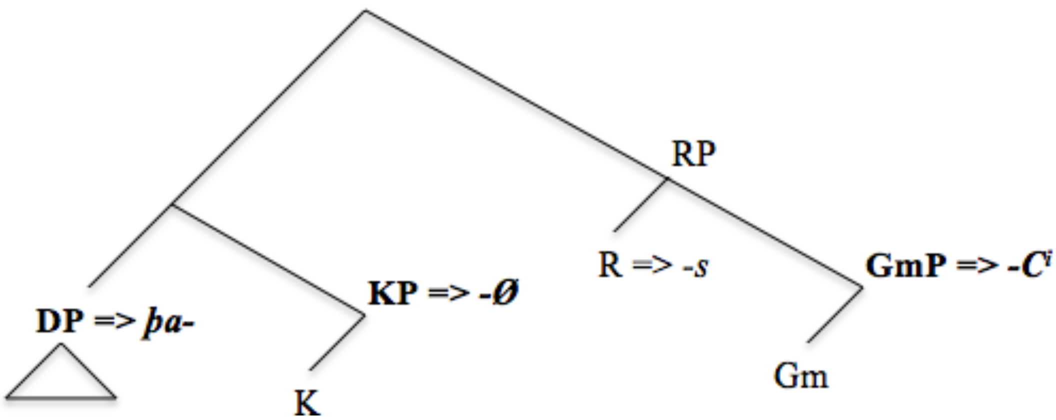


Figure 135 Preliminary nanosyntactic structure of constant form:
F.NOM.SG / N.NOM/ACC.PL þa-Ø-s-C^i > *þessi*

Again, a major question that arose in relation to the derivations elaborated in Section 5.1.2 was why the geminator -C^i would have to spell out as a terminal in the K-final forms but as a phrase in the constant forms. Exploring this question ultimately led us to the conclusion that there are no morphemes in the RDem paradigm that have L-trees consisting of a single terminal.

Though I have come to the conclusion in the preceding section that the structures in Figures 134 and 135 are in fact inadequate, they still have something important to tell us. The terminal vs. phrasal spellout of -C^i is actually the result of a generalization about pied-piping. In the K-final forms (Figure 134), DP does not pied-pipe KP when it moves to the top of the tree. Since KP is left at the bottom of the tree, it entails that GmP cannot constitute a constituent all by itself, and thus the phrase GmP (to the exclusion of KP) cannot be targeted for spellout but the terminal Gm would have to be targeted instead. In the constant forms (Figure 135), DP does pied-pipe KP on its way to the top of the tree.

Because KP has been moved, GmP ends up being a constituent that can be targeted for spellout.

Even though we have divided both Gm and R into two separate features, the position of KP in the K-final forms will still block G/m and R_2/R_1 from being independent constituents, as illustrated in Figure 136.

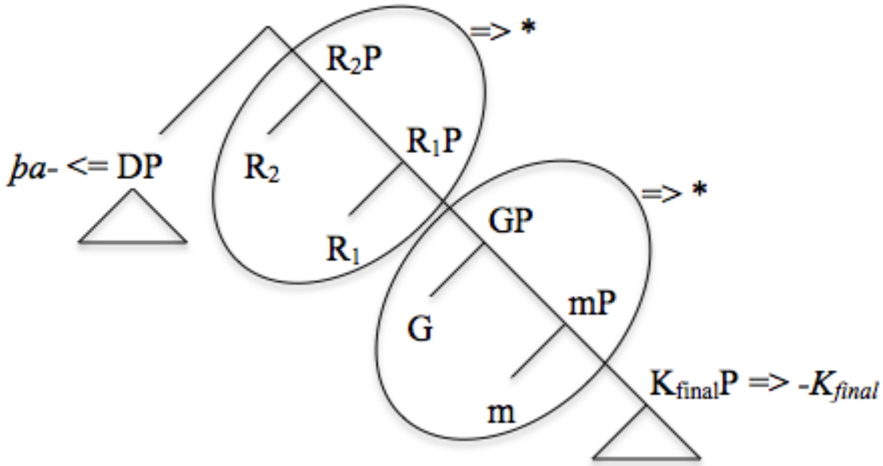


Figure 136 Phrasal non-constituents in the K-final forms

In order to make R_2/R_1 and G/m into constituents which can be properly spelled out, we need to posit that these are complex heads, i.e. [R_2 R_1] and [G m], as shown in Figure 137.

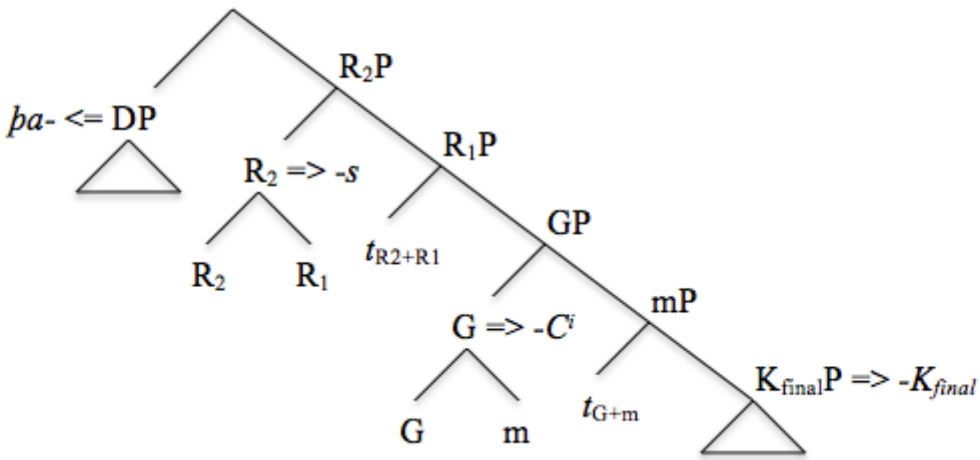


Figure 137 Complex heads in K-final forms

We have postulated that in the K-final forms the sigmatic reinforcer *-s* and the geminator $-C^i$ are complex heads. The lexical entries for these two morphemes are shown in (135).

(135) < *-s* ⇔ [R_2 R_1] >

< $-C^i$ ⇔ [G m] >

These morphemes must also be active in the constant forms, which just like the K-final forms display the sigmatic reinforcer and the geminator. The structure of a constant form is seen in Figure 138.

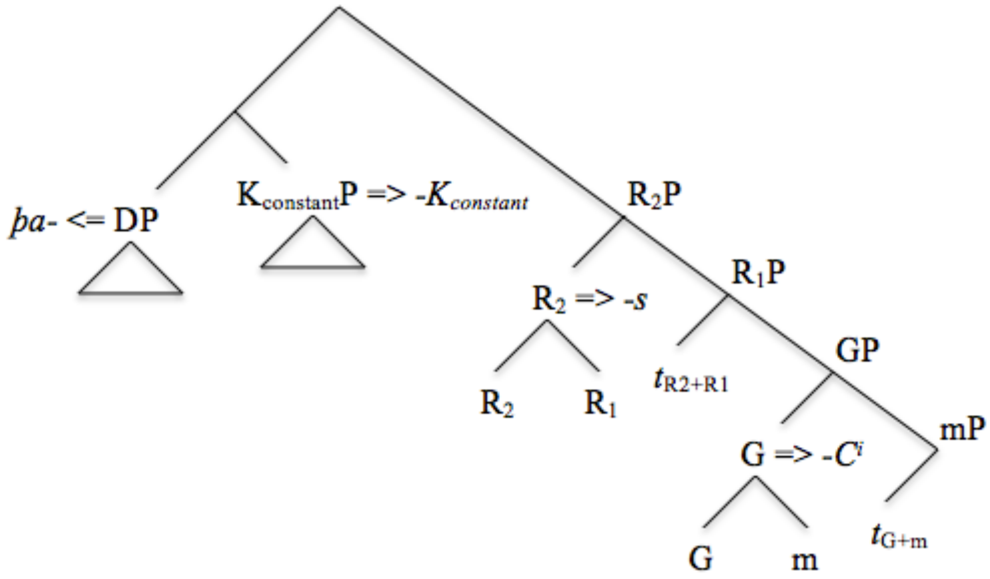


Figure 138 Constant form with complex heads

We can sum up the basic structures of the K-final and the constant forms by the packaging schema in Figure 139. As mentioned in Section 2.2.10, I indicate complex head status in a packaging diagram by a double-edged border.

R_2	R_1	G	m	K	D
-s		$-C^i$		$-K_{final}$ $-K_{constant}$	þa-

Figure 139 Packaging in K-final and constant forms

Recall from Chapter 2 that complex heads are prefix-like elements: they spell out autonomously without the need for structure underneath to undergo spellout-driven movement. Thus, the derivation of a K-final or constant form is expected to look like Figure 140.

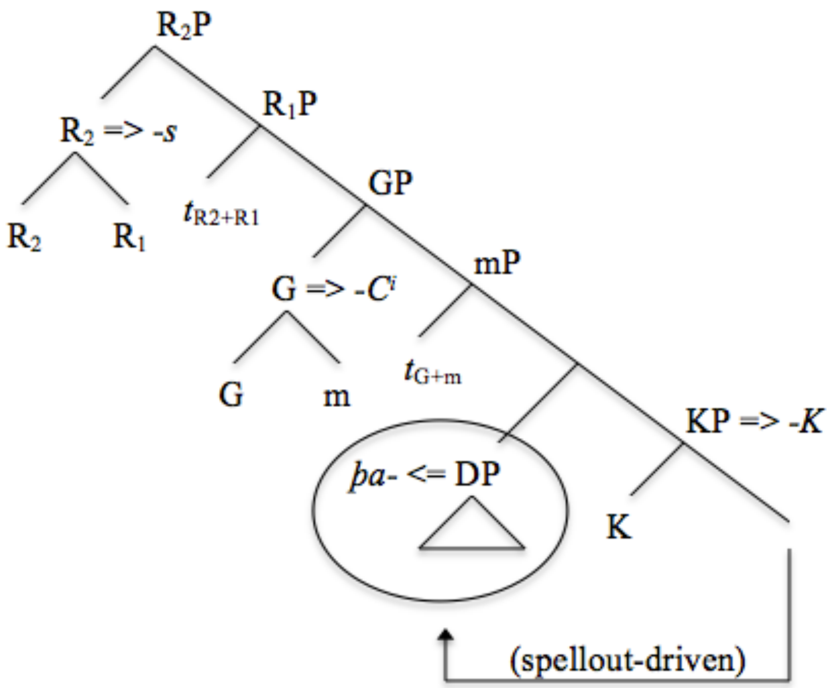


Figure 140 Spellout-driven movement of DP reaches only past KP

As illustrated in Figure 140, the movement of DP to the left of KP is needed for the purposes of spellout. On the other hand, the complex heads -*s* and -C^i can spell out to the left of DP and KP without the need for any spellout-driven evacuations from below. In other words, DP will be stranded in a position below the complex heads. While in principle this is a perfectly admissible structure in terms of spellout, it obviously is not appropriate for our purposes. As seen in Figure 140, we would derive the ominous form **s*-C^i-*þa*-K for the K-final and constant forms. In order to obtain the observed forms with -*s* and -C^i morphemes as suffixes, we still need the movements sketched in Figure 141.

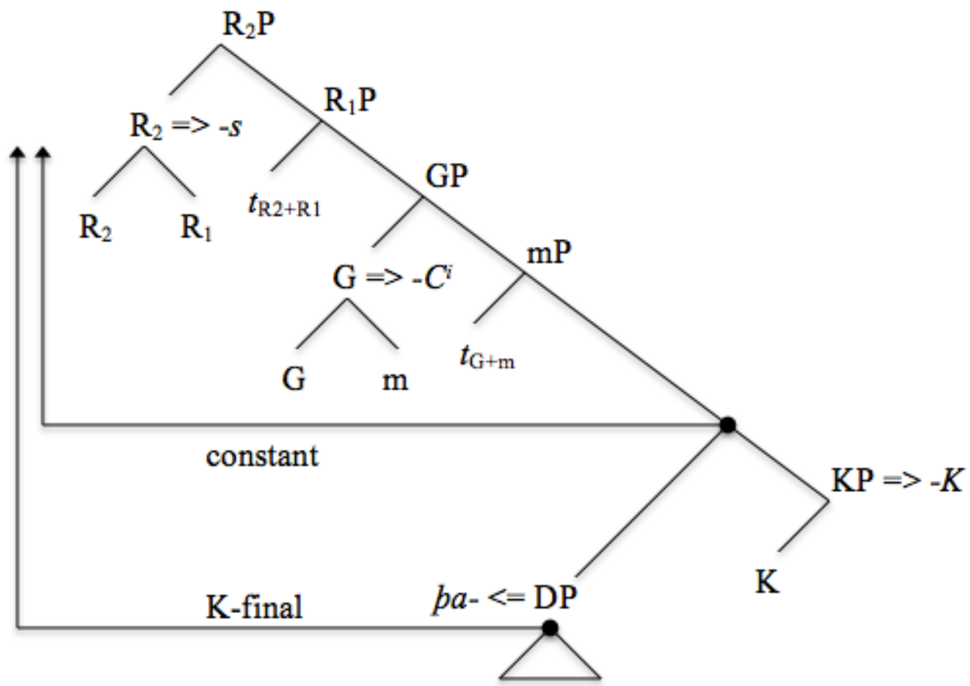


Figure 141 Movements still needed for K-final and constant forms

In other words, in the K-final forms we still need DP to move to the left of *-s*, and in the constant forms we still need [[DP] KP] to move to the left of *-s*.

The movements in Figure 141 are most likely *not* spellout-driven, since *-s* and *-C^i^* can spell out where they are without the need for spellout-driven movement from below. Instead we need to view the movement in Figure 141 as a type of movement that is driven by feature identity or, as it is called in Minimalism (Chomsky 1995, 2000, 2001), Agree. In Minimalism, syntactic movement is the result of a feature *probe* that searches within the structure it c-commands for a matching feature, a *goal*, with which it establishes a so-called Agree relation. This featural relationship can then result in syntactic movement of the goal (if the so-called EPP feature is also present). Since individual features do not move, this will usually involve a 'generalized pied-piping' of additional material outside of just the goal feature (Aboh 2004b). The classic case of syntactic movement is *wh*-movement, where a [wh] feature on a head in the C-domain causes the displacement of a lower phrase also carrying a [wh] feature.

(136) (a) [$_{CP}$ C$_{\mathbf{wh}}$ [$_{IP}$ the Vikings discovered [which continent]$_{\mathbf{wh}}$ before Columbus]]?

(b) [$_{CP}$ [which continent] [$_{C'}$ did [$_{IP}$ the Vikings discover ~~[which continent]~~ before Columbus]]]?

In (136) there is feature identity between C and *which continent*. The resulting probe-goal or Agree relation between the two results in movement of *which continent* into Spec-CP in languages like English.

For the problem at hand, it is important to notice that demonstratives are closely aligned with *wh*-phrases. According to Starke (2001), who builds on Rizzi (1990, 2001), the class of quantificational elements contains *wh*-words, negation, and focus, which all behave similarly with respect to movement. Now, focus has been argued to exist within nominals.[94] For instance, in Gungbe there is a nominal focus marker *tɛ́* (distinct from the clausal focus marker *wɛ̀*) that is part of the internal structure of the *wh*-word 'what'.

(137) é-tɛ́ wɛ̀ Kòfi xɔ̀?
3SG-foc$_N$ foc$_V$ K. buy
'What did Kofi buy?' (Aboh 2004c: 8, modified slightly)

In (137) I assume that the 3SG *é* has moved up to the left of focal *-tɛ́*. That is, focP attracts the D-element *é*, or in traditional notation: [$_{focP}$ [$_{DP}$ é] [$_{foc}$ -tɛ́ [… t_{DP}]]].

Consider also the contrast in (138), from Spanish (Bernstein 2001: 2-3).

(138) (a) este libro interesante
this book interesting
'this interesting book'

(b) el libro interesante este
the book interesting this
'THIS interesting book'

In (138a) the demonstrative *este* has a neutral interpretation, but in (138b) it has a focused interpretation. Again, a plausible analysis is that focP attracts the DP *el libro interesante*, or in traditional notation: [$_{focP}$ [$_{DP}$ el libro interesante] [$_{foc}$ Ø [$_{DemP}$ este [… t_{DP}]]]].

---

[94] Demonstratives in particular are known to develop into focus particles historically speaking (Diessel 1999: §6.6.2, Heine & Kuteva 2002: 95-96, 108-109, 111-112), and in many languages demonstratives and focus particles are morphologically related or even identical. In Ambulas, for instance, Diessel (1999: 149) reports that the demonstratives *kén* 'this' and *wan* 'that' act as focus particles:

(ix) véte dé wak a **wan** méné kaapuk yéménén
see.and he said ah ***FOC*** *you not you.went*
'He saw him and said, "Ah, so *you* did not go."'

In other words, we would like to say that a demonstrative-internal focP (more precisely foc_DP) attracts a D-element of some kind. This drives syntactic movements resembling *wh*-movement in our RDem structures.[95] In Figure 142 we see that focP attracts DP. This is exactly why we see extraction of a DP constituent to a higher position in Figure 142.

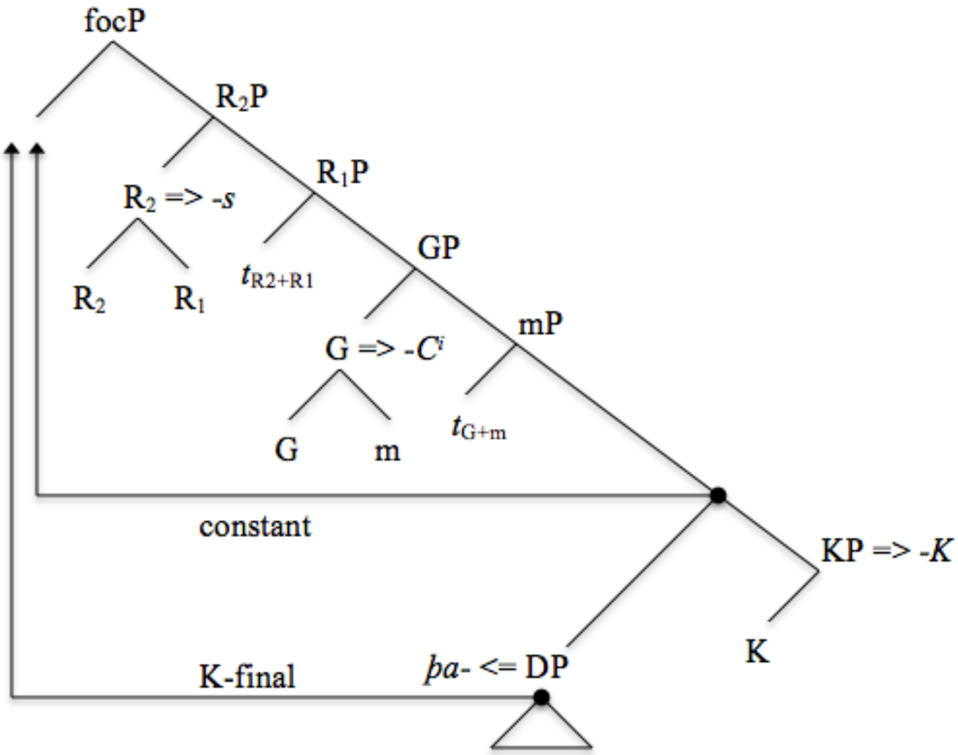


Figure 142 Foc-movement

In other words, there is long-distance movement of DP (± KP), which is a feature-driven, as opposed to a spellout-driven, kind of movement. Foc-movement, then, underlies the

[95] I recognize that this focus position must be different from the position mentioned for Gungbe and Spanish. Aboh (2004c: 10-11) argues that a DP-internal topic can feed topicalization within the clause, since C and D may, as phases, interact at the interface. Consider along these lines the possibility of DP-internal *wh*-movement in Greek (Horrocks & Stavrou 1987), which can then feed movement within CP as well.

(x) Greek DP-internal focus movement feeding movement within CP (Alexiadou, Haegeman & Stavrou 2007: 82)

[$_{CP}$ [$_{DP}$ Tinos$_{wh}$	to	vivlio t_{wh}]	mu	ipes	[$_{CP}$ t_{DP}	pos	dhjavases t_{DP}]].
whose	*the*	*book*	*me.*GEN	*said.*2SG		*that*	*read.*2SG

Analogously we can imagine that a demonstrative-internal foc feature interacts with focus at the level of the entire DP (which then interacts with focalization at the level of CP). This kind of radical nesting is a natural consequence of the highly fine-grained approach taken in nanosyntax.

fact discussed in Section 4.2.1 that all RDem forms display movement of DP to the very top of the structure, i.e. the requirement that all forms begin with *þa-*.

Leaving the cases of the constant and K-final forms aside, we should also ask how foc-movement works in the K-internal forms. I would like to propose what I think is the simplest solution, namely that DP is attracted to focP in these forms as well, but that the movement happens to be string-vacuous. That is, (sub)extracting DP to focP at the last stage of the derivation will not result in a difference in the order of constituents within a K-internal RDem form (whereas in a K-final form it will). This can be seen in Section 5.4.3. In sum, then, foc-movement happens in every single RDem form, but only in certain forms does the movement actually affect the linear order of morphemes.

Though foc-movement may at this point strike the reader as a rather convenient *deus ex machina* for turning prefix-like morphemes in ON (i.e. *-s* and *-C^i*) into suffixes again, we will see in the next chapter that foc-movement of D-elements can explain a robust generalization in the WGmc paradigms (see Section 6.1.5 and 6.2.2 on the ProK generalization), making it essential to our understanding of RDem.

5.3 Pied-piping KP in the constant forms: Two possibilities

As illustrated in Figure 142, my hypothesis is that the K-final forms and the constant forms involve foc-movement of DP. DP, moreover, pied-pipes KP in the case of the constant forms, but does not pied-pipe KP in the K-final forms. The movement of DP alone in the K-final forms is straightforward. As discussed, focP requires a D-element, and since DP is such an element, it moves to focP to satisfy this requirement. Thus the movement of DP, without pied-piping KP, to focP in the K-final forms is easily accounted for. In Figure 143 and henceforth I will mark D-elements with red.

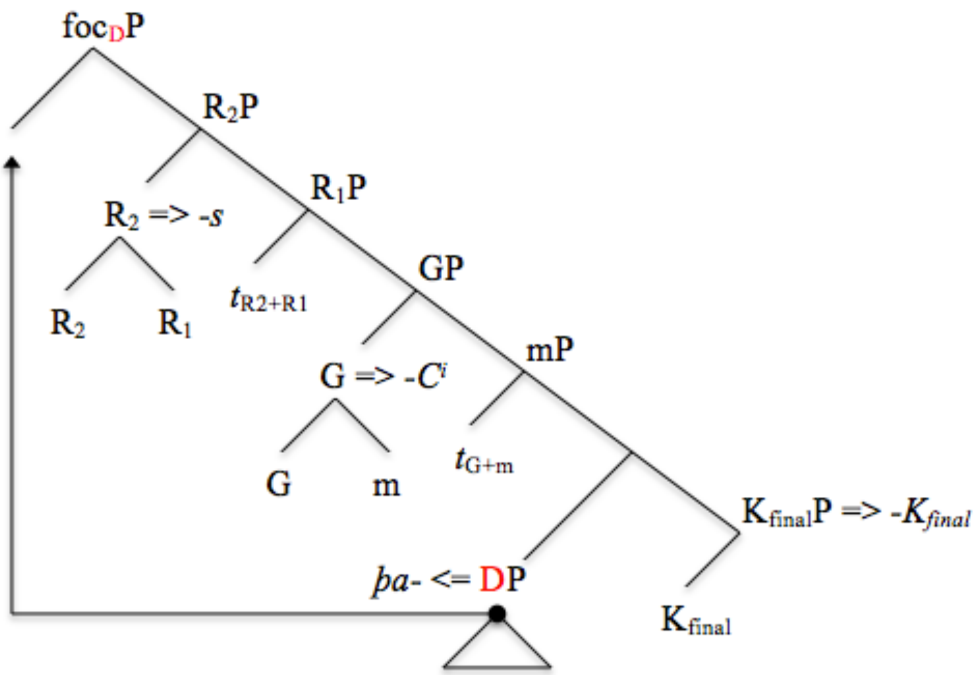


Figure 143 FocP attracts DP in K-final forms

In the constant forms, however, KP needs to move along with DP to focP. In the next chapter we will see that in many of the West Germanic RDem forms, the K ending is pronominal and thus qualifies as a D-element, giving us a clear reason why, in those precise cases, DP would pied-pipe K_DP on its way to focP. In ON, however, K is not pronominal but always adjectival, so KP does not qualify as a D-element. We need another way, then, to handle the facts about KP in the constant forms. In this section I present two possibilities.

5.3.1 A spellout-motivated account

The first account takes a somewhat simplistic approach. The main idea is that there is a special lexical entry (i.e. a phonological idiom) which corresponds to the node containing both DP and KP in the constant forms, i.e. precisely the node that is targeted for movement to focP. The hypothesis is sketched in Figure 144.

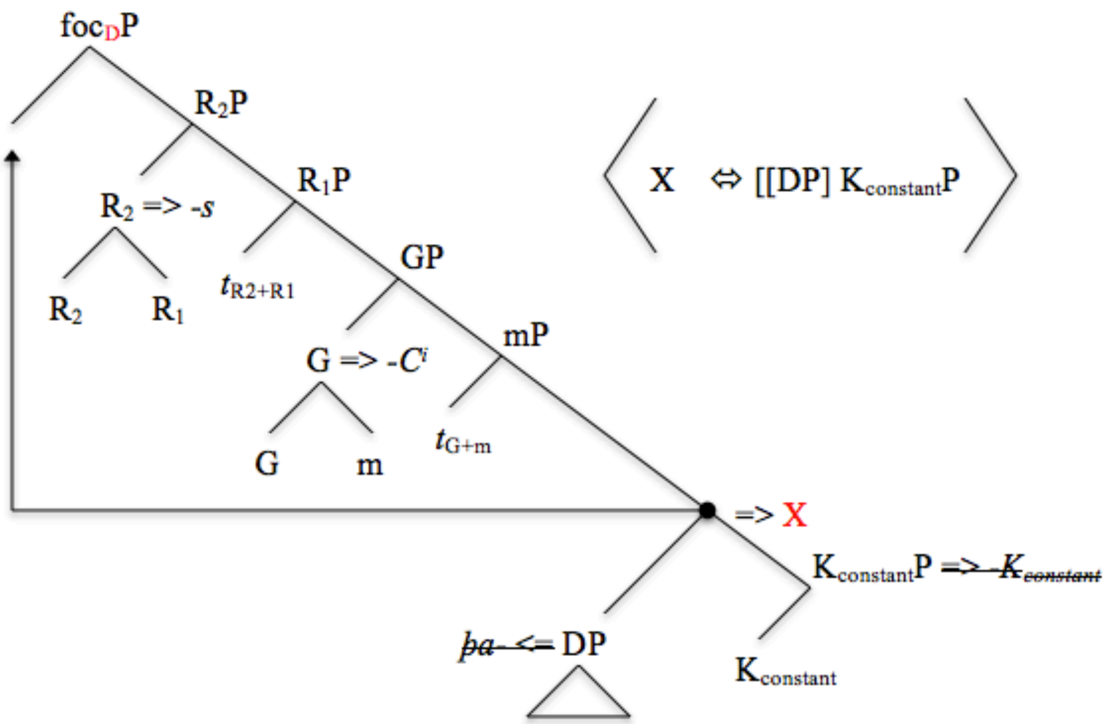


Figure 144 Node at [[DP] KP] is lexicalized as a unit

A good guess for the identity of X in Figure 144 might be the *i*-umlauted base *þe-*. Thus, after the movement in Figure 144, the correct final form is produced: *þe-s-C^i* > *þessi*. (Note that the identity of X cannot be *þa-* since this is already taken by the lexical entry < *þa-* ⇔ DP >.)

We should ask, however, if the existence of a lexical entry like the one in Figure 144 is a good enough reason for focP to move both DP and KP, rather than just DP alone. On the one hand, we might invoke Pantcheva's (2011: 140) principle that direct lexicalization is better than lexicalization by inheritance. In Figure 144, the node containing both DP and KP is lexicalized directly as X, making this option in some sense better than lexicalizing DP and KP separately. By this logic, the reason DP pied-pipes KP when it moves to focP is at least partially due to reasons of spellout. On the other hand, we should keep in mind that focus movement is *not* spellout-driven movement. An argument against the hypothesis sketched in Figure 144, then, is that it gives a spellout-motivated reason for a movement which is driven by feature identity, thereby conflating two types of considerations which should in all likelihood be kept distinct. That is, it should not matter to focP how the constituents it attracts are lexicalized. All that should matter to focP is a constituent's properties, in this case whether or not it is a D-element. This may be reason enough to reject the hypothesis in Figure 144 in favor of the hypothesis presented in the next section.

5.3.2 The nominative stem allomorphy generalization (McFadden 2014)

The second account is inspired by a generalization formulated by McFadden (2014) about the nature of nominative case.

The starting point of the argument is the observation that the constant forms are found in a very restricted area in the paradigmatic space of RDem. In Table 54 the constant forms are those in the shaded cells. As seen in Table 54, the constant forms are nominative forms (cf. the lightly shaded cells) with the minor caveat that there is a NOM/ACC syncretism in the N.PL, giving us a constant form also in the N.ACC.PL (dark shading in Table 54).

Table 54 The constant forms are (essentially) nominative

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	þessi	þessi	þetta	þessar	þessir	þessi
ACC	þessa	þenna	þetta	þessar	þessa	þessi
GEN	þessar	þessa	þessa	þessa	þessa	þessa
DAT	þessi	þessum	þessu	þessum	þessum	þessum

I think this fact about the constant forms is telling, and in fact it fits in well with recent work by McFadden (2014), a basic outline of which I will now give.

A morphological phenomenon observed in various languages is that the nominative forms of certain nouns display a stem (or stem complex) which is distinct from that found in the non-nominative cases. To take two examples, consider Finnish (139) and Latin (140). For additional examples see also Appendix IV. The regular forms are given in (139a) and (140a); the special forms are given in (139b) and (140b).

In Finnish, the NOM.SG of *ihmi-* 'person' (139b) is formed from *ihmi-* plus the special nominative element *-nen* (cf. the regular NOM.SG ending *-Ø* in (139a)). Moreover, the non-nominative forms of *ihmi-* 'person' in (139b) insert a stem-forming morpheme *-se* between *ihmi-* and the K ending. This morpheme is absent in NOM.SG *ihmi-nen*.

(139) Finnish (McFadden 2014: 3, modified slightly)

(a) regular

	'house'	'street'
NOM	talo-Ø	katu-Ø
GEN	talo-n	kadu-n
PART	talo-a	katu-a
INESS	talo-ssa	kadu-ssa

(b) special NOM

	'person'
NOM	**ihmi-nen**
GEN	ihmi-se-n
PART	ihmi-s-tä
INESS	ihmi-se-ssä

In Latin, the NOM.SG of *hom-* 'man' (140b) is formed from *hom-* plus the special nominative element *-ō* (cf. the regular NOM.SG ending *-s* in (140a)). In the non-NOM forms of *hom-*, a special stem-forming element *-in* is inserted between *hom-* and the K ending. The root *gen-* 'kind' in (140b) follows the same pattern (i.e. an irregular NOM.SG ending, with a stem-forming element *-er* inserted in the non-NOM forms), except that there is a syncretism between the NOM and ACC (indeed, neuter nouns in Indo-European are always syncretic between these two cases). The syncretism in N.NOM/ACC.SG *genus* should be familiar from our discussion of the constant forms in Table 54, with the syncretic N.NOM/ACC.PL *þessi*. Due to syncretism, what is normally a nominative-based irregularity surfaces in the accusative as well.

(140) Latin (McFadden 2014: 5, modified slightly) third declension

(a) regular

	'chief'
NOM	princep-s
ACC	princep-em
GEN	princep-is
DAT	princep-ī

(b) special NOM

	'man'	'kind'
NOM	**hom-ō**	**gen-us**
ACC	hom-in-em	**gen-us**
GEN	hom-in-is	gen-er-is
DAT	hom-in-ī	gen-er-ī

On the basis of data like (139) and (140), McFadden (2014) posits the following generalization.

(141) Nominative stem-allomorphy generalization (NSAG) (McFadden 2014: 8)

When there is stem allomorphy based on case, it distinguishes the nominative (along with any cases systematically syncretic with the nominative) from all other cases.

McFadden's (2014) account for the NSAG is couched in Distributed Morphology and inspired primarily by Moskal (2013). The basic idea is that locality (with regard to # and K morphemes) can explain how little *n* (the stem-forming ingredient) surfaces. In some morphological environments, *n* surfaces as Finnish *-nen* or Latin *-ō*; in others it surfaces as Finnish *-se* or Latin *-in*. This is sketched in Figure 145.

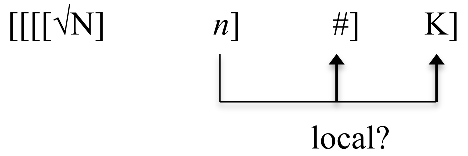


Figure 145 Local relation between *n* and #/K

In Appendix IV I discuss his proposal in much more detail and the various issues arising from it. I also provide my own nanosyntactic account of the NSAG.

For our purposes here, it will be sufficient to present the basic idea behind my own analysis, namely that the NSAG should be thought of as a tendency for lexical packaging. More specifically, I would like to propose that the NSAG is a reflection of the fact that, in line with Caha's (2009) theory of K, *n* is merged closer to the feature K_1 than to the other K-layers, as shown in Figure 146.

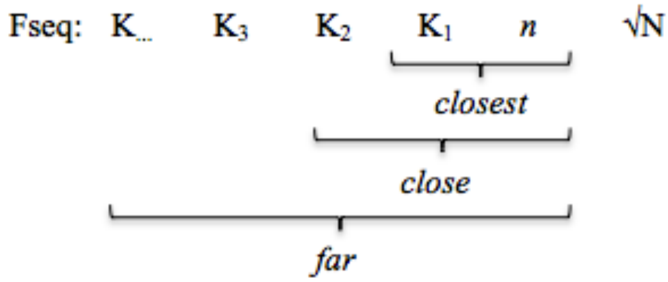


Figure 146 Distance between *n* and K heads

Since *n* is closest to the feature K_1 (the feature responsible for nominative case), it is more likely that *n* will be packaged with K_1 than with the other K features. Moreover, K_2 (responsible for accusative case) is closer to *n* than the K features above it, but still not as close to *n* as K_1 is. Thus it is again more likely that K_2 will be packaged with *n* than that K_3 or above will be. In other words, an 'irregular' *n* will be most likely to implicate nominative (K_1) in its irregularity, a little less likely to implicate the accusative ($K_1 + K_2$) as well, and even less likely to implicate cases beyond the accusative ($K_1 + K_2 + K_{...}$).

My hypothesis, then, is that the reason we observe 'irregular' nominative forms in many languages is because these forms display portmanteau morphemes consisting of *n* and K_1. My analysis of Finnish and Latin is summarized in Figures 147 and 148.

K_1	*n*	N
-nen -ō		ihmi- hom-

Figure 147 Finnish and Latin NOM.SG

K...	K_1	*n*	N
non-nominative endings		-se -in	ihmi- hom-

Figure 148 Finnish and Latin non-nominative

In sum, the NSAG can be understood as the tendency for nominative case, the least complex case structurally speaking, to be lexically packaged with its neighbors in a bigger chunk than the non-nominative cases tend to be packaged in. This gives us the split between bimorphemic nominatives such as NOM.SG *ihmi-**nen*** or *hom-**ō*** vs. trimorphemic non-nominatives such as GEN.SG *ihmi-**se-n*** or *hom-**in-is***. Sometimes, furthermore, the structural chunk consisting of *n* and K_1 also includes the next feature up, K_2, which is responsible for accusative case. This happens in those forms where nominative and accusative are syncretic. Since K_2 is a little bit farther from *n* than K_1 is, though, this kind of packaging is less common. Basically then, the fact that nominative forms are more often irregular than the other case forms follows from its location in the fseq as the lowest case in Caha's (2009) K hierarchy.

The question now is what this might tell us about the ON RDem paradigm, more specifically about the constant forms. Three of the four constant forms in the ON paradigm are nominative forms, with the fourth one an accusative syncretic with the nominative. This pattern conforms to McFadden's (2014) NSAG very well, so I pursue the idea that the constant forms show some manifestation of the kind of lexical packaging shown in Figures 147 and 148.

My main proposal for the constant forms is to package K with the reinforcer features, as shown in Figure 149. That is to say, the nominative (and syncretic accusative) constant forms are the result of a large chunk of structure being stored as a single unit *-ss^i^*, analogous to Finnish *-nen* or Latin *-ō*. In Figure 149 I include Φ(P) as a shorthand for the complex of person, number, and gender features which in a nanosyntactic approach would have to be present somewhere in the vicinity of K, most likely below it.[96]

[96] It is likely that there is even an fseq each for gender, number, and person features. Φ paradigms are notoriously complicated crosslinguistically, however, so time is needed to formulate a nanosyntactic theory of Φ. See

<table>
<tr><th>R_2</th><th>R_1</th><th>G</th><th>m</th><th>$K_{NOM/ACC}$</th><th>$Φ_{constant}$</th><th>D</th></tr>
<tr><td colspan="6">-ss^i</td><td>þa-</td></tr>
</table>

Figure 149 Packaging in the constant forms

Applying the Roll-up Shortcut to Figure 149 gives the result *þa-ss^i*, which correctly gives *þessi*.

The lexical entries for the constant forms are given in more detail in (142).

(142) Lexical entries for constant forms

$<_{324}$ -*Ø* ⇔ K_2P K_1P [$_{ΦP}$ F.SG / N.PL] $>$

$<_{325}$ -*r* ⇔ K_1P [$_{ΦP}$ M.SG] $>$

$<_{869}$ -*ss^i* ⇔ R_2P R_1P GP mP [324 / 325] $>$

As seen in (142), the entry for -*ss^i* will have to refer to the K and Φ features of M.NOM.SG (entry 325) and F.NOM.SG / N.NOM/ACC.PL (entry 324), i.e. it will be a pointer entry. This is shown in entry 869.[97] For pointer entries see Section 2.2.3.2.

The basic final structure of a constant form is shown in Figure 150.

Taraldsen (2009), Caha & Pantcheva (2012), Van Craenenbroeck (2012), Rocquet (2013), and Vanden Wyngaerd (2014) for some discussion.

[97] I will gloss over the not entirely innocent assumption that entry 869 can point to 324 *or* 325. Since various questions about pointers still need to be resolved, I will allow myself this liberty.

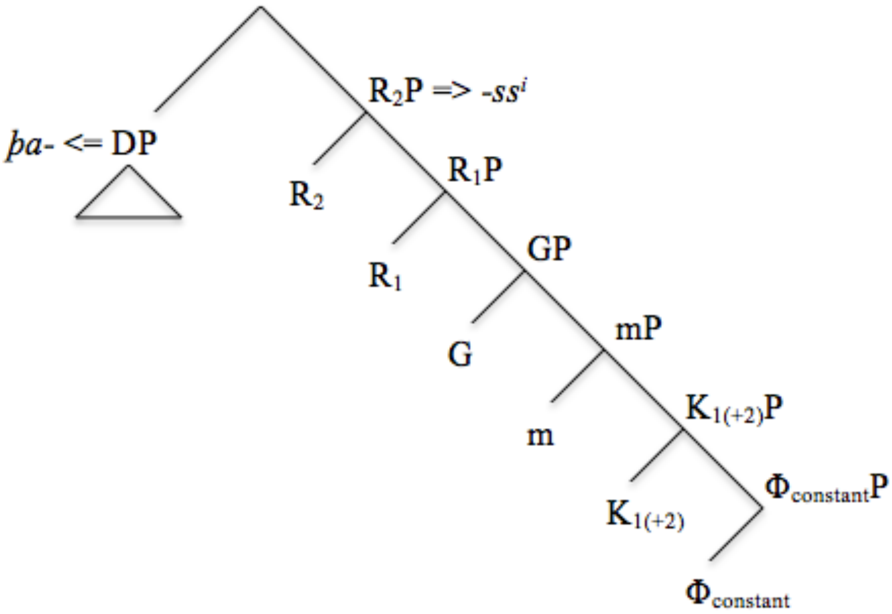
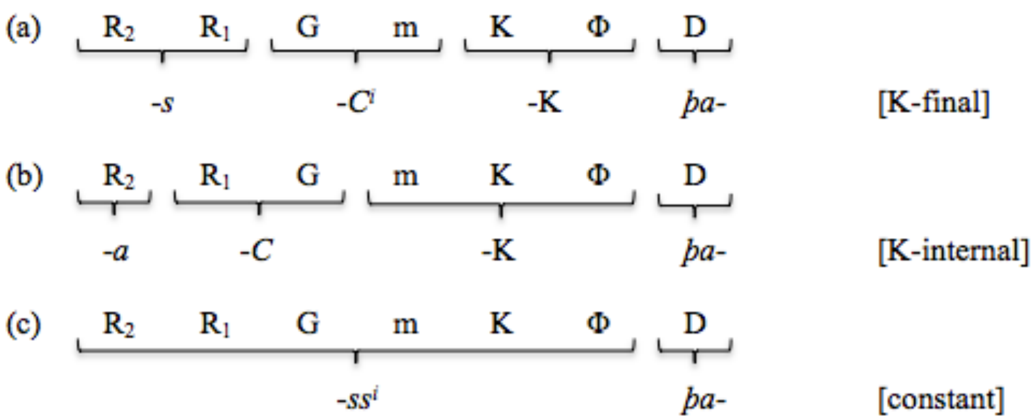


Figure 150 Final structure of a constant form

One remaining issue is that we now seem to be abandoning the template D-K-R-Gm discussed at length in Chapters 3 and 4. The structure in Figure 150 looks more like the K-final template, in fact: D-R-Gm-K. There are two important points to be made regarding this result. First, it is not entirely true that the 'template' in Figure 150 is D-R-Gm-K. The concept of templates that was introduced in Chapter 3, and expanded upon in Chapter 4 using a Cinquean framework, is not directly transferrable to a formal system that allows for phrasal spellout, such as nanosyntax. It is not clear how useful it is to speak of a template D-R-Gm-K when in nanosyntactic terms we just have two phrasal chunks (one for *þa-* and one for *-ss^i*). In nanosyntax, we are more concerned with patterns of lexical storage. Looking at it this way, we have not, in fact, reduced our templates from three to two – we still have three 'templates' in the RDem paradigm because there are still three distinct ways to package the fseq, as seen in (143).

(143) Lexical entries



Three patterns still exist in the RDem paradigm, nanosyntactically speaking.[98]

Observe at this point that the K components in all three templates are anchored at Φ, with each template's K component lexically extending up to a different layer in the sequence. It is because each template's K component has a different shape that the three distinct RDem patterns emerge within the same paradigm: as discussed in Section 2.2.10, the shape of a morpheme which is lower in the fseq will determine which morphemes are spelled out later on. As an example, consider the difference between the F.GEN.SG form *þess-ar*, which is a K-final form, and the M/N.GEN.SG form *þessa*, which is a K-internal form.

(144) Lexical entries for K endings

(a) < *-rar* ⇔ K_3P K_2P K_1P $[_{\Phi P}$ F.SG$]$ > (K-final ending)

(b) < *-s* ⇔ mP K_3P K_2P K_1P $[_{\Phi P}$ M/N.SG$]$ > (K-internal ending, which packages m with K)

Now, if the syntax builds $[_{\Phi P}$ F.SG$]$, then the lexical entry for *-s* (144b) will not be a proper match since it contains $[_{\Phi P}$ M/N.SG$]$ instead. Hence the entry for *-rar* (144a) will be selected over the one for *-s*. Once the syntax has built up to K_3P, the next layer (m) will trigger the new morpheme -C^i, because the entry for *-rar* ends at the K/m boundary (see (143a)). After -C^i, the next morpheme to be triggered will be the sigmatic reinforcer *-s*, since -C^i ends at G and *-s* begins at R_1. Thus the anchoring of the K morpheme has a cascade effect all the way up the sequence.

Imagine instead that the syntax builds $[_{\Phi P}$ M/N.SG$]$. Now the lexical entry for *-rar* (144a) will not be a proper match since it is anchored at $[_{\Phi P}$ F.SG$]$. Hence the entry for *-s* (144b) will be selected over the one for *-rar*. The syntax again builds up to K_3P, matching the structure with the entry for the K ending *-s* all along the way. When it reaches the m layer, moreover, *-s* is still a match because, as a K-interal ending, *-s* packages m with its K

---

[98] Still, something should be said about the arguments in Chapter 3 that the constant forms have internal inflection, i.e. *þa-r-s-*C^i > *þessi*, where *rs* assimilates to *ss*. This is still true, but within a more refined historical chronology. The phonological rule that *rs* assimilates to *ss* was regular in PN (cf. Chapter 1). As discussed in Chapter 3, it was also active in (some grammatical domains of) ON, as seen in the passive/middle forms *finnr-sk* > *finnssk* > *finnsk* 'is found' (Barnes 2004: 144). By the time we reach modern Icelandic, however, this assimilation process does not apply anymore (see Anderson 1988: 6-7), yet we still observe NOM.SG *þessi* in this language. My nanosyntactic proposal is advantageous, then, by obviating the need to posit a completely regular process of *rs* > *ss*, since in my analysis the underlying form is simply *þa-ss*^{i} > *þessi*. Furthermore, it is important to note that the 'loss' of the template D-K-R-Gm does not weaken the way I deduced the correct fseq in Chapter 4. There is no candidate fseq from Chapter 4 that is excluded solely on the basis of the D-K-R-Gm template. In fact we need the template D-K-R-Gm in any case, since this is what the older Dem-*si* forms amount to, e.g. M.DAT.SG $[_{DP}$ *þei-* $[_{KP}$ *-m* $[_{RP/GmP}$ *-si*$]]]$.

features. At the next layer (G), however, the new morpheme -*C* will be triggered, which extends up to R_1. And because -*C* ends at R_1, the next morpheme to be triggered will be the asigmatic reinforcer -*a*, since -*C* ends at R_1 and -*a* begins at R_2. Again, the anchoring of the K morpheme and the way it packages the fseq resonates all the way up the sequence.

The second thing to be said about Figure 150 and my proposed portmanteau composed of the R, Gm, and K ingredients is that it nicely captures certain Old Norwegian forms. In Old Norwegian, the constant forms developed into M.NOM.SG *þesser* and the F.NOM.SG / N.NOM/ACC.PL *þessor* (the form *þessor* was also found in Old Icelandic/ON in the N.NOM/ACC.PL) (Axelsdóttir 2003: 68, Katrín Axelsdóttir p.c.). I will discuss these two forms in turn.

In the case of M.NOM.SG *þesser* it is clear that this is a K-final form, with the M.NOM.SG ending -*r* overtly visible, i.e. *þesse*-***r***. My hypothesis about the structure of -*ss*^{i} can make sense of this diachronic shift from constant to K-final, since KP is at the very bottom of the portmanteau [R_2P R_1P GP mP KP] already. The change from a constant form to a K-final form, then, is simply a matter of spelling out a subtree of the portmanteau separately, as sketched in Figure 151.

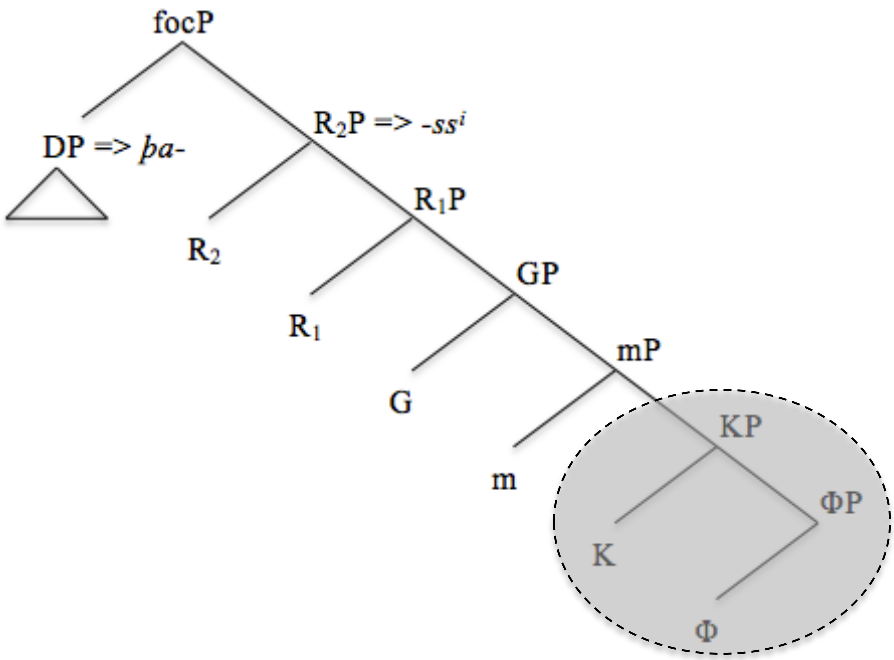


Figure 151 Constant form with incipient change

Spelling out KP separately, moreover, changes the anchoring of the morpheme -*ss*^{i}, completely disrupting its spellout. This leads to a wave of change in the rest of the structure, essentially splitting -*ss*^{i} into the parts -*s* and -*C*^{i} and ultimately resulting in a K-final form, as seen in Figure 152.

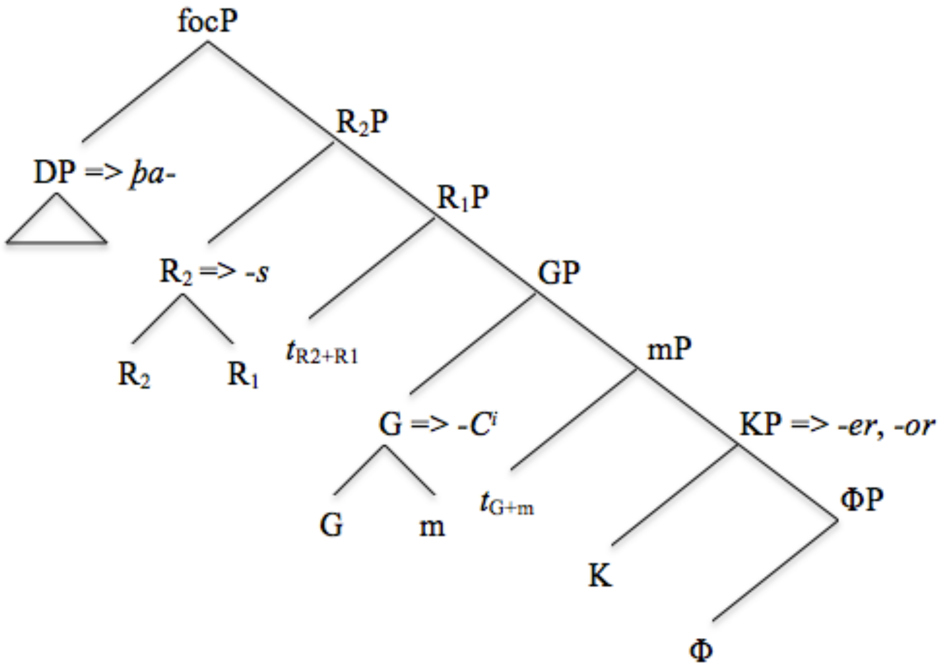


Figure 152 Old Norwegian K-final form

Thus the lexical structure of *-ss*^{i} that I have proposed here puts KP (and ΦP) in such a position that we can easily explain the shift of the constant form *þessi* into the K-final form *þesser*.

Let us now turn to the Old Norwegian variant F.NOM.SG / N.NOM/ACC.PL *þessor*. Importantly, this form shows the ending *-r* despite the fact that we expect a null ending *-Ø* in the F.NOM.SG / N.NOM/ACC.PL. I will pursue the hypothesis that *-r* in these Old Norwegian dialects corresponds purely to K – rather than being a spellout of both K *and* Φ, which is normally the case in the K endings encountered so far. In support of this idea, consider the fact that we also see *-r* in the F.NOM/ACC.PL case ending *-ar* as well as the M.NOM.PL ending *-ir*. Thus we are dealing with bimorphemic case endings: M.NOM.SG *-e-r*, F.NOM.SG / N.NOM/ACC.PL *-o-r*, F.NOM/ACC.PL *-a-r*, and M.NOM.PL *-i-r*. The second component *-r* can be taken to correspond to nominative (and sometimes syncretic accusative), while the first component, the vowel, can be taken to correspond to Φ features. More specifically, *-e* in *-e-r* is the spellout of [$_{\Phi P}$ M.SG], *-o* in *-o-r* is the spellout of [$_{\Phi P}$ F.SG / N.PL], *-a* in *-a-r* is the spellout of [$_{\Phi P}$ F.PL], and *-i* in *-i-r* is the spellout of [$_{\Phi P}$ M.PL]. As illustrated in Figure 153, we are essentially ‘factoring out’ K in these four endings.

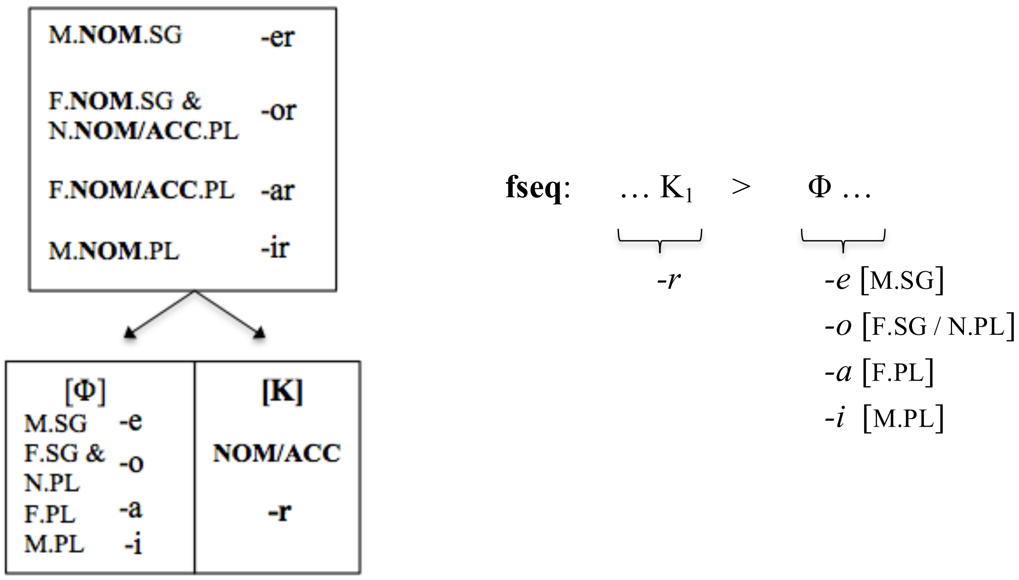


Figure 153 Decomposing some Old Norwegian case endings

This idea – that these particular case endings can be decomposed into separate K and Φ components – may very well also apply to other case endings. I will not pursue a full decomposition along these lines, beyond these endings from Old Norwegian.

Consider now the derivation of the Old Norwegian variants in Figure 154. In the course of spellout-driven movement, DP first moves to the left of ΦP and then the constituent [[DP] ΦP] moves to the left of KP.

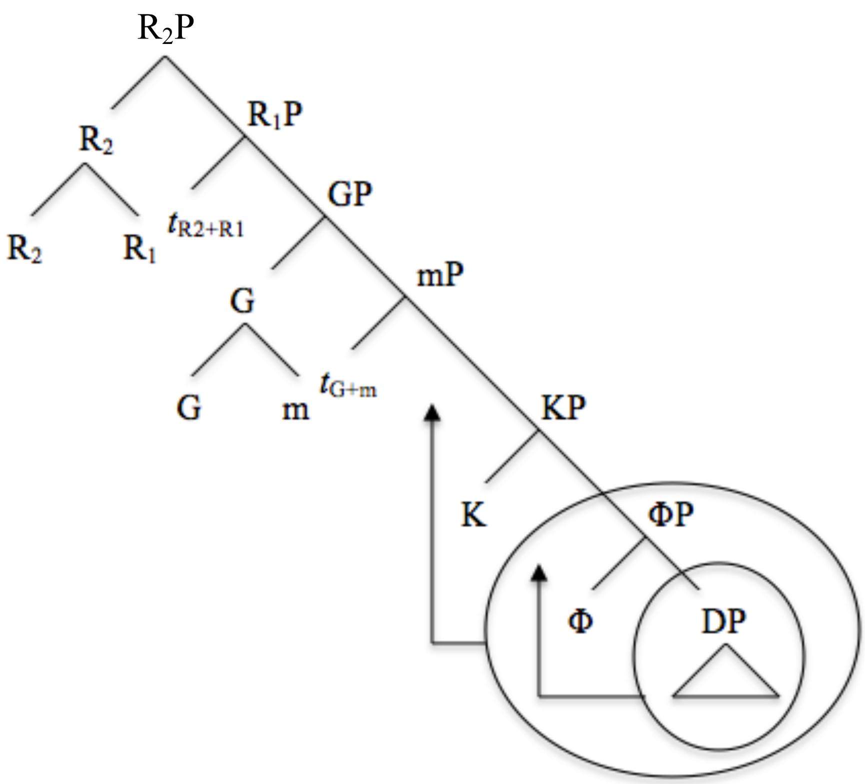


Figure 154 Spellout-driven movements in lower part of the tree

The spellout-driven movements in Figure 154 result in constituents DP, ΦP, and KP, each one corresponding to its own morpheme. Next, focP attracts the D-element DP. As seen in Figure 155, this amounts to subextraction of DP out of [[DP] ΦP].

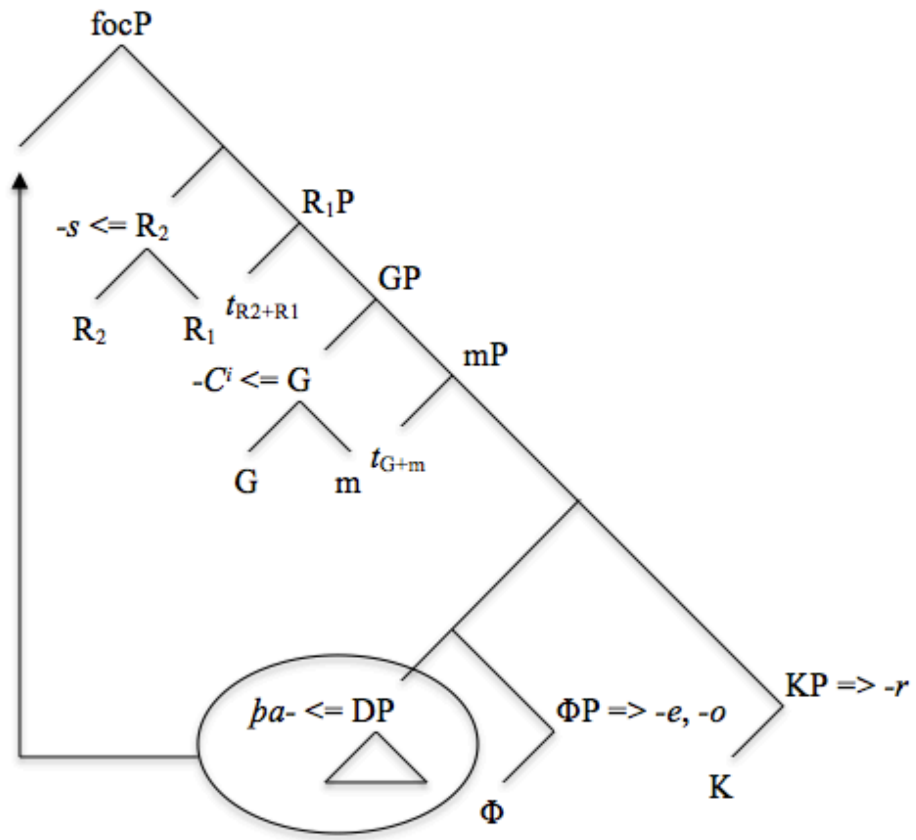


Figure 155 Foc-movement requires subextraction of DP

In Section 4.3 I pointed out that subextraction structures appear to be unattested in ON. On the one hand this is not a bad outcome since Cinque (2005) also finds very few good candidates for subextraction in his crosslinguistic survey of Dem, Num, A, and N. On the other hand, as discussed in Section 2.2.9, subextraction makes up an important part of nanosyntactic 'peeling' analyses (e.g. Caha 2009: Ch. 4, Rocquet 2013). So it is perhaps not unexpected that Old Norwegian provides evidence for a derivation requiring subextraction. As we shall see shortly, in Section 5.4.3, foc-movement of DP in the ON K-internal forms is also a case of subextraction.

To summarize the present section, I conclude that the portmanteau hypothesis for the constant forms, inspired by McFadden's (2014) NSAG, also provides us with an elegant account for the transition from *þessi* to *þesser*/*þessor* in Old Norwegian. In the next section I return to ON and provide explicit derivations of its RDem forms.

5.4 Derivations step by step

In this section I present more detailed derivations of the ON RDem forms using the lexical entries posited in the preceding sections. In Section 5.4.1 I show the derivation of a constant form. In Section 5.4.2 I show the derivation of a K-final form. In Section 5.4.3 I show the derivation of a K-internal form.

5.4.1 Derivation of a constant form

For the derivation of the constant form, let us take the M.NOM.SG form as an example. To enable us to illustrate the idea of competition between multiple entries in a single lexicon, the forms discussed in Sections 5.4.2 and 5.4.3 will also be drawn from the M.SG column and included in the lexicon in all of the examples below. The lexicon, of course, is not complete, having been simplified for the purposes of exposition. Note that I also include the projection ΦP in the structures below. Though there is most certainly internal featural structure to the Φ domain, I abstract away from it here by representing this feature complex as a single head.

(145) **Lexicon**

<$_{210}$ *þa-* ⇔ DP >

<$_{325}$ *-r* ⇔ K_1P [$_{\Phi P}$ M.SG] >

<$_{327}$ -^{i}*n* ⇔ mP K_2P K_1P [$_{\Phi P}$ M.SG] >

<$_{332}$ *-um* ⇔ K_4P K_3P K_2P K_1P [$_{\Phi P}$ M.SG] >

<$_{865}$ *-a* ⇔ R_2P >

<$_{866}$ *-C* ⇔ R_1P GP >

<$_{867}$ *-C*^{i}*-* ⇔ [G m] >

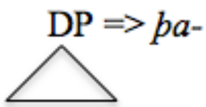
<$_{868}$ *-s-* ⇔ [R_2 R_1] >

<$_{869}$ *-ss*^{i} ⇔ R_2P R_1P GP mP [324 / 325] >

The first steps in the constant derivation involve building DP. At STAY entry 210 matches the syntactic structure, spelling out *þa-*, as seen in (146).

(146) Built up to DP

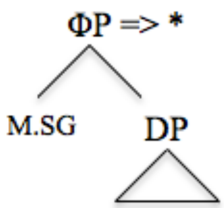
STAY



Next the Φ features are added, which here we will represent simply as [$_{\Phi P}$ M.SG]. The first step in the algorithm, STAY (147a), does not lead to a match with anything in the lexicon, because the constituent [ΦP [DP]] does not exist as such in the lexicon in (145). The next step, CYCLIC, is not available, since there is no sister of the daughter of Φ in the structure (147b). The final step, SNOWBALL, results in the availability of a new constituent, namely ΦP. This structure is a subset in the L-trees of entries 325, 327, 332, and 869. By the Superset Principle, then, there are four lexicalization possibilities. By the Elsewhere Condition, these entries are narrowed down to entry 325, since the L-tree in this entry has the least amount of superfluous structure when compared to the S-tree ΦP. Entry 325 is linked to the phonology *-r*, so ΦP is spelled out as such (147c).

(147) Add Φ

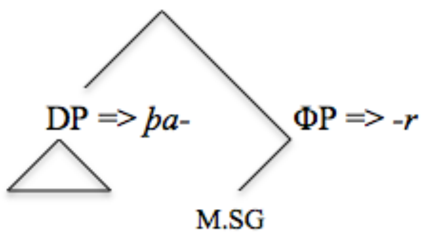
(a) *STAY



(b) *CYCLIC

N.A.

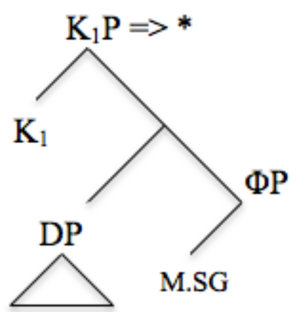
(c) SNOWBALL



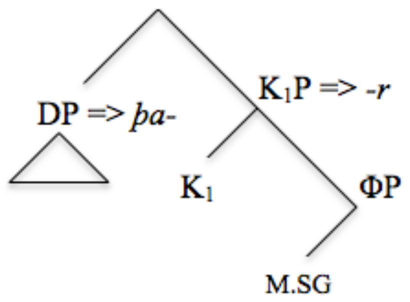
The next feature to be added is K_1. STAY again does not lead to a match, because there is no lexical structure that corresponds to [K_1P [DP ΦP]] (148a). CYCLIC does lead to a match though, because after DP has moved to the left of K_1 there is a constituent [K_1P [ΦP]] that can be matched by entries 325, 327, 332, and 869. Once again, 325 is the best match and [K_1P [ΦP]] spells out as -*r* (148b).

(148) Add K_1

(a) *STAY



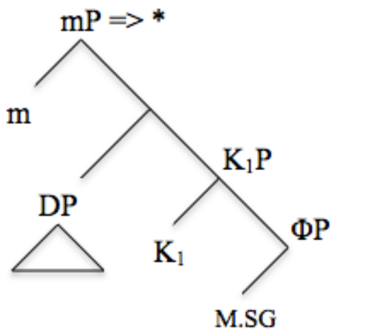
(b) CYCLIC



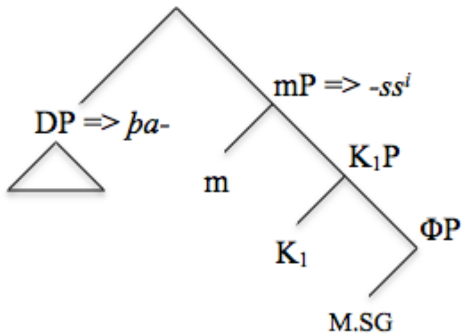
The next feature added to the structure is m. STAY does not produce a match, because there is no L-tree matching [mP [[DP] K_1P ΦP]] (149a). CYCLIC does result in a match, since entry 869 contains the structure [mP [$_{325}$ K_1P [$_{ΦP}$ M.SG]]], and thus -*ss*^{i} is spelled out (149b).

(149) Add m

(a) *STAY



(b) CYCLIC

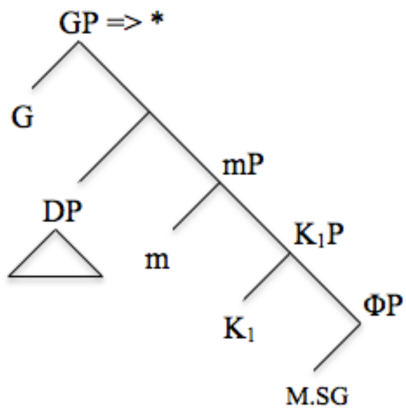


By Cyclic Override, the spellout of *-ss*^{i} here overrides the previous spellout of *-r*.

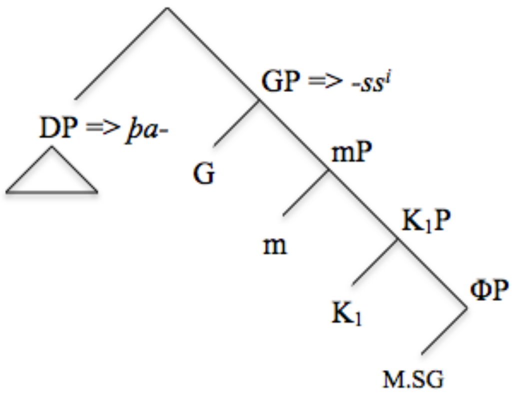
Next G is added. STAY fails to lexicalize G because there is no L-tree corresponding to the entire structure [GP [DP mP K_1P ΦP]] (150a). At CYCLIC, however, entry 869 is a match once again, this time for [GP [mP [K_1P [ΦP]]]] (150b).

(150) Add G

(a) *STAY



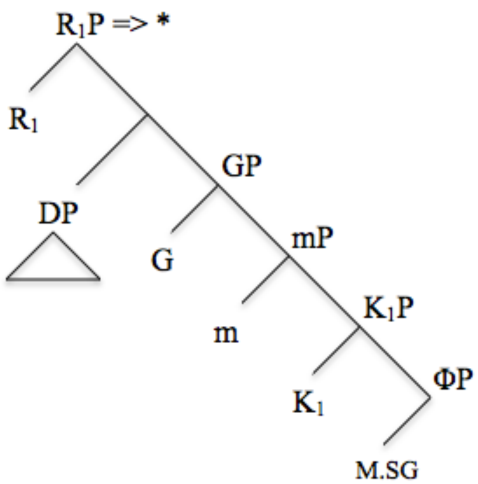
(b) CYCLIC



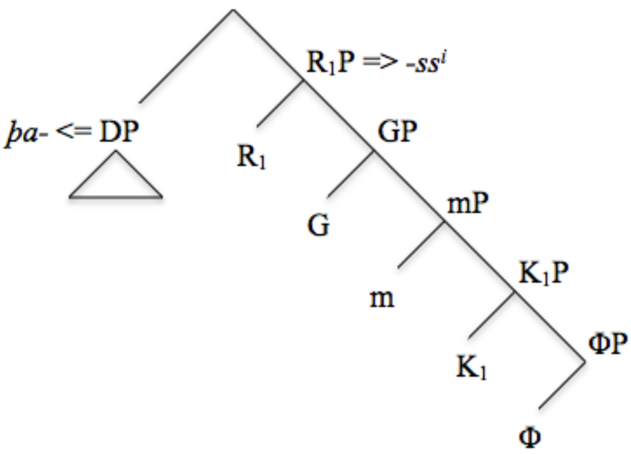
Next R_1 is added. Spelling out R_1 at STAY is not successful because there is no L-tree that matches [R_1P [DP GP mP K_1P ΦP]] (151a). By moving DP at CYCLIC, though, the constituent [R_1P [GP [mP [K_1P [ΦP]]]]] is created, which is matched by entry 869 (151b).

(151) Add R_1

(a) *STAY



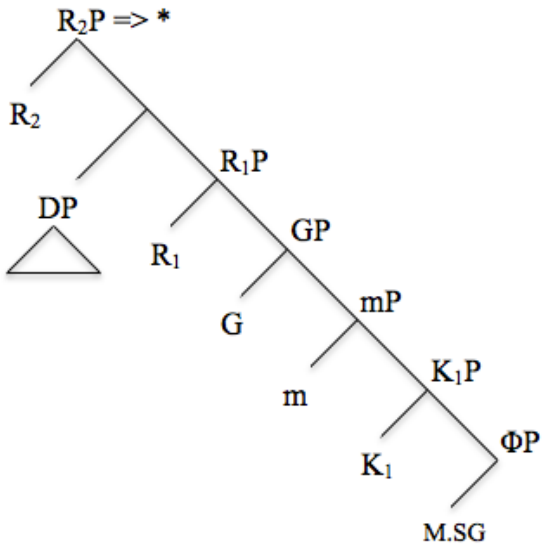
(b) CYCLIC



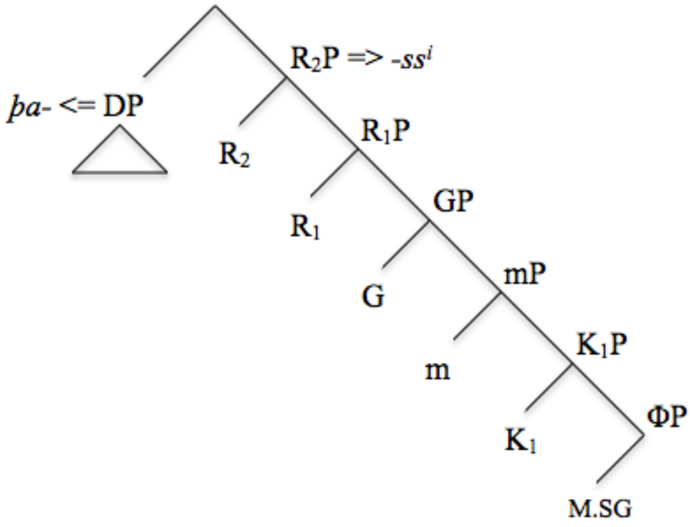
Next R_2 is added. At STAY there is no spellout available, since the entire constituent [R_2P [DP R_1P GP mP K_1P ΦP]] does not have a match in the lexicon (152a). At CYCLIC, a constituent [R_2P [R_1P [GP [mP [K_1P [ΦP]]]]]] is created and finds a match with entry 869 again (152b).

(152) Add R_2

(a) *STAY

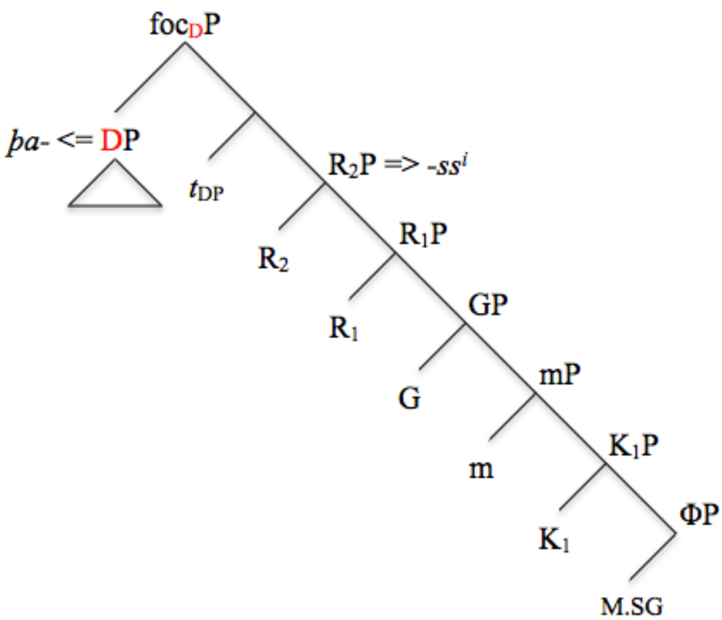


(b) CYCLIC



The final step of the derivation involves filling the foc$_D$P slot. In the case of the constant forms, the movement of DP is string-vacuous, but to keep the constant forms on a par with the K-final forms (where movement of DP to foc$_D$P is not string-vacuous), I posit that DP nevertheless does move (153). Again, the D in DP is red to point out that focP attracts a D-element.

(153) foc-movement



The final result is *þa-ss^i* > *þessi*.

5.4.2 Derivation of a K-final form

To illustrate the derivation of an ON K-final form let us take the M.DAT.SG form. The simplified lexicon already given above is repeated in (154).

(154) **Lexicon**

<$_{210}$ *þa-* ⇔ DP >

<$_{325}$ *-r* ⇔ K_1P [$_{ΦP}$ M.SG] >

<$_{327}$ *-^{i}n* ⇔ mP K_2P K_1P [$_{ΦP}$ M.SG] >

<$_{332}$ *-um* ⇔ K_4P K_3P K_2P K_1P [$_{ΦP}$ M.SG] >

<$_{865}$ *-a* ⇔ R_2P >

<$_{866}$ *-C* ⇔ R_1P GP >

<$_{867}$ *-C^{i}-* ⇔ [G m] >

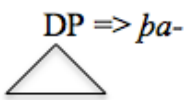
<$_{868}$ *-s-* ⇔ [R_2 R_1] >

<$_{869}$ *-ss^{i}* ⇔ R_2P R_1P GP mP [324 / 325] >

As in the constant forms, DP is first built up and matched at STAY by entry 210, as seen in (155).

(155) Build up to DP

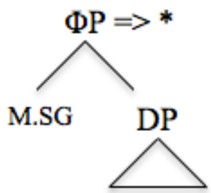
STAY



Next the Φ features M.SG are added. STAY does not produce a match (156a) and CYCLIC does not apply (156b). At SNOWBALL (156c), though, there is a match at ΦP with entries 325, 327, and 332. Entry 325 wins by the Elsewhere Condition, giving the spellout *-r*.

(156) Add Φ

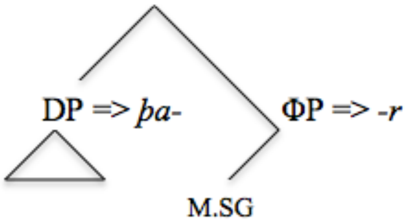
(a) *STAY



(b) *CYCLIC

N.A.

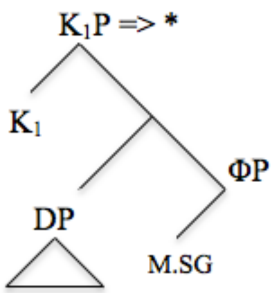
(c) SNOWBALL



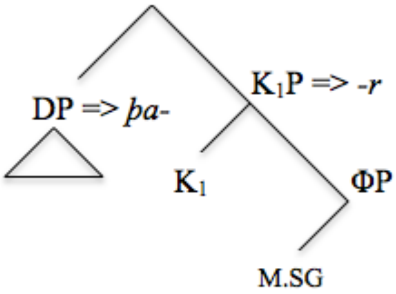
Next K_1 is added. STAY fails to produce a match (157a), but CYCLIC (157b) does result in a match with entries 325, 327, and 332. Once again 325 wins out (157b).

(157) Add K_1

(a) *STAY



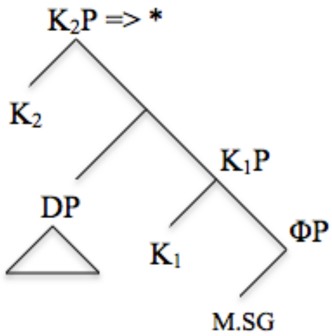
(b) CYCLIC



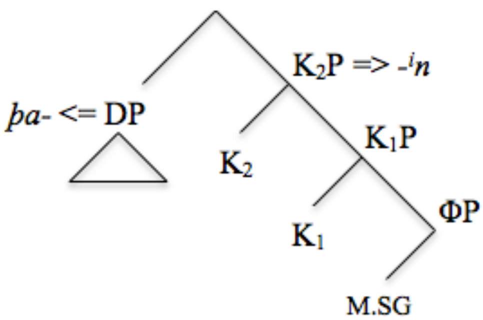
Next K_2 is added. STAY does not result in a match (158a), but CYCLIC does give a match with entries 327 and 332, of which 327 is the better match (158b).

(158) Add K_2

(a) *STAY



(b) CYCLIC



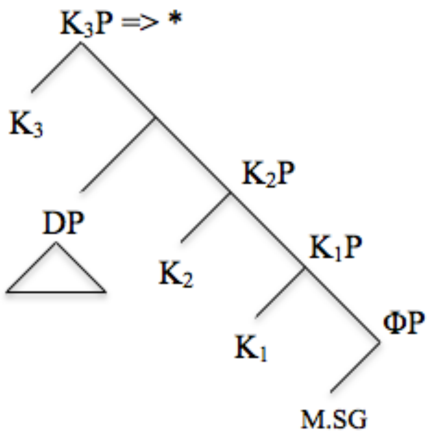
Next K_3 is added. Again, STAY does not produce a match (159a). At CYCLIC there is a match with entry 332.[99] The resulting spellout is *-um* (159b), which overrides the previous

[99] Remember that our lexicon has been artificially reduced for the purposes of this example. In reality there is a tailormade entry for the genitive constituent here.

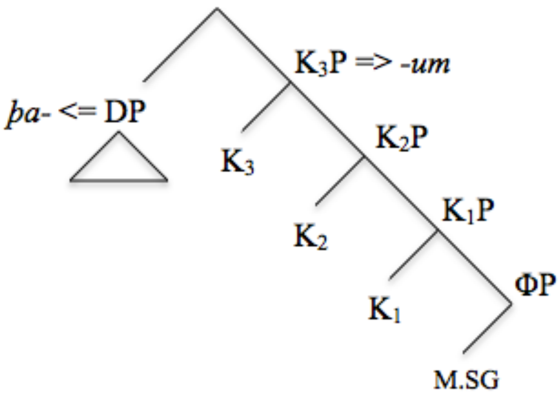
spellout -^{i}n in (159b). Note that entry 327 is not a match anymore since it does not contain the feature K_3.

(159) Add K_3

(a) *STAY



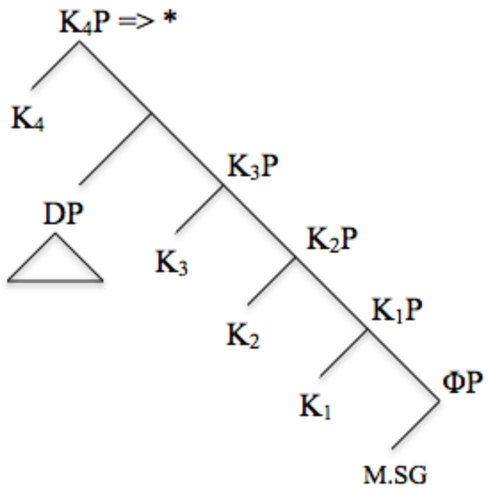
(b) CYCLIC



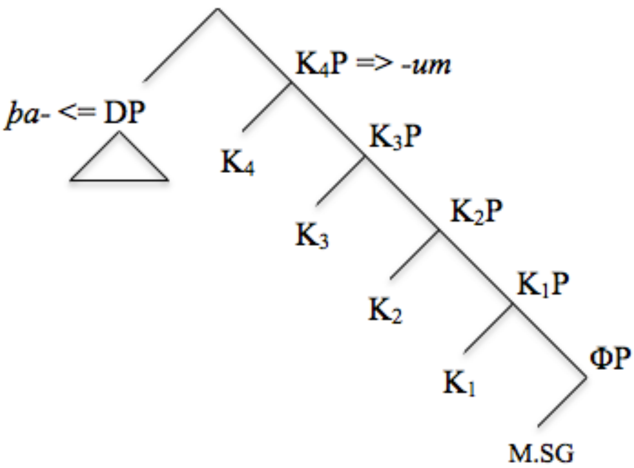
Next K_4 is added. At this point STAY does not succeed (160a), but once again CYCLIC does produce a match, with entry 332 (160b).

(160) Add K_4

(a) *STAY



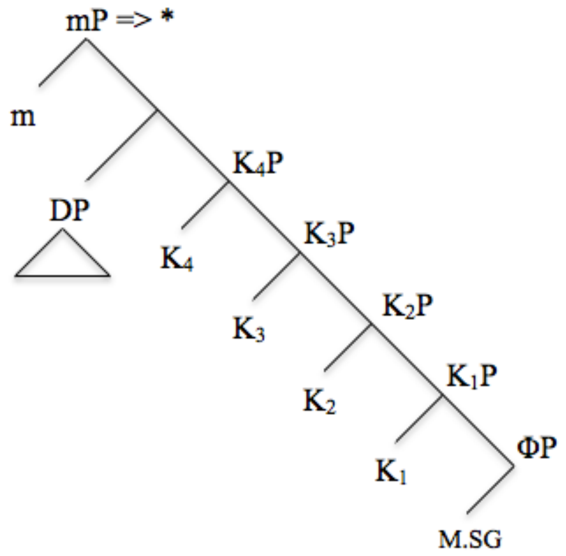
(b) CYCLIC



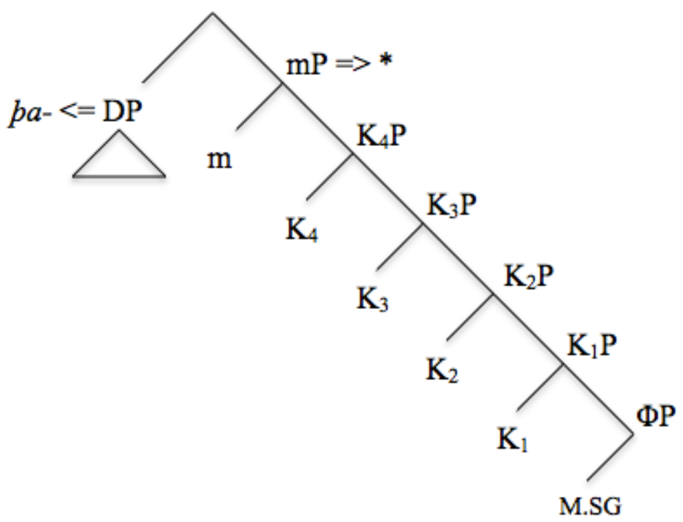
Next m is added. Here we arrive at a complication: STAY does not lead to a match (161a), nor does CYCLIC (161b) or SNOWBALL (161c). Ultimately, the reason CYCLIC fails is that the K-final endings do not package m and K together (this is something only the K-internal endings do). The reason SNOWBALL fails is because there is no entry in the lexicon with the phrase mP at the bottom of its L-tree.

(161) Add m

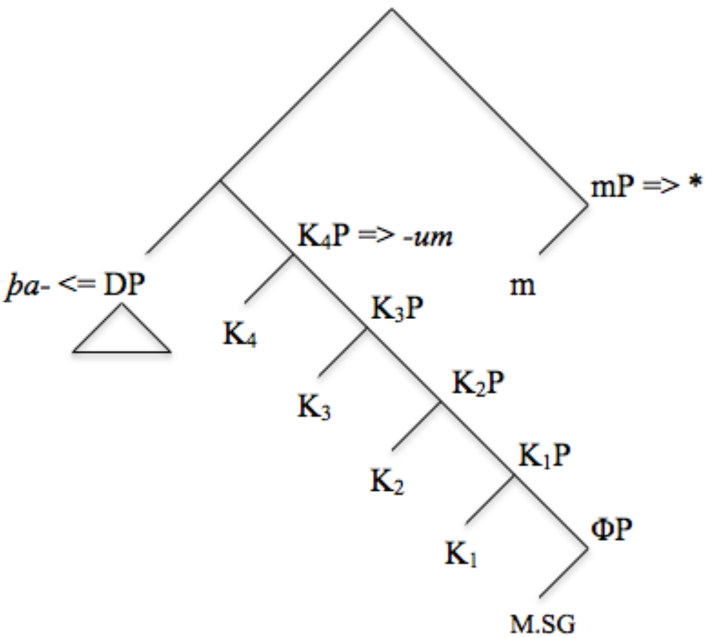
(a) *STAY



(b) *CYCLIC

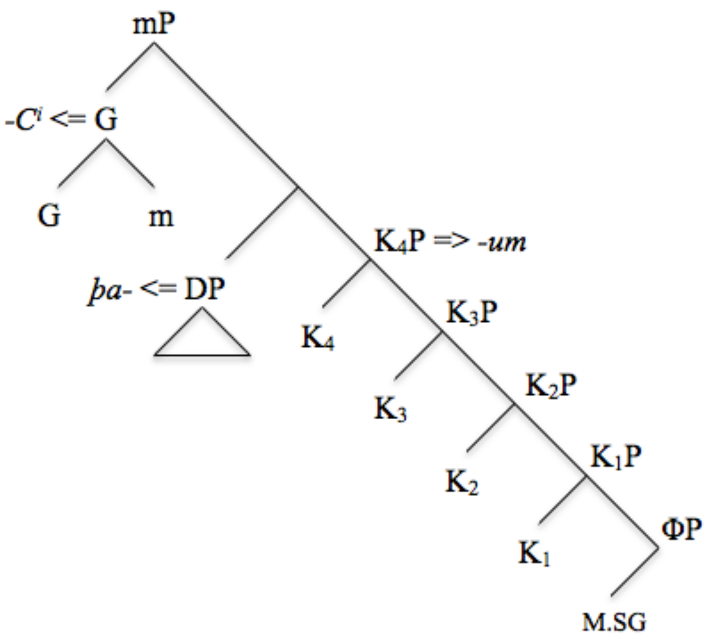


(c) *SNOWBALL



Since all three steps fail at lexicalizing m, the derivation turns to a second workspace, where the complex head [G m] is built. This complex head is then merged as the head of mP. It spells out as -C^i, since it matches entry 867, as seen in (161d).

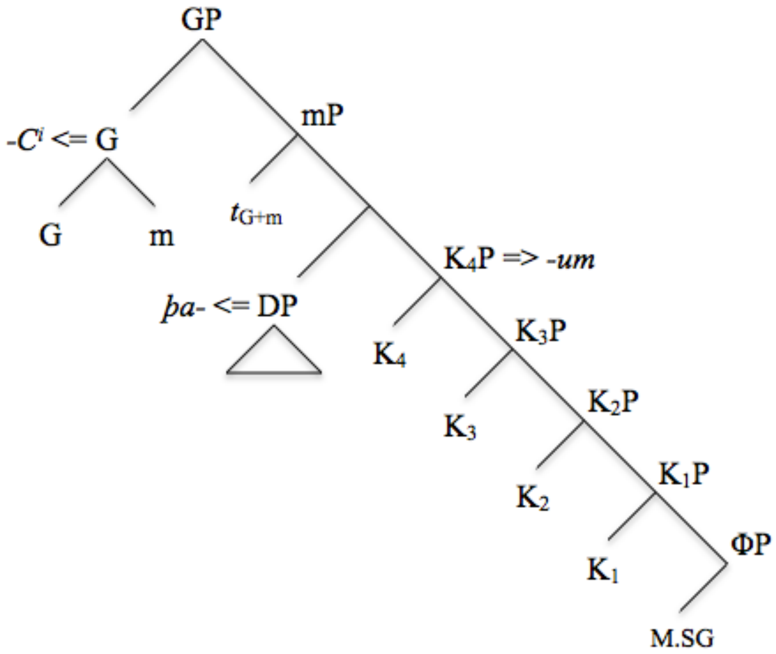
(161) (d) CONSTRUCT [G m]



The next layer to be built is GP. Since the head G has already been built, the complex head [G m] moves to GP and acts as its head. Spellout succeeds at STAY because [G m] is a constituent (162).

(162) Add G = [G m] moves to GP

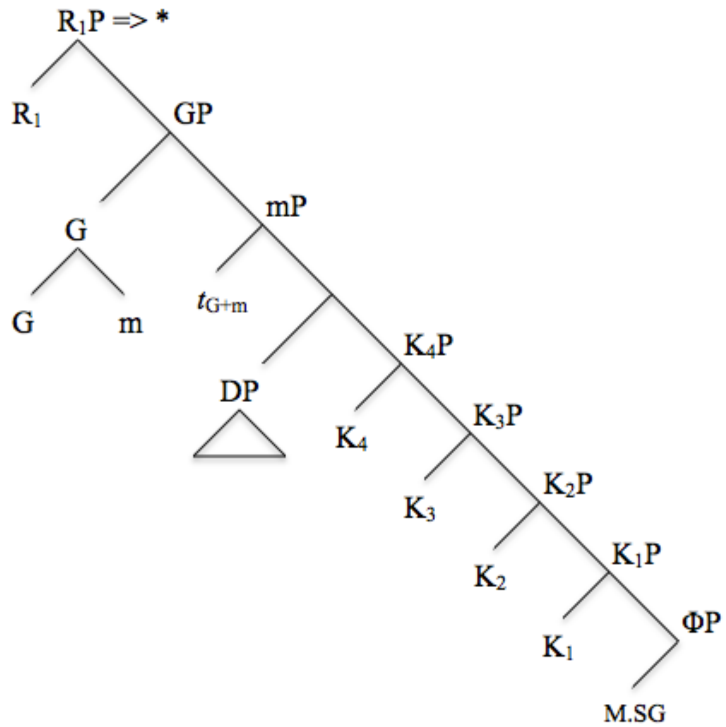
STAY



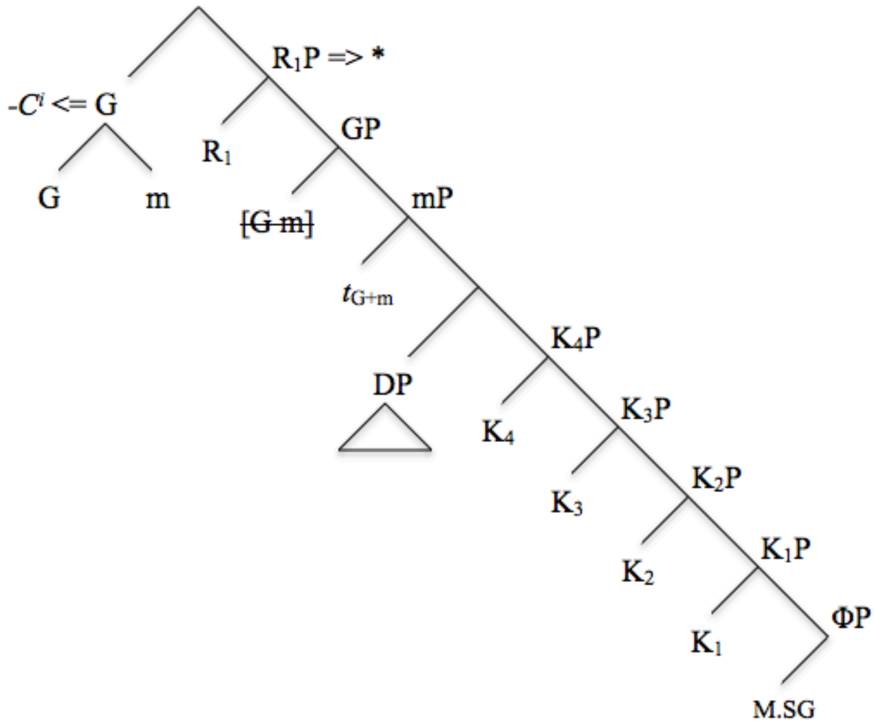
R_1 is the next feature in the fseq to be added. As with the feature m, match fails at STAY (163a), CYCLIC (163b), and SNOWBALL (163c). Note for CYCLIC that the daughter of the sister of R_1 (see Section 2.2.7) is the complex head [G m] merged under GP. Thus [G m] undergoes spellout-driven movement to the left of R_1P, but the constituent this creates is not matchable (especially given that the projection GP is technically deleted after this movement). SNOWBALL, furthermore, does not succeed because there is no lexical entry with the phrase R_1P at the bottom of its L-tree.

(163) Add R_1

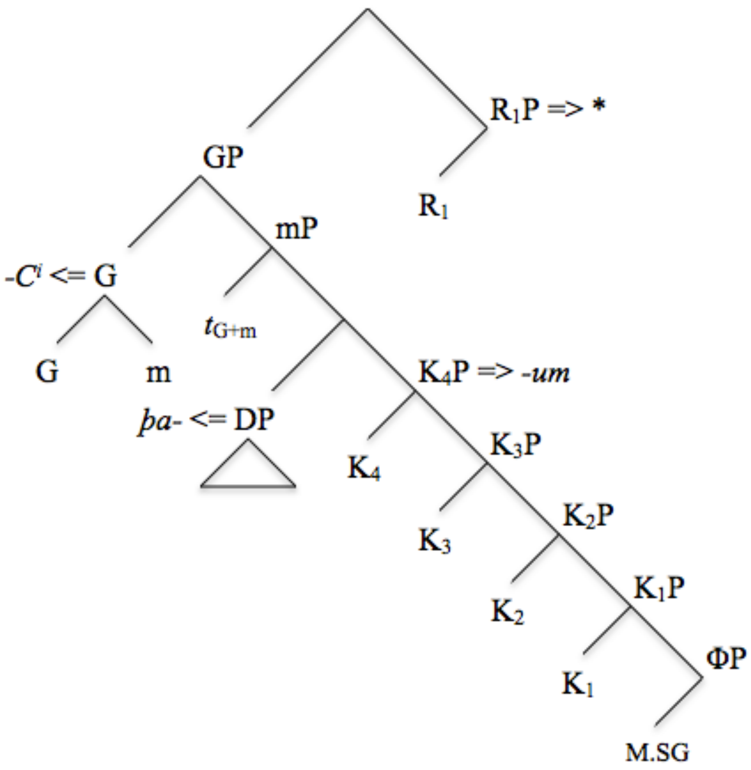
(a) *STAY



(b) *CYCLIC

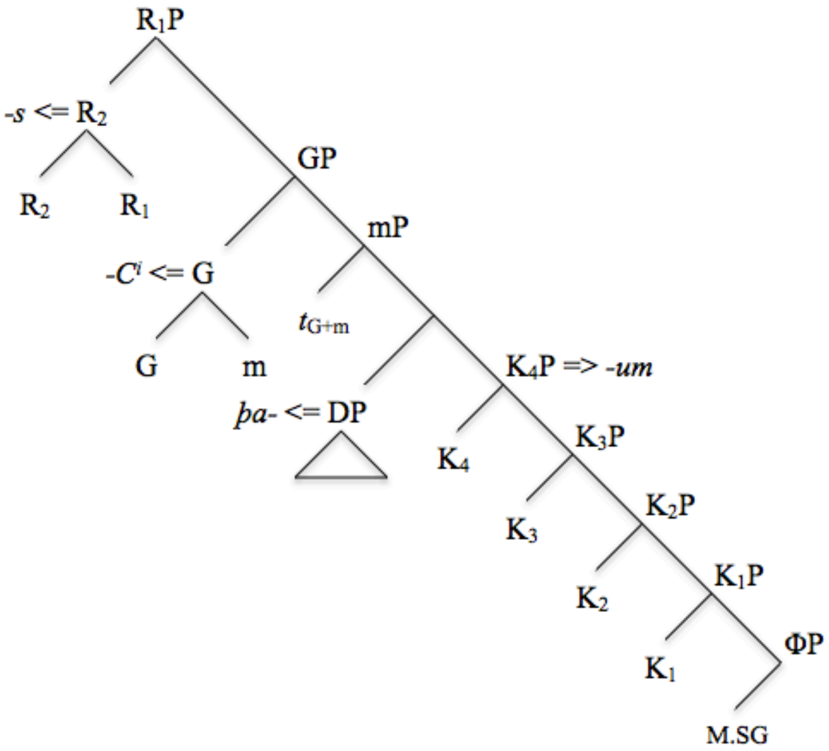


(c) *SNOWBALL



Since all three steps in the algorithm have failed, the system again turns to a second workspace, where the complex head [R_2 R_1] is built. This structure is then inserted in the primary workspace as the head of R_1P. The constituent [R_2 R_1] is lexicalized as *-s* according to entry 868 (163d).

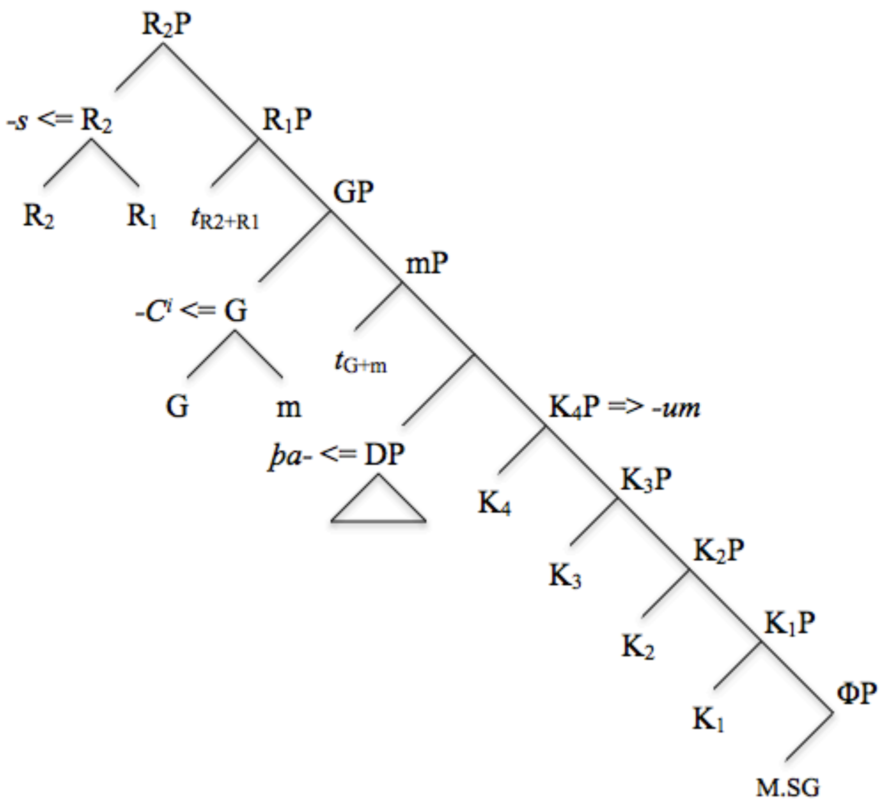
(163) (d) CONSTRUCT [R_2 R_1]



The next feature in the fseq is R_2. Since this feature has already been built, as part of the complex head [R_2 R_1] in (163d), this complex head moves up to serve as the head of R_2P. Again the complex head is lexicalized as *-s* at STAY (164).

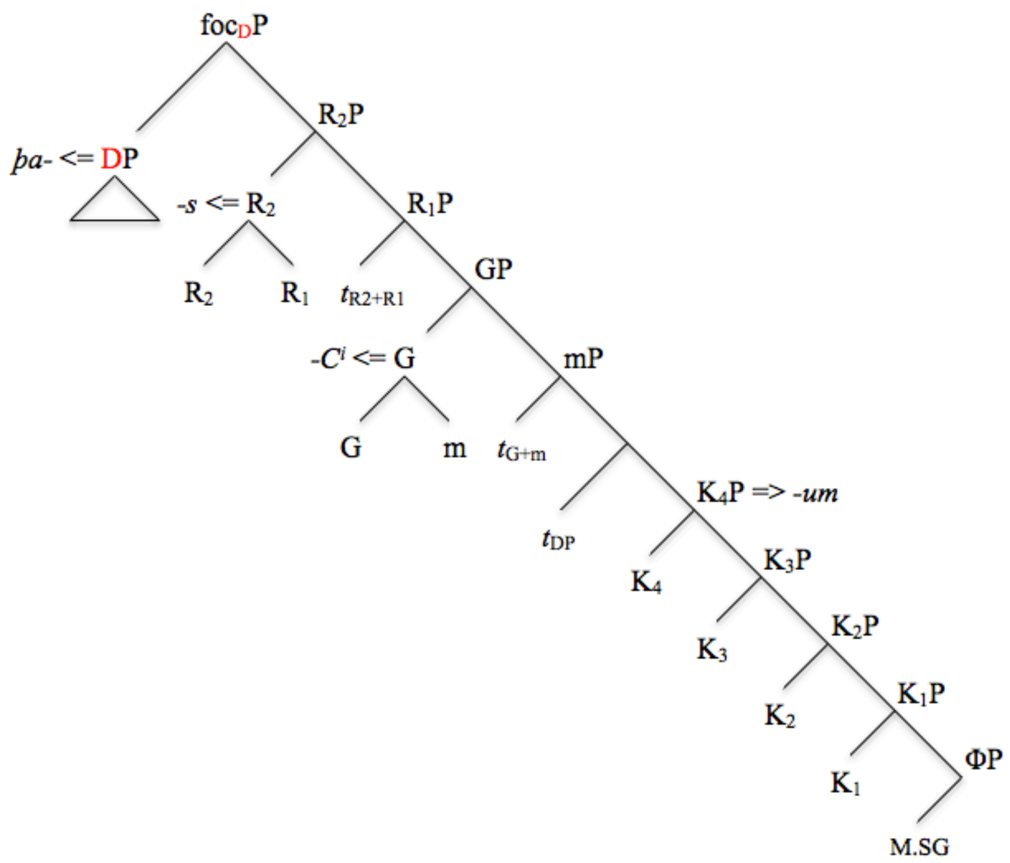
(164) Add R_2 = [R_2 R_1] moves to R_2P

STAY



The final step in the derivation is the addition of foc_DP and the attraction of a D constituent, namely DP (165).

(165) foc-movement



The final result is *þa-s-C^{i}-um* > *þessum*.

5.4.3 Derivation of a K-internal form

For the derivation of a K-internal form let us take the M.ACC.SG. The lexicon is repeated for convenience in (166).

(166) **Lexicon**

<$_{210}$ *þa-* ⇔ DP >

<$_{325}$ *-r* ⇔ K_1P [$_{ΦP}$ M.SG] >

<$_{327}$ *-^{i}n* ⇔ mP K_2P K_1P [$_{ΦP}$ M.SG] >

<$_{332}$ *-um* ⇔ K_4P K_3P K_2P K_1P [$_{ΦP}$ M.SG] >

<$_{865}$ *-a* ⇔ R_2P >

<$_{866}$ *-C* ⇔ R_1P GP >

<$_{867}$ *-C^{i}-* ⇔ [G m] >

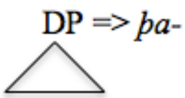
<$_{868}$ *-s-* ⇔ [R_2 R_1] >

<$_{869}$ *-ss^{i}* ⇔ R_2P R_1P GP mP [324 / 325] >

As in the constant and K-final forms, DP is first built and spelled out as *þa-* (167).

(167) Built up to DP

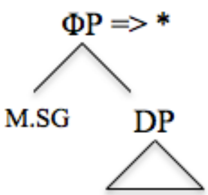
STAY



Next the Φ features M.SG are built. STAY fails to produce a match (168a), and so does CYCLIC (168b). At SNOWBALL, though, there is a match with entries 325, 327, and 332. Entry 325 wins because it has the least amount of 'junk' (168c).

(168) Add Φ

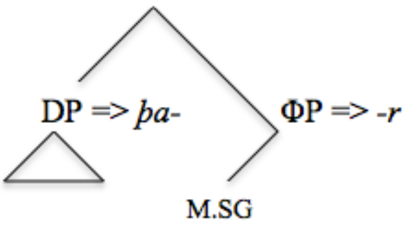
(a) *STAY



(b) *CYCLIC

N.A.

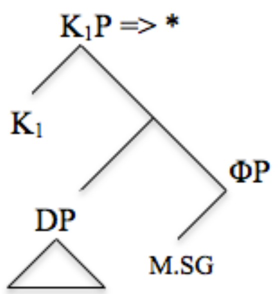
(c) SNOWBALL



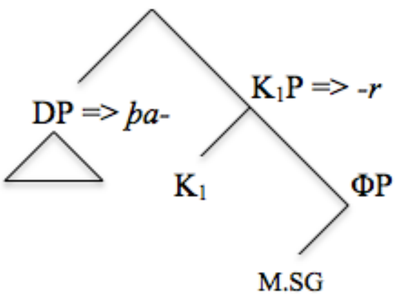
Next K_1 is added. STAY fails to produce a match (169a), but CYCLIC does result in a match with entries 325, 327, and 332. Once again 325 wins by the Elsewhere Condition (169b).

(169) Add K_1

(a) *STAY



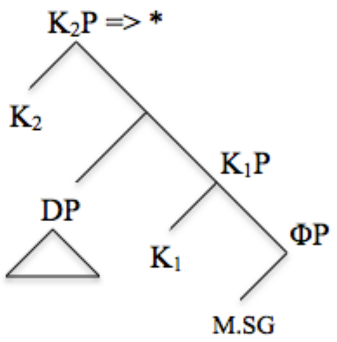
(b) CYCLIC



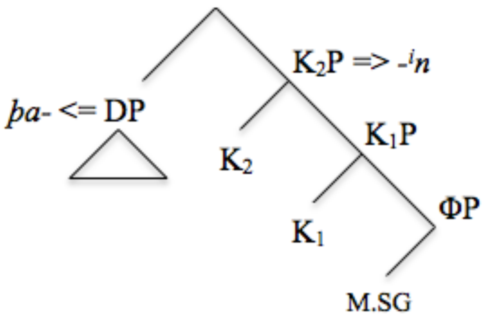
Next K_2 is added. STAY does not give a match (170a), but CYCLIC does produce a match with entries 327 and 332. Entry 327 is a perfect fit (170b).

(170) Add K_2

(a) *STAY



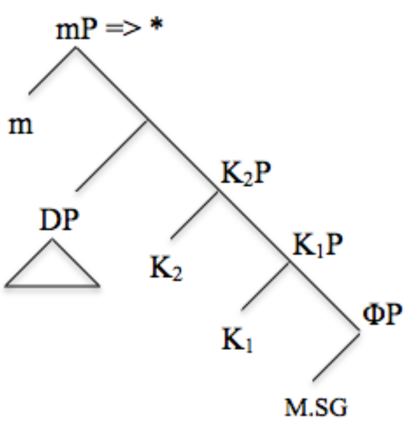
(b) CYCLIC



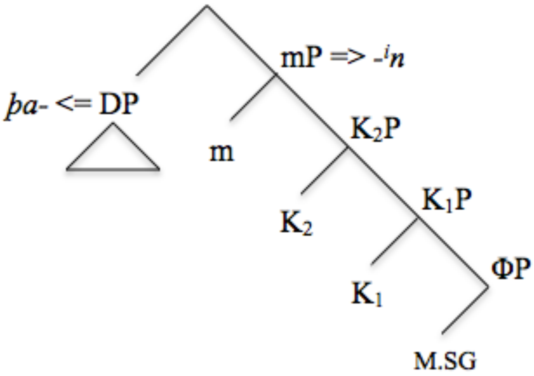
Next m is added. STAY does not lead to a successful spellout (171a). CYCLIC, however, does give a match (171b).

(171) Add m

(a) *STAY



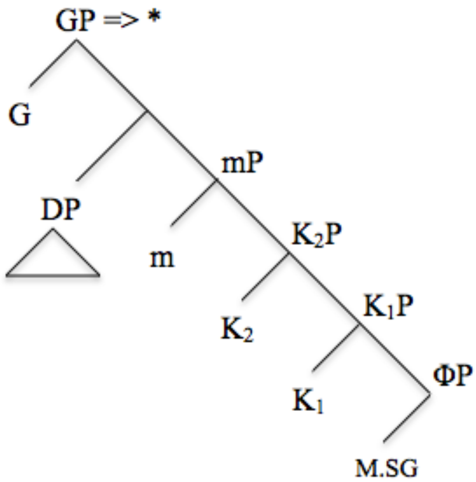
(b) CYCLIC



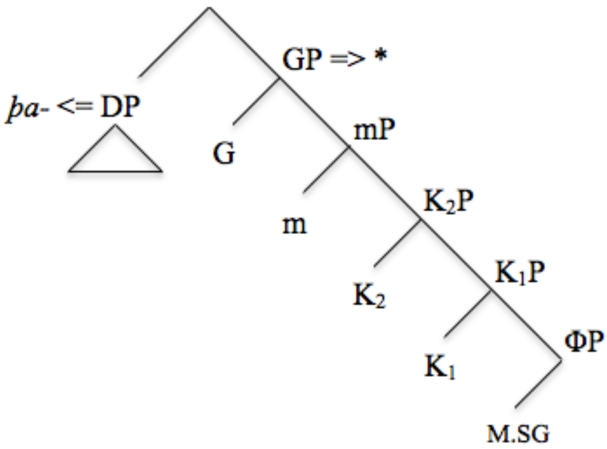
Next G is added to the sequence. STAY does not lead to a match (172a), and neither does CYCLIC (172b). After SNOWBALL, however, the constituent GP finds a match with entry 866 (172c).

(172) Add G

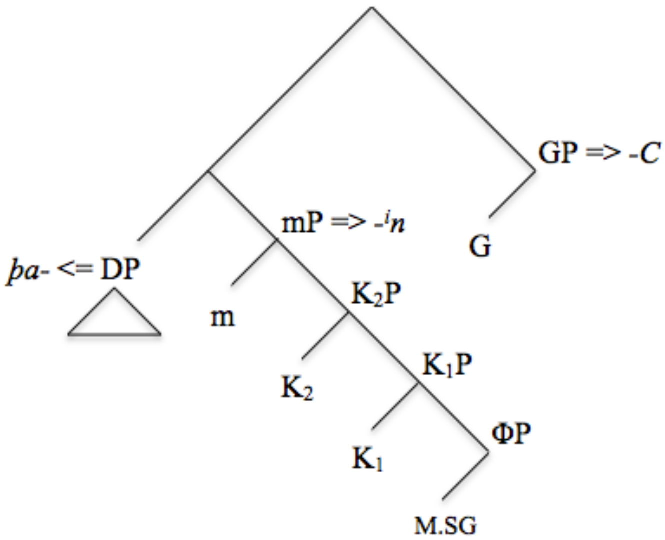
(a) *STAY



(b) *CYCLIC



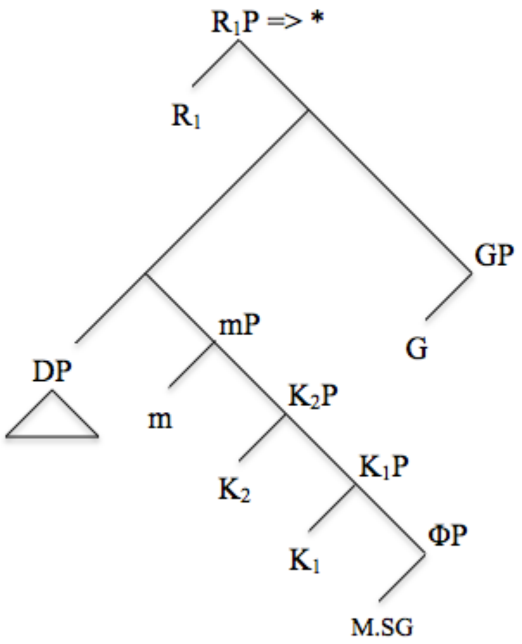
(c) SNOWBALL



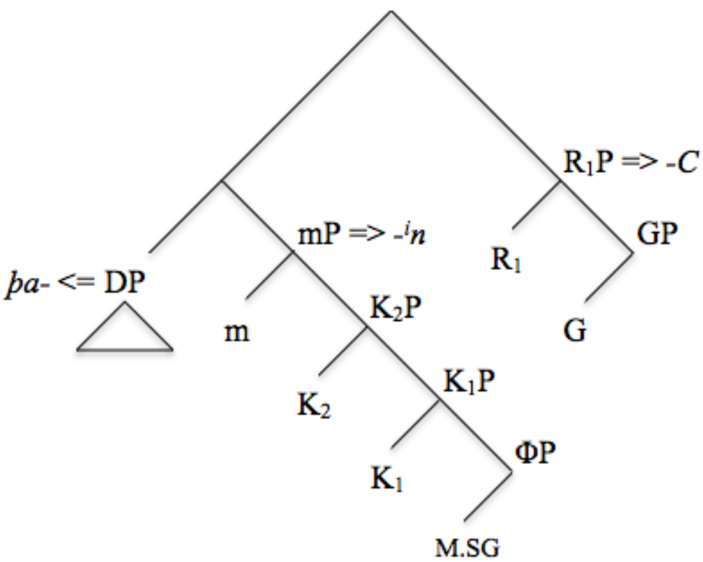
The next feature to be added is R_1. There is no suitable spellout at STAY (173a), but CYCLIC does deliver a match, again with entry 866 (173b).

(173) Add R_1

(a) *STAY



(b) CYCLIC

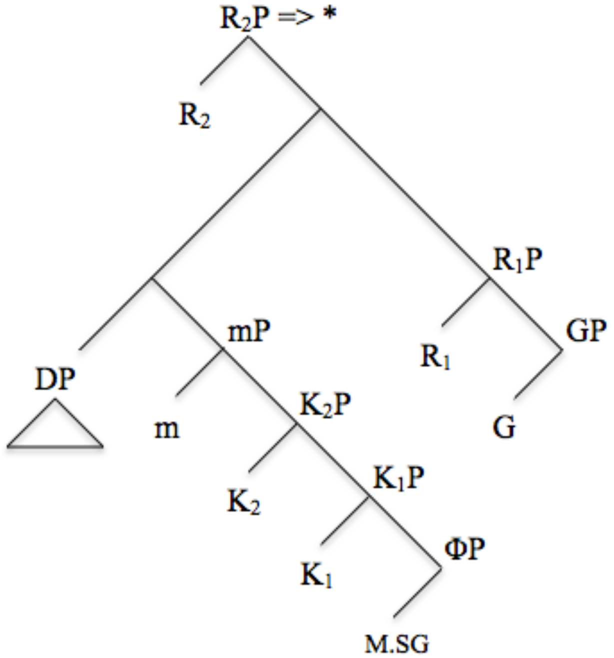


Note that in (173b) the constituent undergoing CYCLIC movement is [[DP] mP K_2P K_1P ΦP], which results from SNOWBALL having succeeded in (172c).

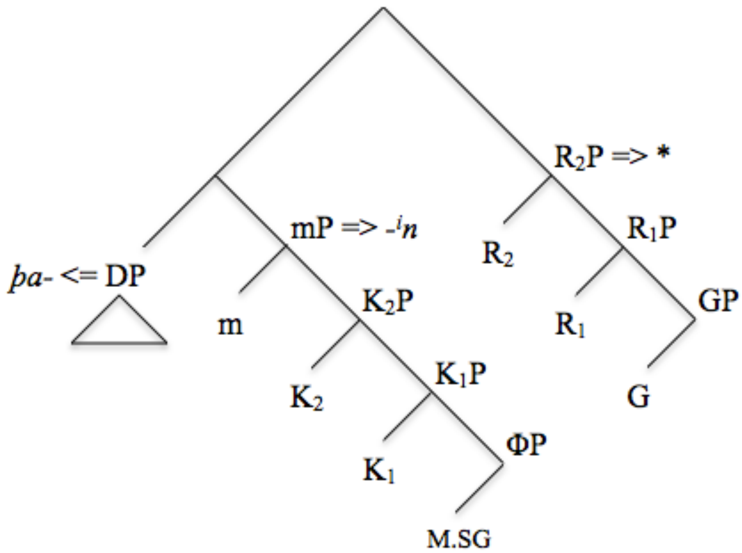
Next R_2 is added to the structure. STAY does not lead to a match (174a), and neither does CYCLIC (174b). SNOWBALL, however, allows R_2P to match with entry 865 and spell out as *-a* (174c).

(174) Add R_2

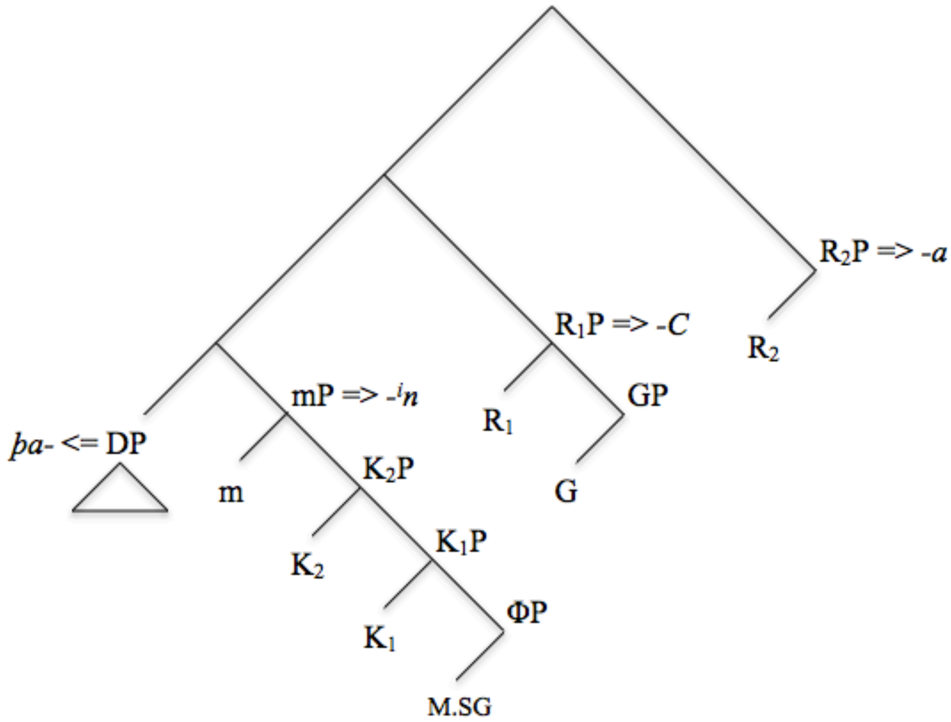
(a) *STAY



(b) *CYCLIC

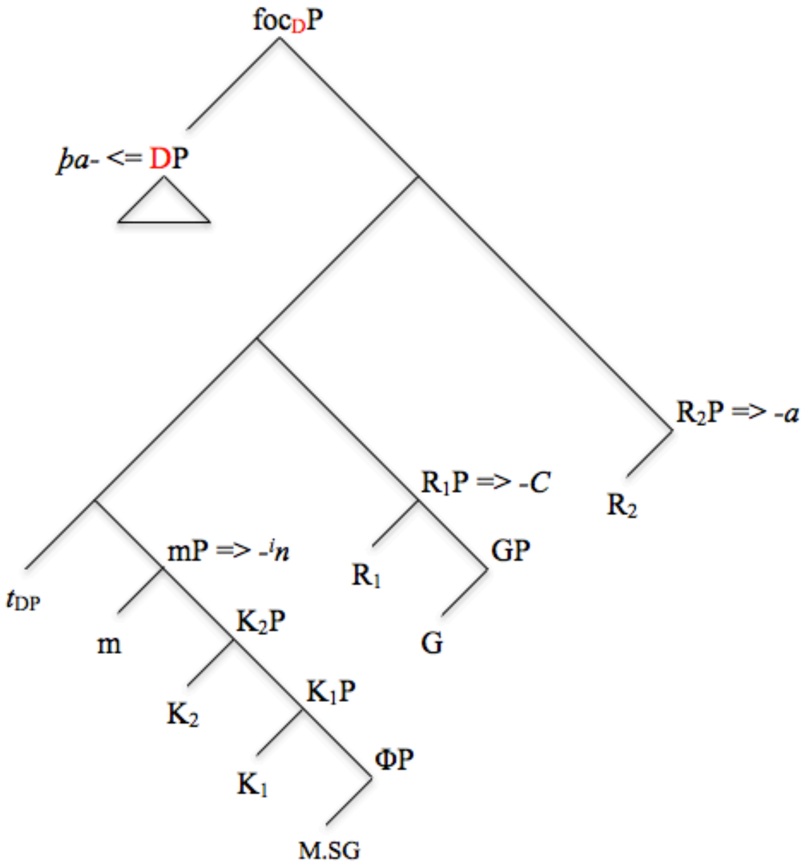


(c) SNOWBALL



Finally, to keep the K-internal forms on a par with the constant and K-final forms, let us posit focus movement as the last step in the derivation. In the K-internal forms, DP will subextract out of the [[DP] KP] constituent within which it is embedded and move to the foc_D position.

(175) foc-movement



The final result is *þa-ⁱn-C-a* > *þenna*.

5.5 Summary

In this chapter I have provided a nanosyntactic theory of the ON RDem paradigms. The main goal was to account for the 'allomorphy' between the sigmatic and asigmatic reinforcers (*-s* and *-a*, respectively) in a precise, lexically principled way.

The first step towards a nanosyntactic analysis was to adapt the U20 RDem structures according to a system allowing not only for terminal spellout but also for phrasal spellout. The preliminary attempt at doing so brought up numerous problems related to the issue of terminal spellout.

To solve the problems encountered I proposed that the features R and Gm be decomposed into two features each. This more fine-grained functional sequence allowed us to develop a more satisfactory account of the *-s* ~ *-a* alternation. Other refinements were proposed as well, including the need for a type of feature-driven movement in the K-final forms and an account of the constant forms which was inspired by McFadden's (2014) NSAG (for which see Appendix IV also).

Finally, step-by-step derivations of a constant, K-final, and K-internal form were provided.

6 Parameterized microvariation: West Germanic and more Norse

The reinforced demonstrative belongs not just to NGmc but also to WGmc. In this chapter I integrate facts from WGmc (OF, OE, OS, and OHG) into the analysis developed in Chapter 5. Some important archaic forms from Norse will also be discussed and accounted for. By the end of the chapter, a rich crosslinguistic picture of parameterized variation will have emerged, which finds a natural explanation in the nanosyntactic approach.

6.1 The West Germanic paradigms

Thus far only ON has been studied in any detail, though we have also looked into important aspects of the OE paradigm as they relate to ON (Section 4.1.3). In this section I integrate the WGmc (OF, OE, OS, OHG) facts into the nanosyntactic account developed so far. We shall see that there are three kinds of RDem forms in WGmc. The *K-final forms* (template (i) in (176)) show adjectival K in the rightmost position. The *D-internal forms* (template (ii) in (176)) have pronominal K endings (labeled K_D) to the left of the sigmatic reinforcer. The *Direct forms* (template (iii) in (176)) consist of the base directly inflected with K_D, with no sigmatic reinforcer present at all. The term *Direct* is capitalized throughout to remind the reader that K is pronominal (i.e. K_D) in these forms.

(176) WGmc templates

(i)	K-final	D-R(Gm)-K	e.g. OS M.ACC.SG *the-s-an*
			e.g. OF M.ACC.SG *thi-s-s-en*
(ii)	D-internal	D-K_D-R	e.g. OS F.NOM.SG *th-iu-s*
(iii)	Direct	D-K_D(Gm)	e.g. OS N.NOM/ACC.SG *thi-t(t)*

As seen in (176), WGmc does not have the straightforward four-slot templates we have come to expect from RDem forms (cf. Chapters 3 and 4). This aspect of the WGmc RDem forms will be discussed in detail below. First the RDem paradigm of each WGmc language will be described in terms of the templates in (176).

6.1.1 Old Frisian

The OF RDem paradigm is given in Table 55. The relevant inflectional endings are provided in Table 56. Irregular items are shaded.

Table 55 OF RDem (Bremmer 2009: 55, Markey 1981: 136)

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	th-iu-s	th-i-s	thi-t	thi-s-s-e	thi-s-s-e	thi-s-s-e
ACC	thi-s-s-e	thi-s-s-en	thi-t	thi-s-s-e	thi-s-s-e	thi-s-s-e
GEN	thi-s-s-er	thi-s-s-es	thi-s-s-es	thi-s-s-er	thi-s-s-er	thi-s-s-er
DAT	thi-s-s-er	thi-s-s-em	thi-s-s-em	thi-s-s-em	thi-s-s-em	thi-s-s-em

Table 56 OF strong adjective endings with relevant pronominal inflection (Markey 1981: 125-6, 135; Bremmer 2009: 54)

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	-Ø [Dem th-**iu**]	-Ø [Dem th-**ī**]	-Ø [Dem the-**t**]	-e	-e	-e
ACC	-e	-en	-Ø [Dem the-**t**]	-e	-e	-e
GEN	-er(e)	-es	-es	-era	-era	-era
DAT	-er(e)	-a/-e [*-em* < PL]	-a/-e [*-em* < PL]	-a/-e [*-em*]	-a/-e [*-em*]	-a/-e [*-em*]

(i) K-final forms

Most of the forms in Table 55 are K-final and thus boxed, on a par with the way the ON paradigm was decomposed and represented in previous chapters. The OF K-final forms, like the ON analogues, consist first of the base (in OF *thi-*), then a geminated reinforcer component *-ss-* (i.e. the sequence of the sigmatic reinforcer plus the geminator: *-s-C*), and finally strong adjectival K. Since not all K-final forms in WGmc display gemination of the sigmatic reinforcer, a more precise designation for the K-final forms in OF is *K-final-SS* (highlighting that *-s* is doubled).

There are two discrepancies (already mentioned in Section 1.2.2) between the boxed RDem forms in Table 55 and the endings in Table 56. First, the M/N.DAT.SG /

DAT.PL form *thissem* displays an older adjectival ending *-em*. Note that the ending *-em* is not pronominal (K_D) but adjectival, and furthermore it belongs to an older stage of the language. I think it is best, then, to think of *-em* as synchronically opaque. Accordingly, we should posit an irregular form, such that whenever the regular M/N.DAT.SG / DAT.PL RDem form **thissa/-e* is built, it is replaced by the special phonological idiom *thissem*.

The second discrepancy is that the GEN.PL RDem forms in Table 55 show the K ending *-er* even though they are expected to show *-era* according to Table 56. This discrepancy can be explained in terms of syncretism with the F.GEN/DAT.SG ending *-er(e)*. Indeed, in modern West Frisian, where only remnants of the old case system survive, there is a GEN/DAT syncretism simply in *-er* (cf. Sipma 1913: 60, 62; Tiersma 1999: 44). The OF RDem form *thisser* apparently anticipates this eventual conflation. I will leave open the issue of how to formalize a case ending which is in historical flux.

(ii) D-internal forms

The rest of the forms in Table 55 should not be considered K-final and hence are not boxed. The forms F.NOM.SG *th-iu-s* and M.NOM.SG *th-i-s* clearly show internal inflection, since K inflection (*-iu, -i*) is to the left of the sigmatic reinforcer *-s*. Furthermore, these forms have pronominal case inflection (K_D), meaning that their K is taken from the Dem paradigm, as indicated in Table 56. These forms, then, are D-internal.

(iii) Direct forms

Finally we come to the N.NOM/ACC.SG form *thi-t*, which does not have a sigmatic reinforcer at all. It consists simply of the base *thi-* with the pronominal ending *-t* directly suffixed to it. By my terminology introduced in (176) above, this is a Direct form.

At this point I would like to emphasize that being being boxed (i.e. K-final) means having adjectival inflection (K), while being non-boxed (i.e. non-K-final) means having pronominal inflection (K_D). The non-K-final forms can be traced back to the older Dem-*si* stage, when K was both pronominal and to the left of the old reinforcer *-si*, while the K-final forms are ‘newer’ in the sense that they have a special RDem stem and display adjectival endings. Below we shall see that this generalization holds not only for OF but for the whole of WGmc.

6.1.2 Old English

The OE RDem paradigm is repeated in Table 57, with the relevant inflectional endings given in Table 58.

Table 57 OE RDem (Campbell 2003: 291, Lass 1994: 145)

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	þ-ēo-s	þ-e(:)-s	þis-Ø	þ-ā-s	þ-ā-s	þ-ā-s
ACC	þ-ā-s	þis-ne	þis-Ø	þ-ā-s	þ-ā-s	þ-ā-s
GEN	þis-re > þisse	þis(s)-es	þis(s)-es	þis-ra > þissa	þis-ra > þissa	þis-ra > þissa
DAT	þis-re > þisse	þis(s)-um	þis(s)-um	þis(s)-um	þis(s)-um	þis(s)-um

Table 58 OE strong adjective endings with relevant pronominal inflection (Campbell 2003: 262, 290)

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	-Ø [Dem s-**ēo**]	-Ø [Dem s-**e(:)**]	-Ø	-e, -a [Dem þ-**ā**]	-e [Dem þ-**ā**]	-Ø [Dem þ-**ā**]
ACC	-e [Dem þ-**ā**]	-ne	-Ø	-e, -a [Dem þ-**ā**]	-e [Dem þ-**ā**]	-Ø [Dem þ-**ā**]
GEN	-re	-es	-es	-ra	-ra	-ra
DAT	-re	-um	-um	-um	-um	-um

(i) K-final forms

Most of the forms in Table 57 are K-final and thus boxed. The OE K-final forms, like the ON analogues, consist first of the base (in OE *þi-*), then the sigmatic reinforcer *-s*, and finally strong adjectival K. In some of the K-final forms there is no gemination of *-s*, which I will label the *K-final-S forms* (e.g. M.ACC.SG *þi-s-ne*), while in other K-final forms there is gemination of *-s*, which I will label the *K-final-SS forms* (e.g. M/N.GEN.SG *þi-s-s-es*).

(ii) D-internal forms

The rest of the forms are non-K-final and thus non-boxed: F.NOM.SG *þ-ēo-s*, M.NOM.SG *þ-e(:)-s*, and F.ACC.SG / NOM/ACC.PL *þ-ā-s*. In these forms the K component is internal, meaning that it precedes the sigmatic reinforcer. In these forms, moreover, the form that K takes originates not from the strong adjective system but from the Dem system instead. These are the OE D-internal forms. OE, then, supports the descriptive generalization formulated above that non-boxed forms in WGmc have pronominal inflection (K_D) rather than strong adjectival inflection.

(iii) Direct forms

OE does not have any Direct forms, that is, forms with K_D inflection attached directly to the base and without the sigmatic reinforcer. We might have expected the N.NOM/ACC.SG form to be a Direct form (as it is in OF, OS, and OHG), but OE *þis* does not fit the pattern, since it shows the reinforcer *-s* and does not show pronominal inflection (N.NOM/ACC.SG K_D being *-t*).

Let us examine OE N.NOM/ACC.SG *þis* in a bit more detail. It was mentioned in Section 1.2.2 that its internal structure could be either *þi-Ø-s* or *þi-s-Ø* since the inflection is null. In fact, the generalization formulated above according to which internal inflection originates in the Dem paradigm and external inflection originates in the adjectival paradigm suggests an answer to this small puzzle. Kluge ~~(1920: §213)~~ reconstructs the proto-form of *þis* as **þi-t-se*, where **-t* is pronominal inflection and, as predicted by our generalization, internally positioned. In the later form *þis*, however, the internal K_D element *-t* has clearly been lost and there is null inflection *-Ø* instead. Since *-Ø* is adjectival, the generalization tells us that this component should be external. Thus I have represented *þis* as *þi-s-Ø* in Table 57, making it a member of the K-final forms.

6.1.3 Old Saxon

The OS RDem paradigm is seen in Table 59, with the relevant inflectional endings given in Table 60. The M.NOM.SG slot is shaded due to the form being a reconstruction.

Table 59 OS RDem (Rauch 1992: 196, Cathey 2000: 37; also Gallée 1910: 240; see Nielsen 2000: 158 for *thitt*)

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL	
NOM	th-iu-s	*the-s-e	thi-t-t (OS_A) thi-t (OS_B)	the-s-e	the-s-e	the-s-e	th-iu-s
ACC	the-s-a	the-s-an	thi-t-t (OS_A) thi-t (OS_B)	the-s-e	the-s-e	the-s-e	th-iu-s
GEN	the-s-ara	the-s-es	the-s-es	the-s-aro	the-s-aro	the-s-aro	
DAT	the-s-aru	the-s-umu	the-s-umu	the-s-um	the-s-um	the-s-um	

Table 60 OS strong adjective endings with relevant pronominal inflection (Rauch 1992: 199, 194; Cathey 2000: 38; also Gallée 1910: 221, 238)

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL	
NOM	-Ø [Dem th-**iu**]	-Ø [Dem th-**e**(:)]	-Ø [Dem tha-**t**]	-e	-e	-e	-Ø [Dem th-**iu**]
ACC	-a	-an	-Ø [Dem tha-**t**]	-e	-e	-e	-Ø [Dem th-**iu**]
GEN	-era (<e> ~ <a>)	-es	-es	-aro	-aro	-aro	
DAT	-eru (<e> ~ <a>)	-umu	-umu	-um	-um	-um	

(i) K-final forms

Once again most of the OS RDem forms are K-final and thus boxed. The OS boxed forms consist of the base (in OS *the-*), then the sigmatic reinforcer *-s*, and finally adjectival K. Note here that gemination of *-s* is suppressed, as opposed to ON and OF (but much like OHG in Section 6.1.4). Thus the K-final forms in OS are more precisely K-final-S forms.[100]

(ii) D-internal forms

The rest of the forms in Table 59 are not K-final and thus not boxed. In the case of F.NOM.SG / N.NOM/ACC.PL *th-iu-s*, the inflectional element *-iu* precedes the sigmatic reinforcer *-s*. The element *-iu*, moreover, is pronominal (K_D). Once again this fits in with the descriptive generalization that non-boxed forms have pronominal inflection. To exemplify the generalization again, consider the N.NOM/ACC.PL slot in the paradigm, where two variants are attested in OS. The variant with K_D displays internal inflection (*th-iu-s*), making it a D-internal form, while the variant with K displays external inflection (*the-s-e*), making it a K-final form.

[100] Certain varieties or dialects of OS apparently did have gemination of the sigmatic reinforcer. Holthausen (1921: 119) reports the form *thessemo* (Trier seg. B) instead of *thesemo* in the M/N.DAT.SG. Whether this is simply a feature arising from analogy which should not be given a more regular structural analysis or not is a question I leave open.

(iii) Direct forms

Next consider the N.NOM/ACC.SG form *thi-**t**(t)*. In this form we see no sign of the sigmatic reinforcer. It is also notable that the base of this form is *thi-*, with a raised root vowel, rather than the expected base *the-*. The base *thi-*, moreover, is directly inflected with the K_D ending *-t*. In other words this is a Direct form. As indicated by the parenthesized *(t)*, there is some dialectal variation in the OS Direct form: sometimes the K_D inflection *-t* is geminated, i.e. *thi-t-C* > *thitt*, and sometimes it is not, i.e. *thi-t*. I will assume that this reflects two dialectal varieties of OS. I have designated *thitt* an OS_A form and *thit* an OS_B form.

The M.NOM.SG form **these* (shaded in Table 59) would appear to be a counterexample to my generalization about non-K-final forms taking pronominal inflection. In this form, the ending *-e* must come from the Dem paradigm and not from the strong adjective paradigm (where the expected ending is *-Ø*), yet the K_D ending *-e* appears to the right of *-s*, making **the-s-e* a K-final form. Crucially, though, the form in question is not directly attested in OS and has only been reconstructed. Perhaps, then, the generalization proposed here is reason to reconstruct the form as **the(:)s* instead, on a par with OE *þe(:)s*. In any case, since the data are inconclusive, I will ignore the M.NOM.SG form in OS.

6.1.4 Old High German

The OHG RDem paradigm is given in Table 61, with the relevant inflectional endings in Table 62. Irregular items have been shaded.

Table 61 OHG RDem (Braune & Reiffenstein 2004: 250; Wright 1906: 67; see EWAhd II for *dizzi/dezzi*)

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	dë-s-iu	dë-s-ē(r)	diz < *þi-t-t (dizzi/dezzi)	dë-s-o	dë-s-e	dë-s-iu
ACC	dë-s-a	dë-s-an	diz < *þi-t-t (dizzi/dezzi)	dë-s-o	dë-s-e	dë-s-iu
GEN	dë-s-era	dë-s-s-es	dë-s-s-es	dë-s-ero	dë-s-ero	dë-s-ero
DAT	dë-s-eru	dë-s-emu	dë-s-emu	dë-s-ēm	dë-s-ēm	dë-s-ēm

Table 62 OHG strong adjective endings (Braune & Reiffenstein 2004: 220; also Wright 1906: 55-6)

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	-iu	-ēr	-aʐ < -at [Dem da-ʐ < *þa-t]	-o	-e	-iu
ACC	-a	-an	-aʐ < -at [Dem da-ʐ < *þa-t]	-o	-e	-iu
GEN	-era	-es	-es	-ero	-ero	-ero
DAT	-eru	-emu	-emu	-ēm	-ēm	-ēm

(i) K-final forms

Most of the OHG forms in Table 61 are K-final and thus boxed. The OHG boxed forms consist of the base (in OHG *dë-*), then the sigmatic reinforcer *-s*, and finally adjectival K. Like OS, the OHG RDem stem does not have a geminated sigmatic reinforcer, except in the M/N.GEN.SG form *dësses* (i.e. *dë-s-C-es*). In OHG, then, we must distinguish between K-final-S forms (e.g. M.ACC.SG *dë-s-an*) and K-final-SS forms (M/N.GEN.SG *dë-s-s-es*).

(ii) D-internal forms

There are no D-internal forms in OHG.

(iii) Direct

There is only a single form in the OHG paradigm that could be considered non-K-final, and, in keeping with our generalization, it takes (took) pronominal inflection (K_D). As discussed in Section 1.2.5, this is the form *diz* [dits], which ultimately comes from **þitt*. This proto-form form can be parsed with a consonant geminator, i.e. *þi-t-C*, where *-t* is the K_D element. As in OS, the root vowel in the base of the N.NOM/ACC.SG Direct form is raised, from *ë* to *i*.

Nonetheless, the fact of the matter is that the form *diz* is synchronically opaque, since **þitt* and *diz* are separated by several centuries. I think the most likely analysis is in fact that *diz* is a phonological idiom: underlyingly it is the K-final form /dë-s-aʐ/, but there is a special idiomatic entry in the OHG lexicon which specifies that *dë-*

s-az is replaced by the phonology *dits* <diz>. Thus I have boxed this form in Table 61 but also shaded it to indicate irregularity.[101]

6.1.5 The ProK generalization

In sum, there are three types of RDem patterns in WGmc: K-final (sometimes with a geminated sigmatic reinforcer and sometimes not), D-internal, and Direct. Moreover, there is an important generalization about the WGmc RDem forms which I will call the *ProK generalization*. Thus far the ProK generalization about pronominal inflection has been discussed in terms of internal or external inflection: in forms with a sigmatic reinforcer, K_D appears to the left of *-s* (see also Ringe & Taylor 2014: 102 who make a similar observation), while K appears to the right of *-s*. However, this way of formulating the generalization leaves out the Direct forms, which take K_D but do not have a sigmatic reinforcer. Thus I will formulate the ProK generalization as in (177), which refers to positioning relative to the RDem base (the D component) rather than to the sigmatic reinforcer.

(177) The ProK generalization

K_D always appears to the immediate right of D.

In the D-internal forms – with the template D-K_D-R (e.g. OS F.NOM.SG *th-iu-s*) – D and K_D are adjacent, with K_D to the immediate right of the base D. Also in the Direct forms – with the template D-K_D(Gm) (e.g. OS N.NOM/ACC.SG *thi-t(t)*) – D and K_D are also adjacent, with K_D to the immediate right of the base D. However, in the K-final forms – with the template D-R-K (e.g. OS M.ACC.SG *the-s-an*) – D and K are, crucially, not adjacent. In sum, if an RDem form has K_D, then it will be either a D-internal form or a Direct form. If it has adjectival K, then it will be a K-final form.

Note here that the ProK generalization also applies (though rather uninterestingly so) to the Dem paradigms of both NGmc and WGmc. This is because Dem forms consist simply of the Dem stem plus pronominal inflection, i.e. D-K_D (much like the Direct forms).[102]

---

[101] A possibility I think is overly creative but which nevertheless deserves to be mentioned is that *dits* should be parsed as *di-t-s*, where *-s* is the sigmatic reinforcer and *-t* the pronominal inflection. This would make the form an internally inflecting item on a par with OF *th-iu-s*, OE *þ-ēo-s*, etc. It is more likely, though, a pure coincidence that the affricate *ts* in OHG (< **tt*) contains a sibilant that happens to look like the sigmatic reinforcer.

[102] It is not enough to say that the ProK generalization is just a fact about Dem, a view according to which the RDem forms with K_D are essentially Dem forms plus the sigmatic reinforcer (e.g. OE Dem NOM/ACC.PL *þā* vs. RDem NOM/ACC.PL *þā-s*), meaning that the K_D placement is unsurprising. Such an approach is insufficient because there are some clear differences between Dem and RDem with K_D. In OE, for instance, the Dem forms

6.1.6 A possible reinterpretation of the Old Norse K-internal forms

Since the ProK generalization as formulated in (177) makes direct reference to K_D, it may at first glance appear not to have anything to do with ON, since the ON RDem paradigm has been described as involving only adjectival (K) inflection. However, there is a possible reinterpretation of the ON data which I would like to mention at this juncture. If this reinterpretation is ultimately correct or not is an issue that will have to be left to future research, but I would like to mention it here since its validity would mean that the ProK generalization applies to the RDem paradigms of both NGmc and WGmc (rather than just to WGmc).

The ON K-internal forms are the following: M.ACC.SG *þe-n-n-a*, N.NOM/ACC.SG *þe-t-t-a*, and M/N.GEN.SG *þe-s-s-a*. The K endings here are M.ACC.SG *-n*, N.NOM/ACC.SG *-t*, and M/N.GEN.SG *-s*. My parsimonious assumption so far has been that these endings are adjectival, since all of the other endings in ON are adjectival. However, these endings are in fact syncretic between K and K_D (see (9) in Section 1.3). If *-n*, *-t*, and *-s* are in fact instantiations of K_D in the RDem forms, then the ProK generalization straightforwardly applies to the ON K-internal forms as well.

One interesting consequence of this alternative analysis is that it potentially dovetails with the variant Dem forms with fronted vowels cited above in Table 53 in Section 5.2.2: M.ACC.SG *þæn(n)*/*þen(n)* and N.NOM/ACC.SG *þæt*/*þet* (Noreen 1923: 314). With the addition of the already-unexpected fronted vocalism in M/N.GEN.SG *þes(s)*, I proposed that these Dem forms were evidence for my idea that the *i*-mutator is packaged with the K ending in the K-internal forms, giving the K_D endings M.ACC.SG -^{i}*n*, N.NOM/ACC.SG -^{i}*t*, and M/N.GEN.SG -^{i}*s*. The proposal at that point was essentially that there is a syncretism between K and K_D in these slots, allowing the K-internal RDem forms – which according to the analysis so far take adjectival K – to make use of the same phonological forms as the K_D endings. Now, however, instead of the proposal that the endings M.ACC.SG -^{i}*n*, N.NOM/ACC.SG -^{i}*t*, and M/N.GEN.SG -^{i}*s* are syncretic between K and K_D, we could imagine instead that these are just K_D endings and that the K-internal forms simply inflect with K_D. The ON K-internal forms, then, take *bona fide* pronominal endings (K_D), and in perfect obedience to (177) show adjacency of D and K_D (i.e. M.ACC.SG ***þe-n***-*C-a*, N.NOM/ACC.SG ***þe-t***-*C-a*, and M/N.GEN.SG ***þe-s***-*C-a*).[103]

F.NOM.SG *sēo* and M.NOM.SG *se(:)* coexisted with the RDem forms F.NOM.SG *þēo-s* and M.NOM.SG *þe(:)-s*. In OF and OS, the Direct forms *thi-t* and *thi-t(t)* differ from their Dem equivalents *the-t* and *tha-t*. On the NGmc side, morevoer, the ON F/M.NOM.SG RDem form *sjá* should, as we shall see below, be analyzed as a ProK form, yet it is not simply the Dem equivalent plus the sigmatic reinforcer (i.e. the hypothetical forms F.NOM.SG **sú-s(si)* and M.NOM.SG **sá-s(si)* are totally unheard of).

[103] As we shall see below (Section 6.2.2), this would mean that the foc-movement in (175) is not a case of subextraction.

6.2 Nanosyntactic structures of the West Germanic forms

In this section I present packaging schemas and tree structures for the K-final, D-internal, and Direct forms in WGmc. It will be seen that the same fseq used for ON can be used to derive the WGmc structures. Moreover, foc-movement of D-elements, introduced in Chapter 5, will be shown to be a crucial device for explaining the ProK generalization encountered throughout Section 6.1.

6.2.1 The K-final forms

6.2.1.1 The K-final-S forms: Old English, Old Saxon, and Old High German

In Section 4.1.3 we saw that the OE RDem paradigm presents us with an alternation between non-gemination of the sigmatic reinforcer (e.g. M.ACC.SG *þi-s-ne*) and gemination of the sigmatic reinforcer (e.g. M/N.GEN.SG *þi-s-s-es*). I accounted for this by hypothesizing that Gm is packaged with K in the cases where gemination of *-s* is suppressed, but that Gm is not packaged with K when gemination of *-s* is expressed. Though now we have split Gm into G and m, the same basic hypothesis still applies. It is summed up by the lexical entries in Figure 156. (The complex head status of *-C* in Figure 156 will be discussed below.)

E.g. M/N.GEN.SG *þi-s-C-es* > *þisses* vs. M.ACC.SG *þi-s-ne*

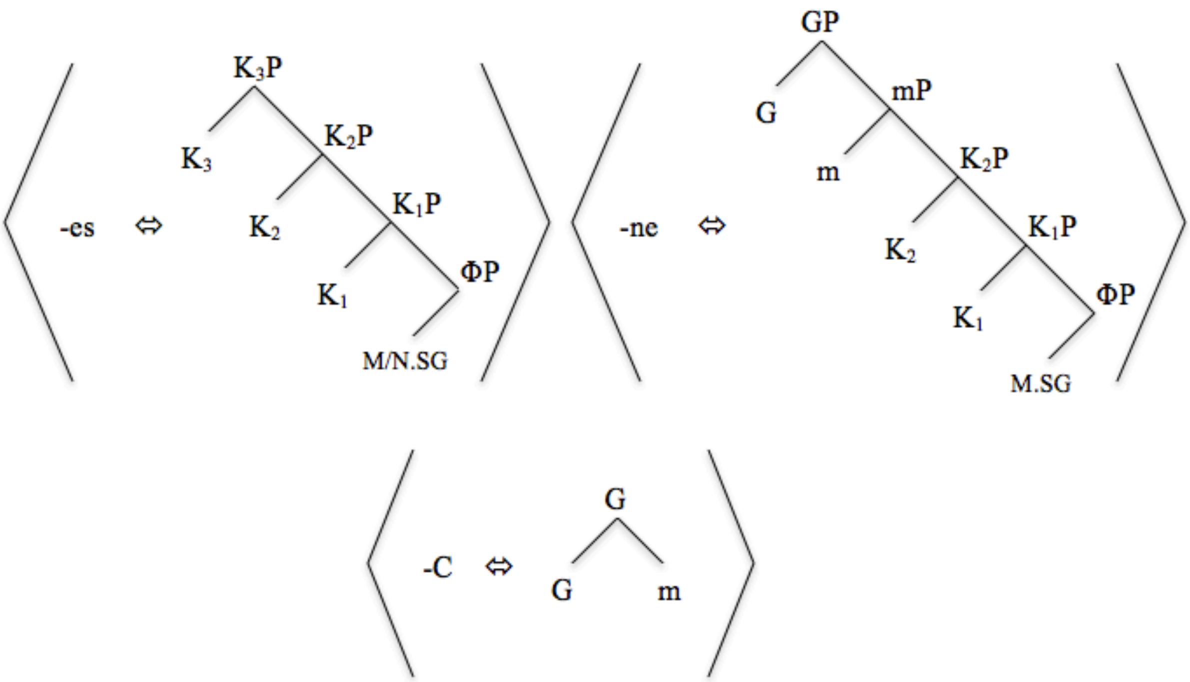


Figure 156 Lexical entries for (non-)gemination in OE

In Figure 156, the M/N.GEN.SG ending *-es* is not lexically packaged with G and m. Thus G and m will be lexicalized separately as the *-C* morpheme, resulting in gemination in the M/N.GEN.SG form *þi-s-**C-es*** > *þisses*. The M.ACC.SG ending *-ne*, on the other hand, is lexically packaged with G and m. Since G and m are swallowed up by the inflectional element *-ne*, the geminator *-C* will not surface, and gemination is suppressed in the M.ACC.SG form *þi-s-**ne***.

The suppression of gemination is extensively manifested in OS and OHG. In OS all of the K-final and D-internal forms show non-gemination of the sigmatic reinforcer. In OHG all of the D-internal forms and most of the K-final forms (all except for M/N.GEN.SG *dësses*) show non-gemination of the sigmatic reinforcer. In other words, K is packaged with G and m in the OS and OHG K-final forms.

The lexical status of the sigmatic reinforcer in the K-final-S forms, moreover, must be that of a complex head. The sigmatic reinforcer *-s* is a complex head in OE, OS, and OHG for the same reason it is a complex head in ON: R_2 and R_1 cannot spell out as the phrasal constituent [R_2P R_1P] if it is nestled between DP and the K ending, as illustrated in Figure 157.

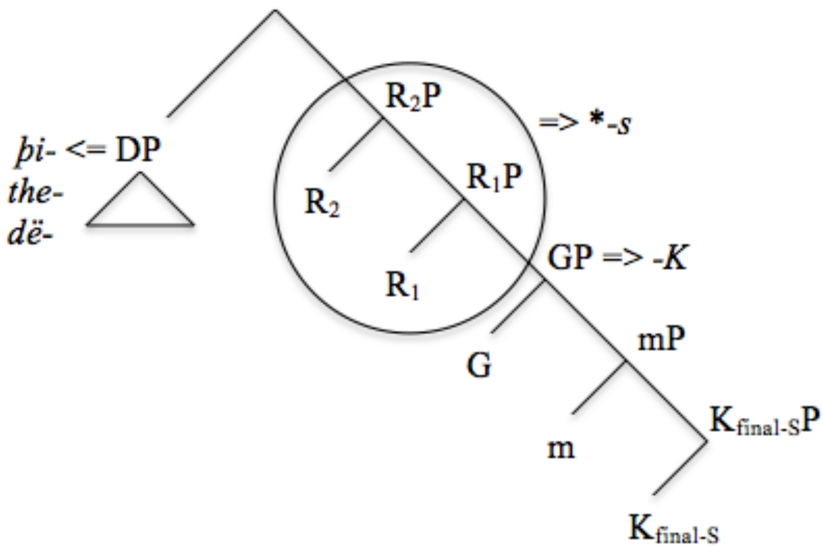


Figure 157 Sigmatic reinforcer cannot be a phrasal constituent

In Figure 157, [R_2P R_1P] is not a constituent to the exclusion of GP, mP, and KP.

The packaging schema for the K-final-S forms in OS and OHG is seen in Figure 158. The double-edged border around *-s* indicates complex head status.

R_2	R_1	G	m	K	D
-s		$-K_{final-S}$			þi- (OE) the- (OS) dë- (OHG)

Figure 158 K-final-S packaging in OE, OS, and OHG

By the Roll-up Shortcut (Section 2.2.10), spellout-driven movement in Figure 158 gives **s-þi-*$K_{final-S}$. That is to say, spellout-driven movement only gets DP to the left of KP. The complex head *-s* can spell out in situ, to the left of [[DP] KP]. Thus, as in the ON K-final forms, DP movement to the left of *-s* is still needed in order to produce the correct result. As in ON, we assume that foc-movement of DP accomplishes this, as shown in Figure 159.

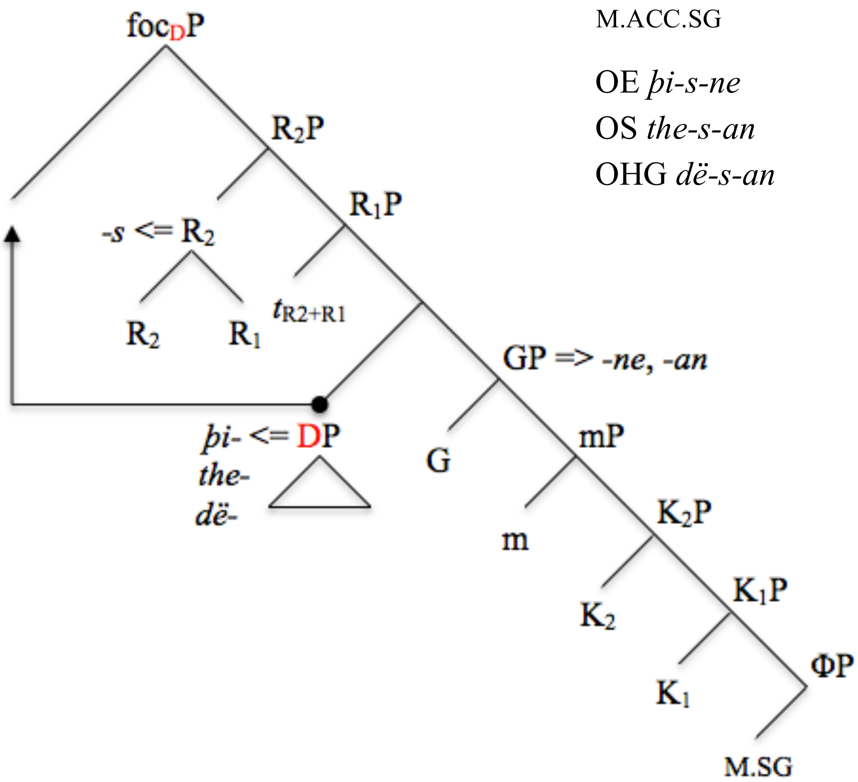


Figure 159 Foc-movement in WGmc K-final-S forms

All the forms which are shaded in Tables 63-65 are accounted for by the kind of packaging in Figure 158 and the foc-movement in Figure 159. Reconstructed and irregular forms are omitted in Tables 63-65.

Table 63 OE K-final-S forms

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	þ-ēo-s	þ-e(:)-s	þi-s-Ø	þ-ā-s	þ-ā-s	þ-ā-s
ACC	þ-ā-s	þi-s-ne	þi-s-Ø	þ-ā-s	þ-ā-s	þ-ā-s
GEN	þi-s-re > þisse	þi-s-s-es	þi-s-s-es	þi-s-ra > þissa	þi-s-ra > þissa	þi-s-ra > þissa
DAT	þi-s-re > þisse	þi-s-s-um	þi-s-s-um	þi-s-s-um	þi-s-s-um	þi-s-s-um

Table 64 OS K-final-S forms

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	th-iu-s		thi-t-t (OS$_A$) thi-t (OS$_B$)	the-s-e	the-s-e	the-s-e / th-iu-s
ACC	the-s-a	the-s-an	thi-t-t (OS$_A$) thi-t (OS$_B$)	the-s-e	the-s-e	the-s-e / th-iu-s
GEN	the-s-ara	the-s-es	the-s-es	the-s-aro	the-s-aro	the-s-aro
DAT	the-s-aru	the-s-umu	the-s-umu	the-s-um	the-s-um	the-s-um

Table 65 OHG K-final-S forms

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	dë-s-iu	dë-s-ē(r)		dë-s-o	dë-s-e	dë-s-iu
ACC	dë-s-a	dë-s-an		dë-s-o	dë-s-e	dë-s-iu
GEN	dë-s-era	dë-s-s-es	dë-s-s-es	dë-s-ero	dë-s-ero	dë-s-ero
DAT	dë-s-eru	dë-s-emu	dë-s-emu	dë-s-ēm	dë-s-ēm	dë-s-ēm

6.2.1.2 The K-final-SS forms: Old Frisian and Old English, with some Old High German

OE displays a subset of K-final-SS forms, indicated in Table 66 by dark shading. In OF, moreover, *all* of the boxed/K-final forms have a gemination of the sigmatic reinforcer, as shown in Table 67 by dark shading. There is also a single K-final-SS form in OHG, namely the M/N.GEN.SG form *dësses*, darkly shaded in Table 68. Irregular forms have been omitted in Tables 66-68.

Table 66 OE K-final-SS forms

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	þ-ēo-s	þ-e(:)-s	þi-s-Ø	þ-ā-s	þ-ā-s	þ-ā-s
ACC	þ-ā-s	þi-s-ne	þi-s-Ø	þ-ā-s	þ-ā-s	þ-ā-s
GEN	þi-s-re > þisse	þi-s-s-es	þi-s-s-es	þi-s-ra > þissa	þi-s-ra > þissa	þi-s-ra > þissa
DAT	þi-s-re > þisse	þi-s-s-um	þi-s-s-um	þi-s-s-um	þi-s-s-um	þi-s-s-um

Table 67 OF K-final-SS forms

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	th-iu-s	th-i-s	thi-t	thi-s-s-e	thi-s-s-e	thi-s-s-e
ACC	thi-s-s-e	thi-s-s-en	thi-t	thi-s-s-e	thi-s-s-e	thi-s-s-e
GEN	thi-s-s-er	thi-s-s-es	thi-s-s-es	thi-s-s-er	thi-s-s-er	thi-s-s-er
DAT	thi-s-s-er					

Table 68 OHG K-final-SS form

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	dë-s-iu	dë-s-ē(r)		dë-s-o	dë-s-e	dë-s-iu
ACC	dë-s-a	dë-s-an		dë-s-o	dë-s-e	dë-s-iu
GEN	dë-s-era	dë-s-s-es	dë-s-s-es	dë-s-ero	dë-s-ero	dë-s-ero
DAT	dë-s-eru	dë-s-emu	dë-s-emu	dë-s-ēm	dë-s-ēm	dë-s-ēm

The K-final-SS forms in Tables 66-68 are analogous to the ON K-final forms: there is a base, then the sigmatic reinforcer, next the geminator, and finally adjectival K. The packaging schema for the ON K-final forms from Section 5.2.3 is repeated in Figure 160.

R_2	**R_1**	**G**	**m**	**K**	**D**
-s		$-C^i$		$-K_{final}$	þa-

Figure 160 ON K-final packaging

To summarize the main lexical properties of the ON K-final forms:

- G and m are not packaged with K, since gemination is expressed.
- Both the sigmatic reinforcer and the geminator are complex heads.
- Foc-movement shifts DP to the left of the sigmatic reinforcer.

The derivation of an ON K-final form from Section 5.3 is repeated in Figure 161.

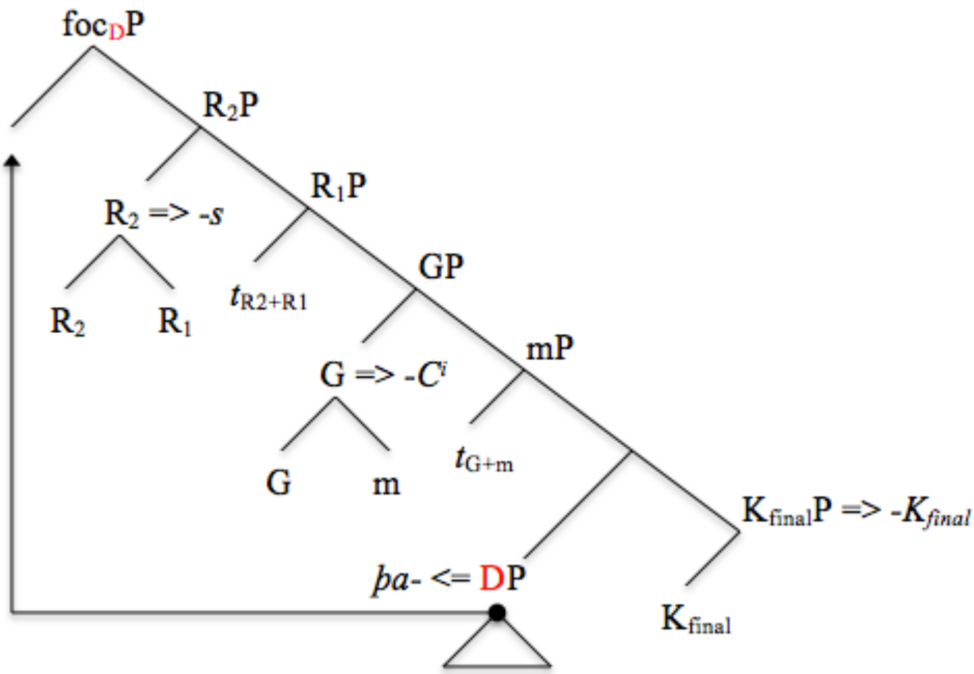


Figure 161 Derivation of the ON K-final forms

The K-final-SS forms in OE (Table 66), OF (Table 67), and OHG (Table 68) are exactly structurally parallel to the ON K-final forms. As seen in Figure 162, the only differences between ON and WGmc are some minor phonological ones (i.e. *-C* in WGmc vs. *-C^i* in ON; the base *þi-/thi-/dë-* in WGmc vs. *þa-* in ON; K endings, of course, will also show various phonological differences between languages).

R_2	R_1	G	m	K	D
-s		-C		$-K_{final\text{-}SS}$	þi- (OE) thi- (OF) dë- (OHG)

Figure 162 K-final-SS forms in OE and OF

The derivation of a K-final-SS form in WGmc is seen in Figure 163, which, again, is structurally parallel to the ON derivation in Figure 161 above.

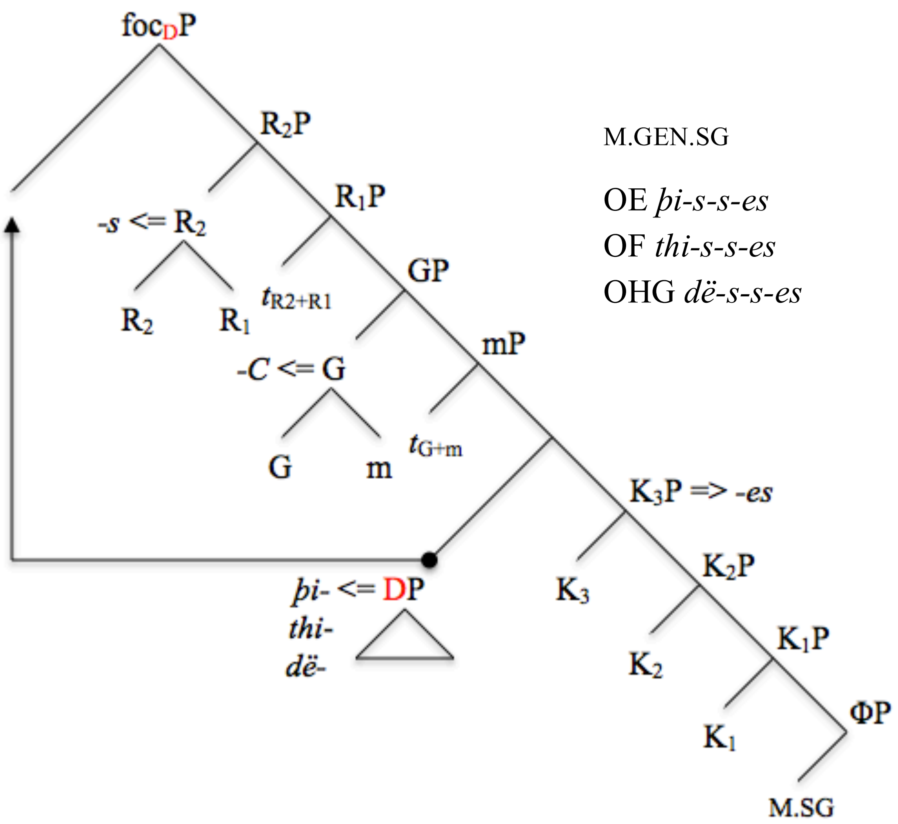


Figure 163 Derivation of the WGmc K-final-SS forms

All of the boxed/K-final forms in WGmc have now been accounted for. In the next section I turn to the D-internal forms.

6.2.2 The D-internal forms and movement to focP

The forms with dark shading in Table 69 (OF), Table 70 (OE), and Table 71 (OS) are classified as D-internal. Recall that OHG does not have any D-internal forms. Irregular and reconstructed forms have been omitted in the tables.

KEY

light shading = accounted for in Section 6.2.1
dark shading = to be accounted for in Section 6.2.2

Table 69 OF D-internal forms

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	th-iu-s	th-i-s	thi-t	thi-s-s-e	thi-s-s-e	thi-s-s-e
ACC	thi-s-s-e	thi-s-s-en	thi-t	thi-s-s-e	thi-s-s-e	thi-s-s-e
GEN	thi-s-s-er	thi-s-s-es	thi-s-s-es	thi-s-s-er	thi-s-s-er	thi-s-s-er
DAT	thi-s-s-er					

Table 70 OE D-internal forms

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	þ-ēo-s	þ-e(:)-s	þi-s-Ø	þ-ā-s	þ-ā-s	þ-ā-s
ACC	þ-ā-s	þi-s-ne	þi-s-Ø	þ-ā-s	þ-ā-s	þ-ā-s
GEN	þi-s-re > þisse	þi-s-s-es	þi-s-s-es	þi-s-ra > þissa	þi-s-ra > þissa	þi-s-ra > þissa
DAT	þi-s-re > þisse	þi-s-s-um	þi-s-s-um	þi-s-s-um	þi-s-s-um	þi-s-s-um

Table 71 OS D-internal forms

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL	
NOM	th-iu-s		thi-t-t (OS_A) thi-t (OS_B)	the-s-e	the-s-e	the-s-e	th-iu-s
ACC	the-s-a	the-s-an	thi-t-t (OS_A) thi-t (OS_B)	the-s-e	the-s-e	the-s-e	th-iu-s
GEN	the-s-ara	the-s-es	the-s-es	the-s-aro	the-s-aro	the-s-aro	
DAT	the-s-aru	the-s-umu	the-s-umu	the-s-um	the-s-um	the-s-um	

These forms consist of the base, plus pronominal inflection (K_D), which appears to the left of the sigmatic reinforcer. In other words, the D-internal forms are internally inflected, which by my ProK generalization above is due to the nature of their inflection (pronominal K_D rather than adjectival K).

(178) The ProK generalization

(a) K to the right of sigmatic reinforcer (K-final)

the-	*-s*	*-an*	(OS M.ACC.SG)
D	R	K	

(b) K_D to the left of sigmatic reinforcer (D-internal)

th-	*-iu*	*-s*	(OS F.NOM.SG)
D	K_D	R	

We can put the ProK generalization in more structural terms, namely in terms of pied-piping: in the K-final forms DP does not pied-pipe KP to the left of the sigmatic reinforcer, stranding KP at the bottom of the structure, while in the D-internal forms DP pied-pipes K_DP to the left of the sigmatic reinforcer (see Figure 164 below). The crucial detail is that K_DP (pronominal K) is pied-piped to focP, but KP (adjectival K) is not.

This is where foc-movement, introduced in Section 5.2.3 for the ON K-final forms, is useful for understanding the ProK generalization. We can understand the pied-piping of K_DP as following from the requirement on focP to move D-elements. To be precise, I propose that focP moves the *largest possible* constituent with D-like features. Figure 164 illustrates; again, feature identity is highlighted in red.

(a) DP movement (stranding KP) (b) [[DP] K_DP] movement (pied-piping K_DP)

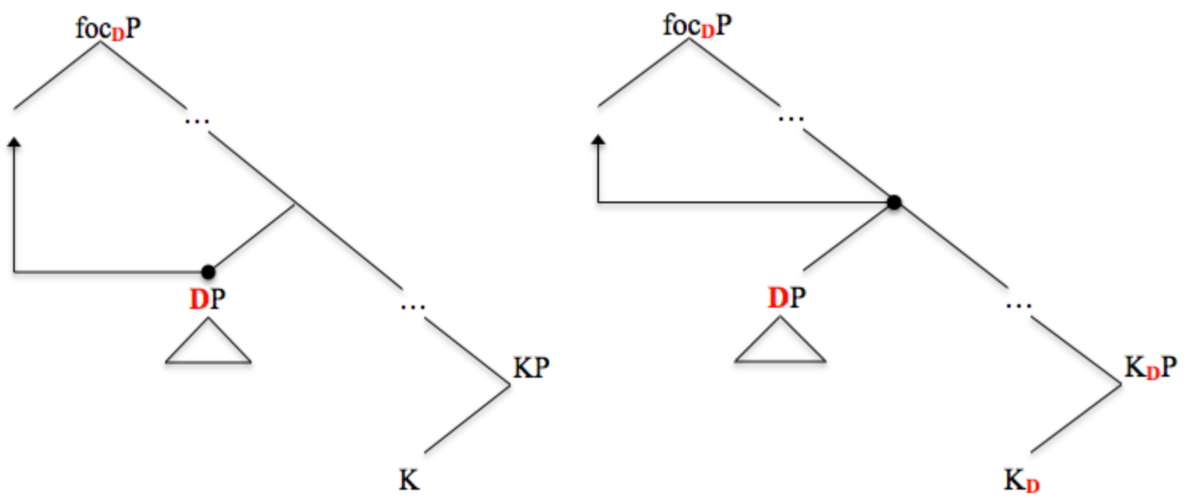


Figure 164 foc$_D$P attracts largest D-like element

When KP is adjectival, it is not a D-like element; thus it is not pied-piped along with DP to focP, as shown in part (a) of Figure 164. When KP is pronominal in nature, it is pied-piped along with DP to focP, as in part (b) of Figure 164. Recall that foc-movement is not spellout-driven but driven rather by feature identity, in this case D(-like) features. I leave the further theoretical details of such movement in the context of nanosyntax to future work.[104]

Since the WGmc D-internal forms do not display gemination of the sigmatic reinforcer, we know that K is packaged with the geminator features, as sketched in Figure 165.

R_2	R_1	G	m	K_D	D
-s		$-K_{D\text{-internal}}$			þ- (OE) th- (OF)

Figure 165 D-internal packaging

The packaging in Figure 165 results in the kind of structure shown in Figure 166. As discussed above, foc-movement will apply to the largest D-like element in the structure, resulting in the movement of the entire constituent [[DP] K_DP] to focP.

[104] A valid question which arises at this point is why my version of pied-piping is so strictly tied to feature identity when more traditional cases of pied-piping, well known in the literature, can drag along pieces of structure which are not featurally related, for instance in **[*A picture* [*of who*]$_{wh}$]$_i$** *did you buy* t_i *yesterday?*, where *a picture* does not at first glance have a *wh*-feature (Jaehoon Choi, p.c.). To go out on a bit of a limb, I would wager that the correct formulation of pied-piping has not yet been discovered, and that ultimately it *should* involve some notion of strict feature identity. Moreover, the granularity of pied-piping in this dissertation is very fine, targeting word-internal phrases. Pied-piping at the sentence-level should operate on the same principles, of course, but it is not yet clear how features and their interrelationships at the word level are 'inherited' into sentence-level phrases. In other words, the relationship between the domain of nanosyntax at this stage (words) and the domain of traditional syntactic research (sentences) needs to be clarified before we can be sure that the version of pied-piping identified in this dissertation really is so different from sentence-level pied-piping. Finally it is very plausible that certain kinds of 'pied-piping' are actually the result of spellout-driven movement, in which case the two types need to be carefully distinguished.

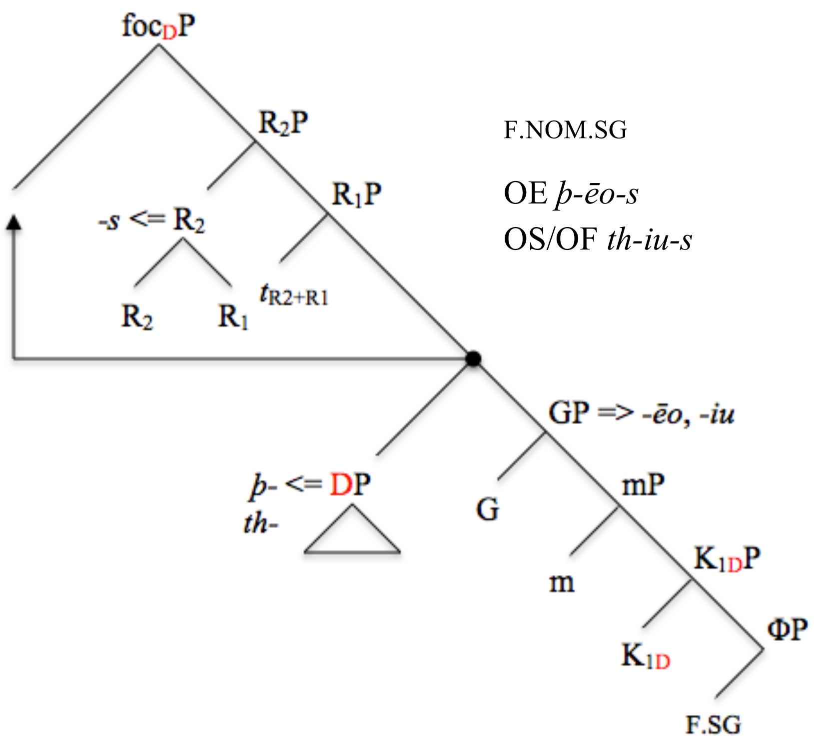


Figure 166 Derivation of the WGmc D-internal forms

This accounts for all of the D-internal forms in WGmc. I now turn to the Direct forms.

6.2.3 The Direct forms

The Direct forms are those WGmc forms consisting of K_D inflection directly suffixed to the base. Importantly, the sigmatic reinforcer is absent in these forms. The Direct forms are indicated by dark shading in Tables 72 and 73 (where the lightly shaded forms have already been accounted for in the previous sections). Recall that OE does not have any Direct forms. Irregular and reconstructed forms have been omitted, so the proto-form **þitt* for OHG *diz* [dits] does not need to be discussed. However, note that **þitt* is also the ancestor of OS_A *thitt*, for which see Section 6.2.3.2.

Table 72 OF Direct forms

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	th-iu-s	th-i-s	thi-t	thi-s-s-e	thi-s-s-e	thi-s-s-e
ACC	thi-s-s-e	thi-s-s-en	thi-t	thi-s-s-e	thi-s-s-e	thi-s-s-e
GEN	thi-s-s-er	thi-s-s-es	thi-s-s-es	thi-s-s-er	thi-s-s-er	thi-s-s-er
DAT	thi-s-s-er					

Table 73 OS Direct forms

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL	
NOM	th-iu-s		thi-t-t (OS_A) thi-t (OS_B)	the-s-e	the-s-e	the-s-e	th-iu-s
ACC	the-s-a	the-s-an	thi-t-t (OS_A) thi-t (OS_B)	the-s-e	the-s-e	the-s-e	th-iu-s
GEN	the-s-ara	the-s-es	the-s-es	the-s-aro	the-s-aro	the-s-aro	
DAT	the-s-aru	the-s-umu	the-s-umu	the-s-um	the-s-um	the-s-um	

There are two subtypes of Direct forms. Each subtype will be discussed in turn.

6.2.3.1 The Direct forms without gemination

The Direct forms in OF and OS_B do not have gemination of the K_D suffix *-t*: OF *thi-t* and OS_B *thi-t*. In other words, these Direct forms suppress both the sigmatic reinforcer and the geminator.

Recall also that OS_B *thit* displays raising of the root vowel. As seen in the rest of the OS RDem paradigm, the regular base is actually *the-*, with root vowel *e*, while the Direct form *thit* has the root vowel *i* instead. As was done for ON, I propose to make use of a floating bundle of phonological features (cf. Gibson & Ringen 2000) in order to trigger this shift of *e* to *i*.[105] I will refer to this as a '*j*-mutator'. Thus the OS_B Direct form actually has the underlying phonological form /the-t^j/. Unlike the *i*-mutator, which surfaces in final position, let us assume that the *j*-mutator is deleted in final position. Thus *the-*t^j surfaces as *thit* after vowel-raising and deletion of the *j*-mutator.

Above I have hypothesized that the lack of gemination of the sigmatic reinforcer in many WGmc RDem forms can be captured by packaging the G and m features with the relevant K ending. The same strategy can be pursued here, by hypothesizing that G and m are part of the lexical makeup of the K_D ending. In order to account for the suppression of the sigmatic reinforcer, we can take this strategy one step further, by hypothesizing that R_2 and R_1 (the features responsible for *-s*) are also packaged with the K_D ending. This is summed up in Figure 167.

[105] As discussed below, the OHG Direct form *diz* also shows raising of the unmarked root vowel *ë* (cf. the OHG RDem base *dë-*) to *i*. Indeed, *e*-vocalism is also dominant in the Dem paradigms of OS and OHG. Thus OS and OHG seem to use the same base (i.e. *the-* and *dë-*, respectively) for their Dem and RDem paradigms. In the Dem paradigms of OE and OF, however, *ā*-vocalism is more prevalent. This may be a clue as to why OE and OF have special RDem stems in *þi-/thi-*. Since their Dem paradigms were more of a hodgepodge of root vocalisms, there was less of a synchronic rationale for taking a Dem stem as the basis for building RDem.

<table>
<tr><th>R_2</th><th>R_1</th><th>G</th><th>m</th><th>K_D</th><th>D</th></tr>
<tr><td colspan="5">-t</td><td>thi- (OF)</td></tr>
<tr><td colspan="5">$-t^j$</td><td>the- (OS_B)</td></tr>
</table>

Figure 167 Packaging in the Direct forms without gemination

The Roll-up Shortcut shows that spellout-driven movement alone produces the correct result in Figure 167, giving OF *thi-t* and OS_B *the-t*^{j} > *thi-t*. In other words, foc-movement is not crucial in getting the correct linear order of morphemes in the Direct forms. Even though foc-movement is string-vacuous in the Direct forms, I will assume that foc-movement nevertheless takes place in these forms as well, keeping them on a par with the rest of the RDem forms. Since K is pronominal in the Direct forms, furthermore, DP will pied-pipe K_DP on its way to focP, as seen in Figure 168.

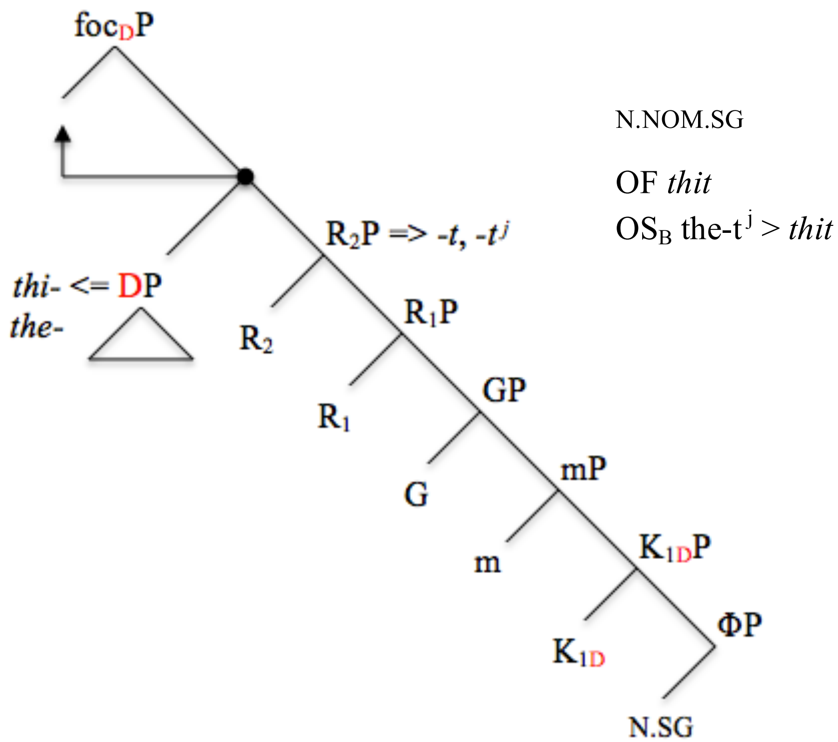


Figure 168 Derivation of the Direct forms without gemination

Next I turn to the second subtype of Direct form.

6.2.3.2 The Direct forms with gemination

The sigmatic reinforcer is suppressed in all of the Direct forms. The Direct forms just discussed in Section 6.2.3.1, moreover, also suppress the consonant geminator. In the OS_A Direct form *thi-t-t*, however, we observe gemination of the K_D component *-t*. This is the second subtype of Direct form.

The consonant geminator in the OS_A Direct form cannot be the same consonant geminator which is present in the K-final-SS forms (e.g. OHG M/N.GEN.SG *dësses*). If it had been the same geminator – i.e. with the structure [G m] – then we would expect also the sigmatic reinforcer to surface in the Direct forms, since the features R_1 and R_2 would still be left over to be lexicalized as *-s*. Instead we need to posit a different kind of consonant geminator for the OS_A form.

Recall from above that the OS_B Direct form displays raising of the root vowel. I assumed above that the underlying form of OS_B *thit* is actually /the-t^j/, where the *j*-mutator causes raising of *e* to *i* and then deletes word-finally. The same phenomenon is observed in OS_A *thitt*, which has the root vowel *i* rather than *e*. In other words, consonant gemination in the OS_A Direct form correlates with vowel-raising. Hence I propose that the consonant geminator in question contains a *j*-mutator, meaning that the phonological shape of the Direct consonant geminator is *-C^j*. Thus the underlying form for OS_A N.NOM/ACC.SG *thitt* is actually /the-t-C^j/.

Furthermore, the structure of the OS_A consonant geminator must account for the suppression of the sigmatic reinforcer. Following a by now familiar strategy, I propose that the structure of *-C^j* contains R_2 and R_1 (the features which normally spell out as *-s*). We may be tempted to posit that the lexical structure of *-C^j*, then, corresponds to the packaging schema in Figure 169, where G and m are also included.

R_2	**R_1**	**G**	**m**	**K**	**D**
-C^j				-t	OS_A the-

Figure 169 Packaging for OS_A *-C^j*

The Roll-up Shortcut gives us the correct result, namely *the-t-C^j*. The derivation of the OS_A form, including foc-movement, is summarized in Figure 170.

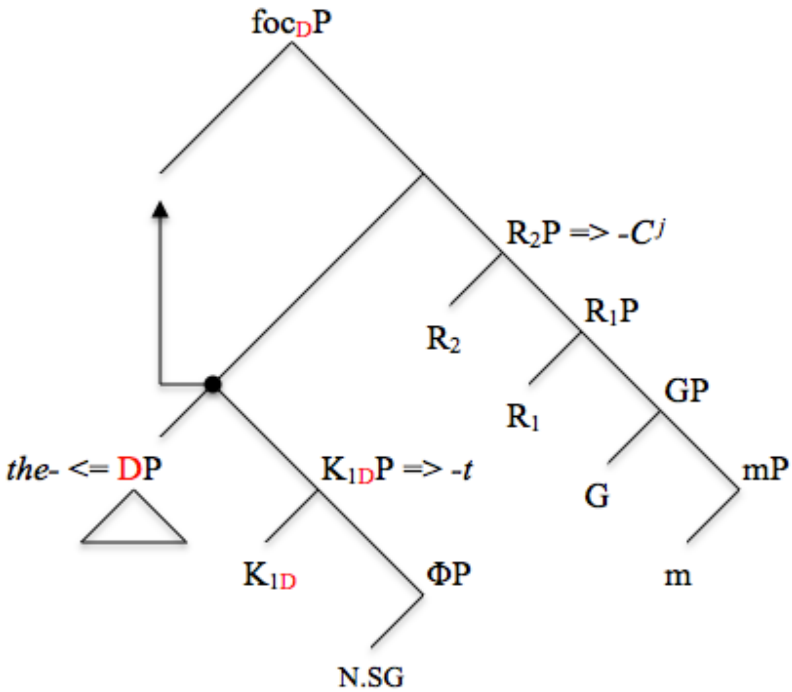


Figure 170 Derivation of OS_A N.NOM.SG the-t-C ^{j} > *thitt*

Since the K_D component is a D-item, there is (string-vacuous) pied-piping of K_DP to foc_DP in Figure 170.

It is interesting to observe here that there is no other consonant geminator in OS which competes with the geminator morpheme -C^j. That is, there is no consonant geminator -*C* corresponding to [G m] in OS (as there is in OF, OE, and OHG). Had there been a geminator like -*C* which was also anchored at the m-layer in OS, then the two geminators would interfere with each other's spellouts (in fact the -C^j geminator would block the -*C* geminator from ever surfacing).

Finally, the reader may have observed that all of the Direct forms fall into McFadden's (2014) NSAG generalization, in that they are 'special' nominative forms, with a systematic accusative syncretism. As discussed in the previous chapter, the NSAG can be thought of in terms of a large chunk of structure being stored with K_1 (and sometimes K_2). This is in fact consistent with my packaging schemas for the Direct forms.

6.2.4 Summary of West Germanic packaging

We have encountered several patterns in WGmc. I have summarized the various ways in which the fseq is packaged in WGmc in Figures 171-174.

<table>
<tr><th>R_2</th><th>R_1</th><th>G</th><th>m</th><th>K</th><th>D</th></tr>
<tr><td colspan="2" rowspan="2">-s</td><td colspan="2">-C</td><td>-$K_{\text{final-SS}}$</td><td rowspan="3">th(i)-</td></tr>
<tr><td colspan="3">-$K_{\text{D-internal}}$</td></tr>
<tr><td colspan="5">-K_{Direct} (-t)</td></tr>
</table>

Figure 171 Old Frisian

<table>
<tr><th>R_2</th><th>R_1</th><th>G</th><th>m</th><th>K</th><th>D</th></tr>
<tr><td colspan="2" rowspan="3">-s</td><td colspan="2">-C</td><td>-$K_{\text{final-SS}}$</td><td rowspan="3">þ(i)-</td></tr>
<tr><td colspan="3">-$K_{\text{final-S}}$</td></tr>
<tr><td colspan="3">-$K_{\text{D-internal}}$</td></tr>
</table>

Figure 172 Old English

<table>
<tr><th>R_2</th><th>R_1</th><th>G</th><th>m</th><th>K</th><th>D</th></tr>
<tr><td colspan="2" rowspan="2">-s</td><td colspan="3">-$K_{\text{final-S}}$</td><td rowspan="4">th(e)-</td></tr>
<tr><td colspan="3">-$K_{\text{D-internal}}$</td></tr>
<tr><td colspan="4">-C^j</td><td>-K_{Direct} (-t) (OS_A)</td></tr>
<tr><td colspan="5">-K_{Direct} (-t^j) (OS_B)</td></tr>
</table>

Figure 173 Old Saxon

<table>
<tr><th>R_2</th><th>R_1</th><th>G</th><th>m</th><th>K</th><th>D</th></tr>
<tr><td colspan="2" rowspan="2">-s</td><td colspan="2">-C</td><td>-$K_{\text{final-SS}}$ (-es)</td><td rowspan="2">dë-</td></tr>
<tr><td colspan="3">-$K_{\text{final-S}}$</td></tr>
</table>

Figure 174 Old High German

It may seem that the WGmc structures discussed in this section are quite different from the ON structures. The next section, however, will show that if we dig a bit deeper into NGmc, similar structures begin to emerge there too, both enriching and binding together our crosslinguistic picture of RDem.

6.3 Older and/or dialectal Norse forms

In this section I continue to expand the empirical scope of my account by discussing forms which lie outside the standard RDem paradigm of classical ON.

6.3.1 *Þenni*

In East Norse, there was a M.NOM/ACC.SG variant *þenni* (Haugen 1982: 101; cf. RN M.ACC.SG **þini**[106]) which underwent a brief period of popularity but subsequently declined in favor of the M.NOM/ACC.SG form *þenna*.[107]

At first glance *þenni* looks like it decomposes as in (179).

(179) M.NOM.SG þa-n-C^{i} > *þenni*
D-K-Gm

The decomposition in (179) cannot be correct, however, since the R features are missing. It is important to recognize that all forms in the RDem paradigm are equally reinforced. The RDem forms may differ along parameters of case, gender, and number, but they do not differ along any parameter of reinforcement. That is, all RDem forms must consist of D, K, G, m, R_1, and R_2.

A better alternative is to put *þenni* on a par with the OS_B form *thit*, as shown by the packaging schema in Figure 175.

R_2	R_1	G	m	K_D	D
-t^{j}					the-
-nn^{i}					þa-

Figure 175 Packaging in OS *thit* and East Norse *þenni*

Thus I am proposing that the underlying form of *þenni* is *þa-nn^{i}* (and not *þa-n-C^{i}*), meaning that *þenni* is composed of two structural chunks just like OS *the-t^{j}* (or the ON constant form *þa-ss^{i}*). The derivation of *þenni* is seen in Figure 176.

[106] Ög 192, Sm 136, U 759.

[107] Originally *þenna* (> *denna*/*denne* in Swedish, Danish, and Norwegian) was only the M.ACC.SG form but due to syncretism it became the M.NOM.SG form as well. The Dem paradigm also underwent such a change: F.NOM.SG *sú* and M.NOM.SG *sá* merged with the M.ACC.SG *þæn(n)*, *þen(n)* (> *den* in Swedish, Danish, and Norwegian).

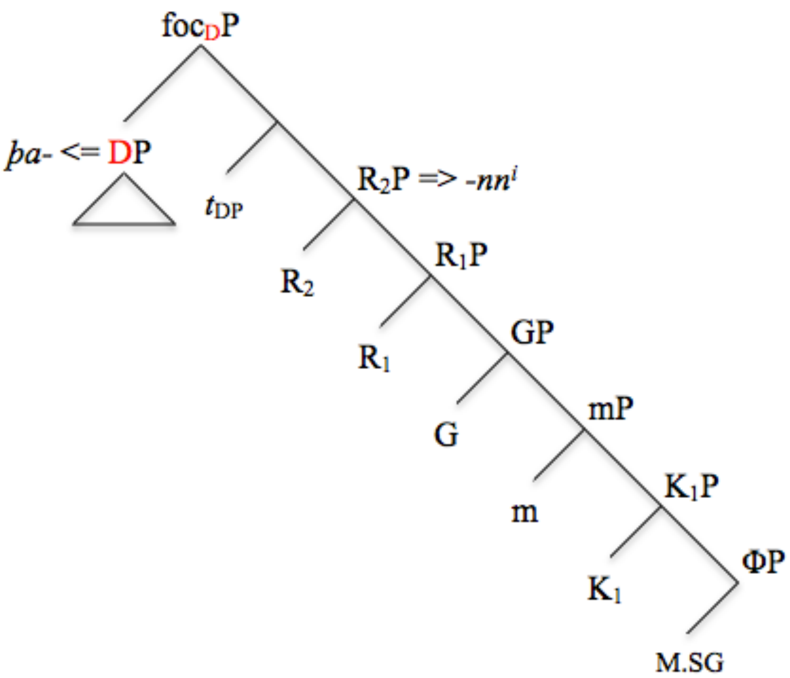


Figure 176 East Norse þa-nn^{i} > *þenni*

Assuming that KP is not pronominal, there is no pied-piping of KP to focP in this structure.

6.3.2 *Þeima*

The next form I will discuss is M.DAT.SG / DAT.PL *þeima*, a skaldic variant of the classical form *þessum* (Heusler 1962: 77, Þórólfsson 1925: 47). This item can be decomposed as *þei-m-a*, namely the regular M.DAT.SG / DAT.PL Dem form *þei-m* plus the asigmatic reinforcer -*a*. The sigmatic reinforcer -*s* is not present.

Again there is a simple explanation for this form if we pay attention to what we have learned about WGmc. In the WGmc K-final-S forms, gemination of the sigmatic reinforcer is suppressed because G and m are packaged together with K. Suppose then that the ON K_D ending -*m* has a lexical structure that contains G and m. If -*m* additionally includes R_1, then only R_2 is left over, which as we know corresponds to the asigmatic reinforcer -*a*. This packaging hypothesis is summarized in Figure 177.

R_2	**R_1**	**G**	**m**	**K**	**D**
-a	-m				þei-

Figure 177 Packaging for *þeima*

Performing the Roll-up Shortcut will produce the attested order of morphemes, namely *þei-m-a*. Foc-movement will move both DP and KP (since *-m* is pronominal), but it will not alter the linear order of the morphemes.

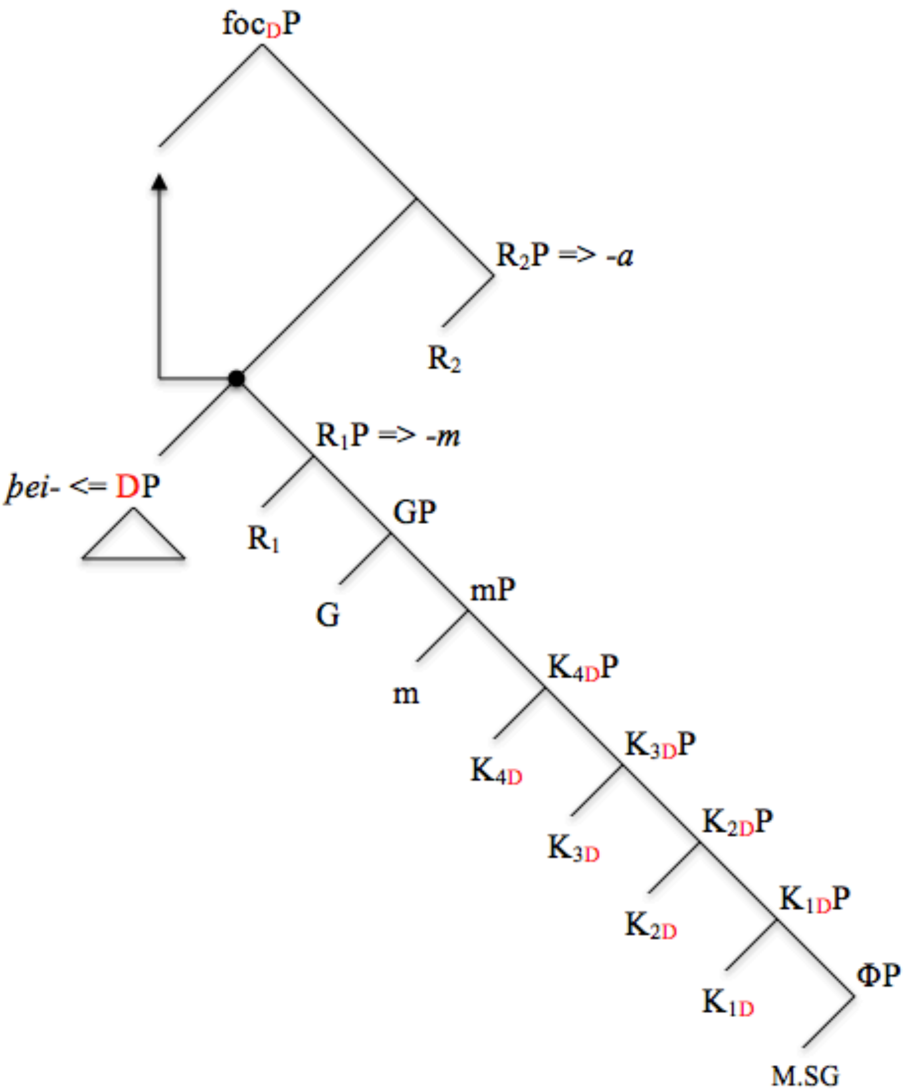


Figure 178 Foc-movement in M.DAT.SG *þeima*

Once again the WGmc perspective provides a useful insight into the structure of a Norse variant which does not resemble the classical ON forms.[108]

[108] Old Swedish also shows M.DAT.SG / DAT.PL *þæmma* (Noreen 1904: 397-398), with gemination of the K component: *þæ-m-C-a*. Here we can simply assume the same packaging as in the ON K-internal forms.

6.3.3 The diachronic lead-up to *þenna* and *þetta*

The ON K-final forms M.ACC.SG *þenna* and N.NOM/ACC.SG *þetta* have an interesting history which can be quite clearly traced in the RN inscriptions. All the stages in (180) are well attested (see Appendix I).

(180) (i) þa-t-si [Dem-*si* without *i*-umlaut]
þa-n-si

(ii) þe-t-si [Dem-*si* with *i*-umlaut]
þe-n-si

(iii) þe-t-s-a [coexistence of sigmatic and asigmatic reinforcers]
þe-n-s-a

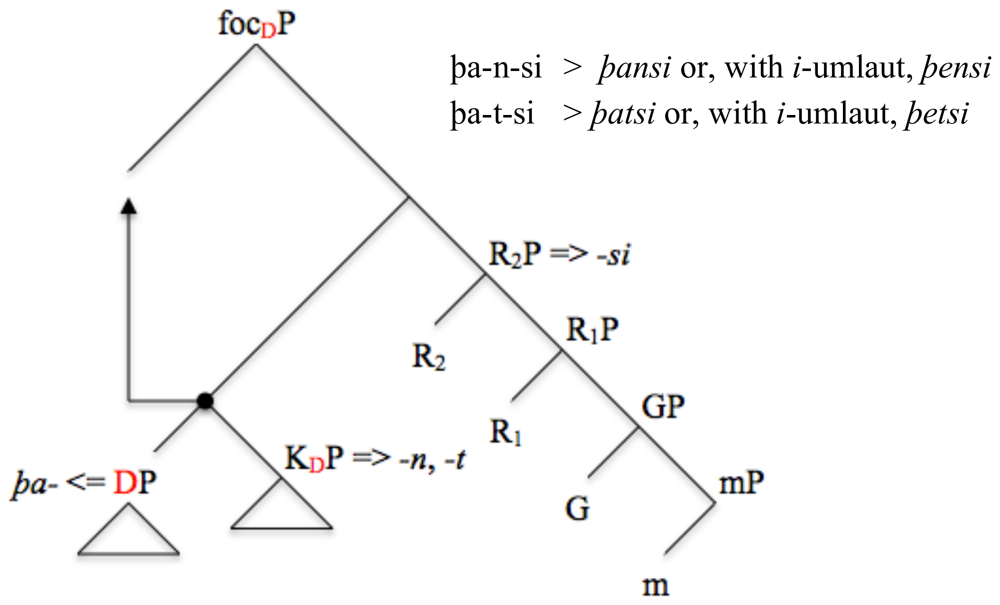
(iv) þe-t-t-a [classical K-internal forms]
þe-n-n-a

The old reinforcer *-si* in stages (i) and (ii) must correspond to the entire span of reinforcer features, as shown by the packaging in Figure 179.

R_2	R_1	G	m	K	D
-si				-n, -t	þa-

Figure 179 Packaging in stage (i)/(ii)

The Roll-up Shortcut results in *þa-n-si* and *þa-t-si*. Since stages (i) and (ii) are Dem-*si* forms, the K component is pronominal and thus it is pied-piped to focP, as seen in Figure 180.



where KP = [K_2P K_1P M.SG] for *-n* or [K_2P K_1P N.SG]/[K_1P N.SG] for *-t*

Figure 180 Derivation of stage (i)/(ii) RN forms

There is no structural difference between stages (i) and (ii), only a phonological one. In stage (i) *-si* does not condition *i*-umlaut of the root vowel *a*, but in stage (ii) it does, fronting *a* to *e*.

In stage (iii), the asigmatic reinforcer *-a* arises. Recall that *-a* is a NGmc innovation which by all accounts is a later development than the old sigmatic reinforcer *-si* (which is common to both NGmc and WGmc). Now, what is interesting about the stage (iii) forms *þetsa* and *þensa* is that they do not yet display consonant gemination. Instead of *-C* we observe the *-s* from *-si.* At this pre-classical stage of Norse, in other words, *-a* and *-s(i)* were not in complementary distribution but actually coexisted. This development is naturally explainable in nanosyntactic terms. Consider Figure 181.

(a) Stage (i) *þat-**si*** / stage (ii) *þet-**si*** > (b) Intermediate stage > (c) Stage (iii) *þet-**s**-**a***

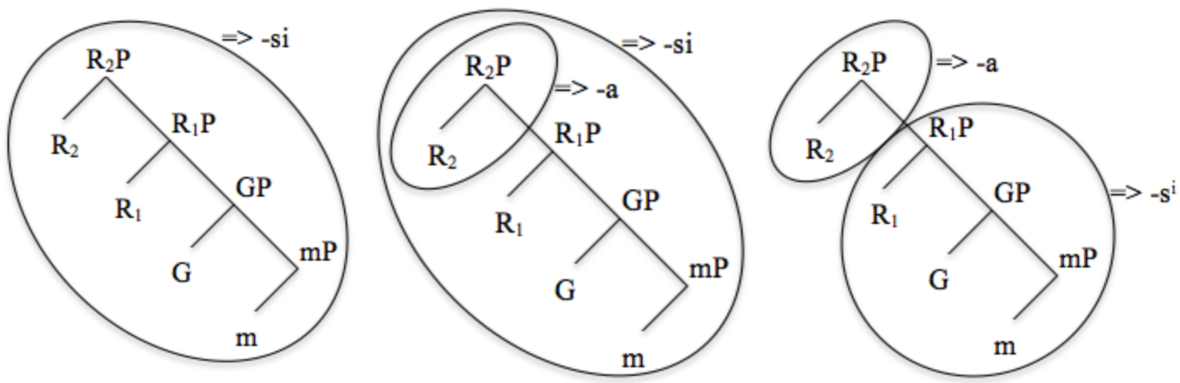


Figure 181 Development of reinforcer L-trees from stage (i)/(ii) to stage (iii)

The three stages of development in Figure 181 can be described as follows:

(a) In stages (i) and (ii), the old sigmatic reinforcer -*si* corresponds to the entire chunk of structure [R_2P [R_1P [GP [mP]]]].

(b) Next, the new asigmatic reinforcer -*a* arises. Its lexical structure corresponds to R_2P. The lexical structure of the sigmatic reinforcer -*si* continues to be [R_2P [R_1P [GP [mP]]]].

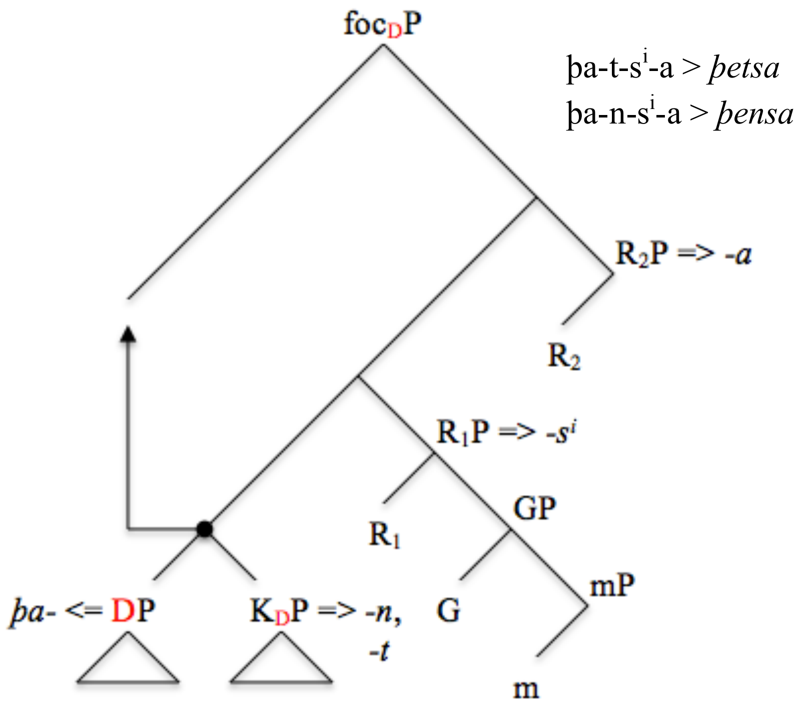
(c) The introduction of the new reinforcer -*a* into the system puts pressure on -*si*, ultimately causing the structural domain of -*si* to shrink. Eventually the lexical structure of -*si* changes. As seen in stage (iii), the sigmatic reinforcer -*s^i^* has been pushed down a layer. The structure of the sigmatic reinforcer is now [R_1P [GP [mP]]].[109]

Floating *i* in -s^i fronts the root vowel but is then deleted before vowels: *þa-n-*s^i*-a* > *þensa* and *þa-t-*s^i*-a* > *þetsa*. The packaging of a stage (iii) form is given in Figure 182, and its derivation is summarized in Figure 183.

R_2	**R_1**	**G**	**m**	**K**	**D**
-a	-s^i			-n, -t	þa-

Figure 182 Packaging in stage (iii)

[109] Observe that the phonological shape of the -*si* reinforcer has undergone a change to -s^i. This is an intermediate form between the old reinforcer -*si* and the classical sigmatic reinforcer -*s* (i.e. -*si* > -s^i > -*s*).



where KP = [K$_2$P K$_1$P M.SG] for *-n* or [K$_2$P K$_1$P N.SG]/[K$_1$P N.SG] for *-t*

Figure 183 Stage (iii) structure[110]

Observe that at this stage the RDem form is still a Dem-*si* form, so K is pronominal. Hence DP and KP both move to focP, though this movement is string-vacuous.

Finally we reach stage (iv), namely the classical K-internal forms *þetta* and *þenna* with which we are already familiar. By this stage *-s^i* has disappeared in the K-internal forms and been supplanted by the geminator. This development is sketched in (181).

(181)	Stage (iii)	>	Stage (iv)
	þe-t-**s**-a	>	þe-t-**C**-a
	þe-n-**s**-a	>	þe-n-**C**-a

The development as it is sketched in (181) gives the impression that *-C* in stage (iv) is a direct development of *-s* from stage (iii). Indeed, as mentioned there is a tempting – yet spurious – view in the literature that the change from *þetsa* to *þetta* and *þensa* to *þenna* can be explained purely in terms of a phonological change (as opposed to a structural change). In their discussion of these forms, Armitage (1911: 207) proposes the rule *ts* > *tt*

[110] Archaic N.DAT.SG *þvísa* (i.e. *því-s-a*) also has this structure, modulo the fact that Dem *því* is an irregular form involving a pointer entry covering the entire [[DP] KP] constituent.

in order to account for *þetsa* > *þetta*, and Haugen (1982: 101) suggests the rule *ns* > *nn* in order to account for *þensa* > *þenna*. Put simply, these rules posit a simple phonological shift from *s* to *C*. However, these rules are completely *ad hoc*: there is not nearly enough evidence in early Germanic for general rules like *ts* > *tt* or *ns* > *nn* (H.F. Nielsen, p.c.).

The analysis developed in Chapter 5 fits the facts better than the spurious view that *ts* becomes *tt* and *ns* becomes *nn*. The fine-grained view in Section 5.2.1 of the ON K-internal forms involves a K morpheme which is packaged with m, and a geminator morpheme corresponding to the R_1 and G layers. With this in mind, the change from stage (iii) to stage (iv) is now seen in Figure 184.

Stage (iii) > Stage (iv)

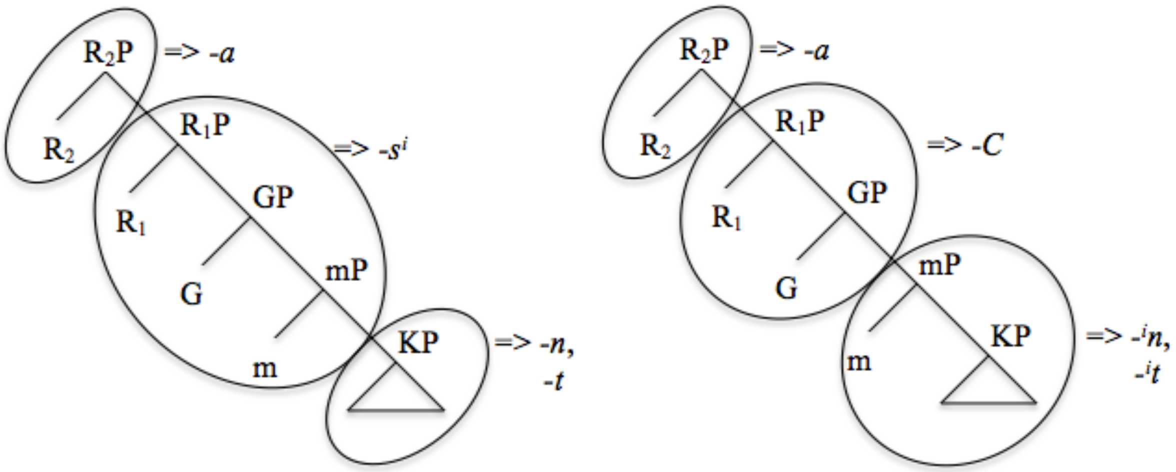


Figure 184 Stage (iii) to stage (iv)

Figure 184 shows a structural change that breaks up the $-s^i$ morpheme: the m layer is lexically repackaged with the K suffix, and the G and R_1 layers become the lexical structure for -*C*. Such an account correctly reflects the fact that the change from stage (iii) to (iv) must have been structural and not phonological.

6.3.4 *Sjá*

Readers familiar with ON will know that *þessi* is not the prototypically cited F/M.NOM.SG form, rather *sjá* is. The origins of this more archaic form becomes clear once the rule in (182) is introduced.

(182) Hiatus vowels from Proto-Norse to West Norse (Haugen 1982: 35)

*ēa > jā

*ēu > jō

In addition to this historical rule, there is also a synchronic rule of ‘hiatus resolution’ in ON. In the words of Faarlund (2004: 14): “When a root ending in *é* is followed by a suffix starting with a back vowel, the stress and the length are shifted to the vowel of the suffix, and the root vowel becomes a semivowel.” The examples he gives are *fé-ar* ‘money-GEN’ > *fjár* and *sé-um* ‘see-1PL’ > *sjóm*. See also Gordon (1956: 277).

The rule in (182) is the key to *sjá*: PN **sē-a* (where *-a* is the asigmatic reinforcer) goes to *sjā*, or, synchronically put, underlying /sē-a/ surfaces as *sjā*. The first component here, namely *sē-*, is an *i*-umlauted allomorph of the M.NOM.SG Dem form *sā* (< *saR*). The development in (183) gave rise to this umlauted stem. The stage (i) form *saR-si* (RN **saRsi**) corresponds to the Dem-*si* stage of the M.NOM.SG RDem. Stage (ii) involves the loss of pleonastic *-R* in *saR-si*, which gives rise to compensatory lengthening of *a*, yielding *sā-si* (RN **sasi**). In stage (iii) there is *i*-umlaut of *ā* to *ē*, yielding the new stem *sē-*, which then takes the asigmatic reinforcer *-a*. The underlying form *sē-a* yields *sjā*, which persists into the early stages of classical ON as well (i.e. stage (iv)). By this stage the F.NOM.SG has also become syncretic with the M.NOM.SG.

(183) The stem *sē-*

(i) M.NOM.SG saR-si > (ii) M.NOM.SG sā-si > (iii)/(iv) M/F.NOM.SG sē-a (*sjā*)

Sjá can be considered a K-internal form: in Table 74 we see that the *sjá* forms are adjacent to the K-internal forms. The hypothesis that *sjá* is a K-internal form also makes sense of the fact that the *sjá* makes use of the asigmatic reinforcer *-a*, something that only K-internal forms do (see also Klingenschmitt 1987: 188).

Table 74 *Sjá* as a K-internal form

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	sjá	sjá	þetta	þessar	þessir	þessi
ACC	þessa	þenna	þetta	þessar	þessa	þessi
GEN	þessar	þessa	þessa	þessa	þessa	þessa
DAT	þessi	þessum	þessu	þessum	þessum	þessum

Traditional accounts have no way of explaining why some forms take the asigmatic reinforcer while others do not. The approach developed in this dissertation is that the shape of the K-internal lexical entries can account for such paradigmatic patterns. It would of course be desirable to bring *sjá* into the fold as well.

First consider stages (i) and (ii). These forms are simply Dem-*si* forms. Their packaging is straightforward, as seen in Figure 185. The only minor complication to keep in mind is that the M.NOM.SG Dem form *sā(R)* is a phonological idiom. That is, the item *sā(R)* overrides the regularly formed but unattested M.NOM.SG structure **þa-r*. The reinforcer *-si* corresponds to all of the features in the R/Gm domain.

R_2	R_1	G	m	K	D
-si				~~-r~~	~~þa-~~
				sā(R)	

Figure 185 Packaging in stages (i) and (ii)

The same basic packaging schema was seen in the previous section for the Dem-*si* forms *þansi*/*þatsi* and *þensi*/*þetsi*.

By stage (iii), the *i*-umlauted allomorph of *sā*, namely *sē*-, arises. The question is what the lexical structure of this new stem is. Recall from the previous section the packaging schema for the stage (iii) forms *þensa* and *þetsa*, repeated in Figure 186.

R_2	R_1	G	m	K	D
-a	-s^i			-n, -t	þa-

Figure 186 Packaging for stage (iii) forms *þensa* and *þetsa*

Even though *sē*- is a stage (iii) form, it does not show the sigmatic reinforcer -*s*^{i} (**sē*-*si* is unattested). By the logic followed thus far, this means that the features for -*s*^{i} – namely m, G, and R_1 – should be packaged along with some other morpheme. I will propose the following.

- First, there existed pronominal K endings M.NOM.SG -^{i}*r* and F.NOM.SG -^{i}*u* (on a par with the *i*-mutated endings of the K-internal forms as discussed in Sections 5.2.1 and 5.2.2).
- Second, these endings lexically correspond to K, m, G, and R_1.
- Finally, these endings participate in a pointer entry for the stem *sē*-.

The pointer entry for *sē*- is illustrated in lexical entry 2388 in (184). The basic idea is that the configurations **þa*-^{i}*r* in the M.NOM.SG and **þa*-^{i}*u* in the F.NOM.SG are overridden by the special phonology *sē*-.

(184) **Lexicon**

<$_{210}$ *þa*- ⇔ DP >

<$_{324A}$ -^{i}*u* ⇔ R_1P GP mP K_{1D}P [$_{\Phi P}$ F.SG] >

<$_{325A}$ -^{i}*r* ⇔ R_1P GP mP K_{1D}P [$_{\Phi P}$ M.SG] >

<$_{2388}$ *sē*- ⇔ [[210] 324A / 325A] >

<$_{865}$ -*a* ⇔ R_2P >

The packaging schema is shown in Figure 187. Since *sē-* corresponds to the structure [[DP] R_1P GP mP KP], R_2P is left over to be lexicalized as the asigmatic reinforcer *-a*.

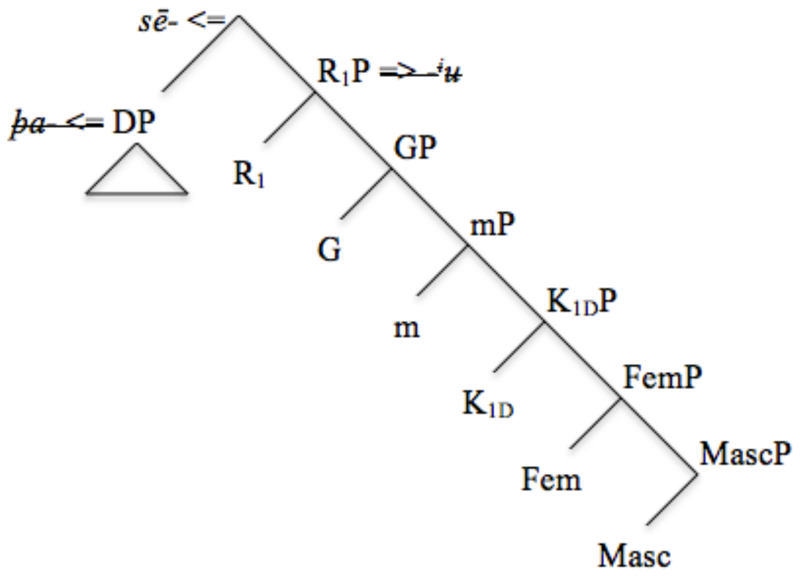
<table>
<tr><th>R₂</th><th>R₁</th><th>G</th><th>m</th><th>K</th><th>D</th></tr>
<tr><td rowspan="2">-a</td><td colspan="4">~~-^{i}u, -^{i}r~~</td><td>~~þa-~~</td></tr>
<tr><td colspan="5">sē-</td></tr>
</table>

Figure 187 Packaging for *sjá*

The result after spellout-driven movement is correct: *sē-a* > *sjā*.

Entry 2388 in (184) refers to *either* 324A (-^{i}*u*) *or* 325A (-^{i}*r*). If we venture a guess at the internal structure of gender features – here I hypothesize that they are below the K domain and that FemP is above MascP – then we can say that both the S-trees in Figure 188 are matched by entry 2388.

(a) F.NOM.SG



(b) M.NOM.SG

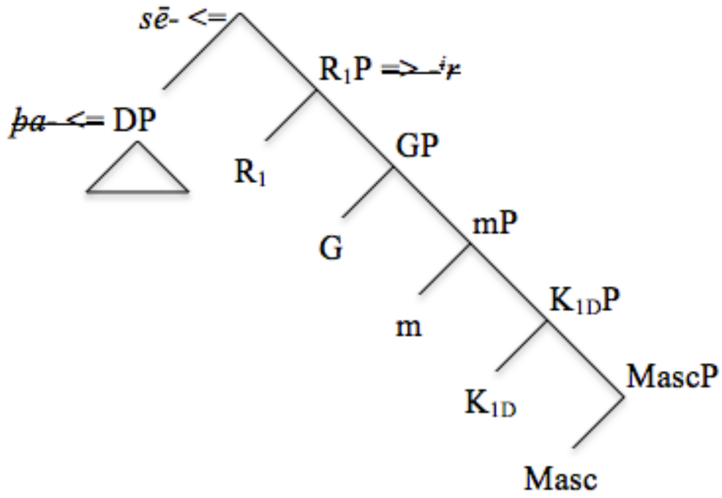


Figure 188 S-trees which may be matched by entry 2388

Observe that the only difference between the two S-trees in Figure 188 is the presence or absence of Fem. The F.NOM.SG structure in (a) shows the full gender structure [FemP [MascP]], while the M.NOM.SG structure in (b) shows a subset of this structure, namely MascP. In other words, the L-tree for *sē-* requires the "Superset [Principle] to restart 'in the middle' of the sequence" as Caha & Pantcheva (2012) have put it (see also Vanden Wyngaerd 2014). According to the approach of Caha & Pantcheva (2012) and Vanden Wyngaerd (2014), a pointer can point to a piece of structure which is calculated by the Superset Principle separately from the rest of the structure. Thus, we can write entry 2388 as in Figure 189. The pointer to [FemP [MascP]] means that this part of the L-tree can match either [FemP [MascP]] (for -^{i}*u*) or the subset MascP (for -^{i}*r*).

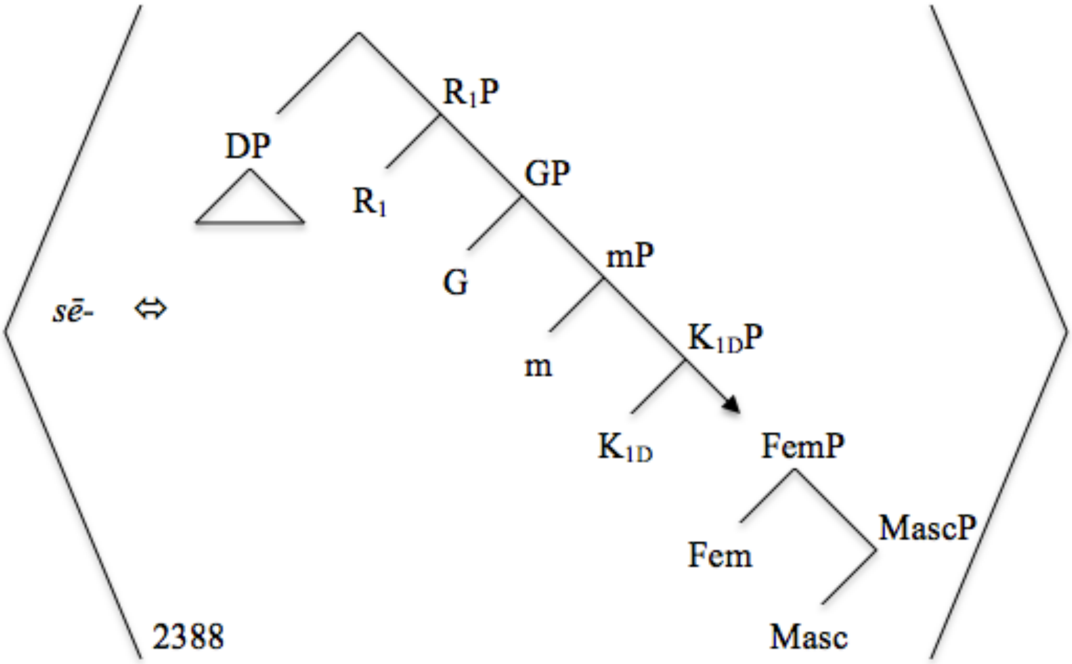


Figure 189 Entry 2388 with a syncretic pointer

The pointer to [FemP [MascP]], then, is a way to capture the syncretism between feminine and masculine (in the NOM.SG).

Movement to focP will target the constituent corresponding to *sē-*, since K is pronominal (K_D) in the structure of *sjá*. Thus DP pied-pipes KP on its way to focP, as seen in Figure 190.

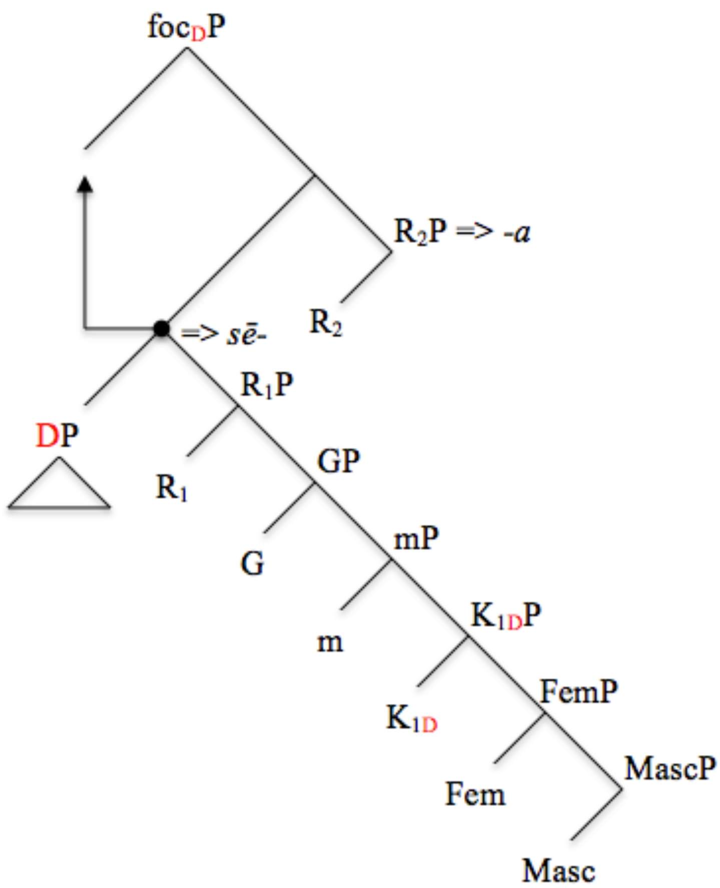


Figure 190 Foc-movement in F.NOM.SG *sjá*

On this note, it is interesting that *sjá* takes pronominal K (K_D) yet still persists into the early stages of classical ON (stage (iv)), where inflection is on its way to becoming solely adjectival in the RDem paradigm. Indeed, *sjá* cannot involve adjectival K: if it did have adjectival K, then foc-movement would attract DP only, which in turn would break up the structure of the idiomatic *sē*-stem and prevent *sē-* from spelling out at all. See Figure 191.

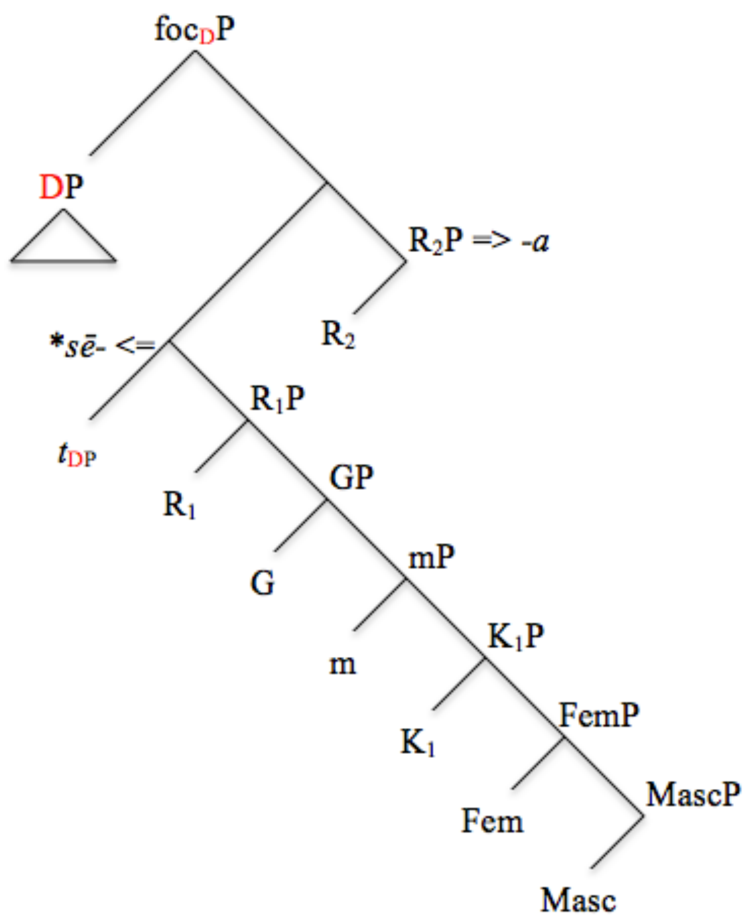


Figure 191 Incorrect foc-movement (*sē-* cannot spell out due to subextraction of DP)

This means that in the early classical ON RDem paradigm, *sjá* is a pronominal form in the midst of adjectival forms. This brings early classical ON a bit closer to its WGmc sisters, whose RDem paradigms regularly display a mixture of pronominal and adjectival K, as seen above. When *sjá* disappears from the RDem paradigm and is replaced by *þessi*, then the paradigm becomes entirely adjectival.

6.4 Pronominal vs. adjectival classification

6.4.1 Leveling out the RDem paradigm

The *sjá* discussion highlights an important historical shift that took place in the history of RDem. As just mentioned in the previous section, *sjá* is a development of the M.NOM.SG Dem-*si* form *sā-si*, which involves the M.NOM.SG phonological idiom *sā*. However, the form *sjá* is eventually replaced by the later classical form *þessi*, which is built on the regular stem *þa-* instead. Indeed, the shift from the Dem-*si* stage to the classical RDem stage is characterized by a gradual smoothing out of irregularities. In its very earliest stages, the RDem paradigm is inextricably linked to the Dem paradigm and all the irregularities thereof. In fact, at this stage RDem simply consists of a Dem form – be it a regular form like F.GEN.SG *þei-rar* or N.NOM/ACC.SG *þa-t*, or an irregular form like M.NOM.SG *sá* or F.NOM.SG *sú* – plus the reinforcer *-si*. Gradually, however, a disconnect develops between Dem and RDem, releasing RDem from the idiosyncrasies of the Dem paradigm. Most notably, RDem develops a base *þe-* throughout the entire paradigm, eliminating older remnants of the Dem paradigm, such as *sjá*. The shift from pronominal inflection (K_D) in the Dem-*si* stage to strong adjective inflection (K) is part of this general development.

I believe that class membership is the key to understanding this change. Recall that there is a disparity between the K endings of Dem and RDem. Above we have encountered two main patterns. In the WGmc languages, RDem displays a mixture of pronominal K_D and adjectival K. In NGmc (ON), on the other hand, RDem displays adjectival K only (with the exception of early classical *sjá*). We know that the WGmc forms with pronominal inflection (K_D) reveal an older layer of RDem history, and that the forms with adjectival inflection (K) represent later stages in the history of RDem.

Let us posit a set of class features belonging to the K domain. Call the pronominal class feature Cl_D and the adjectival class feature Cl_A.[111] This is seen in (185).

(185) (a) Cl_D Cl_A = pronominal

(b) Cl_A = strong adjectival

A cross-class syncretism between D and A, such as ON N.NOM/ACC.SG *-t* (cf. Dem *þa-**t***, RDem *þe-**t**-t-a*, strong adjective *jarp-**t***), is shown in the lexical entry in (186).

(186) < *-t* ⇔ K_2P K_1P → Cl_DP Cl_AP >

The pointer in the L-tree in (186) ensures that all of the S-trees in (187) can be matched.

[111] We probably need more than just the class features Cl_D and Cl_A. For instance, the Gothic third person pronoun paradigm already shows that we also need something like Cl_3, since the F.ACC.SG / N.NOM/ACC.PL ending *-a* of *ij-a* patterns with adjectives (*blind-a*) against the Dem ending *-ō* (*þ-ō*). On the other hand, the adjectival F.DAT.SG ending *-ái* (*blind-ái*) patterns against Dem and third person pronominal *-zái* (*þi-zái*, *i-zái*). This points to a three-way distinction (at least), and syncretisms show us that the order of these three features must be $Cl_D > Cl_3 > Cl_A$ or $Cl_A > Cl_3 > Cl_D$. The fact that *-zái* morphologically contains *-ái* suggests that the former hierarchy is correct. Related to this line of reasoning, the fact that Dem endings are typically more morphologically complex than adjective endings suggests that Dem endings contain more structure and therefore that Cl_D is higher than Cl_A: F.NOM.SG OS *-iu*, OE *-ēo*, ON *-u* (on the view that ON *sú* is underlyingly *þa-u*) vs. OS *-Ø*, OE *-Ø*, ON *-Ø*$^{(u)}$.

(187) Dem NOM.SG $[K_1P \quad Cl_DP\ Cl_AP]$

Dem ACC.SG $[K_2P \quad K_1P \quad Cl_DP\ Cl_AP]$

Adj NOM.SG $[K_1P \quad Cl_AP]$

Adj ACC.SG $[K_2P\ K_1P \quad Cl_AP]$

In other words, both [KP [Cl_DP [Cl_AP]]] (for Dem) and [KP [Cl_AP]] (for RDem/strong adjectives) can spell out as *-t*.

Syncretism is not always the case, however. For example, consider the ON K_D ending M.DAT.SG *-m* vs. the K ending *-um*, given in (188).

(188) (a) < *-m* ⇔ $K_4P\ K_3P\ K_2P\ K_1P\ Cl_DP\ Cl_AP$ > (M.DAT.SG / PL.DAT Dem *þei-**m***)

(b) < *-um* ⇔ $K_4P\ K_3P\ K_2P\ K_1P\ Cl_AP$ > (M.DAT.SG / PL.DAT RDem *þess-**um***)

Here there are two separate lexical entries, one for Dem structures with [Cl_DP [Cl_AP]] and one for adjectival structures with Cl_AP.

In the early stages of RDem's history, RDem is still classified as a pronoun (class D), hence the feature Cl_D must be involved whenever the syntax builds an RDem form. Thus, in this period the syntactic structure of the K ending will contain [**Cl_DP** [Cl_AP]]. This means that only inflectional endings with Cl_D in their lexical structures will be able to match RDem structures in the syntax. Inflectional endings without Cl_D (purely adjectival endings) will not be able to match the RDem structures. In the later stages, however, RDem becomes classified as an adjective (class A), hence the feature Cl_D will not be built during the syntactic derivation anymore (as this is what it means to be an adjectival structure). Thus the syntactic structure of the K ending will contain simply Cl_A. Lexical entries with only Cl_A will be better matches than those lexical entries with both Cl_D and Cl_A (by the Elsewhere Principle). This is summarized in Figure 192.

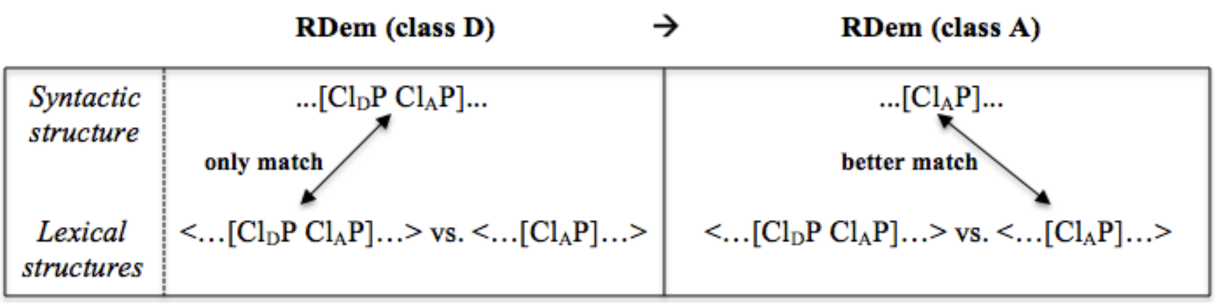


Figure 192 The shift from class D to class A and its effect on matching

Endings which are syncretic across the D and A classes will match either way, as needed.

As a concrete example, take the old N.DAT.SG form *því-s-a*, which was ousted by the later form *þe-s-s-u*. The N.DAT.SG Dem form is *því*, a phonological idiom with the basic structure [[DP] KP], plus the *-s*^{i} and *-a* reinforcers. In other words, *því-s-a* is a stage (iii) form. See Figure 193 for its packaging schema.

<table>
<tr><th>R_2</th><th>R_1</th><th>G</th><th>m</th><th>K</th><th>D</th></tr>
<tr><td rowspan="2">-a</td><td colspan="3" rowspan="2">-s^i</td><td>~~-u~~</td><td>~~þ(a)-~~</td></tr>
<tr><td colspan="2">því</td></tr>
</table>

Figure 193 Packaging for stage (iii) form *því-s-a*

Let us zoom in on the lower part of Figure 193. Since *því* is a Dem form, by definition it must contain Cl_D. Thus, within *því*, there is a KP within which the class feature Cl_D is embedded. Now, my hypothesis states that the switch from class D to class A amounts to the syntax not building Cl_D anymore during the derivation of an RDem form. This has consequences for a form like *þvísa*, where a Dem form is an integral part of the structure. If RDem structures systematically lack Cl_D, then the match between RDem structures and lexical entries containing Cl_D will be less than perfect. Instead, lexical entries without Cl_D will become more suitable matches for RDem structures. Since the lexical structure of *því* necessarily involves a Cl_D feature, it will only be a suitable match during the stage when the syntactic structure of RDem contains Cl_D. When RDem's structure stops containing Cl_D, the lexical structure of *því* becomes less of a suitable match. Thus the irregular Dem form *því* will not enter into the RDem derivation anymore, and instead the regular D and K ingredients (*þa-* and *-u*, respectively) will surface individually. Ultimately, this results in the regularly formed stage (iv) form *þessu*. In this way, the RDem paradigm is leveled out over time, in the sense that it becomes less and less 'reliant' on the Dem paradigm and its class feature Cl_D.

6.4.2 Choosing the simpler stem

Something which has been taken for granted so far is that *þa-* (and not *þei-*) is the RDem stem in ON. Here I will briefly touch on why this might be the case.

Many aspects of the history of RDem involve a simplification of structure (plausibly some form of grammaticalization). In the previous section, for instance, I hypothesized that the loss of Cl_D is one part of this process. Another potential characteristic of the process is the overall shift over time from internal to external inflection. In ON, at least, external inflection in the boxed forms is due to 'successive cyclic' movement of DP to focP. This would mean that the boxed/K-final forms are less 'morphological' (i.e. spellout-driven) and more 'syntactic' (cf. the two types of movement discussed by

Jayaseelan 2010). This can be thought of in terms of the historical tendency to shift from synthetic (more ‘morphology’-based) to analytic (more ‘syntax’-based).[112]

I think the choice of *þa-* over *þei-* is also due to a process of structural simplification. It can be seen in the Dem paradigm, for instance, that the *þa-* stem gravitates towards the top and the *þei-* stem more toward the bottom. This can be seen in Table 75, where *þa-* is in light gray and *þei-* is in dark gray.

Table 75 *þa-* is lighter than *þei-*

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	sú	sá	þa-t	þ-ær, þ-ár	þei-r	þa-u
ACC	þ-á	þa-nn	þa-t	þ-ær, þ-ár	þ-á	þa-u
GEN	þei-r(r)ar	þe-ss	þe-ss	þei-r(r)a	þei-r(r)a	þei-r(r)a
DAT	þei-r(r)i	þei-m	þ-ví, þ-ȳ	þei-m	þei-m	þei-m

We know that the K domain is cumulative, meaning the lower cases genitive and dative are more structurally complex than the upper cases nominative and accusative (Caha 2009). Thus – whatever the correct analysis for tracking *þa-* and *þei-* in the paradigm turns out to be – we can say with some confidence that the simpler and less structurally complex stem is *þa-*. Thus, the absence of *þei-* in the later stages of the RDem paradigm can be seen as a choice of the simpler stem *þa-* over the more complex one *þei-*. The precise implementation I leave for future work.

6.5 The cross-Germanic picture

In this section I summarize the crosslinguistic picture which has emerged. While a fair amount of variation has been revealed across the Old Germanic sister languages, all of it can be explained by positing a single fseq and hypothesizing that this sequence is packaged differently in the languages’ lexicons. The crosslinguistic variation observed across the NWGmc RDem paradigms is summarized in Table 76.

[112] I use scare-quotes because in nanosyntax there is no absolute distinction between morphology and syntax.

Table 76 Crosslinguistic variation in the Northwest Germanic reinforced demonstrative

	WEST GERMANIC					NORTH GERMANIC
Lexical entry	OE	OF	OS_B	OS_A	OHG	ON
< -s ⇔ [R_2 R_1] >	✓	✓	✓		✓	< -s ⇔ [R_2 R_1] >
						< -a ⇔ R_2P >
< -C ⇔ [G m] >	✓	✓			✓	< -C^i ⇔ [G m] >
< -C^j ⇔ R_2P R_1P GP mP >				✓		< -C ⇔ R_1P GP >
-K_D & -K	✓	✓	✓		✓	-K
< -$K_{final\text{-}S}$ ⇔ GP mP KP >	✓		✓		✓	
< -$K_{final\text{-}SS}$ ⇔ KP >	✓	✓			✓	< -$K_{final\text{-}SS}$ ⇔ KP >
< -$K_{D\text{-}internal}$ ⇔ GP mP K_DP >	✓	✓	✓			< -ss^i ⇔ R_2P R_1P GP mP $K_{constant}P$ >
< -K_{Direct} ⇔ R_2P R_1P GP mP K_DP >		✓	✓			< -$^iK_{internal}$ ⇔ mP KP >
< -K_{Direct} ⇔ K_DP >				✓		
RDem base	< þ(i)- ⇔ DP >	< th(i)- ⇔ DP >	< th(e)- ⇔ DP >		< dë- ⇔ DP >	< þa- ⇔ DP >

Various differences and similarities between the West and North branches become clear in Table 76. Let us briefly discuss each category in turn.

6.5.1 Reinforcers

In both branches we see a sigmatic reinforcer -*s* (< NWGmc *-*si*) whose lexical structure is a complex head consisting of the features R_2 and R_1.

In each branch we see two kinds of consonant geminator reinforcers: one with a mutator and one without. The mutator-less geminator -*C* in WGmc (present in all of WGmc except OS) corresponds to the complex head [G m], while the geminator with a mutator -C^j (present only in OS_A) corresponds to the phrase [R_2P [R_1P [GP [mP]]]]. In NGmc, the mutator-less geminator -*C* corresponds to the phrase [R_1P [GP]], while the geminator with a mutator -C^i corresponds to the complex head [G m].

Finally, there is one reinforcer that is specific to NGmc. This is the asigmatic reinforcer -*a*, with the lexical structure R_2P.

6.5.2 Case

There is an important difference between the RDem paradigms of WGmc and NGmc when it comes to the nature of their case endings. The WGmc RDem paradigms take mostly strong adjectival case endings (K), but in every language there is also a subset of forms with pronominal inflection (K_D) instead. The NGmc RDem paradigm makes use exclusively of strong adjectival endings (though some exceptions have been noted throughout this dissertation).

There are various ways in which reinforcer features can be packaged with K features. In all of the WGmc languages except for OF we see K-final forms without gemination of the sigmatic reinforcer -*s*. This is accounted for by packaging the geminator features G and m with K. In the K-final forms with gemination of -*s*, K is not packaged with G and m. In the NGmc K-final forms there is always gemination of -*s*, and thus G and m are not packaged with K. ON and OF are alike, then, in that their K-final forms always have gemination of the sigmatic reinforcer and thus do not package G and m with K.

In the D-internal forms of WGmc, gemination of -*s* is also suppressed, so K is packaged with G and m. An important detail is that K in these forms is pronominal (K_D), as opposed to the K-final forms where K is adjectival. This has consequences for foc-movement, as discussed above in relation to the ProK generalization.

In the Direct forms of WGmc there is some variation. In OS_A the case ending -*t* corresponds simply to K_DP. This allows for the geminator -C^j to arise later in the derivation, giving OS_A *thitt*. In OS_B and OF, on the other hand, there is no such gemination of *-t*. In these languages, *-t* corresponds to the entire span from R_2 down to K. As with the D-internal forms, K in the Direct forms is pronominal (K_D).

In the constant forms of NGmc, there is a reinforcer morpheme *-ss*^{i} which corresponds to the entire span of layers from R_2 down to K. In Chapter 5 I discussed how this morpheme can be understood in terms of McFadden's (2014) NSAG.

In the K-internal forms of NGmc, there is packaging of the *i*-mutator feature m with K. This kind of lexical packaging guarantees the expression of *-C* and *-a* at later stages of the derivation.

6.5.3 Base

A clear commonality between WGmc and NGmc is an RDem base beginning with a voiceless interdental fricative plus a vowel, except in OHG *dë-* where there is a dental stop instead (due to the sound change *þ/ð* > *d* in German).

In OS, OHG, and ON, moreover, the RDem base can be traced back to the Dem paradigm. In OE and OF, however, the RDem base is not identical to a Dem base, but is specific to the RDem paradigm.

6.6 Summary

By taking into consideration the RDem paradigms of WGmc alongside NGmc, a rich panorama of variation has emerged. Importantly, this variation is constrained in ways that we can understand quite precisely. The functional sequence is the same for all languages, but it can be packaged differently in each one, and different lexical entries give rise to different syntactic structures. Though there is a significant degree of morphological variation across the RDem paradigms of NWGmc, all of that variation is captured neatly by Table 76. This table would not be possible but for the fine-grained morphological decomposition that nanosyntax allows us to perform.

Moreover, traditional genetic relations within the Germanic family tree (see Figure 1) are reflected in the RDem data. OE and OF (< Anglo-Frisian) share various characteristics between them (such as the shape and distribution of their K-final-SS and D-internal forms, as well as their identical RDem base *þi-*/*thi-*) and in opposition to OS and OHG, which themselves group together in some instances (e.g. the existence of Direct forms with gemination and vowel-mutation). At the same time, the continental WGmc languages OF, OS, and OHG pattern together against insular OE, most notably in the fact that OE does not have Direct forms of any kind. OHG, being an eastern outlier in WGmc, is also distinct from the rest of WGmc, most clear from its lack of D-internal forms of any kind. Finally, of course, it is clear that WGmc and NGmc are separate branches, but it is equally clear that their RDem forms are closely related, all stemming from the Dem-*si* stage of Proto-NWGmc. While there is a diverse array of RDem structures to be found across Old Germanic, nanosyntax allows us to capture and explain this diversity in a way that is both meticulous and simple at the same time.

7 Conclusion and further issues

7.1 Summary of the dissertation

The primary goal of this dissertation was to understand the internal structure of RDem in Northwest Germanic, focusing on the RDem paradigm of ON, discussed in detail in Chapters 3, 4, and 5. A good understanding of ON allowed us then to incorporate data from OE, OF, OS, OHG, and even RN in Chapter 6. The perspective was primarily synchronic, though various aspects of the diachrony of RDem came into focus as well.

Chapter 1 was an overview of the main empirical data considered in this dissertation and some philological background. Chapter 2 was an introduction to Cinque's (2005) U20 program and to the theory of nanosyntax. Chapters 3 through 6 elaborated an answer to the main research questions with which this dissertation was concerned. These may be summarized as follows:

(i) What are the morphological ingredients that make up RDem?

(ii) How many combinations of these ingredients (i.e. structures) are attested/possible?

(iii) What is the fseq corresponding to these ingredients?

(iv) How is this fseq lexically packaged (i.e. divided up in the lexicon)?

(v) How are the RDem structures derived?

These questions constitute the central focus of this dissertation, and my main goal was to give a coherent and internally consistent answer to each of them. Not unexpectedly, numerous additional issues have arisen along the way. For some of these this final chapter will try to formulate tentative answers. In the present section I will first summarize the findings of this work by answering the questions in (i-v) above.

The answer to question (i) *What are the morphological ingredients?* was developed in Chapter 3, which presented a fine-grained morphological decomposition of RDem in ON. This resulted in five separate morphological ingredients, seen in (189).

(189) Five morphological ingredients

þa-	base (D)	
-K	strong adjective ending (K)	
*-C^i^*	consonant geminator with *i*-mutator (Gm)	
-s	the sigmatic reinforcer	(R)
-a	the asigmatic reinforcer	(R)

Chapter 3 also provided a partial answer to question (ii) *How many possible structures?* It did so by proposing that there are three kinds of RDem structures in ON. The three RDem structures in ON are shown in (190).

(190) Three templates

(i) the K-final forms

D-R-Gm-K
þa-s-C^i-K

(ii) the K-internal forms

D-K-Gm-R
þa-K-C^i-*a*

(iii) the constant forms

D-K-R-Gm
þa-K-*s*-C^i

Importantly, the sigmatic and asigmatic reinforcers are in complementary distribution. Thus, while five separate ingredients can be identified, there are in fact only four syntactic heads available for each instantiation of RDem, because *-s* and *-a* are realizations of the same syntactic head (the reason they are both labeled R in (189)).

Question (iii) *What is the fseq for RDem?* was answered in Chapter 4, where I showed how the correct functional sequence for RDem could be deduced using Cinque's (2005) U20 principles. The correct fseq was found to be R > Gm > K > D. The Cinquean answer to question (iv) *How is the fseq packaged?* is very straightforward: each morphological ingredient corresponds to a single syntactic head (D = *þa-*, K = -K, Gm = -C^i, R = *-s* ~ *-a*). Chapter 4 provided a partial answer to question (v) *How are the RDem structures* *derived?* as well. Not only are there three kinds of RDem structures attested, but based on the functional sequence I postulated and due to a combination of syntactic, phonological, and morphological constraints, only three structures are *possible*. In line with Cinque's U20 system, the K-final forms can be said to have a fully cyclic derivation, the constant forms a partially cyclic derivation, and the K-internal forms a roll-up derivation. An interesting generalization arises with respect to the realization of the R head: the cyclic-type derivations always realize R as sigmatic *-s*, while the non-cyclic (i.e. roll-up) derivations always realize R as asigmatic *-a*. The *-s* ~ *-a* allomorphy, then, can be seen as morphological support for Cinque's derivational system.

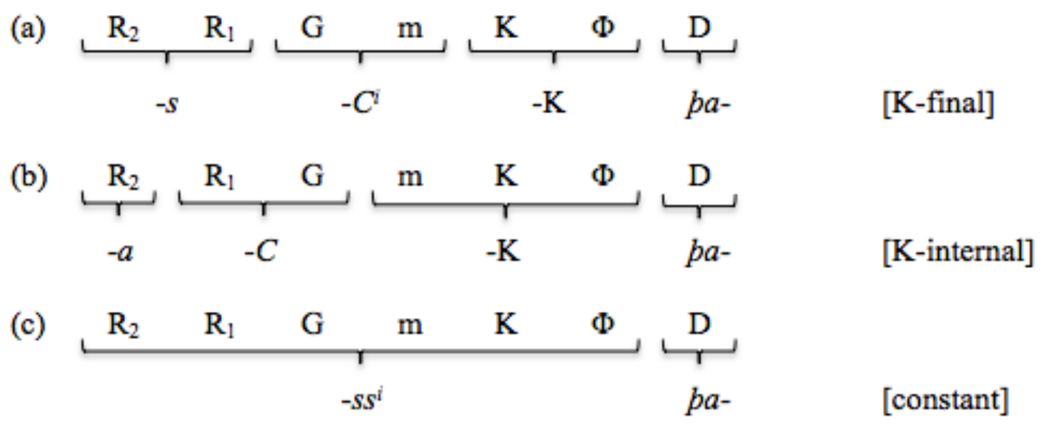
In a Cinquean system this allomorphy was seen to be governed by the type of syntactic derivation at stake, rather than lexical content. But in fact we want the opposite, namely a system whose morphosyntactic derivations are governed by the content of the lexicon. Thus a more nuanced answer to question (v) was needed.

Chapter 5 provided a nanosyntactic analysis of RDem in ON. First I pointed out that the Cinque-style account of the *-s* ~ *-a* allomorphy developed in Chapter 4 cannot be quite right. According to the Cinquean account, it would be the derivation which determines whether R is lexicalized by *-s* or by *–a*. However, according to the Principles and Parameters framework, it is the content of the lexicon which should determine how the derivations proceeds. To remedy this problem, I proceeded to refine the functional sequence by decomposing both the reinforcer feature R into two features and the geminator head Gm into two features. Following Caha (2009), moreover, K can be decomposed into its component K features. This decomposition led us to the hierarchy in (191).

(191) $R_2 > R_1 > G > m > K_{...} > K_2 > K_1 > D$

Making the fseq more fine-grained allowed us to capture the *-s* ~ *-a* alternation in terms of lexical structure rather than in terms of derivational type. The alternation between *-s* and *-a* is accounted for if we assume that *-s* corresponds to the complex head [R_2 R_1] and that *-a* corresponds to the phrasal layer R_2P.

(192) Packaging in ON



A number of other facts also happened to fall out from this more fine-grained approach, as seen in Chapters 5 and 6, which pursued questions (iv) and (v) further. These chapters attempted to capture crosslinguistic variation in the RDem paradigms in terms of lexical packaging. The way each type of RDem morpheme is packaged in both NGmc and WGmc was discussed, and the kind of derivation that results from each type of packaging schema is explained as well. It emerges that an important part of the derivation of RDem is focus movement, a non-spellout-driven movement that targets D-like elements. Foc-movement accounts for the pied-piping of KP with DP in the WGmc forms that have pronominal K (K_D). In NGmc, this pied-piping does not take place since K is adjectival; thus only DP moves, without KP.

All of the observed variation in the RDem structures of NGmc and WGmc can be reduced to the way lexical entries are structured in these languages (see Table 76). That is, while the functional sequence itself is universal, the components of the functional sequence happen to be packaged slightly differently in each language, which leads also to different derivations during the spellout process. This conclusion is fully in line with one of the major tenet of the Principles and Parameters framework: crosslinguistic variation is superficial; all of the variation observed across languages can be reduced to the lexicon, where the arbitrary or exceptional facts of linguistic knowledge are stored (as opposed to the unexceptional functional sequence, which is universal and innate) (see Starke 2011a and also Chomsky 2001: 2).

Various questions remain, of course, and in this chapter I formulate these questions explicitly so that they may constitute the basis for further work.

7.2 Further issues

In this section I will briefly summarize some points of further research.

7.2.1 Dem stems

It is clear that the various Dem stems attested in the Germanic languages and how they relate to the RDem bases (sometimes identical, sometimes not) need to be investigated further. While the RDem base *þe-* in ON, the base *the-* in OS, and the base *dë-* in OHG all come from their respective Dem paradigms, OE *þi-* and OF *thi-* cannot be said to come from the Dem paradigms due to their *i*-vocalism. Based on the *i*-vocalism in these bases, it may be that these bases are somehow related to the third person personal pronoun stem *hi-* instead. The question is, then, how these differences are encoded in the fine-grained internal structure of DP. For example, in the Dem paradigm of ON, how does the system know that the Dem stem *þei-* shows up in a subset of the paradigm but that in another subset of the paradigm the stem *þa-* does? Presumably some combination of K, Φ, and perhaps some species of Agr cooperate to make such patterns emerge.

7.2.2 Restarting the fseq

In Section 4.1.6 I pointed out that Bernstein's (1997, 2001) adverbial reinforcers are word-external while the reinforcers of RDem are word-internal (more like those of Leu 2007, 2008, 2015 and Kayne 2005). Interestingly, however, the same features seem to appear on both the word-internal and word-external level. That is, certain features appear to be repeated at different points in the structure, producing a nesting effect of sorts.

Roehrs (2010) has done a detailed study of reinforcer-demonstrative constructions in Germanic and Romance. He also endorses a fine-grained decomposition of demonstratives and reinforcers in his work, though his approach differs quite markedly from what I have done in this dissertation. Many interesting questions arise from taking his data (and analysis) into consideration. His Type 1 pattern, for instance, which is observed in Yiddish, can be seen in (193).

(193) Type 1 (D > R) (Roehrs 2010: 226-227, 243; Jacobs 2005)

der	doziker	guter	man	(Yiddish)
Det	*this*	*good*	*man*	

'this good man'

In (193) a D-like element precedes a reinforcer-like element (which both precede the adjective and the head noun). However, at a more fine-grained level, these elements have internal structure which look to involve a sort of repetition, as seen in (194).

(194) d-er d-oz-ik-er
D-K [[D-K_{frozen}]-R-K]

The determiner *der* is made up of D (*d-*) and K (*-er*) features, and the item *doziker* is made up of D (*d(-oz)*), R (*-ik*), and K (*-er*) features. Indeed, even *doz* has some internal structures, since it comes from the Yiddish N.NOM/ACC.SG determiner *dos* (Roehrs 2010: 248, fn. 25). Thus there is a frozen NOM/ACC K_D morpheme embedded within *doziker*.

This resembles the modern Icelandic data mentioned in Chapter 4, where it was seen that the roll-up forms display an extra K (perhaps Agr) morpheme, as seen in (195).

(195) (a) þe-**t**-C-a-**ð**
þe-**n**-C-a-**n**
þe-**s**-C-a-**s**

(b) [[D-**K**-Gm-R]-**K/Agr**]

In (195) we see agreement between the internal K morpheme and the external K marker. In (194) we see K agreement between *d-er* and *dozik-er*. The Yiddish example is more complex, however, since D is also doubled. In fact, the entire sequence looks to be replicated in certain Yiddish Type 2 cases, such as in (196) (from Roehrs 2010: 243; Jacobs 2005).

(196) Type 2 (R > D) (Roehrs 2010: 226-227, 243; Jacobs 2005)

ot-o	d-i	d-ozik-e	froj
R-R	D-K	D-R-K	*woman*

'this woman'

Both *ot* and *-o* are reinforcer particles which cooccur with the D-K element *di* and the additional element *dozike* 'this'.[113] While *o-to* and *d-i* account for the R, D, and K features, *dozike* apparently replicates the entire sequence again (i.e. *doz-ik-e* = D-K-R).

A similar phenomenon is present in some well known data from English. Kayne (2005), Roehrs (2010: 259-260), and Henry (2010) have observed that there is an asymmetry in English between prenominal and postnominal adverbial reinforcers.

[113] *Ot(-o) der guter man* appears to alternate also with *der-o guter man*, with the reinforcer particle to the right of D-K (Roehrs 2010: 243, Jacobs 2005: 186). Thus *-o* is able, under certain circumstances, to lexicalize more or less the same span that *ot(-o)* does.

(197) (a) this here book (cf. Swedish *den här bok-en* 'this book', lit. 'the here book-the'; Leu 2007, 2008, 2015)

(b) this book (right/over) here

In (197a), prenominal *here* is a (non-locative) reinforcer element. This is supported by facts such as (i) prenominal *here* cannot be stressed (198a) (Kayne 2005: 66) and (ii) prenominal *here* cannot be modified by a prepositional intensifier like *right* or *over* (or *right over*) (198b) (Kayne 2005: 66, Roehrs 2010: 260).

(198) (a) *This HERE letter is more important than that THERE one.

(b) *this right/over here letter

However, postnominal *here* is a locative element, as seen by the grammaticality of (199).

(199) (a) This letter HERE is more important than that one THERE.

(b) this letter right/over here

Both kinds of *here* may even cooccur (my own judgments), as in (200).

(200) (a) This here letter HERE is more important than that there letter THERE.

(b) this here letter right/over HERE

Thus (non-standard) English shows a repetition of R-like elements, and importantly the semantics of these elements can differ slightly depending on their position.

A possible solution to such issues could be formulated in terms of 'gapping' (Starke 2013; see also Caha 2009: §9.3). The basic idea in a gapping approach is that structures are sometimes built in the syntax which are featurally impoverished somehow. For instance, imagine that the syntax builds the structure [A [C]], even though the sequence dictates A > B > C. In other words there is a gap at B, but if the lexicon has a match for the structure [A [C]] then a legitimate spellout can be produced. However, if the semantic import of B is still necessary in the derivation, then the syntax must start over and try to construct a structure with B again. Syntax cannot just start at B, though, so it begins at the bottom of the sequence again, building, say, [B [A]]. The result is that A has been doubled. We could imagine, then, that some of the items in Yiddish (and English) are featurally impoverished in this way, necessitating a second attempt at 'filling in' the gaps by syntax. Feature-doubling would be a symptom of this process.

7.2.3 Partial pronominal inflection in RDem

As for the RDem paradigms of WGmc, an obvious question is why a subset of forms – the D-internal and Direct forms – take pronominal K endings (or put differently, why a subset of forms must be built with the Class-D feature Cl_D), such as OE *þ-ēo-s*, *þ-ē-s*, or *þ-ā-s*, as discussed in Section 6.4.1. Now, it would be decidedly odd to posit that because these RDem forms share the inflectional endings of Dem they are somehow 'more demonstrative' than the other forms in the paradigm which take adjectival endings. Instead I think a solution to this problem must rely on the maxim that (idiosyncratic) properties of the lexicon drive syntactic derivations and variation. This is a core lesson of nanosyntax and of the Principles and Parameters program in general.

Imagine that an RDem structure requires some feature X at some stage of its derivation. Any lexical entry lacking X would therefore not be usable at this stage in the RDem derivation. Take OE, for instance, which has the following K_D endings in its D-internal RDem forms: F.NOM.SG *-ēo*, M.NOM.SG *-ē*, and F.ACC.SG / NOM/ACC.PL *-ā*. Let us say that these case endings contain our hypothetical feature X and can therefore be used in the RDem paradigm (201a), but that the rest of the K_D endings do not contain the feature X and therefore cannot be used in the RDem paradigm (201b). Imagine now that the lexical structures of the adjectival K endings in OE normally contain the feature X (202b), *except* in the F.NOM.SG, M.NOM.SG, and F.ACC.SG / NOM/ACC.PL (202a). In other words, there is a very specific complementary distribution of the feature X across the OE case systems.

(201) **Dem (K_D)**

(a)	F.NOM.SG	-ēo	**X** Cl_D Cl_A
	M.NOM.SG	-ē	**X** Cl_D Cl_A
	F.ACC.SG / NOM/ACC.PL	-ā	**X** Cl_D Cl_A
(b)	Rest of Dem	...	Cl_D Cl_A

(202) **Strong adjectives (K)**

(a)	F.NOM.SG	-Ø	Cl_A
	M.NOM.SG	-Ø	Cl_A
	F.ACC.SG / NOM/ACC.PL	-e	Cl_A
(b)	Rest of strong adjectives	...	**X** Cl_A

In the kind of system sketched in (201) and (202), only lexical entries which happen to contain the feature X will be available for spelling out a K ending in RDem, since an RDem structure by hypothesis absolutely requires the feature X. It just so happens, due to idiosyncratic properties of the OE lexicon, furthermore, that in a certain subset of slots in the paradigm, K_D endings are the only available entries for spelling out structures with X. In the other slots, K endings are the only available entries for spelling out structures with X. This, abstractly speaking, is how the WGmc hybrid system can be modeled by using the idea that syntax is dependent on what kinds of lexical entries are available in the lexicon. In this respect, syntax cannot be picky: the derivation will crash unless a matchable structure is built. And it just so happens that a matchable structure in the F.NOM.SG, M.NOM.SG, and F.ACC.SG / NOM/ACC.PL of OE will require a Cl_D feature to be present.

7.3 Concluding remarks

In this chapter I have summarized the main findings of this thesis and then addressed some ideas about how fine-grained nanosyntactic functional sequences can result in multi-word constituents, taking us a bit farther on the road from internal ('morphological') word structure towards external ('syntactic') structure. Nanosyntax teaches us that morphology and syntax form a single spectrum, fading into each other imperceptibly. As mentioned, previous studies on reinforcers (Bernstein 1997, 2001; Brugè 1996, 2002; Roehrs 2010; even Kayne 2005 and Leu 2007, 2008) belong more to the 'syntactic' side of the spectrum, while this dissertation belongs more to the 'morphological' side.

It is important to note that highly restricted, empirically focused studies – such as this dissertation's look at a single paradigm – are a necessary starting point when using a theory like nanosyntax, which proposes a radically detailed and fine-grained decomposition of syntax and morphology. Jumping straight to patterns at the multi-word level without having a full understanding of the word-internal structures will almost inevitably lead to unreliable and inconsistent results. Some caution is therefore warranted when trying to reinterpret the results of Bernstein (1997, 2001), Brugè (1996, 2002), Roehrs (2010), Kayne (2005), and Leu (2007, 2008) from a nanosyntactic point of view. For now, though, we would do well to continue intensive work on morphological decomposition.

References

Abels, Klaus. 2011. Hierarchy-order relations in the Germanic verb cluster and in the noun phrase. *Groninger Arbeiten zur germanistischen Linguistik* 53(2): 1-28.

Abels, Klaus. 2003. Successive cyclicity, anti-locality, and adposition stranding. PhD dissertation, University of Connecticut.

Abels, Klaus and Peter Muriungi. 2008. The focus particle in Kîîtharaka: Syntax and semantics. *Lingua* 118: 687-731.

Abels, Klaus and Ad Neeleman. 2009. Universal 20 without the LCA. In *Merging Features: Computation, Interpretation and Acquisition* (J.M. Brucart, A. Gavarrò, J. Solà, eds): 60-79. Oxford: OUP.

Abels, Klaus and Ad Neeleman. 2012. Linear asymmetries and the LCA. *Syntax* 15(1): 25-74.

Abney, Steven. 1987. *The English Noun Phrase in its Sentential Aspect*. PhD dissertation, MIT.

Aboh, Enoch. 2004a. *The Morphosyntax of Complement-Head Sequences*. New York: OUP.

Aboh, Enoch. 2004b. Snowballing movement and generalized pied-piping. In *Triggers* (A. Breitbarth, H. van Riemsdijk, eds): 15-48. Berlin: Mouton de Gruyter.

Aboh, Enoch. 2004c. Topic and focus within D. *Linguistics in the Netherlands 2004* 21: 1-12.

Adams, J.N. 2013. *Social Variation and the Latin Language*. Cambridge: CUP.

Alexiadou, Artemis, Liliane Haegeman, and Melita Stavrou. 2007. *Noun Phrase in the Generative Perspective*. Berlin: Mouton de Gruyter.

Anderson, Stephen R. 1969. *West Scandinavian Vowel Systems and the Ordering of Phonological Rules*. PhD dissertation, MIT.

Anderson, Stephen R. 1988. The grammar of Icelandic verbs in *-st*. Available at http://cowgill.ling.yale.edu/sra/icel_st.pdf.

Antonsen, Elmer H. 1975. *A Concise Grammar of the Older Runic Inscriptions*. Tübingen: Niemeyer.

Antonsen, Elmer H. 2002. *Runes and Germanic Linguistics*. New York: Mouton de Gruyter.

Armitage, L. 1911. *An Introduction to the Study of Old High German*. Oxford: Clarendon.

Axelsdóttir, Katrín. 2003. Saga ábendingarfornafnsins *sjá* [The history of the demonstrative pronoun *sjá*]. *Íslenskt mál og almenn málfræði* 25(1): 41-77. http://timarit.is/view_page_init.jsp?pageId=5414324&requestedFileType=Web%20display&pdfView=FitH&lang=en (accessed March 7, 2013).

Baker, Mark. 1985. The mirror principle and morphosyntactic explanation. *Linguistic Inquiry* 16(3): 373-415.

Baker, Mark. 1988. *Incorporation: A Theory of Grammatical Function Changing*. Chicago: University of Chicago Press.

Bandle, Oskar. 1956. *Die Sprache der Guðbrandsbiblía: Orthographie und Laute*. Copenhagen: Formen (Bibliotheca Arnamagnæana XVII).

Barnes, Michael. 2004 [1999]. *A New Introduction to Old Norse, Part I: Grammar*. University College London: Viking Society for Northern Research.

Barnes, Michael. 2005. Language. In *A Companion to Old Norse-Icelandic Literature and Culture* (R. McTurk, ed.): 173-189. Malden: Blackwell Publishing.

Beekes, Robert S.P. 2011 [1995]. *Comparative Indo-European linguistics: An Introduction*, 2nd ed (Michiel de Vaan, rev. ed.). Amsterdam: John Benjamins.

Beghelli, Filippo and Tim Stowell. 1997. Distributiviy and negation: The syntax of *each* and *every*. In *Ways of Scope Taking* (A. Szabolcsi, ed.): 71-107. Dordrecht: Kluwer.

Belletti, Adriana. 1990. *Generalized verb movement: Aspects of verb syntax*. Torino: Rosenberg & Sellier.

Belletti, Adriana. 2001. Agreement projections. In *The Handbook of Contemporary Syntactic Theory* (M. Baltin, C. Collins, eds): 483-510. Oxford: Blackwell.

Belletti, Adriana (ed.). 2004. *The Cartography of Syntactic Structures, Vol. 3: Structures and Beyond*. New York: OUP.

Benincà, Paola and Cecilia Poletto. 2004. Topic, focus, and V2: Defining the CP sublayers. In *The Cartography of Syntactic Structures, Vol. 2: The Structure of CP and IP* (L. Rizzi, ed.): 52-75. New York: OUP.

Bernstein, Judy. 1997. Demonstratives and reinforcers in Romance and Germanic languages. *Lingua* 102: 87-113.

Bernstein, Judy. 2001. Focusing the "right" way in Romance determiner phrases. *Probus* 13: 1-29.

Bianchi, Valentina and Mara Frascarelli. 2009. Is topic a root phenomenon? Ms., University of Siena and University of Roma 3.

Blake, Barry J. 1994 [2004]. *Case*. Cambridge: CUP.

Bobaljik, Jonathan. 2007. On comparative suppletion. Ms., University of Connecticut.

Bobaljik, Jonathan. 2012. *Universals in Comparative Morphology: Suppletion, Superlatives, and the Structure of Words*. Cambridge, MA: MIT Press.

Bošković, Željko. 2005. On the locality of left branch extraction and the structure of NP. *Studia Linguistica* 59: 1-45.

Bošković, Željko. 2008. What will you have, DP or NP? *Proceedings of NELS* 37: 101-114.

Bošković, Željko. 2009. More on the no-DP analysis of article-less languages. *Studia Linguistica* 2: 187-204.

Bošković, Željko. 2010. On NPs and clauses. Ms., University of Connecticut.

Braune, Wilhelm, with Ingo Reiffenstein. 2004 [1886]. *Althochdeutsche Grammatik I*, 15th ed. Tübingen: Max Niemeyer Verlag.

Bremmer, Rolf H., Jr. 2009. *An Introduction to Old Frisian: History, Grammar, Reader, Glossary*. Philadelphia: John Benjamins.

Brugè, Laura. 1996. Demonstrative movement in Spanish: A comparative approach. *University of Venice Working Papers in Linguistics* 6(1): 1-61.

Brugè, Laura. 2002. The positions of demonstratives in the extended nominal projection. In *The Cartography of Syntactic Structures, Vol. 1: Functional Structure in DP and IP* (G. Cinque, ed.): 15-53. New York: OUP.

Caha, Pavel. 2009. *The Nanosyntax of Case*. PhD dissertation, University of Tromsø.

Caha, Pavel. 2010. The parameters of case marking and spell out driven movement. *Linguistic Variation Yearbook* 10: 32-77.

Caha, Pavel. 2013. Explaining the structure of case paradigms by the mechanisms of Nanosyntax: The Classical Armenian nominal declension. *Natural Language and Linguistic Theory* 31(4): 1015-1066.

Caha, Pavel and Marina Pantcheva. 2012. Contiguity beyond linearity: Modeling cross-dimensional syncretisms. Workshop on the representation and selection of exponents. June 7. Slides available online.

Campbell, Lyle. 1999. *Historical Linguistics: An Introduction*. Cambridge, MA: MIT Press.

Campbell, Alistair. 2003 [1959]. *Old English Grammar*. New York: OUP.

Cardinaletti, Anna. 1997. Subjects and clause structure. In *The New Comparative Syntax* (L. Haegeman, ed.): 33-63. London: Addison, Wesley, Longman.

Cardinaletti, Anna. 2004. Towards a cartography of subject positions. In *The Cartography of Syntactic Structures, Vol. 2: The Structure of CP and IP* (L. Rizzi, ed.): 115-165. New York: OUP.

Cathey, James E. 2000. *Old Saxon*. LINCOM Publishers.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht : Foris Publications.

Chomsky, Noam. 1986. *Barriers*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In *Step by Step – Essays in Minimalist Syntax in Honor of Howard Lasnik* (R. Martin, D. Michaels, J. Uriagereka, S.J. Keyser): 89-155. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A Life in Language* (M. Kenstowicz, ed.): 1-50. Cambridge, MA: MIT Press.

Cinque, Guglielmo. 1995. The 'antisymmetric' program: Theoretical and typological implications. *University of Venice Working Papers in Linguistics* 5(2): 1-21.

Cinque, Guglielmo. 1999. *Adverbs and Inflectional Heads*. Oxford: OUP.

Cinque, Guglielmo (ed.). 2002. *The Cartography of Syntactic Structures, Vol. 1: Functional Structure in DP and IP*. New York: OUP.

Cinque, Guglielmo. 2004. A phrasal movement analysis of the Romanian DP. *Studia Linguistica et Philologica in Honorem D. Irimia* (A. Minut, E. Munteanu, eds): 129-142. Iaşi: Editura Universitatii A.I. Cuza.

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

Cinque, Guglielmo. 2009. The fundamental left-right asymmetry of natural languages. In *Universals of Language Today* (S. Scalise, E. Magni, A. Bisetto, eds): 165-184. Dordrecht: Springer.

Cinque, Guglielmo. 2010a. *The Syntax of Adjectives: A Comparative Study*. Cambridge, MA: MIT Press.

Cinque, Guglielmo. 2010b. Word order typology: A change of perspective. Lingbuzz/001656.

Cinque, Guglielmo. 2012. Cognition, Universal Grammar, and typological generalizations. Lingbuzz/001658.

Cinque, Guglielmo and Luigi Rizzi. 2008. The cartography of syntactic structures. In *CISCL Working Papers on Language and Cognition* (V. Moscati, ed.) 2: 43-59.

Cleasby, Richard and Guðbrandur Vigfússon. 1874. *An Icelandic-English Dictionary*. Widely available online.

Craigie, William, James Murray, and John Simpson (eds.). 1971. *The Compact Edition of the Oxford English Dictionary: Volume II, P-Z*. Oxford: OUP.

Danckaert, Lieven. 2012. *Latin Embedded Clauses: The Left Periphery*. Amsterdam: John Benjamins.

Déchaine, Rose-Marie and Martina Wiltschko. 2002. Decomposing pronouns. *Linguistic Inquiry* 33: 409-443.

De Clercq, Karen. 2013. *A Unified Syntax of Negation.* PhD dissertation, Ghent University.

Dékány, Éva. 2009. The nanosyntax of Hungarian postpositions. *Nordlyd, Tromsø University Working Papers on Language and Linguistics* 36 [Special issue on Nanosyntax]: 41-76. http://septentrio.uit.no/index.php/nordlyd/ (accessed April 20, 2015).

Derolez, René and Anne-Marie Simon-Vandenbergen. 1988. Gothic *saihw* and *sai*, with some notes on imperative interjections in Germanic. In *Languages and Cultures: Studies in Honor of Edgar C. Polomé* (M.A. Jazayery, E.C. Polomé, W. Winter, eds): 97-110. Berlin: Mouton.

D'Hulst, Yves. 2006. Romance plurals. *Lingua* 116: 1303-1329.

Diessel, Holger. 1999. *Demonstratives: Form, Function, and Grammaticalization*. Amsterdam: John Benjamins.

Dryer, Matthew. 2009. On the order of demonstrative, numeral, adjective, and noun: An alternative to Cinque. Handout available at http://allegatifac.unipv.it/caterinamauri/sitovecchio/1898313034_cinqueH09.pdf.

Embick, David and Rolf Noyer. 2007. Distributed Morphology and the syntax-morphology interface. In *The Oxford Handbook of Linguistic Interfaces* (G. Ramchand, C. Reiss, eds): 289-324. Oxford: OUP.

EWAhd II = Lloyd, Springer, and Lühr (1998).

Faarlund, Jan Terje. 2004. *The Syntax of Old Norse: With a Survey of the Inflectional Morphology and a Complete Bibliography*. Oxford: OUP.

Fábregas, Antonio. 2007. An exhaustive lexicalisation account of directional complements. Lingbuzz/00930.

Fábregas, Antonio. 2009. An argument for phrasal spellout: Indefinites and interrogatives in Spanish. *Nordlyd, Tromsø University Working Papers on Language and Linguistics* 36 [Special issue on Nanosyntax]: 129-168. http://septentrio.uit.no/index.php/nordlyd/ (accessed February 4, 2011).

Feist, Sigmund. 1939 [1923]. *Vergleichendes Wörterbuch der Gotischen Sprache*, 3rd ed. Leiden: E.J. Brill.

Fillmore, Charles J. 1971. *Deixis Lectures* (Santa Cruz). Available online at www-personal.umich.edu/ ~jlawler/Deixis.html.

Fillmore, Charles J. 1982. Towards a descriptive framework for spatial deixis. In *Speech, Place, and Action* (R.J. Jarvella, A.M. Zwicky, eds): 31-59. Chichester: John Wiley.

Fillmore, Charles J. 1997. *Lectures on Deixis*. Stanford: CSLI Publications.

Fortson, Benjamin W. IV. 2004. *Indo-European Language and Culture: An Introduction*. Malden: Blackwell Publishing.

Fortuna, Marcin. 2013. Vowel length in Icelandic compounds and the role of FENs. Slides available at https://btk.ppke.hu/uploads/articles/51961/file/Fortuna%202013%20Vowel%20length%20in%20Icelandic%20compounds%20and%20the%20role%20of%20FENs%20%28slides%29.pdf.

Galleé, Johan Hendrik. 1910. *Altsächsische Grammatik*. Halle: Max Niemeyer.

Gibson, Courtenay St. John and Catherine O. Ringen. 2000. Icelandic umlaut in optimality theory. *Nordic Journal of Linguistics* 23(1): 49-64.

Giusti, Giuliana. 1997. The categorical status of determiners. In *The New Comparative Syntax* (L. Haegeman, ed.): 95-123. London: Addison, Wesley, Longman.

Gordon, E.V. 1956 [1927]. *An Introduction to Old Norse* (Taylor, ed.). New York: OUP.

Greenberg, Joseph. 1963. Some universals of grammar with particular reference to the order of meaningful elements. *Universals of Language* (Greenberg, ed.): 73-113. Cambridge, MA: MIT Press.

Grimm, Jakob. 1831. *Deutsche Grammatik: Dritter Theil*. Göttingen bei Dieterich.

Grimshaw, Jane. 1979. Complement selection and the lexicon. *Linguistics Inquiry* 10: 279-325.

Grimshaw, Jane. 1991. Extended projection. Ms., Brandeis University.

Grimshaw, Jane. 2000. Locality and extended projection. In *Lexical Specification and Insertion* (P. Coopmans, M. Everaert, J. Grimshaw, eds): 115-135. Amsterdam: John Benjamins.

Grønvik, Ottar. 1981. *Runene på Tunestenen: Alfabet – Språkform – Budskap*. Oslo: Universitetsforlaget.

Haegeman, Liliane. 1995. *The Syntax of Negation.* Cambridge: CUP.

Haegeman, Liliane. 2014. West Flemish verb-based discourse markers and the articulation of the Speech Act layer. *Studia Linguistica* 68(1): 116-139.

Haegeman, Liliane and Virginia Hill. 2013. The syntacticization of discourse. In *Syntax and Its Limits* (Folli, Sevdali, Truswell, eds): 370-390. Oxford: OUP.

Haegeman, Liliane and Raffaella Zanuttini. 1991. Negative heads and the Neg criterion. *The Linguistic Review* 8: 233-51.

Hale, Kenneth and Jay Keyser. 1993. On argument structure and the lexical expression of syntactic relations. In *The View from Building 20: A Festschrift for Sylvain Bromberger* (K. Hale, S.J. Keyser, eds): 53-108. Cambridge : MIT Press.

Halle, Morris and Alec Marantz. 1993. Distributed Morphology and the pieces of inflection. In *The View from Building 20: A Festschrift for Sylvain Bromberger* (K. Hale, J. Keyser, eds): 111-176. Cambridge: MIT Press.

Hansson, Gunnar Ólafur. 2013. Alternations, allomorphy and abstractness: Icelandic *u*-umlaut revisited. Slides from SCL 25. May 13-15. Reykjavík.

Haugen, Einar. 1976. *The Scandinavian Languages: An Introduction to Their History*. Cambridge, MA: Harvard University Press.

Haugen, Einar. 1982. *Scandinavian Language Structures*. Minneapolis: University of Minnesota Press.

Heine, Bernd and Tania Kuteva. 2002. *World Lexicon of Grammaticalization*. Cambridge: CUP.

Henry, Alison. 2010. Demonstratives in Belfast English: Microvariation and externalization. LAGB handout. September 1-4. Leeds.

Heusler, Andreas. 1962 [1913]. *Altisländisches Elementarbuch*, 5th ed. Heidelberg: C. Winter.

Hewett, Waterman Thomas. 1879. *The Frisian Language and Literature: A Historical Study*. Ithaca: Finch & Apgar.

Himmelmann, Nikolaus. 1997. *Deiktikon, Artikel, Nominalphrase: Zur Emergenz syntaktischer Struktur*. Tübingen: Max Niemeyer.

Himmelmann, Nikolaus. 2001. Chapter 62: Articles. In *Language Typology and Language Universals* (M. Haspelmath, E. König, W. Oesterreicher, W. Raible, eds): 831-841. Berlin: De Gruyter.

Holthausen, Ferdinand. 1921. *Altsächsisches Elementarbuch*. Heidelberg: Carl Winter Universitätsverlag.

Horrocks, Geoffrey and Melita Stavrou. 1987. Bounding theory and Greek syntax: Evidence from *wh*-movement in NP. *Journal of Linguistics* 23: 79-108.

Ingason, Anton Karl. 2013. Icelandic umlaut as morpheme-specific phonology. Slides available online at http://www.ling.upenn.edu/~ingason/papers/ingason2013umlautslides.pdf.

Jacobs, Neil G. 2005. *Yiddish: A Linguistic Introduction*. Cambridge: CUP.

Jayaseelan, K.A. 2010. On two types of movement. *Proceedings of GLOW in Asia VIII 2010: Universals and Variation* (G. Ming-le, ed.): 28-34. Beijing: Beijing Language and Culture University Press.

Jayaseelan, K.A. 2015. Parallel work spaces in syntax and the inexistence of internal merge. Lingbuzz/002371.

Johnsen, Sverre Stausland. 2012. Syncope of long *ī in Old Norse nouns. In *Germansk filologi og norske ord: Festskrift til Harald Bjorvand* (J.O. Askedal, T. Schmidt, R. Theil, eds.): 33-51. Oslo: Novus forlag.

Kaye, Jonathan. 1990. ‘Coda’ licensing. *Phonology* 7: 301-330.

Kayne, Richard S. 1984. *Connectedness and Binary Branching*. Dordrecht: Foris.

Kayne, Richard S. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

Kayne, Richard S. 2005. *Movement and Silence*. New York: OUP.

Kayne, Richard S. 2010. Why are there no directionality parameters? Lingbuzz/001100.

Kayne, Richard and Jean-Yves Pollock. 2010. Notes on French and English demonstratives. In *Structure Preserved: Studies in Syntax for Jan Koster* (J.W. Zwart, M. de Vries, eds): 215-228. Amsterdam: John Benjamins.

Kieckers, Ernst. 1917-1920. Verschiedenes. *Indogermanische Forschungen* 38: 209-19.

Kiparsky, Paul. 1973. ‘Elsewhere’ in phonology. In *A Festschrift for Morris Halle* (P. Kiparsky, S. Anderson, eds): 93–106. New York: Holt, Rinehart & Winston.

Klingenschmitt, Gert. 1987. Erbe und Neuerung beim germanischen Demonstrativpronomen. In *Althochdeutsch, Band I: Grammatik. Glossen und Texte* (Herbert Kolb, Klaus Matzel & Karl Stackmann; Rolf Bergmann, Heinrich Tiefenbach & Lothar Voetz, eds.): 169-189. Heidelberg: Carl Winter Universitätsverlag.

Kluge, Friedrich. ~~1920. *Deutsche Sprachgeschichte*. Leipzig: Quelle & Meyer.~~

Koopman, Hilda and Anna Szabolcsi. 2000. *Verbal Complexes*. Cambridge, MA: MIT Press.

Krahe, Hans. 1969. *Germanische Sprachwissenschaft, Band II: Formenlehre* (Wolfgang Meid, ed.), 7th ed. Berlin: Walter de Gruyter.

Krause, Wolfgang. 1971. *Die Sprache der urnordischen Runeninschriften*. Heidelberg: Carl Winter Universitätsverlag.

Kuhn, Hans. 1952-1953. Review of Ernst Schwarz, *Goten, Nordgermanen, Angelsachsen*. In: *Anzeiger für deutsches Altertum und deutsche Literatur* 66: 45-52.

Laenzlinger, Christopher. 1998. *Comparative Studies in Word Order Variations: Pronouns, Adverbs and German Clause Structure*. Amsterdam: John Benjamins.

Lander, Eric T. 2013. What Gothic can teach us about the Northwest Germanic reinforced demonstrative. Paper presented at IGG39 (February 21-23, Modena) and SCL25 (May 13-15, Reykjavík).

Lander, Eric T. Forthcoming. Intraparadigmatic cyclic and roll-up derivations in the Old Norse reinforced demonstrative. *The Linguistic Review*.

Lander, Eric T. and Liliane Haegeman. 2014. Old Norse as an NP language: With observations on the Common Norse and Northwest Germanic runic inscriptions. *Transactions of the Philological Society* 112(3): 279–318.

Lander, Eric T. and Liliane Haegeman. 2015. The nanosyntax of spatial deixis. Handout from IGG 41. February 26-28. Perugia.

Lass, Roger. 1994. *Old English: A Historical Linguistic Companion*. Cambridge: CUP.

Lehmann, Winfred P. 1966. The grouping of the Germanic languages. In *Ancient Indo-European dialects* (Birnbaum, Puhvel, eds.): 13-27. Berkeley & Los Angeles: University of California Press.

Leu, Thomas. 2007. These *HERE* demonstratives. *U.Penn Working Papers in Linguistics* 13(1): 141-154.

Leu, Thomas. 2008. *The Internal Syntax of Determiners* (GAGL 47). PhD dissertation, New York University.

Leu, Thomas. 2015. *The Architecture of Determiners*. New York: OUP.

Levinson, Stephen C. 1983. *Pragmatics*. Cambridge: CUP.

Lloyd, Albert Larry, Otto Springer, and Rosemarie Lühr (eds.). 1998. *Etymologisches Wörterbuch des Althochdeutschen: Band II, bî-ezzo*. Göttingen: Vandenhoeck & Ruprecht.

Lundquist, Björn. 2008. *Nominalizations and Participles in Swedish.* PhD dissertation, University of Tromsø.

Lyons, John. 1977. *Semantics*. Cambridge: CUP.

Markey, Thomas L. 1981. *Frisian*. The Hague: Mouton.

McCarthy, John J. 2008. The serial interaction of stress and syncope. *Natural Language and Linguistic Theory* 26: 499-546.

McCawley, James D. 1968. Lexical insertion in a transformational grammar without Deep Structure. In *Papers from the Fourth Regional Meeting of the Chicago Linguistic Society* (B.J. Darden, C.-J.N. Bayley, A. Davidson, eds): 71-80. Chicago: University of Chicago.

McFadden, Thomas. 2004. On the pronominal inflection of the Germanic strong adjective. *U.Penn Working Papers in Linguistics* 10(1): 123-136. Also published in *Münchener Studien zur Sprachwissenschaft* 63 (2003) [2009]. 53-82.

McFadden, Thomas. 2014. Why nominative is special: Stem-allomorphy and case structures. Handout from GLOW 37. HUBrussel/KULeuven. Brussels, Belgium.

Moskal, Beata. 2013. A case study in nominal suppletion. Ms., University of Connecticut.

Muriungi, Peter. 2008. *Phrasal Movement Inside Bantu Verbs: Deriving Affix Scope and Order in Kîîtharaka*. PhD dissertation, University of Tromsø.

Neeleman, Ad and Kriszta Szendrői. 2007. Radical pro-drop and the morphology of pronouns. *Linguistic Inquiry* 38(4): 671-714.

Nielsen, Hans Frede. 1989. *The Germanic Languages: Origins and Early Dialectal Interrelationships*. Tuscaloosa: University of Alabama Press.

Nielsen, Hans Frede. 2000. *The Early Runic Language of Scandinavia*. Heidelberg: Carl Winter.

Nilsen, Øystein. 2003. *Eliminating Positions: Syntax and Semantics of Sentence Modification*. PhD dissertation, Utrecht University.

Noreen, Adolf. 1923 [1884]. *Altnordische Grammatik I: Altisländische und altnorwegische grammatik, unter Berücksichtigung des Urnordischen*, 4th ed. Halle (Saale): Max Niemeyer.

Noreen, Adolf. 1904. *Altnordische Grammatik II: Altschwedische grammatik, mit Einschluss des Altgutnischen*. Halle (Saale): Max Niemeyer.

Norris, Mark. 2010. The architecture of derivational OT: Evidence from Icelandic syncope. Available at http://people.ucsc.edu/~mnorris/research/syncope_nels_paper.pdf.

OED, P-Z = Craigie, Murray, Simpson (1971).

Pantcheva, Marina. 2011. *Decomposing Path: The Nanosyntax of Directional Expressions*. PhD dissertation, University of Tromsø.

Pollock, Jean-Yves. 1989. Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20(3): 365-424.

Project Wulfila. www.wulfila.be. Antwerp University.

Prokosch, Eduard. 1939. *A Comparative Germanic Grammar*. Philadelphia: Linguistic Society of America.

Puskás, Genoveva. 2000. *Word Order in Hungarian: The Syntax of A'-Positions*. Amsterdam: John Benjamins.

Ramchand, Gillian. 2008. *Verb Meaning and the Lexicon: A First-Phase Syntax*. Cambridge: CUP.

Rask, Rasmus Kristian. 1976 [1811]. *A Grammar of the Icelandic or Old Norse Tongue*. Amsterdam: John Benjamins.

Rauch, Irmengard. 1992. *The Old Saxon Language: Grammar, Epic Narrative, Linguistic Interference*. New York: Peter Lang.

Rauch, Irmengard. 2011. *The Gothic Language: Grammar, Genetic Provenance and Typology, Readings*, 2nd ed. New York: Peter Lang.

Ringe, Donald. 2006. *A Linguistic History of English: Vol. I, From Proto-Indo-European to Proto-Germanic*. Oxford: OUP.

Ringe, Donald and Ann Taylor. 2014. *A Linguistic History of English: Vol. II, The Development of Old English*. Oxford: OUP.

Rizzi, Luigi. 1990. *Relativized Minimality*. Cambridge, MA: MIT Press.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of Grammar* (L. Haegeman, ed.): 281-337. Kluwer: Dordrecht.

Rizzi, Luigi. 2001. Relativized Minimality effects. In *The Handbook of Contemporary Syntactic Theory* (M. Baltin, C. Collins, eds): 89-110. Oxford: Blackwell.

Rizzi, Luigi. 2004a. Locality and left periphery. In *The Cartography of Syntactic Structures, Vol. 3: Structures and Beyond* (Belletti, ed.): 104-131. New York: OUP.

Rizzi, Luigi (ed.). 2004b. *The Cartography of Syntactic Structures, Vol. 2: The Structure of CP and IP*. New York: OUP.

Rizzi, Luigi. 2013. Syntactic cartography and the syntacticisation of scope-discourse semantics. In *Mind, Values and Metaphysics – Philosophical Papers Dedicated to Kevin Mulligan* (A. Reboul, ed.). Dordrecht: Springer.

Robinson, Orrin W. 1992. *Old English and Its Closest Relatives: A Survey of the Earliest Germanic Languages*. Stanford: Stanford University Press.

Rocquet, Amélie. 2013. *Splitting Objects: A Nanosyntactic Account of Direct Object Marking*. PhD dissertation, Ghent University.

Roehrs, Dorian. 2009. *Demonstratives and Definite Articles as Nominal Auxiliaries*. Amsterdam: John Benjamins.

Roehrs, Dorian. 2010. Demonstrative-reinforcer constructions. *Journal of Comparative Germanic Linguistics* 13: 225-268.

Samnordisk runtextdatabas. www.nordiska.uu.se/forskn/samnord.htm. Uppsala University. Accessed February 13, 2013.

Sandøy, Helge. 1994. The nature of "overlong syllables" in the Scandinavian languages. In *Phonologica: Proceedings of the 7th International Phonology Meeting 1992* (W.U. Dressler, M. Prinzhorn, J.R. Rennison, eds): 233-242. Torino: Rosenberg & Sellier.

Schwarz, Ernst. 1951. *Goten, Nordgermanen, Angelsachsen: Studien zur Ausgliederung der germanischen Sprachen*. Bern: Francke.

Shlonsky, Ur. 2010. The cartographic enterprise in syntax. In *Language and Linguistics Compass,* 4(6): 417-429.

Sievers, Eduard. 1876. Kleine Beiträge zur deutschen Grammatik III. Die starke Adjectivdeclination. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 2: 98-124.

Sipma, Pieter. 1913. *Phonology and Grammar of Modern West Frisian: With Phonetic Texts and Glossary*. London: OUP.

Starke, Michal. 2001. *Move Dissolves into Merge: A Theory of Locality*. PhD dissertation, University of Geneva.

Starke, Michal. 2004. On the inexistence of specifiers and the nature of heads. In *The Cartography of Syntactic Structures, Vol. 3: Structures and Beyond* (Belletti, ed.): 252–268. New York: OUP.

Starke, Michal. 2009. Nanosyntax: A short primer to a new approach to language. *Nordlyd, Tromsø University Working Papers on Language and Linguistics* 36 [Special issue on Nanosyntax]: 1-6. http://septentrio.uit.no/index.php/nordlyd/ (accessed March 12, 2012).

Starke, Michal. 2011a. Towards an elegant solution to language variation: Variation reduces to the size of lexically stored trees. LingBuzz/001183.

Starke, Michal. 2011b. Course on nanosyntax. November 7-10. Ghent University.

Starke, Michal. 2012. Talk on the internal structure of verbs. December. Ghent University

Starke, Michal. 2013. Auxiliaries and structural gaps: Current issues in Nanosyntax. Lecture series presented at CRISSP. 18, 20, 22 March. Hogeschool-Universiteit Brussel.

Streitberg, Wilhelm. 2000 [1910]. *Die Gotische Bibel* (2 volumes). Heidelberg: Carl Winter Universitätsverlag.

Sturtevant, Albert Morey. 1943. Notes on Old Norse phonology. *Scandinavian Studies* 17(5): 157-166. Society for the Advancement of Scandinavian Study.

Svenonius, Peter. 2006. The emergence of axial parts. *Nordlyd, Tromsø Working Papers in Linguistics* 33(1): 49-77.

Svenonius, Peter. 2010. Spatial prepositions in English. In *The Cartography of Syntactic Structure, Vol. 6* (G. Cinque and L. Rizzi, eds): 127-160. Oxford: OUP.

Szemerényi, Oswald J.L. 1996. *Introduction to Indo-European Linguistics*. Oxford: Clarendon Press.

Tanghe, Sanne and Marlies Jansegers. 2014. Marcadores del discurso derivados de los verbos de percepción: Un análisis comparativo entre el español y el italiano. *Revue Romane* 49(1): 1-31.

Taraldsen, Knut Tarald. 2009. The nanosyntax of Nguni noun class prefixes and concords. LingBuzz/000876.

Tiersma, Pieter Meijes. 1999 [1985]. *Frisian Reference Grammar*, 2nd ed. Ljouwert: Fryske Akademy.

Travis, Lisa. 1984. *Parameters and Effects of Word Order Variation*. PhD dissertation, MIT.

Twaddell, W.F. 1963. Graphical alternation in Old Saxon suffixes. *Monatshefte* 55(4): 225-228.

Þórólfsson, Björn K. 1925. Um íslenskar orðmyndir á 14. og 15. öld og breytingar fleirra úr fornmálinu með viðauka um nýjungar í orðmyndum á 16. öld og síðar. Reykjavík.

Van Craenenbroeck, Jeroen. 2009. Alternatives to cartography: An introduction. In *Alternatives to Cartography* (J. van Craenenbroeck, ed.): 1-14. Berlin: Mouton de Gruyter.

Van Craenenbroeck, Jeroen. 2012. Three challenges for Nanosyntax. Talk given at nanosyntax workshop. CASTL Decennium 2012. September 14. Tromsø. Slides available at http://jeroenvancraenenbroeck.net/talks/.

Vanden Wyngaerd, Guido. 2014. Pointers and the ABA diagnostic. Talk given at nanosyntax workshop. SLE 47. September 12. Poznan.

Vangsnes, Øystein. 2013. Syncretism and functional expansion in Germanic wh-expressions. *Language Sciences* 36: 47-65.

Vangsnes, Øystein. 2014. Indexicals by nanosyntax: *wh* and D items apart. Talk given at the nanosyntax workshop. SLE 47. September 12. Poznan.

Wackernagel, Jacob. 1892. Über ein Gesetz der indogermanischen Wortstellung. *Indogermanische Forschungen* 1: 333-436.

Watkins, Calvert (ed.). 2000. *The American Heritage Dictionary of Indo-European Roots*. Boston & New York: Houghton Mifflin.

Weerman, Fred and Jacqueline Evers-Vermeul. 2002. Pronouns and case. *Lingua* 112: 301-338.

Wright, Joseph. 1906 [1888]. *An Old High German Primer, with Grammar, Notes, and Glossary*, 2nd ed. Oxford: Clarendon Press.

Wurmbrand, Susanne. 2006. Verb clusters, verb raising, and restructuring. In *The Blackwell Companion to Syntax, Vol. 5* (M. Everaert, H. van Riemsdijk, eds): 229-343. Oxford: Blackwell.

Zanuttini, Raffaella. 1991. *Syntactic Properties of Sentence Negation: A Comparative Study of Romance Languages*. PhD dissertation, University of Pennsylvania.

Appendix I

Runic Norse data (Samnordisk runtextdatabas)

Stage (i) Dem-*si* forms

susi	DR 229
þasi	Ög 10, 157, 162, 212; Sm 17, 80; Sö 101; U 69, 126, 345
saR:si	Sö 137, Sö 340
sasi	DR 189
þan:si	Sö 158
þansi	extremely prevalent, e.g. DR 40; Ög 44
þat:si	Sö 47
þatsi	Sö 46
þaimsi	Öl 1
þaʀsi	Sö 40
þiʀsi	Sö 346
þausi	DR 4, 42, 133, 143, 209, 277, 293, 294; Nä 3; Ög Fv1970;310; Sö 173, 296; Vg 67, 115

Stage (ii) *i*-mutated Dem-*si* forms

þansi	extremely prevalent, e.g. DR 40, 53, 291; Ög 81, 165; Sö 45, 131, 154; Sm 42, 78; Vg 47, 51; U 342, 394
þan:si	Sö 158
þensi	Br E2; DR 83; Ög 201, 203, 207, 211; Vg 73, 127, 175
þinsi	Br SC14; DR 220, EM85;239, EM85;265; N 84, 237; Ög 47, 103, 104; Öl 15, 42; Sm 125; Sö 258, 296, 367; U 319, 818, 1143, Fv1990;32A; Vg 4, 79, 122, 135

þatsi	Sö 46
þat:si	Sö 47

þitsi	DR 383

þasi	Ög 10, 157, 162, 212; Sm 17, 80; Sö 101; U 69, 126, 345

þesi	Ög 68; Sm 100; U 617
þisi	DR 229, 269; Ög 214; Sö 127; U 335, 947, Fv1992;157; Vg 183; X ByNT1984;32

Stage (iii) Forms with both -*s* and -*a*

þitsa	Sö 188
þensa	U Fv1983;228; Ög 86; Sö 19, 20, 187, 350, Fv1969;298; U 19, 379, 478, 668, 720; Vs 18; Nä 11; DR 345, 389
þinsa	extremely prevalent, e.g. DR 387; Sö 25, 28, 179; Öl 28; Ög 13, 225; U 25, 35, 37; Vs 19; M 1, 17; Hs 6

Asigmatic reinforcer

-eka	DR 261
-ika	DR IK41,1 and DR IK98
-ekA	Ög KJ59 and DR 357
-kA	Ög KJ59

Appendix II

Intrusive *r* in Norse

Here I explain a potentially confusing fact about the assimilation or deletion of *r* in ON which was discussed in connection with the *r*-initial endings of the boxed forms in Section 3.1.2.

Observe that an adjective like *hvass-* 'sharp', which is crucially similar to RDem in that its stems ends in *ss*, could in fact display the *r* of the *r*-initial endings sometimes; see Table 77.

Table 77 *hvass-* 'sharp' (Barnes 2004: 103)

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	hvǫss	hvass	hvasst	hvassar	hvassir	hvǫss
ACC	hvassa	hvassan	hvasst	hvassar	hvassa	hvǫss
GEN	hvass(r)ar	hvass	hvass	hvass(r)a	hvass(r)a	hvass(r)a
DAT	hvass(r)i	hvǫssum	hvǫssu	hvǫssum	hvǫssum	hvǫssum

As seen in Table 77, we observe that F.GEN.SG *hvass(r)ar*, F.DAT.SG *hvass(r)i*, and GEN.PL *hvass(r)a* apparently alternate between *r*-endings and *r*-less endings, the latter version being the one that we encounter in RDem.

The availability of the variants with *r* (*hvassrar*, *hvassri*, *hvassra*) is problematic for the application of the rule *ssr* > *ss*, since it suggests that the sequence *ssr* could survive in ON, meaning that for some reason this rule is not exceptionless. If *ssr* is a legitimate consonant cluster, then the absence of this sequence in RDem (i.e. the predicted but unattested forms **þessrar*, **þessri*, **þessra*) would mean that we are not dealing with a phonological rule after all, but rather that something else is at stake in the RDem paradigm, such as a different class of endings. For the idea that there is a different class of endings at stake, we may note that (i) all the main classes of strong feminine nouns (*a*-, *i*-, and *r*-classes) take *-ar* instead of *-rar* in the F.GEN.SG (e.g. *helg-ar*), (ii) the *a*1-class of strong feminine nouns takes *-i* instead of *-ri* in the F.SG.DAT (e.g. *helg-i*), and (iii) all classes of strong nouns show *-a* instead of *-ra* in the PL.GEN (*hest-a*, *gest-a*, etc.) (the names of these noun classes come from Faarlund 2004). One hypothesis, then, could be that the *r*-initial adjective endings found in *hvassrar*, *hvassri*, and *hvassra* are the primary forms, while the *r*-less RDem forms are due to contamination from the *r*-less endings of the nominal classes just mentioned.[114]

For this to be true, though, forms like *hvassrar*, *hvassri*, *hvassra* should be the primary forms both synchronically and historically. That is, they should not be byproducts of an analogical change or of a later development in post-classical Norse. Indeed, it can be shown quite easily that the forms without *r* are primary and that the ones with *r* are later developments, meaning that the consonant cluster *ssr* belongs to a post-classical stage of ON. In this dissertation I am concerned with classical ON. As we will see in the rest of this section, the cluster *ssr* is not allowed at this stage in the language, and thus the rule *ssr* > *ss* is accurate for our purposes.

Return to the Dem paradigm, repeated in Table 78.

Table 78 ON Dem 'that'

<table>
<tr><th></th><th>F.SG</th><th>M.SG</th><th colspan="2">N.SG</th><th colspan="2">F.PL</th><th>M.PL</th><th>N.PL</th></tr>
<tr><td>NOM</td><td>sú</td><td>sá</td><td colspan="2">þat</td><td>þær</td><td>þār</td><td>þeir</td><td>þau</td></tr>
<tr><td>ACC</td><td>þá</td><td>þann</td><td colspan="2">þat</td><td>þær</td><td>þār</td><td>þá</td><td>þau</td></tr>
<tr><td>GEN</td><td>þeir(r)ar</td><td>þess</td><td colspan="2">þess</td><td colspan="2">þeir(r)a</td><td>þeir(r)a</td><td>þeir(r)a</td></tr>
<tr><td>DAT</td><td>þeir(r)i</td><td>þeim</td><td>því</td><td>þȳ</td><td colspan="2">þeim</td><td>þeim</td><td>þeim</td></tr>
</table>

Note that we see an alternation also in the Dem paradigm between single-*r* and double-*r* forms: F.SG.GEN *þeir(r)ar*, F.SG.DAT *þeir(r)i*, and PL.GEN *þeir(r)a*. In other words, there seems to be an intrusive *r* in the Dem paradigm in exactly the same slots as in the paradigm of *hvass-*.

In addition, there are a number of other lexical items which have both single-*r* and double-*r* forms in their paradigms. These items include *fá-* 'few', *grá-* 'gray', *blá-* 'blue', etc. In the shaded forms in Table 79 we see the relevant alternation again.

[114] Indeed, this hypothesis would find some common ground in Sievers' (1876) influential idea that some of the strong adjective suffixes are the result of 'contamination' between classes, since according to him some of these endings were taken from the demonstrative paradigm, e.g. F.GEN.SG **-zōz* > ON *-rar*, F.DAT.SG **-zai* > *-ri*, GEN.PL **-zǭ* > *-ra* (Haugen 1982: 93-4, though his morpheme boundaries are misleading).

Table 79 *fá-* 'few'

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	fá	fár(r)	fátt	fár	fáir	fá
ACC	fá	fán	fátt	fár	fá	fá
GEN	fár(r)ar	fás	fás	fár(r)a	fár(r)a	fár(r)a
DAT	fár(r)i	fám	fá	fám	fám	fám

Note in Table 79 that the M.NOM.SG form also has an intrusive *r*. The reason Dem in Table 78 does not have this option is that its M.NOM.SG form *sá* does not have an *r* to begin with.

The appearance of intrusive *r* has been said to be a phonological phenomenon in which *r* is geminated after a long accented vowel (see Sturtevant 1943: 157-9 and references cited there). But, as pointed out by Sturtevant (1943), this is unlikely considering that F.NOM/ACC.PL *þær*, *fár* and M.NOM.PL *þeir, fáir* are predicted, incorrectly, to display the variants **þærr*, **fárr*, **þeirr*, and **fáirr*. According to Sturtevant, then, a better hypothesis is that intrusive *r* is a morphological irregularity due to an analogical change with adjectives like *stór-* (i.e. stems ending in *r*) as model. In Table 80, *rr* appears exactly where intrusive *r* appears in the paradigm for *fá-* in Table 79; crucially, *rr* does *not* appear in F.NOM/ACC.PL *stórar* and M.NOM.PL *stórir*.

Table 80 *stór-* 'big'

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	stór	stór-r	stór-t	stór-ar	stór-ir	stór
ACC	stór-a	stór-an	stór-t	stór-ar	stór-a	stór
GEN	stór-rar	stór-s	stór-s	stór-ra	stór-ra	stór-ra
DAT	stór-ri	stór-um	stór-u	stór-um	stór-um	stór-um

I conclude with Sturtevant (1943) that intrusive *r* is an analogical phenomenon and as such an irregularity of sorts. Extending this proposal to the *r*-initial endings in *hvass-*, however, requires further discussion, since *hvass-* does not have a long stem vowel and as such does not fall into the same category as adjectives like *fá-*, *grá-*, and *blá-*. But with *hvass-* too an appeal to analogy can be made to account for the appearance of intrusive *r*. This requires further discussion of adjectival inflection.

Adjectives in ON with stem-final *s*, *l*, or *n* caused a following inflection-initial *r* to assimilate to *s*, *l*, or *n*, respectively. This type of assimilation was extremely prevalent in ON and many adjectives followed this pattern. This is shown in Tables 81-83 (Barnes 2004: 106).

Table 81 *laus-* ‘loose’

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	laus	laus-s	laus-t	laus-ar	laus-ir	laus
ACC	laus-a	laus-an	laus-t	laus-ar	laus-a	laus
GEN	laus-sar	laus-s	laus-s	laus-sa	laus-sa	laus-sa
DAT	laus-si	laus-um	laus-u	laus-um	laus-um	laus-um

Table 82 *gamal-* ‘old’

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	gǫmul	gamal-l	gamal-t	gaml-ar	gaml-ir	gǫmul
ACC	gaml-a	gaml-an	gamal-t	gaml-ar	gaml-a	gǫmul
GEN	gamal-lar	gamal-s	gamal-s	gamal-la	gamal-la	gamal-la
DAT	gamal-li	gǫml-um	gǫml-u	gǫml-um	gǫml-um	gǫml-um

Table 83 *komin-* ‘come’ (past participle)

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	komin	komin-n	komi-t	komn-ar	komn-ir	komin
ACC	komn-a	komin-n	komi-t	komn-ar	komn-a	komin
GEN	komin-nar	komin-s	komin-s	komin-na	komin-na	komin-na
DAT	komin-ni	komn-um	komn-u	komn-um	komn-um	komn-um

At a later point in history (the 1500s, sometimes earlier), however, the adjectives with such assimilations began reintroducing the underlying *r*, resulting in phonologically pleonastic forms like F.GEN.SG *gamall**r**ar*, F.DAT.SG *gamall**r**i*, and GEN.PL *gamall**r**a* (importantly, a new *r* was not introduced in M.NOM.SG **gamall**r***). The forms with this additional intrusive *r* remained popular in Icelandic for a couple of hundred years afterwards, but they were eventually replaced by the *r*-less forms again (present-day Icelandic uses the *r*-less forms, though its Dem forms have *rr*: *þeirrar*, *þeirri*, and *þeirra*). The cause of this *r*-intrusion was, once again, analogy with the great many adjectives in ON which did not display *r*-assimilation, e.g. *sjúk-rar*, *sjúk-ri*, *sjúk-ra* and *harðast-rar*, *harðast-ri*, *harðast-ra* (from *sjúk-* ‘sick’ and *harðast-* ‘hardest’; Barnes 2004: 104, 108). See Noreen (1923: 200-202, 292), Þórólfsson (1925), and Bandle (1956).

Therefore, the forms *hvassrar*, *hvassri*, and *hvassra* are secondary forms which appeared in post-classical ON by the same analogical change that caused *gamallar* to become *gamallrar*. The primary forms are *hvassar*, *hvassi*, and *hvassa*. They are *r*-less because assimilation (or deletion) has taken place. The hypothesis for the parallelism between *hvass-* and *gamal-* is supported also by the fact that the *r*-form **hvassr* is not attested for M.NOM.SG, just as **gamallr* is not attested. In sum, *r*-intrusion tends to be a development in later ON, while the earlier, primary forms of RDem and adjectives like *hvass-* display *r*-assimilation (or deletion). Thus a different class of inflectional endings does not need to be posited for RDem. RDem simply inflects with normal, *n*-type strong adjective endings.

There are also relevant comparative facts from OE that shed light on *r*-intrusion. Consider the shaded forms in the OE RDem paradigm in Table 84.

Table 84 Old English RDem (Lass 1994: 145)

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	þēos	þēs	þis	þās	þās	þās
ACC	þās	þisne	þis	þās	þās	þās
GEN	þis-re > þisse	þis(s)es	þis(s)es	þis-ra > þissa	þis-ra > þissa	þis-ra > þissa
DAT	þis-re > þisse	þis(s)um	þis(s)um	þis(s)um	þis(s)um	þis(s)um

Note that F.GEN/DAT.SG *þisse* and GEN.PL *þissa* are the result of *sr* > *ss* assimilation. The underlying strong adjective endings are F.GEN/DAT.SG *-re* and GEN.PL *-ra*. However, for these forms the variants *þissere* and *þisre* also exist. Once again, these *r*-forms are the result of analogy with the F.GEN/DAT.SG Dem form *þǣre* (Lass 1994: 145) and probably also with regular adjective forms like F.GEN/DAT.SG *gōd-re* (*gōd-* 'good'). In other words, while the OE RDem paradigm prominently shows *sr* > *ss* assimilation, it also displays *r*-intrusion due to later analogical influences, a situation parallel to the one just discussed for ON.

We have now postulated two developments: (i) *r*-deletion as a result of assimilation or deletion, and (ii) subsequent insertion of intrusive *r* as a result of analogies in post-classical ON (which for the most part lies outside the scope of this dissertation). At this point we may wonder if ON RDem also displays (later) *r*-intrusive variants. The answer is yes, but with an interesting catch. The *r*-variants are F.GEN.SG *þessarar*, F.DAT.SG *þessari*, and GEN.PL *þessara*. Consistent with the developments of intrusive *r* described above, these RDem forms were quite rare in older stages of ON but gradually grew in popularity, evidenced by the fact that the modern Icelandic RDem paradigm continues these very forms. Note, however, that these variants are not exactly what might be expected, since it appears that the *r*-initial endings attach to a stem *þessa-* instead of the expected stem *þess-*.

With respect to the development of this stem, Axelsdóttir (2003: 65-8) has argued that the innovation of this bisyllabic stem (perhaps an innovation from the Norwegian variety of Norse) is itself an analogical change. A feature of feminine adjectives of the *n*-type class was that if they had a bisyllabic NOM.SG, then they also had a trisyllabic DAT.SG. Contrast bisyllabic *heiðin* 'pagan' with monosyllabic *góð* 'good' in (203) (based on Axelsdóttir 2003: 66).

(203)	(a)	F.NOM.SG	góð		**hei.ðin**
		F.ACC.SG	góð.a		heið.na
		F.GEN.SG	góð.rar		hei.ðin.nar
		F.DAT.SG	góð.ri		**hei.ðin.ni**
	(b)		*early ON*	>	*classical ON*
		F.NOM.SG	sjá	>	**þessi**
			...		...
		F.DAT.SG	þes.si	>	**þes.sa.ri**

The change in RDem from a monosyllabic F.NOM.SG form *sjá* to a bisyllabic form *þessi*, seen in (203b), is crucial for Axelsdóttir's hypothesis, since this allows for *þessi* to take on a trisyllabic DAT.SG. To make a trisyllabic dative, the vowel *a* is inserted to the left of the *r*-initial F.DAT.SG ending *-ri*, giving *þess-**a**-ri*.[115] The *r*-initial endings in the genitive subsequently fell into the same pattern, giving the trisyllabic F.GEN.SG *þess-a-rar* and then GEN.PL *þess-a-ra*. The *r*-variants, with the stem *þessa-*, are therefore the exception rather than the rule.

[115] As should be clear from Section 3.2.3.3, the innovatory bisyllabic stem taking *r*-initial endings could not have been **þessi-* (i.e. **þessirar*, **þessiri*, **þessira*) since (floating) *i* is deleted in medial open syllables.

Appendix III

A note on ‘morphophonology’

Here I would like to address two important assumptions I have made with regard to the so-called ‘floating *i*’ which some may consider controversial. The first important assumption I have made is that *i*-umlaut is, at least to some extent, still active in ON, even though it is widely assumed that *i*-umlaut was only ‘productive’ in PN and had become ‘opaque’ by the time of ON. By ‘productive’ it is meant that umlaut is an active phonetic process of assimilation/vowel harmony. Eventually this phonetic kind of umlaut gives rise to new phonological distinctions, that is, front vowels are phonemicized upon the loss of conditioning environments for umlaut. Certain morphological alternations, furthermore, arise which are marked by [±back] vowels. This is what is meant by ‘opaque’: that umlauted vowels are not phonetically predictable anymore and have passed into a domain commonly called ‘morphophonology’.

Some perspective on this issue is helpful. Consider *u*-umlaut, which is uncontroversially considered to have been a productive, phonologically conditioned process in ON, while its direct descendant in modern Icelandic, *ʏ*-umlaut, is considered to be non-productive despite a great number of *a* ~ *œ* and *a* ~ *ʏ* alternations in Icelandic, which are hallmarks of *ʏ*-umlaut. There is a great deal of contemporary research which aims to understand the exact nature of *ʏ*-umlaut and its place in Icelandic grammar. In other words, there is still plenty of debate among phonologists about just how opaque *ʏ*-umlaut actually is, and if it belongs to phonology, morphology, or ‘morphophonology’ (see Hansson 2013 for a concise overview). One strategy for understanding the phenomenon comes from Gibson & Ringen (2000). They argue that the phonological structure of certain morphemes contains a floating bundle of [+round, –back] features, thereby explaining why some morphemes induce umlaut while others do not. This account is attractive because it captures a dual intuition about *ʏ*/*u*-umlaut, namely that it is a phonologically grounded phenomenon which has been incorporated into certain morphemes over time. (Indeed, one thing which seems clear is that *ʏ*-umlaut cannot be understood as a purely phonologically conditioned phenomenon in modern Icelandic.)

What I am doing with *i*-umlaut in ON, then, is exactly parallel to what Gibson & Ringen (see also Ingason 2013) have done for *Y*-umlaut in modern Icelandic. My *i*-mutator can be considered a floating phonological diacritic just like their *Y*-mutator, and it is also a way to model the gradual grammaticalization of *i*-umlaut, whereby the phonological forms of certain morphemes have acquired a floating *i*-mutator. Issues remain, but at the very least it would seem that my *i*-diacritic is a legitimate option with parallels in the work of contemporary research in morphophonology.

The second important assumption I have made is that *i*-syncope is still active in ON, while the traditional view is that it was active only in PN. Again some perspective will be helpful. Just as with umlaut, there is debate about the status of vowel-syncope in modern Icelandic, with many phonologists trying to integrate various V ~ 0 alternations into the phonology of modern Icelandic (e.g. *ha.mar-r* 'hammer.NOM.SG' vs. *ha.ma.r-i* 'hammer.DAT.SG' > *hamri*, where *a* goes to zero in the open syllable) even though there are various exceptions which suggest opacity of a previously regular process. Again, I think it is unwise to assume right off the bat that no theory of phonology will be able to identify underlying regularities for vowel-syncope in modern Icelandic (see McCarthy 2008 or Norris 2010 for OT accounts; for some relevant discussion within Government Phonology, see Fortuna 2013). And again, if syncope is regular in modern Icelandic, then surely it is in ON too (assuming that processes like syncope move unidirectionally from productive to non-productive). I am proposing, then, that we not jump the gun on labeling *i*-syncope 'opaque' in ON. If work on modern Icelandic reveals underlying regularities in the operation of syncope, then my assumption about ON is automatically vindicated.

Consider what Anderson (1969: 27-8) writes regarding counterexamples to *i*-umlaut in modern Icelandic where there is no apparent trigger for the vowel-fronting observed (e.g. *bók* 'book' – *bækur* 'books', *mús* 'mouse' - *mýs* 'mice'): "Confronted with these facts, we might be tempted to abandon the attempt to predict the occurrence of umlaut in phonological terms in Icelandic, and claim that it must be indicated in the lexicon by some *ad hoc* feature...such a solution would involve an enormous cost, and some other way would certainly be desirable if one could be found." Anderson then proceeds to present accounts of both *i*- and *u*-umlaut which preserve these phenomena as regular, productive processes in the morphophonology of Icelandic. The details are not crucial here, but it is important to note that this solution is possible in the first place. Furthermore, note that Anderson's approach applies to both *i*- and *u*-umlaut in *modern* Icelandic. If Anderson is on the right track about *i*-umlaut in modern Icelandic, then there is no question that *i*-umlaut in ON is also productive. Much debate and research have been spawned since Anderson's early work on this, and different theories take different views. This is not the place to choose between classical generative phonology, Optimality Theory, and Government Phonology. I am merely pointing out that there are many options available for grounding my assumptions about *i* in a formal theory of phonology.

I am proposing that we take the facts about *i*-umlaut and *i*-syncope at face value until the complexities of umlaut and syncope can be sorted out more carefully. Importantly, recall also that nanosyntax is quite restrictive in that it posits only three potential slots in a lexical entry: < /phonology/, [Syntax], CONCEPTUAL INFO >. This leaves no room for 'morphophonology', the ill-defined domain floating between phonology and syntax. While many researchers like to relegate umlaut and syncope to this domain, this option is unavailable to a nanosyntactician.

Appendix IV

A nanosyntactic approach to the NSAG (McFadden 2014)

In this appendix I provide more detailed discussion of McFadden's (2014) NSAG and my nanosyntactic account of the generalization.

A. McFadden (2014)

I will first illustrate the empirical data underlying the NSAG with data drawn from Tamil, Finnish, Latin, and Icelandic.

As illustrated in Tamil (204), the nominative forms in this language are set apart by two properties. On the one hand, the NOM.SG of *mar-* 'tree' (204b) is formed from *mar-* plus a special nominative element *-am* (cf. the regular NOM.SG ending *-Ø* in (204a)). Moreover, while the non-nominative forms of *mar-* 'tree' in (93b) regularly insert a stem-forming morpheme *-att* between *mar-* and the K ending, this morpheme is absent in NOM.SG *mar-am*. For Tamil *vii-* 'house' in (204b), the special NOM.SG form is *vii-ḍu*, and again the non-nominative forms of *vii-* 'house' insert the stem-forming morpheme *-ṭṭ* between *vii-* and the K ending.

(204) Tamil (McFadden 2014: 1-2, slightly modified)

(a) regular

	'boy'	'mother'
NOM	payyan-Ø	ammaa-Ø
ACC	payyan-ai	ammaav-ai
DAT	payyan-ukku	ammaav-ukku
INSTR	payyan-aale	ammaav-aale

(b) special NOM

	'tree'	'house'
NOM	**mar-am**	**vii-ḍu**
ACC	mar-att-ai	vii-ṭṭ-ai
DAT	mar-att-ukku	vii-ṭṭ-ukku
INSTR	mar-att-aale	vii-ṭṭ-aale

Similar effects are seen in Finnish. In Finnish (205), the NOM.SG of *ihmi-* 'person' (205b) is formed from *ihmi-* plus the special nominative element *-nen* (cf. the regular NOM.SG ending *-Ø* in (205a)). Moreover, the non-nominative forms of *ihmi-* 'person' in (205b) insert a stem-forming morpheme *-se* between *ihmi-* and the K ending.

(205) Finnish (McFadden 2014: 3, modified slightly)

(a) regular

	'house'	'street'
NOM	talo-Ø	katu-Ø
GEN	talo-n	kadu-n
PART	talo-a	katu-a
INESS	talo-ssa	kadu-ssa

(b) special NOM

	'person'
NOM	**ihmi-nen**
GEN	ihmi-se-n
PART	ihmi-s-tä
INESS	ihmi-se-ssä

Latin nominatives are special as well. In Latin (206), the NOM.SG of *hom-* 'man' (206b) is formed from *hom-* plus the special nominative element *-ō* (cf. the regular NOM.SG ending *-s* in (206a)). The non-nominative forms of *hom-* 'man' in (206b) insert a stem-forming morpheme *-in* between *hom-* and the K ending. In the case of *sen-* 'old man' in (206b), however, the irregularity goes in the opposite direction: here it is the NOM.SG which inserts a special stem-forming morpheme *-ec* between the base *sen-* and the regular NOM.SG ending *-s*; in the non-nominative forms of 'old man', the element *-ec* is not present. Finally, observe that *gen-* 'kind' in (206b) follows the same pattern as *hom-* 'man' (that is, an irregular NOM.SG ending, with a stem-forming element *-er* inserted in the non-nominative forms), except that there is a syncretism between the nominative and accusative. Due to this syncretism, what is normally a nominative-based irregularity surfaces in the accusative as well.

(206) Latin (McFadden 2014: 5, modified slightly) third declension

(a) regular

	'chief'
NOM	princep-s
ACC	princep-em
GEN	princep-is
DAT	princep-ī

(b) special NOM

	'man'	'old man'	'kind'
NOM	**hom-ō**	**sen-ec-s** < senex >	**gen-us**
ACC	hom-in-em	sen-em	**gen-us**
GEN	hom-in-is	sen-is	gen-er-is
DAT	hom-in-ī	sen-ī	gen-er-ī

In Icelandic (207), the NOM.SG of *mann-* 'man' (207b) is formed from the special stem *mað-* plus the regular NOM.SG ending *-ur* (cf. (207a)). The non-nominative forms take the regular stem *mann-* instead.

(207) Icelandic (McFadden 2014: 5, modified slightly)

(a) regular

	'horse'
NOM	hest-ur
ACC	hest-Ø
GEN	hest-s
DAT	hest-i

(b) special NOM

	'man'
NOM	**mað-ur**
ACC	mann-Ø
GEN	mann-s
DAT	mann-i

As seen in (204-207), the NSAG manifests itself empirically in a number of ways. McFadden is very clear on distinguishing phonology from morphosyntax. In the regular cases (the (a) examples), any alternations in the stem can be accounted for phonologically. For instance, in the Tamil paradigm for *ammaa-* 'mother', we see the insertion of *v* before a vowel-initial ending. This is a phonological rule, meaning that this is not a reason to posit a structural difference between nominative *ammaa-* and non-nominative *ammaav-*. Similarly, for Finnish *katu-* 'street', we see *-tu-* in open syllables (NOM *katu*, PART *katua*) and *-du-* in closed ones (GEN *kadun*, INESS *kadussa*). Again, this means we should not posit two different stems *katu-* vs. *kadu-*, only that PF may affect certain forms after they have been built by the (morpho)syntax.

What is interesting from a structural/morphosyntactic point of view is that certain nouns in the languages illustrated above have a nominative form or stem which is different from the non-nominative forms. Tamil has nominative *mar-am* vs. non-nominative *mar-att-* 'tree'; Finnish has nominative *ihmi-nen* vs. non-nominative *ihmi-se-* 'person'; Latin has nominative *sen-ec-* vs. non-nominative *sen-* 'old man'; and Icelandic has nominative *mað-* vs. non-nominative *mann-* 'man'. Note in passing that the nominative stem *gen-* (vs. non-nominative *gen-er-*) in Latin also bleeds into the accusative since neuter nouns in Indo-European are always syncretic between these two cases. To capture patterns like these McFadden posits the following generalization.

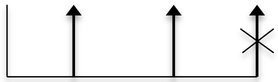
(208) Nominative stem-allomorphy generalization (NSAG) (McFadden 2014: 8)

> When there is stem allomorphy based on case, it distinguishes the nominative (along with any cases systematically syncretic with the nominative) from all other cases.

McFadden proposes various possible ways to explain the NSAG, which I will discuss next. Ultimately, my own interpretation is that the NSAG arises from a tendency to lexically package the fseq in certain ways, which ultimately is due to the merge order in the fseq.

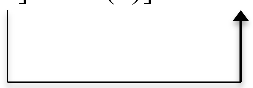
McFadden's main proposal for explaining the NSAG is based on Moskal (2013). According to Moskal's hypothesis, root-allomorphy (i.e. allomorphy of √N) based on K is blocked in case both *n* (a stem-forming element) and # (number) are between the root and K.

(209) [[[[√N] *n*] #] K]



In the NSAG cases that McFadden discusses, the issue is not root-allomorphy but rather stem-allomorphy (i.e. elements involving *n* in addition to √N). The stem-forming element, as mentioned, is taken to be *n*, and its allomorphy is conditioned by K. See (210). For instance in Finnish, *n* corresponds to *-se*, which surfaces in the genitive whether or not # (here plural *-i*) intervenes. The same goes for Latin, except in this language # and K are packaged together (cf. Section 2.2.1).

(210) [[[[√N] *n*] (#)] K]



Finn.	[*ihmi-*	***-se-***]		*-n*	'[person]-GEN.SG'
	[*ihmi-*	***-s-***]	*-i*	*-en*	'[person]-GEN.PL'
Lat.	[*hom-*	***-in-***]		*-is*	'[man]-GEN.SG'
	[*hom-*	***-in-***]		*-um*	'[man]-GEN.PL'

As seen in (210), *n* is not too far away from K to be conditioned by K, since only # intervenes. If the singular is thought of as the total absence of #, moreover, then there is *no* intervening node between *n* and K. Thus we expect both the singular (with no #) and the plural (with #) to be able to affect the realization of *n*. This is indeed the case, as is particularly clear for the Finnish GEN.PL *ihmi-s-i-en*, where # (*-i*) and K (*-en*) are morphologically distinct.

Citing Lamontagne & Travis (1987), Bittner & Hale (1996), among others, McFadden (2014) proposes that nominative is the absence of case. If this is so, then the special nominative forms in (204-207) can be thought of as simply the root √N plus *n*.

(211) √N *n*

Tam.	*mar-*	*-am*	'tree.NOM'
Finn.	*ihmi-*	*-nen*	'person.NOM'
Lat.	*hom-*	*-ō*	'man.NOM'

However, this does not work as cleanly when the plural is taken into account. In Finnish and Latin, the locality hypothesis from Moskal (2013) makes the correct forms fall out, with the additional assumption that nominative is the total lack of K (and singular the total lack of #).

Table 85 Finnish and Latin (McFadden 2014: 15, modified slightly)

	SG	PL	SG	PL
NOM	ihmi-nen [root-*n*]	ihmi-se-t [root-*n*-#]	hom-ō [root-*n*]	hom-in-ēs [root-*n*-#]
GEN	ihmi-se-n [root-*n*-K]	ihmi-s-i-en [root-*n*-#-K]	hom-in-is [root-*n*-K]	hom-in-um [root-*n*-#K]

In Table 85, *n* surfaces as Finnish *-nen* or Latin *-ō* if it is not to the left of anything. This happens only in the NOM.SG, since both K and # are, by hypothesis, absent under these exact circumstances. Once another node intervenes, whether it be # or K, *n* is realized as Finnish *-se* or Latin *-in* instead. In Tamil, however, this general approach does not lead to the correct results.

Table 86 Tamil (McFadden 2014: 15)

	SG	PL
NOM	mar-am [root-*n*]	mar-aŋ-gaɭ [root-*n*-#]
ACC	mar-att-ai [root-*n*-K]	mar-aŋ-gaɭ-ai [root-*n*-#-K]

In Table 86, *-aŋ* is an allophonic version of the NOM.SG *n*-morpheme *-am*. Thus the plural forms in Table 86 take the NOM.SG stem, which is the flipside of the situation in Table 85. As McFadden points out, the Tamil non-NOM.SG *n* (*-att-*) seems to be sensitive to the presence of K, but *-att-* is blocked in the plural by the presence of #. In Finnish and Latin, on the other hand, it is the NOM.SG *n* (*-nen* and *-ō*) which is blocked, by either # or K. Thus the locality conditions are quite different: in Tamil only # blocks, but in Finnish and Latin either # or K can block.

Because of these differences, it is very difficult to deduce a generalization about locality that will work in all three languages. Furthermore, the claim that nominative is the lack of K and singular the lack of # makes the available structure diminishingly sparse, adding to the difficulty in making the correct patterns fall out. The impression one gets is that there are not enough environments in which precise rules can apply to give the observed forms.

Before I propose my own (nanosyntactic) account of the Tamil, Finnish, Latin, and Icelandic facts, I would like to point out another feature of the tendentious nature of the NSAG. As explicitly stated in (208), the NSAG also applies to any cases that are syncretic with the nominative, often the accusative. If we closely inspect the consequences of this fact, it will become clear that when interpreted as an absolute principle, the NSAG is problematic. As McFadden (2014: 11-12) recognizes, the underlying structures of nominative and accusative need to be kept distinct, but insisting on this (an insistence to which I am certainly sympathetic) leads to a very unnatural view on the NSAG. Take Latin NOM/ACC.SG *gen-us* as an example. Recall that *n* is realized as *-us* precisely when there is nothing to the right of *n* (cf. Finnish and Latin above); that is, spellout of *n* as *-us* will be blocked if there is # and/or K. If we want to keep nominative and accusative distinct, however, it is hard to see how this can also be the case in the ACC.SG. Even if nominative really is the lack of K, this surely cannot be the case for accusative as well. That is, presumably some K feature is added in the accusative. This K, of course, would block the expected form.

McFadden realizes this, and his ideas on how to solve this problem reveal the unnaturalness of a Distributed Morphology approach to the NSAG:

> One way or another, the idea should be that, with these nouns [i.e. those with NOM/ACC syncretism], what shows up in syntactic contexts where we expect the accusative are structurally speaking nominatives, at least at the point when the exponent for little n is inserted. (McFadden 2014: 18)

In order for this to be the case, he proposes that what starts out as an accusative structure later has "the relevant head deleted or pruned before vocabulary insertion" or that there is "a kind of Differential Object Marking, so that accusative case assignment rules would simply not apply to nouns of the relevant classes, leaving them caseless" (McFadden 2014: 18). These solutions I find to be unlikely, since they stipulate that some accusatives are somehow morphosyntactically special beyond the relatively minor fact that there is a syncretism with nominative. I think it is more likely that the syncretism is just a basic fact about the lexical packaging for these nouns, and that otherwise these nouns are normal, structurally speaking. That way we can avoid all the additional machinery required for deletion or pruning or differential object marking, which are invoked simply in order to preserve the absolutist nature of the NSAG as it is formulated by McFadden.

B. A nanosyntactic approach

Let me begin this section by singling out two important examples from above. The first is Latin NOM.SG *sen-ec-s* and the Icelandic NOM.SG *mað-ur*. In both cases we see a special nominative stem, *sen-ec-* and *mað-*, but there is also a nominative ending, namely Latin *-s* and Icelandic *-ur*. In the Latin case it is clear that *-ec-* is an instantiation of *n*, which leaves little doubt that *-s* is a K morpheme, especially considering that the normal NOM.SG ending is *-s*, as seen in the completely regular NOM.SG *princep-s* 'chief'. In Icelandic, moreover, it is clear that *-ur* is a NOM.SG ending since it appears throughout a large class of masculine nouns, whether they are regular or not: *hest-ur* 'horse-NOM.SG', *mað-ur* 'man-NOM.SG', etc. Clearly nominative is not simply the absence of K. This is evidence for Caha's (2009) view that nominative corresponds to a K-head, i.e. K_1, rather than the view that nominative is the absence of any K whatsoever.

Let us now try to capture the data in (204-207). Let us first assume that there are in fact features for singular (Sg) and plural (Pl) (rather than Sg being the absence of #), giving the fseq in (212).

(212) $K_{...}$ K_1 Pl Sg *n* √N

Imagine now that the special nominative endings *-nen* in Finnish and *-ō* in Latin are portmanteaus of K_1, Sg, and *n*. This is illustrated in Figure 194.

K_1	**Sg**	***n***	**√N**	**SpO**
-nen			ihmi-	*ihmi-nen*
-ō			hom-	*hom-ō*

Figure 194 NOM.SG packaging

Now recall that the non-nominative case endings cooccur with a stem formant *n*, i.e. *-se-* in Finnish and *-in-* in Latin. Since *n* will need to be lexicalized on its own in these forms, the non-nominative case endings cannot package *n* with K. This gives Figure 195.

$K_{...}$	**K_1**	**Sg**	***n***	**√N**	**SpO**
-$K_{\text{non-NOM, SG}}$			-se-	ihmi-	*ihmi-se-n*
-$K_{\text{non-NOM, SG}}$			-in-	hom-	*hom-in-is*

Figure 195 Non-nominative, singular packaging

Finally, the plural endings will also have the *n* lexicalized as *-se-* or *-in-*, so the plural endings cannot package *n* with them, but they must also allow for *n* to surface overtly in the NOM.PL. This means that *n* must extend all the way up to Sg, as in Figure 196.[116]

[116] Note that Finnish and Tamil are more agglutinative than Latin, so they can lexicalize the Pl layer separately in Finnish GEN.PL *ihmi-s-**i**-en* or Tamil ACC.PL *mar-aŋ-**gaḷ**-ai*. In the NOM.PL *ihmi-se-t* or *mar-aŋ-gaḷ* (and everywhere in Latin) K and Pl are a portmanteau.

(K...)	K_1	Pl	Sg	*n*	√N	SpO
$-K_{PL}$			-se-		ihmi-	*ihmi-se-t*
$-K_{PL}$			-in-		hom-	*hom-in-ēs*

[where $-K_{PL}$ = plural, either nominative or not]

Figure 196 Plural packaging

We end up with the lexical entries in (213), which are crucial for our understanding of the NSAG.

(213) Lexical entries for Finnish

< *ihmi-* ⇔ √NP > < $-K_{PL}$ ⇔ ($K_{...}$P) K_1P PlP >

< *-nen* ⇔ K_1P SgP *n*P > < $-K_{non\text{-}NOM.SG}$ ⇔ $K_{...}$P K_1P SgP >

< *-se-* ⇔ SgP *n*P >

(K...)	K_1	Pl	Sg	*n*	√N	SpO
$-K_{PL}$			**-se-**		ihmi-	*ihmi-se-t*

K...	K_1	Sg	*n*	√N	SpO
$-K_{non\text{-}NOM, SG}$			**-se-**	ihmi-	*ihmi-se-n*

Figure 197 Shrinking of *-se-*

K_1	Sg	*n*	√N	SpO
-nen			ihmi-	*ihmi-nen*

Figure 198 Tailormade entry for *-nen*

As seen in Figure 197, both layers of the *-se*-morpheme – SgP and *n*P – are lexicalized in the presence of the plural endings. The *-se*-morpheme shrinks to *n*P, however, in the presence of the non-nominative singular case endings. These case endings are anchored (see Section 2.2.6) at Sg, meaning that SgP must be spelled out as part of the case endings. By the Superset Principle, though, *n*P can still be spelled out as *-se-*. Finally we see in Figure 198 that the morpheme *-nen* is, according to my proposal, a morpheme tailormade for the NOM.SG, since it lexicalizes the K_1, Sg, and *n* layers as a single chunk. That is, *-nen* will be spelled out just in case the structure [K_1P SgP *n*P] happens to be built. If the syntax builds beyond K_1 in the singular, e.g. [K_2P K_1P SgP *n*P], then the entry for *-nen* will not apply since nothing beyond K_1 is stored in its lexical structure. This captures the fact that *-nen* is suppressed in the non-NOM.SG part of the paradigm. If the syntax builds non-singular case endings, e.g. [K_1P PlP SgP *n*P], then *-nen* will also fail to apply, since its lexical structure does not contain Pl. This captures the fact that *-nen* is suppressed in the plural.

In Tamil, on the other hand, the lexical entries are slightly different, resulting in the mirror-image-style pattern in Table 86 above. In both Finnish/Latin and Tamil, the NOM.SG *n*-allomorph corresponds to a chunk, i.e. K_1, Sg, and *n* are packaged together. The non-NOM.SG *n*-allomorph, however, differs crucially in one respect between Finnish/Latin and Tamil. Recall from (213) that the non-NOM.SG *n*-allomorph *-se-* extends up to Sg in Finnish, meaning that the plural case endings will fit on top of *-se-* perfectly. Thus *-se-* is lexicalized in the plural in Finnish. In Tamil, however, the non-NOM.SG *n*-allomorph *-att-* does not appear in the plural. As seen in (214) and Figure 200, this is because Tamil *-att-* does not extend up to Sg, so it cannot be used to lexicalize Sg. Instead the nominative morpheme *-am* has to jump in to lexicalize Sg, as shown in Figure 199.

(214) Lexical entries for Tamil

< *mar-* ⇔ √NP > < *-K*$_{PL}$ ⇔ (K$_{...}$P) K_1P PlP >

< *-am* ⇔ K_1P SgP *n*P > < *-K*$_{non\text{-}NOM.SG}$ ⇔ K$_{...}$P K_1P SgP >

< *-att-* ⇔ *n*P >

K_1	Sg	*n*	√N	SpO
-am			mar-	*mar-am*

(K...)	K_1	Pl	Sg	*n*	√N	SpO
-K$_{PL}$			**-am-**		mar-	*mar-aŋ-gaḷ*

Figure 199 Shrinking of *-am*

K...	K_1	Sg	*n*	√N	SpO
-K$_{non\text{-}NOM.SG}$			**-att-**	mar-	*mar-att-ai*

Figure 200 *n*P alone corresponds to *-att-*

In other words, in Finnish/Latin we see shrinking of the non-nominative, singular allomorph; in Tamil, on the other hand, we see shrinking of the NOM.SG allomorph instead.

The point of this exercise is to show how the NSAG can be thought of as a tendency for lexical packaging. Recall from Chapter 2 that the nanosyntactic algorithm for matching syntactic structure with lexical structure is STAY > CYCLIC > SNOWBALL. The second step, CYCLIC, can be paraphrased as 'keep building the same morpheme'. The third step, SNOWBALL, can be paraphrased as 'start a new morpheme'. That is, the system is economical in that it prefers to spell out a span of features as one morpheme rather than several. However, a span of features will eventually become too large for a single morpheme to handle, so another one will have to take over.

As already mentioned above in the main text of this chapter, my proposal is that the NSAG is simply a reflection of the fact that *n* is closer to K_1 than it is to the other K-layers, and thus it is more likely that *n* will be packaged with K_1 than the K features above it. Purely in terms of distance within the fseq, CYCLIC will be more likely to apply up to K_1 than to the features beyond K_1. The farther beyond K_1 we go, the more likely it will be that SNOWBALL will have to applied. There is no absolute CYCLIC/SNOWBALL border at K_1, then. This we know because cases syncretic with nominative, such as accusative, can be packaged into the *n*-morpheme; that is to say, CYCLIC can clearly apply at least up to K_2.[117]

[117] McFadden (2014: 9) himself brings up the possibility that the NSAG may very well be the result of more general principles surrounding markedness or even just tendencies of historical change. My proposal, that the NSAG is a tendency for lexically packaging morphemes in certain ways, can be thought of in this light.

Appendix V

The proximal demonstrative

On the basis of extensive crosslinguistic evidence (syncretism and morphological containment), Lander & Haegeman (2015) argue that spatial deixis is arranged in the fseq in the following hierarchy: Distal 'far from speaker and hearer' > Medial 'close to hearer' > Proximal 'close to speaker'.

(215) [Dx_3P	[Dx_2P	[Dx_1P]]]	=	Distal
	[Dx_2P	[Dx_1P]]	=	Medial
		[Dx_1P]	=	Proximal

Ignoring the medial, (215) shows that the distal structurally contains the proximal, which is shown overtly in the morphology of various languages, as shown by Lander & Haegeman (2015).

One point that emerges from the diachronic literature is that RDem seems to have developedinto a proximal demonstrative in (some of) the modern Germanic languages, The 'bar' Dem form functions as the Distal. In the context of the present thesis, this is potentially paradoxical: the question arises from (215) why in the account developed here RDem appears to be larger, i.e. structurally more complex, than the distal Dem, when according to Lander & Haegeman (2015) the proximal should be smaller, i.e. structurally less complex, than the distal. According to (215), the distal is expected to morphologically contain the proximal, but the evidence suggests that the RDem, which has acquired the proximal reading, has the form [[Dem]-*si*] and thus appears to contain the distal [Dem].

While at first glance this seems to be a paradox, the issue dissolves when we take a closer look. First of all, it is rather clear that the older Dem-*si* stage of RDem is *not* a proximal demonstrative, but rather a neutral demonstrative that was reinforced. Dem at this early stage, crucially, meant not only 'that', but *either* 'that' *or* 'this'. Put differently, at the early stages Dem was syncretic between proximal and distal, a situation found, for instance, in present day French. On top of this neutral demonstrative, then, the reinforcer

-*si* could be added. In other words, the addition of -*si* is not at the basis of the spatial-deictic reading; spatial deixis is already built into the neutral Dem. Put differently, the containment relation we see between Dem-*si* and Dem is not one of proximal containing distal, but one of a reinforced demonstrative containing a non-reinforced demonstrative. So reinforcers should not be equated with markers of spatial deixis.

Now let us consider some languages in which RDem seems to have developed into a proximal demonstrative, for instance Swedish *denna*, *detta* 'this' vs. *den*, *det* 'that'. Here there appears to be morphological containment in the opposite direction of Lander & Haegeman's (2015) proposal: [[*den*]-*na*], [[*det*]-*ta*] are the proximals vs. [*den*], [*det*] which are the distal forms, and the proximal seems to contain the distal, in contrast with (215) in which the distal contains the proximal

(216) [[den]-na] → [[distal]-proximal]??
[[det]-ta]

However, here again the pattern is not exactly what it seems. For instance, the containment relation does not hold throughout the entire paradigm: specifically, there is no containment relation visible in the plural: *dessa* 'these' vs. *dom* 'those'. Moreover, it is not completely clear that these forms are what Lander & Haegeman (2015) are interested in. Indeed, rather than *denna*, *detta*, the more common way to express the proximal in Swedish is the configuration *den här*, *det här* 'the here = this'. In this configuration *den*, *det* is unstressed and can plausibly be equated with the prenominal definite article. The distal, moreover, is most commonly expressed by a parallel configuration: *den där*, *det där* 'the there = that' (see Leu 2007, 2008, 2015 for discussion). Thus there is actually no overt containment relation whatsoever between *den här*, *det här* vs. *den där*, *det där*, and thus Swedish is not a counterexample to (215).

On the basis of this we can propose that the real proximal/distal contrast lies in the adverbs *här* 'here' vs. *där* 'there'. In Swedish, these adverbs combine with D to create the proximal vs. distal system: *den här*, *det här* 'this' vs. *den där*, *det där* 'that'.

(217) [Dx_3P [Dx_2P [Dx_1P]]] => *där* + D = *den där*
[Dx_1P] => *här* + D = *den här*

Thus, the true markers of spatial deixis in Swedish are actually *här* for proximal and *där* for distal. Supporting evidence for this analysis comes from the Norwegian dialect of Trøndersk, where the adverbs *her* 'here' and *der* 'there' act alone, without the aid of a definite item, as proximal and distal markers.

(218) Trøndersk (Leu 2008: 23, fn.25)

(a) herre film-en
here movie-the
'this movie'

(b) derre film-en
there movie-the
'that movie'

This is evidence that the adverb is really what is at stake in the proximal/distal distinction, and not the form which is historically related to RDem, namely *denna*, *detta*. The modern Scandinavian morpheme *-Ca/-Ce* in *den-na/den-ne* and *det-ta/det-te* is perhaps more accurately a morpheme containing both spatial-deictic and reinforcement features, explaining its more complex makeup.

Along similar lines, consider the fact that English *this* (< RDem) vs. *that* (< Dem) does not display a containment relation contrary to Lander & Haegeman's prediction. Neither does German, in its system of *d-ies-er* 'this' (< RDem) vs. *j-en-er* 'that'. Nor does Dutch, in its system of *deze*, *dit* 'this' (< RDem) vs. *die*, *dat* 'that' (< Dem). Thus these WGmc languages do not pose a problem for Lander & Haegeman's (2015) generalization about the relative sizes of the proximal and the distal.

Interestingly, the historical development of RDem into Swiss German did *not* result in a proximal demonstrative, as discussed in Section 4.1.6. The data are repeated in (219).

(219) Swiss German (Leu 2008: 36-37)

(a) Dem = neutral: 'this' *or* 'that'

F.SG dɛ
M.SG diä
N.SG das

(b) RDem = discourse-salient, contrastive: 'the other'

F.SG disi
M.SG disä
N.SG dises

As seen in (219b), Swiss German RDem does not have the proximal reading 'this' but rather the contrastive reading 'the other'. Thus, in Swiss German RDem retains the older semantics of reinforcement rather than developing towards a spatial-deictic reading. As seen in (219a), the Dem form in Swiss German is a neutral demonstrative (as it was in Old Germanic), namely a form that means *either* proximal 'this' *or* distal 'that'.

The point of this appendix is to show that the development of RDem into the modern languages is not nearly as straightforward as often claimed. It is simply not the case that RDem was a proximal in Old Germanic. It is also not the case that RDem straightforwardly developed into the modern Germanic proximal. Thus we have to be very careful when we look at Germanic data in the context of hypotheses about spatial deixis, such as Lander & Haegeman (2015). Indeed, in Mainland Scandinavian there is evidence that the adverbs for 'here' and 'there' are better candidates for lexicalizing spatial-deictic features, not the form which is historically derived from RDem. In WGmc, we see that the Swiss German RDem is not a proximal demonstrative at all, rather it has a discourse-salient, contrastive reading. In sum, we must be very cautious with the traditional narrative that RDem gave rise to the modern Germanic proximal.

Summary in English

This dissertation is a detailed study of the internal structure of the reinforced demonstrative pronoun (RDem) of the oldest Northwest Germanic (NWGmc) languages: Runic Norse (RN) *súsi*, *sási*, *þatsi*; Old Norse (ON) *sjá/þessi*, *sjá/þessi*, *þetta*; Old Frisian (OF) *thius*, *this*, *thit*; Old English (OE) *þēos*, *þe(:)s*, *þis*; Old Saxon (OS) *thius*, **these*, *thit*; and Old High German (OHG) *dësiu*, *dësēr*, *diz*. (Forms are given in the order F.NOM.SG, M.NOM.SG, N.NOM/ACC.SG.)

Consider the ON RDem paradigm, given in the table below.

Three patterns in the ON RDem paradigm (Gordon 1956: 294-295)

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	þe-ssi	þe-ssi	þe-tt-a	þe-ss-ar	þe-ss-ir	þe-ssi
ACC	þe-ss-a	þe-nn-a	þe-tt-a	þe-ss-ar	þe-ss-a	þe-ssi
GEN	þe-ss-ar	þe-ss-a	þe-ss-a	þe-ss-a	þe-ss-a	þe-ss-a
DAT	þe-ss-i	þe-ss-um	þe-ss-u	þe-ss-um	þe-ss-um	þe-ss-um

The RDem forms in this table can be decomposed into smaller pieces. For instance, the N.DAT.SG form *þessu* is composed of the base *þe-* plus the geminated reinforcer component *-ss-* plus the strong adjective ending (K) *-u*. In fact, all of the forms which are boxed in the table have this basic template: *þe-* + *-ss-* + K.

The non-boxed forms can be divided into two sorts. The first sort (lightly shaded in the table) consists of the base plus the geminated inflectional ending (K) (M.ACC.SG *-n*, N.NOM/ACC.SG *-t*, M/N.GEN.SG *-s*) plus the reinforcer *-a*. The template for the lightly shaded forms, then, is: *þe-* + -KK- + *-a*. The second sort (with dark shading in the table) consists of the base plus the geminated reinforcer component *-ss-* plus *-i*. The template for the darkly shaded forms, then, is (roughly): *þe-* + *-ss-* + *-i*.

With three different patterns observable in the ON RDem paradigm, it is fair to say that there is a significant amount of intraparadigmatic variation at stake.

After an overview of the data and some philological background in Chapter 1, an introduction to Michal Starke's theory of nanosyntax, an offshoot of the cartographic program, is provided in Chapter 2. In Chapter 3 the fine-grained decomposition of the ON RDem forms is undertaken in more detail. Ultimately, five distinct morphological ingredients are identified and given labels. Since the morphemes *-s* and *-a* are in complementary distribution, they are assumed to be two realizations of one syntactic head, R.

þa-	base (D)
-K	strong adjective ending (K)
-C^i	consonant geminator with *i*-mutator (Gm)
-s	the sigmatic reinforcer } (R)
-a	the asigmatic reinforcer } (R)

As seen above, these components combine into three different structures within the RDem paradigm. The boxed forms are referred to as *K-final forms*, the darkly shaded forms are referred to as *constant forms*, and the lightly shaded forms are referred to as *K-internal forms*.

(i) the K-final forms

D-R-Gm-K
þa-s-C^i-K

(ii) the K-internal forms

D-K-Gm-R
þa-K-C^i*-a*

(iii) the constant forms

D-K-R-Gm
þa-K-*s*-C^i (where ^{i} surfaces word-finally as *-i*)

Chapter 4 treats these discoveries within the formal framework of Cinque (2005). Working within his U20 program, I demonstrate how the correct functional sequence of RDem can be deduced, how the three RDem structures are derived, and why *only* these three structures (of the 24 possible structures) are attested. Interestingly, support for Cinque's system is found in the allomorphy between the reinforcer morphemes *-s* (found in the cyclic-type derivations of the K-final and constant forms) and *-a* (found in the roll-up derivation of the K-internal forms).

Chapter 5 provides a nanosyntactic analysis of the facts. It is shown that a more fine-grained functional sequence allows us to capture the *-s* ~ *-a* alternation in terms of lexical structure. A number of other facts also fall out from this more fine-grained approach. Finally in Chapter 6 I bring the WGmc facts into the scope of the analysis. With a complete understanding of ON under our belts, we are able to capture the various points of morphological variation between the RDem paradigms of OE, OF, OS, OHG (and even RN) in a very simple way.

It is argued that all of the variation observed across the RDem paradigms can be boiled down to the way lexical entries are structured, in line with the nanosyntactic approach. Each language's morphological inventory divides the functional sequence up differently, which then leads to different derivations during the process of spellout (i.e. lexicalization). My findings support the central tenet of the Principles and Parameters framework: crosslinguistic variation is superficial; all the variation observed across languages can be reduced to the lexicon, where the arbitrary or exceptional parts of linguistic knowledge are stored. The functional sequence, on the other hand, remains unexceptional, universal, and innate.

Samenvatting in het Nederlands

Dit proefschrift is een gedetailleerde studie van de interne structuur van het versterkt aanwijzend voornaamwoord (RDem) van de oudste Noordwest-Germaanse (NWGmc) talen: Runisch Noors (RN) *súsi*, *sási*, *þatsi*; Oudnoors (ON) *sjá/þessi*, *sjá/þessi*, *þetta*; Oudfries (OF) *thius*, *this*, *thit*; Oudengels (OE) *þēos*, *þe(:)s*, *þis*; Oudsaksisch (OS) *thius*, **these*, *thit*; en Oudhoogduits (OHG) *dësiu*, *dësēr*, *diz*. (De vormen zijn in volgende volgorde gegeven: vrouwelijk enkelvoudig nominatief (F.NOM.SG), mannelijk enkelvoudig nominatief (M.NOM.SG), onzijdig enkelvoudig nominatief/accusatief (N.NOM/ACC.SG).)

Bekijk de ON RDem-paradigma, gegeven in de tabel hieronder.

Drie patronen in het ON RDem-paradigma (Gordon 1956: 294-295)

	F.SG	M.SG	N.SG	F.PL	M.PL	N.PL
NOM	þe-ssi	þe-ssi	þe-tt-a	þe-ss-ar	þe-ss-ir	þe-ssi
ACC	þe-ss-a	þe-nn-a	þe-tt-a	þe-ss-ar	þe-ss-a	þe-ssi
GEN	þe-ss-ar	þe-ss-a	þe-ss-a	þe-ss-a	þe-ss-a	þe-ss-a
DAT	þe-ss-i	þe-ss-um	þe-ss-u	þe-ss-um	þe-ss-um	þe-ss-um

De RDem-vormen in deze tabel kunnen ontleed worden in kleinere stukjes. Bijvoorbeeld, de onzijdige enkelvoudige datieve (N.DAT.SG) vorm *þessu* is opgemaakt uit de basis *þe-* plus het gegemineerde versterkingscomponent *-ss-* plus de sterke adjectieve uitgang (K) *-u.* Alle omkaderde vormen in de tabel hebben immers dezelfde opbouw: *þe-* + *-ss-* + K.

De niet omkaderde vormen kunnen verdeeld worden in twee soorten. De eerste soort (lichtgrijs in de kader) bestaat uit de basis plus de gegemineerde uitgang (K) (M.ACC.SG *-n*, N.NOM/ACC.SG *-t*, M/N.GEN.SG *-s*) plus de versterker *-a*. De opbouw voor de lichtgrijze vormen is vervolgens: *þe-* + -KK- + *-a*. De tweede soort (donkergrijs in de kader) bestaat uit de basis plus het gegemineerde versterkingscomponent *-ss-* plus *-i.* De opbouw voor de donkergrijze vormen is vervolgens (ruwweg): *þe-* + *-ss-* + *-i*.

Met drie verschillende geobserveerde patronen in het ON RDem-paradigma, kan gezegd worden dat er een significante hoeveelheid intraparadigmatische variatie in het spel is.

Na een overzicht van de data en wat filologische achtergrond in hoofdstuk 1, biedt hoofdstuk 2 een inleiding tot de nanosyntactische theorie van Michal Starke, een aftakking van het cartografische programma. In hoofdstuk 3 wordt de ontleding van de ON RDem-vormen gedetailleerd bekeken. Uiteindelijk zullen vijf distinctieve morfologische ingrediënten geïdentificeerd en benoemd worden. Omdat de morfemen *-s* en *-a* complementair zijn, worden ze verondersteld twee realisaties van een syntactisch hoofd, R, te zijn.

þa-	basis (D)	
-K	sterke adjectieve uitgang (K)	
-C^i	geminator met *i*-mutator (Gm)	
-s	de sigmatische versterker	(R)
-a	de asigmatische versterker	

Zoals hierboven getoond wordt, kunnen deze componenten drie verschillende structuren maken binnen het RDem-paradigma. Ik verwijs naar de omkaderde vormen als *K-finale vormen*, de donkergrijze als *constante vormen*, en tenslotte de lichtgrijze vormen als *K-interne vormen*.

(i) De K-finale vormen

D-R-Gm-K
þa-s-C^i-K

(ii) De K-interne vormen

D-K-Gm-R
þa-K-C^i*-a*

(iii) De constante vormen

D-K-R-Gm
þa-K-*s*-C^i (waar ^{i} verschijnt aan het einde van een woord als *-i*)

Hoofdstuk 4 behandelt deze ontdekkingen in de formele omkadering van Cinque (2005). Werkend binnen zijn U20 programma, toon ik aan hoe de correcte functionele opvolging (fseq) van RDem afgeleid kan worden, hoe de drie RDem-structuren syntactisch opgebouwd worden, en waarom alleen *deze* drie structuren (van de 24 mogelijke structuren) geobserveerd worden. Opmerkelijk genoeg wordt steun voor Cinque's systeem gevonden in de allomorfie tussen de versterkende morfemen -*s* (zoals gezien in de cyclische afleidingen van de K-finale en constante vormen) en -*a* (zoals gezien in de sneeuwbalafleiding van de K-interne vormen).

Hoofdstuk 5 biedt een nanosyntactische analyse van de feiten. Het wordt aangetoond dat een meer fijnkorreligere fseq ons toestaat de -*s* ~ -*a* allomorfie te omvatten aangaande met lexicale structuur. Enkele andere feiten worden duidelijk door deze fijnkorrelige benadering. Tenslotte in hoofdstuk 6 betrek ik de WGmc feiten in de analyse. Met een compleet begrip van de feiten in ON, zijn we in staat de verschillende onderdelen van morfologische variatie tussen de RDem-paradigmas van OE, OF, OS, OHG (en zelfs RN) op een zeer eenvoudige manier te bevatten.

Het is beargumenteerd dat, in lijn met de nanosyntactische benadering, alle variaties gezien in de RDem-paradigmas herleid kunnen worden tot hoe de lexicale vermeldingen zijn gestructureerd. De morfologische inventaris van elke taal verdeeld de fseq anders, wat als gevolg heeft dat er verschillende afleidingen verschijnen tijdens de lexicalisatie. Mijn bevindingen ondersteunen de centrale leerstelling van de omkadering van Principes en Parameters: intertaalkundige variatie is oppervlakkig; al de variaties geobserveerd in talen kunnen herleid worden tot de lexicon, waar de willekeurige of exceptionele onderdelen van taalkundige kennis opgeslagen worden. Anderzijds blijft de fseq universeel en aangeboren.