

Chapter 2

Phrasal compounds in Modern Icelandic with reference to Icelandic word formation in general

Kristín Bjarnadóttir

The Árni Magnússon Institute for Icelandic Studies, University of Iceland

In Icelandic, as in many other languages, phrasal compounds are an interface phenomenon of the different components of grammar. The rules of syntax seem to be preserved in the phrasal component of Icelandic compounds, as they show full internal case assignment and agreement. Phrasal compounds in Icelandic can be divided into two distinct groups. The first group contains common words which are part of the core vocabulary irrespective of genre, and these are not stylistically marked in any way. Examples of these structures can be found in texts from the 13th century onwards. The second group contains more complex compounds, mainly found in informal writing, as in blogs, and in speech. These seem to be 20th century phenomena. Phrasal compounds of both types are relatively rare in Icelandic, but other types of compounding are extremely productive. Traditionally, Icelandic compounds are divided into two groups, i.e., compounds containing stems and compounds containing inflected word forms, mostly genitives, as non-heads. Phrasal compounds in Icelandic also have genitive non-heads, raising questions on the difference between the processes in non-phrasal and phrasal compounding in Icelandic.

1 Introduction

Compounding is extremely productive in Icelandic, and an indication of this can be seen in the proportions of non-compounds (base words) vs. compounds in The Database of Modern Icelandic Inflection (DMII, Bjarnadóttir 2012), a full-form database of inflectional forms produced at The Árni Magnússon Institute



for Icelandic Studies and its forerunner, The Institute of Lexicography.¹ The DMII contains the core vocabulary of Modern Icelandic, with approximately 280,000 paradigms. The vocabulary is not selected by morphological criteria, apart from the self-explanatory fact that only inflected words are included. The sources of the DMII are lexicographic data, both from traditional dictionary archives and corpora. Out of 278,764 paradigms in the DMII on Dec. 15th 2015, 32,118 entries were non-compounds, and the remaining 246,646 entries were compounds. The DMII contains both lexicalized compounds and purely productive ones, but the same rules of word formation pertain to both, i.e., they are morphologically identical.

The DMII only contains compounds written as continuous strings, in accordance with current Icelandic spelling conventions. These spelling conventions are a feature of Modern Icelandic and they do not hold in older forms of the language. To give a very simple and common example, patronyms are written as a continuous string in Modern Icelandic, e.g. *Bjarnadóttir* ‘daughter of Bjarni’, not *Bjarna dóttir* as evidenced in older texts. Residues of the older spelling are still found in some instances in Modern Icelandic, as when the names of the sagas are written discontinuously: *Njáls saga* ‘The Story of Burnt Njáll’. This is traditional in the names of the sagas and recommended in the current spelling rules for Icelandic, but otherwise the continuous string is the norm. Spelling mistakes in present-day Icelandic do, however, very often involve the splitting of compounds, and these are most commonly found in informal texts where phrasal compounds (PCs) are very often found. These problems with spelling make PCs elusive both in traditional lexicographic archives and in automatic word extraction. PCs are here taken to be compounds where the non-head contains any kind of syntactic phrase, from noun phrases and prepositional phrases up to full finite sentences.

Discussion of PCs is largely absent from the linguistic literature on Icelandic, and probably first mentioned in Bjarnadóttir 1996[2005], citing examples not adhering to Botha’s (1981) No Phrase Constraint. The Icelandic examples cited in Bjarnadóttir 1996[2005] are now a part of a private collection of over 200,000 Icelandic compounds, with full analysis of structure and constituent parts. The sources for this collection are to a large extent the same as for the DMII. The following analysis of PCs is based on this collection, with approx. 200 additional

¹ The DMII was initially conceived as a language resource for natural language processing, but was also intended for use in lexicography and linguistic research. The paradigms are accessible online as a reference tool and are used as such by the general public. Downloadable data and website: <http://bin.arnastofnun.is>.

PCs from other sources, such as *Íslenskur orðasjóður* (Wortschatz, University of Leipzig, see Hallsteinsdóttir et al. 2007), a corpus of texts from Icelandic websites, which is a good source of informal language. The total number of PCs used in this study is approx. 900. The problems involved in finding the more informal PCs are described in §3, cf. (16). At the present stage of technology, the data is sparse, and the full picture of PCs in Icelandic therefore awaits a better analysis of multiword lexical items.

In this study, PCs in Modern Icelandic are divided into two groups, based on structure, and usage or genre. The first group (Phrasal Compounds I, PCIs) contains structures which are attested by examples from the 13th century onwards, as in the Dictionary of Old Norse Prose (ONP). These PCIs are very much a feature of Modern Icelandic, they are not marked in any way stylistically, and they may appear in any genre. The most common structures of phrasal non-heads in this group are prepositional phrases (1a), and genitive noun phrases (1b), both showing full inflection or agreement:²

- (1) a. milli ríkja samningur
 between.PREP state.N.NEUT.GEN.PL contract.N.MASC
 ‘international agreement’
- b. tveggja manna far
 two.NUM.GEN.PL man.N.MASC.GEN.PL vehicle
 ‘a boat for two’

The second group (Phrasal Compounds II, PCIIs) contains PCs that are found in certain informal genres, i.e., in blogs, social media, and speech, etc. All the examples are recent, they are often considered a little strange, and the question “Is this really a word?” is sometimes heard in connection with them. The structure of the non-head in PCIIs ranges from nominative noun phrases (2a) to fully-fledged sentences (2b):

- (2) a. maður -á -mann aðferð
 man.NOM to.PREP man.ACC method
 ‘“man to man” method’
- b. ég- er- bara- einn- af- ykkur- strákuðum -brosið
 I am just one of you boy.DAT.DEF.PL smile.NOM.DEF
 ‘the “I am just one of you boys” smile’

² The compounds are aligned to the glosses, but Icelandic spelling conventions stipulate that they are written continuously. Hyphens are shown when part of the spelling.

This study presents a classification of PCs in Icelandic based on their morphological structure. Section 2 contains an introduction to the relevant features of Icelandic compounding in general, demonstrating both stems and inflected forms of nouns and adjectives as non-heads of compounds. Section 3 contains a classification of the PCs, divided into the traditional, not marked in any way by style or genre (PCI), and the more informal (PCII). The PCII constructions do, for the most part, have counterparts in Carola Trips' analysis of English PCs (Trips 2016). Section 4 contains the conclusion, along with a few words on future work.

2 Compounding in Icelandic

A comprehensive description of Icelandic compounding drawing on empirical data has not yet been written, and apart from short chapters in textbooks, the grammatical literature on Icelandic compounds is fragmentary and mostly written in Icelandic. A description of relevant aspects of Icelandic compounding is therefore in order here.

2.1 Binary branching and stress pattern

Following Eiríkur Rögnvaldsson's textbook on Icelandic morphology (Rögnvaldsson 1986), linguists have assumed binary branching for Icelandic compounds. Recent experiments with automatic compound splitting have confirmed this observation (Daðason & Bjarnadóttir 2014). Morphologically, Icelandic compounds are right-headed, and inflectional endings of compounds attach to the word-final element. All word classes can appear both as head and non-head, but noun-noun compounds are by far the most productive (Bjarnadóttir 1996[2005]). As almost all the PCs discussed here are nouns, with a marginal number of adjectives, the topic of this section is limited to compounds with nouns or adjectives as heads.

The word formation rules for lexicalized and productive compounds in Icelandic are morphologically identical, as the data in the DMII shows. That resource is intended for use in language technology and the vocabulary is therefore chosen to be as inclusive as possible, including data both from lexicographic sources and corpora.

As stated above, compounds are expected to be written as continuous strings in Modern Icelandic. As the spelling can fluctuate, this is sometimes not a very definitive criterion, and it would be of great interest to define compounds either with reference to syntactic analysis or by using phonological aspects, such as intonation and stress (Árnason 2011, and references cited there). Empirical data of this kind, however, does not exist as yet.

In spite of this, the basic stress pattern of Icelandic words can be a guideline in determining whether an item is a compound or phrase. This pattern is fairly simple, with word initial stress, and secondary stress, etc., on constituents according to morphological binary trees, as in ¹*kransæðar*³*hjáveitu*²*græðlingur* ‘coronary bypass transplant’ in Figure 1. The prosodic pattern is therefore determined by the morphological structure, with the numbers 1, 2, 3, in the binary tree in Figure 1 signifying primary, secondary and tertiary stress.

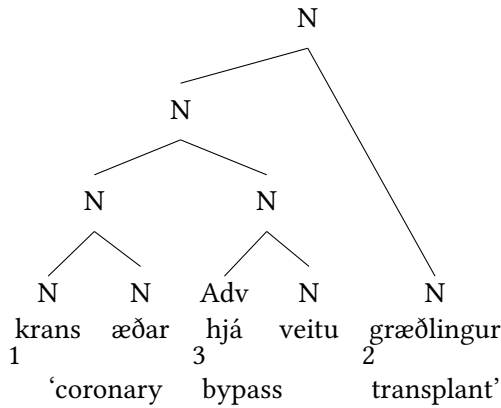


Figure 1: The stress pattern of an Icelandic compound

The compounds discussed in §2 are assumed to conform to this basic stress pattern, as do most of the PCIs in §3.1, but there is still insufficient research on the topic for an exact description of the exceptions. The complex structures in the PCIs in §3.2 below are more of a problem where stress is concerned, as the relatively simple rules of word stress do not apply to syntactic phrases as non-heads. Informally, the observation that the head of the PCIs is stressed has been confirmed by native speakers, but proper experiments have not been carried out. The question whether these are indeed compounds phonologically therefore remains open, but comparative data from other languages shows that similar structures are analysed as PCs in those, as is the case in Trips (2016) for English. As most of the examples of PCIs here are from written texts or transcriptions where the original sound files are unavailable, the question of phonology may be a moot point.

2.2 Recursion

Noun-noun compounds are by far the most common type of compounds in Icelandic, and also the most structurally complex. As stated above, Icelandic compounds are right-headed, but the constituent structure in recursive compounds can be either left- or right-branching, cf. examples in (9–13). Theoretically there is no limit to the length of compounds, and the classic example of this is the frequently quoted word in (3) *Vaðlaheiðarvegavinnuverkfærageymsluskúrsútidyrallyklakippuhringur*, where *Vaðlaheiði* is a compound place name.

- (3) Vaðlaheiðar vega vinnu verk færa geymslu skúrs úti dyra lykla kippu
V. road works work tools storage shed out door key bunch
hringur
ring
'key ring of the key chain of the outer door to the storage tool shed of the road works on the Vaðlaheiði plateau'

Overlong compounds are apt to be split up in Icelandic, using prepositional phrases at need, and in reality more than seven constituents are rare (Snædal 1992; Daðason & Bjarnadóttir 2014). The compound in (3) could be rephrased as

- (4) lyklakippuhringur fyrir útidyrnar á verkfærageymsluskúr vegavinnunnar
á V
'a key chain ring for the outside door of the tool storage shed of the roadworks on V.'

In spite of the trend towards splitting, overlong compounds do sometimes occur, such as *Norðausturatlantshafsfiskveiðinefndin* 'The North East Atlantic Ocean Fisheries (lit. Fish-Catching) Committee'. Long PCs should therefore not cause a problem for Icelanders just because of their length, even if they are not common.

2.3 Inflection or compound markers?

Nouns and adjectives as non-heads in Icelandic compounds appear in different forms, i.e., as stems or inflectional forms, mostly genitive. Dative non-heads are also found in compounds, as in *gyðjumlíkur* 'goddess.N.FEM.DAT.PL like.ADJ' (Bjarnadóttir 2002). A very limited number of non-head combining forms are also found, e.g., *kven*- of the feminine noun *kona* 'woman' where the regular non-head would be *konu* (GEN.SG) or *kvenna* (GEN.PL). Linking phonemes also occur, but

these are rare, with the proportion 0.005% in 38,000 non-heads in compounds in Bjarnadóttir 1996[2005]. The discussion here will be limited to stems and genitive forms as non-heads, as these are very frequent, whereas the other types are very rare.

The analysis of genitives as such in Icelandic compounds is traditional in the Icelandic grammatical literature, dating back to Rasmus Christian Rask's seminal work on Icelandic grammar Rask (1811). According to this analysis, nouns as non-heads appear as stems or genitive forms, singular or plural. Corresponding structures in Faroese and some West Norwegian dialects are analysed in the same manner in Indriðason (2014) and Thráinsson et al. (2004)

The nature of these genitives in Icelandic compounds and the question whether these are true inflectional forms or linking phonemes are matters of debate, especially in theories that specify a strict ordering of derivation, compounding and inflection. The argumentation that these genitives are not a part of morphological structures, but attributes within noun phrases, is difficult to maintain for the following reasons: The stress pattern described in §2.1. can be used to determine whether a structure is a compound or phrase, but additionally, basic Icelandic word order provides clues, as genitive attributes are usually placed after the nominal head in a sentence: *bók Kristínar* 'Kristín's book'. The reverse order, *Kristínar bók*, is usually found with contrastive stress (cf. Thráinsson 2007: 92–96). Furthermore, this analysis would leave almost half of the vocabulary, i.e. the so-called weak inflection, unavailable for compound formation as these can never appear as stems in non-heads, cf. §2.5.³

The case against analysing the genitive non-heads in Icelandic compounds as compound markers or linking phonemes for Icelandic also rests on the fact that the non-heads appear as the correct genitive forms, in spite of the complexity of the inflectional patterns. Inflectional variants are very common, and the paradigms in the DMII reflect this, with 594 inflectional patterns listed for the major word classes, i.e., nouns, adjectives, verbs, and adverbs (Bjarnadóttir 2012). The reason for the high number of inflectional patterns in the DMII is that each paradigm contains all inflectional variants, i.e., a word is not assumed to belong to more than one inflectional class, as in the traditional classification in Icelandic textbooks. The rampant variation found among genitive singular inflectional forms is fully reflected in the form of the non-heads.

³ Further argumentation against level ordering or split morphology can be found in Icelandic derivation, as some suffixes can attach to genitive non-heads: *mannlegur* man.N.MASC.STEM -ly.SUFF.ADJ 'human', *mannslegur* man.N.MASC.GEN.SG -ly.SUFF.ADJ 'manly', *mannalegur* man.N.MASC.GEN.PL -ly.SUFF.ADJ 'pompous, conceited' (Bjarnadóttir 1996[2005]; Indriðason 1994).

The non-heads appear as correct genitive forms, as shown in all the examples in §2.4.⁴ To give an example, the base word *vegur* ‘way, road’ has the genitive singular forms *vegar* and *vegs*, the first of which is much more frequent. Both *-ar* and *-s* appear in the non-heads of compounds, i.e., *vegarendi* ‘end of road’, *vegsauki* ‘increase of way’, i.e., ‘promotion’. (The genitive plural *vega* is also used in compounds: *vegamót* ‘joint of roads, i.e., crossroads’). Compounds with the head *vegur* can exhibit variants in the same way as the base word, but the crux of the matter is that these variants can be reflected in the non-heads of compounds as well, as in (5b–c). However, some compounds with the head *vegur* only have *-s* as a genitive ending, thus exhibiting a different inflectional pattern from the base words, which is interesting in light of Lieber’s theories of percolation (1989). This genitive is always reflected in the non-heads of recursive compounds, as in *útvegur* ‘out-way’, i.e., ‘fishing, fisheries’, and *farvegur* ‘passage way’, i.e., ‘channel, course’ (5d–e). Underscoring marks the genitive endings:

- | | | |
|-----|--|--------------------------------------|
| (5) | Lemma | Gen.sg. |
| | a. <i>vegur</i> ‘way, road’ | <i>vegar</i> , <i>vegs</i> |
| | Compounds: | |
| | <i>vegarendi</i> ‘end of road’ | |
| | <i>vegsauki</i> ‘increase of way’, i.e., ‘promotion’ | |
| | b. <i>reiðvegur</i> ‘(horse) riding road’ | <i>reiðvegar</i> , <i>reiðvegs</i> |
| | Compounds: | |
| | <i>reiðvegarspotti</i> ‘stretch of riding road’ | |
| | <i>reiðvegsframkvæmd</i> ‘riding road construction’ | |
| | c. <i>Laugavegur</i> ‘pool way’ (street name) | <i>Laugavegar</i> , <i>Laugavegs</i> |
| | Compounds: | |
| | <i>Laugavegsapótek</i> ‘Pool Street Drug Store’ | |
| | <i>Laugarvegarganga</i> ‘a walk along Pool Street’ | |
| | d. <i>útvegur</i> ‘out-way’ (‘fisheries, fishing’) | <i>útvegs</i> |
| | Compound: | |
| | <i>útvegsþorp</i> / * <i>útvegarþorp</i> ‘fisheries village’ | |
| | e. <i>farvegur</i> ‘passage way’ | <i>farvegs</i> |
| | Compound: | |
| | <i>farvegsbreyting</i> /* <i>farvegarbreyting</i> ‘change of course’ | |

The conclusion is that *-s* and *-ar* are inflectional endings in Icelandic compounds and not linking phonemes. This is directly opposite to the case of German, where paradigmatically incorrect forms such as *liebesbrief* ‘love letter’ are

⁴ The exceptions are few, and can for the most part be explained by historical changes. These obsolete inflectional forms are only a feature of lexicalized compounds.

analysed as containing a prosodic marker, here *-s-*. With the correct feminine genitive, the compound would be *liebebrief* (Trips, personal communication).

The function of the genitive in compounding is considered in Indriðason (1999; 2014) in the light of the split morphology hypothesis (Perlmutter 1988) and the split inflection theory (Booij 1994), and his conclusion is that “the genitive in Icelandic compounds can formally be categorized as contextual inflection but functionally as inherent inflection. This dual role of the genitive is unique and creates problems for the theories previously mentioned” (Indriðason 2014: 30). The aim here is to present these so-called genitive forms, to be able to compare them with the genitives in the PCs in §3, as these undoubtedly contain inflectional forms. The question is, then, whether the “ordinary” (i.e., non-phrasal) compounds contain true genitives.

2.4 Non-head in compounds: Nouns

Examples of the different forms found in the non-heads of noun-noun compounds are shown in (6) (see Bjarnadóttir 2002). These nouns are all written as continuous strings without hyphens. The lemma forms are shown in parentheses, as in *naglrót* (*nögl+rót*). Underscoring is used for genitive endings and for emphasis, as in *nögl*, to mark the umlaut.

(6) Form of non-head in noun-noun compounds

A. Stem

1. Lemma form⁵

- a. orð myndun (orð+myndun)
 word.N.NEUT formation.N.FEM
 ‘word formation’

2. Without umlaut

- b. nagl röt (nögl+röt)
 nail.N.FEM root.N.FEM
 ‘base of finger/toe-nail’

3. With umlaut (rare)

- c. lög brot (lög+bro^ut)
 law.N.NEUT.PL breaking.N.NEUT
 ‘infraction of law’

⁵ Lemma form without nominative ending where applicable, as in *hest* for the masculine *hestur*, subtracting the masculine nominative ending *-ur*.

4. Irregular (rare change in stem)

- d. mann tal (maður+tal)
man.N.MASC count.N.NEUT
'census'

B. Inflectional forms

1. Genitive singular

- e. borðs horn (borð+horn)
table.N.NEUT.GEN.SG corner.N.NEUT
'corner of a table'

- f. hunds haus (hundur+haus)
dog.N.MASC.GEN.SG head.N.MASC
'head of a dog'

- g. kattar haus (köttur+haus)
cat.N.MASC.GEN.SG head.N.MASC
'head of a cat'

- h. penna strik (penni+strik)⁶
pen.N.MASC.GEN.SG stroke.N.NEUT
'stroke of a pen'

- i. peru tré (pera+tré)
pear.N.FEM.GEN.SG tree.N.NEUT
'pear tree'

- j. bókar kápa (bók+kápa)
book.N.FEM.GEN.SG coat.N.FEM
'dust jacket'

2. Genitive plural⁷

- k. orða bók (orð+bók)
word.N.NEUT.GEN.PL book.N.FEM
'dictionary'

- l. bíla stæði (bíl+stæði)
car.N.MASC.GEN.PL place.N.NEUT
'car place, parking lot'

⁶The use of stems is limited in some inflectional classes, cf. §2.5.

⁷The genitive plural of all nouns ends in *-a* (or *-na* for some feminine and neuter nouns).

m.	bóka	búð	(bók+búð)
	book.N.FEM.GEN.PL	store.N.FEM	
	‘book shop’		
n.	dúfna	kofi	(dúfa+kofi) ⁸
	pigeon.N.FEM.GEN.PL	hut.N.MASC	
	‘pigeon hut’		

The genitive forms of the non-head in compounds are in accordance with the correct genitives, as they occur in the paradigms in the DMII. To give examples, the genitives of the masculine nouns *hundur* ‘dog’ and *köttur* ‘cat’ are *hunds*/**hundar* and *kattar*/**kötts*, and always appear as such when the genitive is used in the non-heads of the compounds of these words (cf. B.1.f and g in (6)). A choice of identical linking phonemes to the correct genitive endings is less than convincing, especially as the choice of endings on individual words is for the most part idiosyncratic. PCs with genitive phrases as non-heads also invariably contain the correct genitive forms.

The choice of stem or inflected form seems to be arbitrary for compounds where the non-head is a base noun, i.e., not a compound (Bjarnadóttir 1995), with the exceptions discussed below (this section). The compounds *bóksala* and *bóka-búð* shown in (7) thus contain the stem and the genitive plural of the word *bók* ‘book’ as non-heads without any discernible reason for the difference, as the compounds are semantically identical with synonyms as heads. The distribution is not phonetically conditioned either, as seen in *blekborði* (*k+b*) ‘ink strip’ (cf. *bók-abúð*), and *bókasafn* (*ka+s*) ‘book museum’, i.e., ‘library’ (cf. *bóksala*) occurring freely on morpheme boundaries:

- (7) a. bók.N.FEM.STEM sala.N.FEM ‘book shop’
 b. bóka.N.FEM.GEN.PL búð.N.FEM ‘book shop’

The choice of stem or genitive construction may be arbitrary in non-recursive compounds, as in (7), but it turns out that it is not free, i.e., the form itself can be lexicalized, so to speak, as users will only accept the expected variant, thus *bóksala*, *bókabúð* vs. **bókasala*, **bókbúð*. The same can apply to the choice between genitive singular and plural, which is often not semantically significant, as in (8a–b) where *barns*/*barna* can refer to one or more children.⁸

⁸ The difference between genitive singular and plural can be significant, as in *bróðr-sonur* ‘brother’s.N.MASC.GEN.SG son.N.MASC.SG’ (‘the son of (your) brother’), *bróðr-synir* ‘brother’s.N.MASC.GEN.SG sons.N.MASC.PL’ (‘the sons of (your) brother’), and *bræðr-synir* ‘broth-

- (8) a. barns meðlag
 child.N.NEUT.GEN.SG support.N.NEUT
 ‘child support’ (paid by parent)
- b. barna lífeyrir
 child.N.NEUT.GEN.PL support.N.MASC.PL
 ‘child support’ (paid by state, etc.)
- c. barns vagger
 child.N.NEUT.GEN.SG crib.N.FEM
 ‘baby’s crib’
- d. barna rúm
 child.N.NEUT.GEN.PL bed.N.NEUT
 ‘baby’s cot’

The choice between stem and genitive appears to be less free in recursive compounding, with left-branching compounds ([[N N] N]) tending to result in genitive constructions (Jónsson 1984), when the corresponding non-recursive compound does not, as in the pairs *skrifborðsfótur* (9a) and *borðfótur* (9b), and *olíuverðshækkun* (9c) and *verðhækkun* (9d):

- (9) a. [skrif borðs] fótur
 write.N.NEUT.STEM desk.N.NEUT.GEN.SG leg.N.MASC
 ‘writing desk leg’
- b. borð fótur
 desk.N.NEUT.STEM leg.N.MASC
 ‘desk leg’
- c. [olíu verðs] hækkun
 oil.N.FEM.GEN.SG price.N.NEUT.GEN.SG rise.N.FEM
 ‘rise in oil price’
- d. verð hækkun
 price.N.NEUT.STEM rise.N.FEM
 ‘price rise’

ers’.N.MASC.GEN.PL sons’.N.MASC.PL (‘the sons of (your) brothers’). The compound *bræðrasonur* ‘brothers’.N.MASC.GEN.PL son’.N.MASC.SG (‘the son of (your) brothers’) is not found. Some nouns exhibit agreement of number between non-head and head, as in the singular *mannsnafn* ‘persons’.N.MASC.GEN.SG name’.N.NEUT.SG (i.e., ‘Christian name’) vs. the plural *mannanöfn* ‘persons’.N.MASC.GEN.PL names.N.NEUT.PL (i.e., ‘Christian names’). It is unclear how extensive number agreement of this type is in compounds and the topic awaits further research.

Left-branching recursive compounds with stems of compounds as non-heads do also occur, although they are much rarer than the corresponding genitive constructions. These are of two kinds, i.e., with a stem compound as first part of the non-head [[N.STEM N].STEM N] (cf. *saltfiskútflutningur*, 10b), and with a genitive compound as first part of the non-head [[N.GEN N].STEM N] (cf. *fjánhúsdyr*, 10c):

- (10) a. [kú fisk] plógur
 cow.N.FEM.STEM fish.N.MASC.STEM plough.N.MASC
 ‘ocean quahog plough’
- b. [salt fisk] útflutningur
 salt.N.NEUT.STEM fish.N.MASC.STEM export.N.MASC
 ‘salt fish export’
- c. [fjár hús] dyr
 sheep.N.NEUT.GEN.SG house.N.NEUT.STEM door.N.FEM.PL
 ‘sheep house door’
- d. [betrunar hús] vist
 betterment.N.FEM.GEN.SG house.N.NEUT.STEM stay.N.FEM.SG
 ‘stay in jail’
- e. [rentu kammer] bréf
 rent.N.FEM.GEN.SG chamber.N.NEUT.STEM letter.N.NEUT.SG
 ‘letter from the (Danish) ministry of finance’ (*renta*: ‘rent, interest’)

The observation in Jónsson 1984 of the strong tendency towards genitive in compound non-heads holds for the most part, but stem compounds as in (10a–b) do also exist in compound non-heads, sometimes even as variant forms, as in (10b) *saltfiskútflutningur* [[N.STEM N].STEM N] where the corresponding *saltfisksútflutningur* [[N.STEM N].GEN N] is also found.⁹ The compounds in (10c–d) are more problematic, as these contain a stem ending in -s where the genitive ending would also be an -s. The syllables containing the genitive are unstressed, moreover, as can be inferred from Figure 1 above, and the difference in vowel length normally occurring in such genitives (i.e., *hús* vs. *húss*) may thus not be discernible (Árnason 2011). This could therefore be a matter of spelling, although the genitive -s is usually preserved in such cases. The compound in (10e), *rentukammerbréf*, contains an undisputed genitive construction in *rentu*.N.GEN.SG.*kammer*, but the first part is in fact a weak feminine noun which can never appear as a

⁹ In this case the stem compound *saltfiskútflutningur* seems to be much more common than the genitive compound *saltfisksútflutningur*. The frequency on timarit.is (The National Library’s corpus of newspapers and journals) is 372/104.

stem, as is the case in the word *olía* in (9c) (cf. §2.5). The evidence for the construction [[N.GEN N]_{STEM} N] therefore does not seem to be very strong.

Right-branching recursive compounds do not exhibit similar restrictions as the left-branching ones do, as stem constructions and genitive constructions mix freely:

- (11) I. [N.STEM [N.STEM N]]
- a. [stál [borð búnaður]]
 steel.N.NEUT.STEM table.N.NEUT.STEM equipment.N.MASC.SG
 ‘steel cutlery’
- b. [stíl [hug sjón]]
 style.N.MASC.STEM mind.N.MASC.STEM vision.N.FEM.SG
 ‘ideal of style’
- c. [her [flug maður]]
 army.N.MASC.STEM flight.N.NEUT.STEM man.N.MASC.SG
 ‘military pilot’
- II. [N.GEN. [N.STEM N]]
- d. [togara [sjó maður]]
 trawler.N.MASC.GEN.SG sea.N.MASC.STEM man.N.MASC.SG
 ‘trawler fisherman’
- e. [bómullar [hand klæði]]
 cotton.N.FEM.GEN.SG hand.N.FEM.STEM cloth.N.NEUT
 ‘cotton towel’
- f. [atvinnu [flug maður]]
 profession.N.FEM.GEN.SG flight.N.NEUT.STEM man.N.MASC.SG
 ‘professional pilot’
- III. [N.STEM [N.GEN. N]]
- g. [plast [hnífa par]]
 plastic.N.NEUT.STEM knife.N.MASC.GEN.PL pair.N.NEUT.SG
 ‘plastic cutlery’ (usually set of knife, fork & spoon)¹⁰
- h. [hör [vasa klútur]]
 linen.N.MASC.STEM pocket.N.MASC.GEN.SG cloth.N.MASC
 ‘linen handkerchief’

¹⁰ The spoon may be optional, but this is emphatically not a pair of two plastic knives, i.e., not [[*plast hnífa*] *par*].

- i. [grunn [fjár festing]]
 base.N.MASC.STEM capital.N.NEUT.GEN.SG fastening.N.FEM
 ‘basic investment’
- IV. [N.GEN.[N.GEN. N]]
- j. [biskups [skjala safn]]
 bishop.N.MASC.GEN.SG document.N.NEUT.GEN.PL collection.N.NEUT
 ‘archives of the bishop’
- k. [blúndu [vasa klútur]]
 lace.N.FEM.GEN.SG pocket.N.MASC.GEN.SG cloth.N.NEUT
 ‘lace handkerchief’
- l. [hernaðar [leyndar mál]]
 warfare.MASC.GEN.SG secret.N.FEM.GEN.SG matter.N.NEUT
 ‘military secret’

The examples in (11g–i) are critical in respect to theories with any kind of ordering of stem and genitive compounds. Imposing a left-branching structure on (11g) would change the meaning of *plasthnifapar* from ‘a set of knife and fork made from plastic’ to ‘a pair of knives ...’. Posing different structures for (11h) *hörvasa-klútur* ‘linen handkerchief’ and (11k) *blúnduvasaklútur* ‘lace handkerchief’ and the corresponding set of towels in (12a–b) seems semantically counterintuitive:

- (12) a. [hör [hand klæði]]
 linen.N.MASC.STEM hand.N.NEUT.STEM cloth.N.MASC
 ‘linen towel’
- b. [bómullar [hand klæði]]
 cotton.N.FEM.GEN.SG hand.N.NEUT.STEM cloth.N.MASC
 ‘cotton towel’
- c. [hör [vasa klútur]]
 linen.N.MASC.STEM pocket.N.MASC.GEN cloth.N.MASC
 ‘linen handkerchief’
- d. [blúndu [vasa klútur]]
 lace.N.FEM.GEN.SG pocket.N.MASC.GEN cloth.N.MASC
 ‘lace handkerchief’

An explanation based on the fact that *handklæði* and *vasaklútur* are lexicalized compounds will not suffice either, as fully productive compounds with these structures are easily made:¹¹

¹¹ These compounds are nonce formations. All nonce formations in this text are clearly marked as such.

- (13) a. [plast [penna dallur]]
 plastic.N.NEUT.STEM pen.N.MASC.GEN tub.N.MASC
 ‘plastic pen container’
- b. [postulíns [penna dallur]]
 porcelain.N.NEUT.GEN.SG pen.N.MASC.GEN tub.N.MASC
 ‘porcelain pen container’

The modifiers *plast* and *postulín* refer to the material of the container, not the pens stored in it.

2.5 Restriction of the use of stems as non-heads

Words from some inflectional classes can never appear as stems in compounds and there the genitive forms are always used. This applies to the so-called weak inflection of feminine and masculine nouns, e.g., feminine nouns ending in *-a* in the nominative singular, as in *olía* in *olíuverðshækkun* ‘a rise in the price of oil’ in (9c), and masculine nouns ending in *-i* in the nominative singular, as in *vasi* in *vasaklútur* ‘(pocket) handkerchief’ in (12). Words of this type are very numerous, as seen in the DMII which contains 27,381 non-compounds. Out of a total of 13,116 masculine and feminine nouns, 6,540 belong to the weak inflection, or just under 50%.

This fact should not be forgotten when the proportions of stem compounds and genitive compounds are considered, as the result is that a large proportion of the vocabulary is unavailable for stem compounds.¹² The consequences of this for any kind of ordering based on the difference of stems and inflected non-heads in compounds are unclear, but the option of specifying that half of the vocabulary is unavailable at any given level seems counter-intuitive.

2.6 Non-heads in compounds: Adjectives

Adjectives as non-heads of compounds exhibit similar variants as nouns do, i.e., stems (*lítill* ‘small’ in *lítilmenni* ‘insignificant character’ (A.b in 14) and genitives (*lítils* in *lítilsverður* ‘insignificant’ (B.c in 14)). Internal inflection is also found in adjectives as non-heads in compounds with nominal heads, where agreement of gender, case, and number is exactly the same within the compounds as in

¹² There are a few exceptions where the combining forms of weak masculine nouns are stems, e.g., *sím-* for *sími* ‘telephone’, e.g., *símhringing* ‘telephone call’, where *síma-* would be expected. These cases are extremely rare and most compounds with *sími* have the genitive non-head *síma*, e.g., *símasamband* ‘telephone connection’.

syntax, as in the nominative *litlifingur* ‘little finger, pinkie’, where the ending *-i-* in the non-head is a portmanteau adjectival ending for masculine, singular, nominative, definite, and the accusative *litlafingur*, where the ending *-a-* is a portmanteau adjectival ending for masculine, singular, accusative, definite, cf. (14C.). A comparison of agreement within a compound and in syntax is shown in Table 1.

(14) *Form of adjectives as non-heads of compounds*

A. Stem

a. blá ber (blár+ber)
 blue.ADJ.STEM berry.N.NEUT
 ‘blueberry’

b. lítil menni (lítill+-menni (*menni*=bound form))
 small.ADJ.STEM man.N.MASC
 ‘insignificant character’

B. Inflection, genitive (indefinite)

c. lítils verður (lítill+verður)
 small.ADJ.GEN.SG.INDEF worthy.ADJ
 ‘insignificant’

d. sjúkra hús (sjúkur+hús)
 sick.ADJ.GEN.PL.INDEF house.N.NEUT
 ‘hospital’

C. Inflection, internal¹³

1. Positive degree

e. litli fingur
 little.ADJ.MASC.DEF finger.N.MASC.INDEF
 (*lítill+fingur*; Acc. *litlafingur*)
 ‘pinkie, little finger’

f. Bratta brekka
 steep.ADJ.FEM.DEF hill.N.FEM
 (*brattur+brekka*; Acc. *Bröttubrekku*)
 ‘Steep Hill’ (placename)

¹³ Degree, as shown in the superlative *hæstiréttur* ‘supreme court’ in (14) (C)g, is not an instance of internal inflection but a contextual feature (Indriðason 2014: 21). Internal inflection in the comparative only appears in place names in Modern Icelandic and is not shown here.

2. Superlative

g. hæsti	réttur
highest.ADJ.MASC.SUP.DEF	court.N.MASC.INDEF
(hár+réttur; Acc. hæstarétt)	
‘supreme court’	

Definiteness is an inflectional feature of Icelandic adjectives (cf. Table 1). The genitives in B in (14) are indefinite forms, but adjectival non-heads in C in (14) are always definite, irrespective of the definiteness of the compound as a whole. For explanation, Table 1 contains the paradigms for the noun phrase *lítill fingur* ‘little finger’ in column 1 and 2, and the internal inflection for the compound *litlifingur* ‘pinkie’ in the compound in column 3.

Table 1: Paradigms for noun phrases and internal adjectival inflection

singular	indefinite	definite	compound
NOM.	lítill fingur	litli fingurinn	litlifingur
ACC.	lítinn fingur	litla fingurinn	litlafingur
DAT.	litlum fingri	litla fingrinum	litlafingri
GEN.	lítils fingurs	litla fingursins	litlafingurs
plural			
NOM.	litlir fingur	litlu fingurnir	litlufingur
ACC.	litla fingur	litlu fingurna	litlufingur
DAT.	litlum fingrum	litlu fingrunum	litlufingrum
GEN.	lítilla fingra	litlu fingranna	litlufingra

Note that the internal inflection in the compound in column 3 in Table 1 is identical to the definite inflectional form in column 2. This is in fact the case in all compounds of this type in the DMII, but the construction is not very common, except in place names. The form of the compound is indefinite, however, and the cliticized definite article can be attached, in the same manner as in other nouns, as seen in the examples in (15):

- (15) a. Hann braut litlafingur, held ég.
 he broke littlefinger.N.MASC.ACC.SG.INDEF, think I
 ‘He broke a pinkie, I think.’

- b. Hann braut á sér litlafingurinn
 he broke on himself (the) littlefinger.N.MASC.ACC.SG.DEF
 ‘He broke his pinkie.’
- c. Litlifingurinn brotnaði.
 the-littlefinger.N.MASC.NOM.SG.DEF broke
 ‘The pinkie broke/was broken.’

The genitive constructions with adjectival non-heads have direct counterparts in PCs, in the same manner as nouns. They are, however, quite rare, cf. §3.1.2.

2.7 The relevant features of non-phrasal compounding for PCs

The salient points in this section in connection with the PCs discussed in the next section are these:

- Genitive non-heads are one of two basic options in forming Icelandic non-phrasal compounds. The other main option is to have non-head stems. Genitive non-heads are also found in PCs, as will be discussed in §3.
- The distribution of genitives and stems as non-heads in compounds is partly dependent on the inflectional class of the non-head, as masculine and feminine words from the so-called weak inflection cannot appear as stems in compounds, with exceptions mentioned in Footnote 13. Right-branching compounds with genitive non-heads in a lower node than stem non-heads are quite common. Therefore, it is difficult to maintain strict ordering of stems and genitives as non-heads of compounds for Icelandic.
- The inflected non-heads are in accordance with the “correct” inflection of the respected unbound forms. This also applies in PCs.
- The internal inflection of adjectival non-heads could perhaps be analysed as a phrase-to-word conversion. §3.1.4 contains PCs with prepositional phrases which could also be analysed as phrase to word conversion, as could some of the PCII structures in §3.2, cf. also Footnote 16. The process of phrase to word conversion (or nominalization) will not be discussed in any detail, as the necessary research is not available.

3 Phrasal compounds

Below, Icelandic PCs are divided into two groups. The first group (§3.1) contains common words which are not stylistically marked in any way, some of which are attested from medieval times to the present day. This group of PCs contains genitive noun phrases and prepositional phrases as modifiers of nouns, as in Table 2. (Examples of all constructions are given in the following sections.)

Table 2: Phrasal Compounds I, from lexicographic sources

a.	NP.GEN. + N	Phrase internal agreement
b.	AdjP.GEN. + N	Phrase internal agreement (rare)
c.	PP + N	Case assignment by preposition

The second group contains more complex PCs, mainly found in informal writing and in speech. The structures are variable, up to full main clauses. The evidence for some of the structures is weak, down to single examples. The classification in §3.2 reflects this. It should be noted that the more traditional PC types shown in Table 2 also appear in the more informal texts used as sources for PCIIs.

Table 3: Phrasal Compounds II, from the web, etc.

a.	[N.NOMINATIVE + PP]NP + N	Case assignment by preposition
b.	Miscellaneous non-predicates:	AdjP, AdvP, negation, etc.
c.	Predicates:	Imperatives, questions, finite S, etc.

As the second type of PCs is very much a feature of informal speech and text, the spelling tends to be varied. In fact, Icelandic spelling rules do not include any indication of the correct form in these cases.¹⁴ The PCs are therefore a free-for-all in Icelandic spelling, which makes them very difficult to extract automatically from text. The examples in (16) show spelling variations with different quotation marks, hyphenation, and spaces, found in data from a corpus of Icelandic websites, *Íslenskur orðasjóður* (Hallsteinsdóttir et al. 2007):

¹⁴ The only indication is the spelling of compounds containing multiword first parts of foreign origin, such as the translation of *New York City*, i.e., *New York-borg*, where the spelling rules place a hyphen before the compound head *borg* ‘city’. The space in *New York* from the English original is maintained. Judging by all the mistakes made, this spelling rule seems to be hard to learn, and extending it to phrasal compounds seems to be counter-intuitive as spelling as in (16f) (*munni við munn-aðferð*) is hardly ever found.

- (16) a. „*allt eða ekkert*“ *aðferðin* ‘the all or nothing method’
 b. “*allt eða ekkert*” *dæmi* ‘[an] all or nothing example’
 c. ‘*allt eða ekkert*’ *týpa* ‘[an] all or nothing type’, i.e., ‘guy’
 d. *allt eða ekkert dæmi* ‘[an] all or nothing example’
 e. *munni-við-munn-öndun* ‘[a] mouth to mouth breathing’
 f. *munni við munn-aðferð* ‘[a] mouth to mouth method’
 g. *allt-eða-ekkert hugsunarháttur* ‘[an] all or nothing way of thinking’

The possible spelling varieties are not exhausted in this search, but at present, tools for an automatic search do not exist. No attempt is made here to normalize the spelling in these examples, resulting in strange quotation marks at times.

The PCs discussed here are a subset from a collection of over 200,000 compounds compiled by the author over a period of over 30 years. The sources are mostly the same as those for the DMII mentioned above (Bjarnadóttir 1996[2005]; 2012), and the analysis contains lemmatization and full analysis of the constituents of the compounds. This resource returned ca. 700 PCs, almost all of which are PCIs. In addition, ca. 200 PCs from the web, from blogs, social media, radio, and TV, were collected from *Íslenskur orðasjóður*, and from miscellaneous sources, personal communication, etc. Finding data for PCIs turned out to be difficult, because of unstandardized spelling. The remainder of this section contains a classification of these 900 PCs.

3.1 Phrasal compounds I

3.1.1 Genitive noun phrase and nominal head

Genitive noun phrases with adjectives are common in any genre as non-heads of compounds. There is full agreement of gender, case and number within the noun phrases, as in *Góðrarvonarhöfði* ‘Cape of Good Hope’ (17c). where the adjective *góður* ‘good’ agrees in gender with the feminine noun *von* ‘hope’, and both agree in case (gen.), and number (sg.). The head *höfði* ‘cape’ is a masculine noun.

- (17) a. [hálfsmánaðar]NP blað
 half.ADJ.MASC.GEN.SG month.N.MASC.GEN.SG paper.N.NEUT.SG
 ‘biweekly journal’
 b. [heilsárs]PP dekk
 whole.ADJ.NEUT.GEN.SG year.N.NEUT.GEN.SG tyre.N.NEUT.SG
 ‘all year tyre’

- c. [Góðrar vonar]NP höfði
 good.ADJ.FEM.GEN.SG hope.N.FEM.GEN.SG cape.N.MASC.SG
 ‘Cape of Good Hope’
- d. [allra sálna]NP messa
 all.ADJ.FEM.GEN.PL soul.N.FEM.GEN.PL mass.N.NEUT.SG
 ‘All Souls’ Day’ Nov. 2nd

PCs of this type are found in Old Icelandic, as in *allramannagisting* ‘all men’s night lodging’ and *allralandamaður* ‘all countries’ man’ (AM 132 fol., AD c1300–1350, cf. ONP). These two PCs do not appear to be lexicalized as an entity, as the head can easily be changed as in the nonce formation *allramannalygi* ‘all men’s lies’ (nonce formation by Jóhannes Bjarni Sigtryggsson). The PCs in (17) are lexicalized, with the possible exception of (17a), and the stress pattern of unlexicalized PCs of this type needs to be investigated as there is a tendency to split them apart in writing.

3.1.2 Genitive adjectival phrase and nominal head

PCs with adjectival phrases as heads are rare. There is agreement for case and number in the example below, but gender is indistinct in the genitive plural:

- (18) [allra heilagra]ADJP messa
 all.ADJ.GEN.PL holie.ADJ.GEN.PL mass.N.FEM.SG
 ‘All Saints’ Day’ Nov. 1st

The construction is similar to *allrasálmamessa* (17d) above. The PC in (18) is also found in Old Icelandic (GKS 1812 4°, cf. ONP), along with variants, e.g., *allraheilagradagur* ‘All Saints’ Day’, *allraheilagrahátíð* ‘All Saints’ Feast’.

3.1.3 Noun phrase with numeral and nominal head

The cardinal numbers 1–4 inflect for gender and case in Icelandic, and these appear in PCs with the same construction as the adjectives in §3.1.1 There is full agreement of numeral and noun for case and number. Gender is distinguished in the singular, but the genitive plural is the same for numerals in all genders, as in adjectives.

- (19) a. [eins manns]NP herbergi
 one.NUM.MASC.GEN.SG man.N.MASC.GEN.SG room.N.NEUT.SG
 ‘single room’

- b. [tveggja manna]NP herbergi
 two.NUM.MASC.GEN.PL men.N.MASC.GEN.PL room
 ‘double room’
- c. [fimm ára]NP áætlun
 five.NUM year.N.NEUT.GEN.PL plan.N.FEM.SG
 ‘five year plan’
- d. [sex liða]NP háttur
 six.NUM parts’.N.MASC.GEN.PL meter.N.MASC.SG
 ‘hexameter’

The phrases in PCs in (19) are fully transparent and not lexicalized, as can be demonstrated by the free replacement of the numerals, cf. (19b–d):

- (20) a. [nín manna]NP herbergi
 nine.NUM men.N.MASC.GEN.PL room.N.NEUT.SG
 ‘a room for nine’
- b. [þriggja ára]NP áætlun
 three.NUM year.N.NEUT.GEN.PL plan.N.FEM.SG
 ‘three years’ plan’
- c. [sjö og hálf árs]NP áætlun
 seven.NUM and half.NEUT.GEN.SG year.N.NEUT.GEN.SG plan.N.FEM.SG
 ‘seven and a half year plan’
- d. [ellefu liða]NP háttur
 eleven.NUM part.N.MASC.GEN.PL meter.N.MASC.SG
 ‘a (hypothetical) meter with 11 parts’

A similar construction in German does not exhibit this agreement, and is in fact used as an argument against a phrasal analysis, as in Pafel (2015) on words like *Zweibettzimmer*, partly because “their parts do not agree as the parts of the corresponding phrase would do”. Icelandic PCs containing NPs with numerals always show agreement, as do those with adjectives. They would therefore seem to point to a different conclusion from Pafel’s and be considered true PCs and not ‘pseudo-phrasal’ compounds like the German construction.

The agreement of numeral and noun within PCs obeys the same rules as in syntax, as can be seen in the following examples of PCs and corresponding sentences:

- (21) a. [einnar hæðar]NP skýjakljúfur
 one.NUM.FEM.GEN.SG storey.N.FEM.GEN.SG skyscraper.N.MASC.SG
 Ég ætla að fá einn [súpu disk]
 I FUT have one.NUM.MASC.SG soup.N.FEM.GEN.SG plate.N.MASC.SG
 ‘I’ll have one plate of soup’
- b. [tveggja hæða]NP hús
 two.NUM.FEM.GEN.PL storey.N.FEM.GEN.PL house.N.NEUT.SG
 Ég ætla að fá tvo [súpu diska]
 I FUT have two.NUM.MASC.PL soup.N.FEM.GEN.SG plate.N.MASC.PL
 ‘I’ll have two plates of soup’
- c. [tuttugu-og-einnar hæðar]NP
 twenty-one.NUM.FEM.GEN.SG storey.N.FEM.GEN.SG
 skýjakljúfur
 skyscraper.N.MASC.SG
 Ég ætla að fá [tuttugu_og_einn] [súpu
 I FUT have twenty-one.num.masc.sg soup.N.FEM.GEN.SG
 disk]
 plate.N.MASC.SG
 ‘I’ll have twenty-one plates of soup’

The peculiarity of the agreement with the last part of the numeral is always observed, i.e., any number that ends in *einn* ‘one’ takes the singular, irrespective of it being 1, 21 or 1001, both in syntax and within compounds: *Ég var að lesa Þúsund og eina nótt*SG. ‘I’ve been reading *Thousand and One Night[s]*’.

3.1.4 Prepositional phrase and nominal head

The most common type of PCs in Icelandic contains a prepositional phrase as a non-head. The prepositions occurring in these PCs govern the genitive, and case assigned by the preposition is always maintained. The stress pattern is regular, as shown in Figure 1. Most of these PCs are easily rephrased as sentences, as in *milliríkjasamningur* (cf. translation in 22a) vs. *samningur milli ríkja* ‘a contract between states’.

- (22) a. [milli ríkja]PP samningur
 between.PREP state.N.NEUT.GEN.PL contract.N.MASC.SG
 ‘international agreement’

- b. [milli landa]PP flug
 between.PREP country.N.NEUT.GEN.PL flight.N.MASC.SG
 ‘international air transport’
- c. [innan dyra]PP þægindi
 inside.PREP door.N.FEM.GEN.PL conveniences.N.NEUT.PL
 ‘indoor conveniences’
- d. [innan lands]PP markaður
 inside.PREP country.N.NEUT.GEN.SG market.N.MASC
 ‘domestic market’
- e. [neðan jarðar]PP lest
 below.PREP earth.N.FEM.GEN.SG train.N.FEM
 ‘subway, underground train’

Prepositional phrases also seem to be converted to adverbials or adjectives (Bjarnadóttir 1996[2005]), as in *innanhúss* ‘indoors’ shown successively in (23a, b, c) as an adverb, an adjective and a full prepositional phrase with the definite article. This type of word formation in Icelandic is a neglected field, but it is very common.¹⁵

- (23) a. mótaröð í frjálsum innanhúss
 series.of.events in free.ADJ.DAT.PL indoors.ADV
 ‘a series of track and field events indoors’ (frjálсар.ADJ.FEM.PL ‘track and field’)
- b. innanhúss knattspyrnskór
 indoors.ADJ soccer.shoes.N.MASC.PL.INDEF
 ‘indoors soccer shoes’
- c. Hópurinn starfar innan.PREP hússins.N.NEUT.GEN.DEF
 ‘The group works inside the house’

The compound *innanhúss* in (23a) is not lexicalized, in the lexicographer’s sense of the meaning being different from the sum of the parts (see Svensén 1993: 42), as demonstrated by nonce compounds such as *innanbókarvísun* ‘inside

¹⁵ The analysis of phrase to word conversion seems to be obvious, and the phenomenon is supported by words like the verb *svei-mér-þá-a* ‘shame.INTERJ.me.PRON.DAT.now.ADV, with infinitival ending -a, as in *Hann svei-mér-þá-aði sér duglega* ‘He said “shame-on-me” with gusto’’. (Bjarnadóttir 1996[2005]). There are not many such cases; *gleym-mér-ei* ‘forget me not’ is probably the most common.

(a) book citation’, which are easily formed.¹⁶ They cannot be modified in the same manner as the sentence in (23c), e.g., **innanbókarinnarvísun* ‘inside-the-book citation’.

This type of PC is quite common in Modern Icelandic, and the construction also exists in Old Icelandic. The word *innanffórðungsmaður* ‘inside the quarter man’ (i.e., ‘an inhabitant of a district (quarter)’) appears in *Grágás* ‘The Gray Goose Law’ (GKS 1157 fol., AD 1260?). The modern term for ‘vagrant’ is the PC *utangarðsmaður* ‘outside garden/wall man’ first attested in a Norwegian diploma in AD 1300 (AM dipl norv facs I 12).¹⁷

3.1.5 Prepositional phrase and adjectival head

PCs with adjectival heads are held to be marginal (Meibauer 2007: 237) or not in accordance with the properties of PCs in Germanic languages, as in Trips (2016: 153).

The adjective *utanríkispólitískur* ‘of foreign politics’ in (24b) is a PCI with a possible adjectival head found in the DMII but not present in the data used for this study. Google returns 58 examples of this PC, from the media and the website of Alþingi, the Icelandic Parliament. (For comparison, about 1,400 instances of the adjective *flokkspólitískur* ‘party political’ are found on Google.) Google returns about 4,140 instances of the corresponding noun *utanríkispólitík* ‘foreign politics’, which has the structure in (24a) (cf. §3.1.4, 22). The parallel analysis of the adjective is shown in (24b). If the possibility of recursive compounding and derivation is considered allowable, the analysis in (24c) is the result, deriving the adjective from the PC.

- (24) a. [utan ríkis]PP pólitík
 outside.PREP state.N.NEUT.GEN.SG politics.N.FEM
 ‘(the) politics of foreign affairs’
- b. [utan ríkis]PP pólitískur
 outside.PREP state.N.NEUT.GEN.SG political.ADJ
 ‘pertaining to foreign politics’
 (as in *utanríkispólitískur veruleiki* ‘(the) reality of foreign politics’)
- c. [[utan ríkis]PP pólitík] +skur
 outside.PREP state.N.NEUT.GEN.SG politics.N.FEM +al.ADJ.SUFF

¹⁶ Spontaneous creation by Jóhannes Bjarni Sigtryggsson referring to citation in the present volume, Feb. 12th 2016.

¹⁷ All examples from Old Icelandic are derived from *The Dictionary of Old Norse Prose* (ONP).

Recursive compounding and derivation have been proposed for Icelandic (Bjarnadóttir 1996[2005]) which solves issues of bracketing paradoxes that are otherwise common for Icelandic, but are of course not compatible with most current models of the language architecture. PCs with adjectival heads are certainly very rare in Icelandic, but two more PCIs are shown in (29) in §3.2.2.

3.2 Phrasal compounds II

Informal speech and texts provide examples of constructions of PCs not found in other genres. These constructions range from types of noun phrases not described in the previous section to full predicate phrases, such as *ég-verð-að-vita-hvað-gerist-næst-bók* ‘I must (to) know what happens next book’. In this section the PCs are divided into non-predicative (§3.2.1–2) and predicative PCs (§3.2.3). The classification in this section is partly based on Carola Trips’ analysis of English PCs (Trips 2016).

These PCs can be humorous, as in *ég-er-bara-einn-af-ykkur-strákunum-brosið* ‘“I’m just one of you boys” smile’, but that is certainly not always the case. The head of the Icelandic Confederation of Labour certainly did not have anything humorous in mind when he used the word *ef-og-þá-kannski-hlutir* ‘if and then maybe things’ of vague offers in negotiations, on the brink of a general strike.¹⁸ The PCs are very often spontaneous ad hoc constructs, but occasionally they do catch on and become a part of everyday language, sometimes as a part of the jargon within small groups as when linguists in Iceland refer to chapters on future work as *gaman-væri-að-kaflinn* ‘the “it would be fun to” chapter’.

The phrasal non-heads in the PCs are not necessarily lexicalized, at least not in the lexicographer’s sense of their meaning being greater than or different from the sum of the parts. They can, for instance, be used to describe any kind of attitude, such as *ég-er-svo-glöð-svipurinn* ‘the “I’m so happy” expression’ and *oj-barasta-hvað-þetta-er-leiðinlegt-mómentið* ‘the “ugh how boring this is” moment’. The semantics of these PCs would be an interesting topic for research, but as yet the Icelandic data is too scarce to warrant further speculation.

3.2.1 Nominative noun phrase and nominal head

The non-head of this construction is a noun phrase in the nominative, and as such a novel feature in Icelandic compounding although the dating of it is difficult for reasons of spelling and lack of analysis of older texts. The noun is followed by a prepositional phrase.

¹⁸ Ásmundur Stefánsson, newscast on Icelandic Radio in the Nineties (Bjarnadóttir 1996[2005]).

- (25) a. [maður -á -mann] aðferð
 man.N.MASC.NOM.SG to.PREP man.N.MASC.ACC.SG method.N.FEM.SG
 ‘man-to-man method’
- b. [skref -fyrir -skref]
 step.N.NEUT.NOM.SG for.PREP step.N.NEUT.ACC.SG
 leiðbeiningar
 instructions.N.FEM.PL
 ‘step-by-step instructions’
- c. [poki -í -öskju] kerfi
 bag.N.MASC.NOM.SG in.PREP box.N.MASC.DAT.SG system.N.NEUT.SG
 ‘bag in a box system’
- d. [korter -í -þrjú] -náungi
 quarter.N.NEUT.NOM.SG to.PREP three.NUM.ACC guy.N.MASC.SG
 ‘a quarter to three guy’¹⁹

The nominalization of the non-head NPs in these PCs seems to be a possibility. This is supported by an anecdotal example from a fellow linguist quoting his young daughter, where the definite article is cliticized onto a noun phrase as a whole, e.g., *bland í poka* ‘mixture in a bag’ (of sweets bought by weight) becomes *bland-í-pokað mitt* ‘my the “mix-in-a-bag”’:²⁰

- (26) a. [bland -í -poka]NP -ð
 mix.N.NEUT.NOM.SG in.PREP bag.N.NEUT.DAT.SG the.DEF.ART.NOM.SG
 mitt
 my.POSS.NEUT.NOM.SG
 ‘my mixture in a bag’

Similar PCs containing foreign noun phrases, mostly English, are easily found (27a–c), as are constructions containing adjectival or even verb phrases (27d–e):

- (27) a. *coast-to-coast skautahlaup* “coast to coast” ice skating’
 b. *step-by-step bók* ‘a “step by step” book’
 c. *point-in-time afritun* ‘a “point in time” backup’

¹⁹ A ‘quarter to three guy’ refers to the now expired closing hours of Icelandic bars, indicating a certain desperation. In the data used here, the head is *náungi*, but the synonym *gæi* is more frequent.

²⁰ Jón Hilmar Jónsson, The Árni Magnússon Institute for Icelandic Studies, personal communication.

- d. *all-in-one prentari* ‘an “all in one” printer’
 e. *cut-to-fit skjákort* ‘a “cut to fit” graphics adapter’

Needless to say, there is no agreement in the foreign phrases in the PC loanwords, but similar Icelandic PCs do exist, perhaps modelled on the loanwords. These show rather interesting agreement, as can be seen in the examples in (28a–b). The preposition *í* ‘in’ governs the dative of *einni/einu* ‘one’ in the non-heads of the PCs *allt-í-einni-tölva* ‘all in one computer’ and *allt-í-einu-tæki* ‘all in one tool’, but the gender is in agreement with the head of the PCs, the feminine *tölva* ‘computer’ in (28a) and the neuter *tæki* ‘tool, instrument’ in (28b).

- (28) a. [allt -í -einni] -tölva
 all.NEUT.NOM.SG in.PREP one.NUM.FEM.DAT.SG computer.N.FEM.NOM.SG
 ‘all in one computer’
 b. [allt -í -einu] tæki
 all.NEUT.NOM.SG in.PREP one.NUM.NEUT.DAT.SG tool.N.NEUT.NOM.SG
 ‘all in one tool’

Because of spelling issues, these PCs are elusive in texts, cf. (16) above.

3.2.2 Miscellaneous non-predicates

The remainder of the data for PCs discussed here contains a miscellany of words that are listed here for completeness, but the data is so scarce that any analysis is bound to be inconclusive. The non-heads of these PCs are seen to be adjectival phrases, with or without negation, and adverbial phrases. All the phrases contain full syntactic agreement; they are lifted straight from syntax and attached in front of nominal or adjectival heads, or “stuck in front of these words” in the rather informal wording straight from the mouth of a non-linguist.

- (29) I. AdjP + N:
 a. [óðýrari -en -í -Frakklandi] [net
 cheaper.ADJ than in France.N.NEUT.DAT Internet
 tenging]
 connection.N.FEM
 ‘A “cheaper than in France” Internet connection’

II. AdjP +N, with negation:

- b. [ekki- ofur- frjálsleg] -kristni
 not super free-like.ADJ.FEM.INDEF Christianity.N.FEM.INDEF
 ‘“not super liberal” Christianity’
- c. ekki- svo- fjarlæg framtíð
 not so distant.ADJ.FEM.INDEF future.N.FEM.INDEF
 ‘A “not so distant” future’
- d. ekki- svo- dapurlegi -dagurinn
 not so sad.ADJ.MASC.DEF day.N.MASC.DEF
 ‘the “not so sad” day’

III. AdvP/PP + Adj:

- e. klukkan -tíu -á -laugardegi -skemmtilegur
 clock ten on Saturday.N.MASC.DAT amusing.ADJ
 ‘“ten o’clock on a Saturday” amusing’
- f. inn-á- hvert- einasta -heimili -s
 [into.PREP [every- one].NEUT.ACC.SG home.N.NEUT ‘s’.GEN]PP
 -frægur
 famous.ADJ
 ‘famous in every single home’

Two of the examples of PCIIs, *klukkan-tíu-á-laugardegi-skemmtilegur* ‘“ten o’clock on a Saturday” amusing’ (29e), and *inn-á-hvert-einasta-heimilisfrægur*²¹ ‘in every single home famous’ (29f), have adjectives as heads and thus contravene one of the properties in Trips (2016), where Germanic PCs are assumed to have only nominal heads (p.154). According to Meibauer (2007: 236–237), adjectival heads in PCs are marginal, as they seem to be in Icelandic where very few examples have been found (cf. §3.1.2 for an example of a PCI adjective). Speakers do seem to accept the PCIs above to the same degree as the other PCIIs in (29). More data is needed to establish the status of PC adjectives; as of now they seem to be as marginal in Icelandic as Meibauer found them to be in German.

²¹ The PC *inn-á-hvert-einasta-heimilis-frægur* contains an unexpected -s- (marked with an * in (29f). This -s- is the correct genitive singular ending for the neuter noun *heimili*, which is out of place in this PC as the preposition takes the accusative. It could possibly be a linking phoneme, as -s- can be (Bjarnadóttir 1996[2005]). This PC is remarkable as it is the only example found to date of a prepositional phrase PC where the preposition does not take the genitive.

3.2.3 Predicates

The last group of Icelandic PCs listed here are verbal predicates, cf. Trips (2016) for similar constructions in English. As in the previous groups in this section, the data is scarce. The examples in this section are divided into four groups, i.e., imperatives, infinite, finite sentences, and questions. These PCs are generally found in blogs, and they are very spontaneous, easily understood, and considered to be more or less odd, incorrect, or at least very strange. These are attested examples, however, and as such seem to be within the capacity of the users, even if the selfsame users often treat them as jokes. It should be noted that the imperative in Icelandic can contain a subject pronoun cliticized onto the verbal form, i.e., *ruglaðu* (*rugla.IMP þú2.PERS.PRON*) Thus *ruglaðu mig*, lit. ‘confuse you me’.

(30) Imperative (directives)

- a. [haltu kjafti brjóstsykur
hold+you.IMP mouth.N.MASC.DAT.SG candy.N.MASC
“hold your mouth” candy’, i.e., “shut up” candy (because of size)
- b. [ruglaðu -mig- í- hausnum] -mynd
confuse+you.IMP me in head.DEF movie
‘a “make me confused” movie’
- c. [gerðu- það- sjálfur] -tónlist
do+you.IMP it yourself music
“do it yourself” music’
- d. [gettu -betur-] liðið
guess+you.IMP better.ADV.COMP team.N.NEUT.DEF
‘the “guess better” team’ (a quiz team)
- e. [‘skrifaðu- í- gestabókina- eða- ég- kýli- þig- í- andlitið’] dagur
write+you in guest+book.DEF or I punch you in face.DEF day
‘a “write in the guestbook or I’ll punch you in the face” day’

One of the imperative PCs above is a common word, *haltukjafitibrjóstsykur* ‘“shut up” candy’ (30a), used of a really large mouthful of candy that makes speech difficult. The parallel *haltukjafitityggjó* ‘“shut up” chewing gum’ also exists. *Gettu-betur-liðið* ‘the “guess better” team’ (30d) is well known and used of the quiz teams taking part in a very popular television program called *Gettu betur* ‘Guess better’. The other imperative PCs are nonce formations, and get gradually stranger the longer they are. The Icelandic PC imperatives have direct

counterparts in English, according to Carola Trips' analysis (2016:160) of English directives.

The infinitive in (31) is awkward and included here for the sake of completeness, as other examples of this type of PC have not been found.

- (31) Infinitive
sprengja- í- loft- upp lemja- vondu- kalla -stórmynd
explode.INF in air up hit.INF bad guys'.NP.ACC.PL.INDEF big-movie
'an "explode into air hit bad guys" blockbuster movie'

The PC in (31) could be a directive, as in *koma svo!* 'come on now', in which case it would have the meaning "'let's explode into the air & hit the bad guys" blockbuster movie'. The use of the infinitive with the function of a directive needs to be investigated. If the analysis holds up, the counterpart is found in Carola Trips' directives with *let* (2016:160).

As in Trips (2016: 159–160) for English, Icelandic PCs may have full clauses and questions as non-heads:

- (32) I. Question
a. hvers-vegna- ekki -samfélagið
why not society.N.NEUT.SG,DEF
'the "why not" society'
- II. Finite S
b. ég- verð- að- vita- hvað- gerist -næst -bók
I must to know what happens next book.N.FEM.SG.INDEF
'An "I must know what happens next" book'
- c. ég- þoli- ekki- illar- tölvur -klúbburinn
I tolerate not bad computers club.N.MASC.SG.DEF
'the "I cannot bear evil computers" club'
- d. ég- rek- þau- í gegnum- þig- horn
I stick them through you horns.N.NEUT.PL.INDEF
'the "I will stick through you" horns'
- e. allt- er- list tímabilið
all is art [time-span]N.NEUT.SG.DEF
'the "everything is art" period'
- f. ég- er- bara- einn- af- ykkur- strákuðum- brosið
I am just one of you boys.N.MASC.DEF smile.N.NEUT.DEF
'the "I am just one of you boys" smile'

All the examples in (32) are nonce formations. These constructions are not common, but they pose no burden of interpretation on native speaker recipient readers or listeners. More data would be helpful here, but as the spelling is completely unpredictable better tools are needed for automatic extraction.

4 Conclusion

In this study, PCs in Icelandic are divided into two groups; those that have formed a part of Icelandic since medieval times (PCIs), and constructions not attested until the 20th century (PCIIs). The first group, PCIs, contains only genitives or prepositional phrases with prepositions taking the genitive as non-heads, but the non-heads in the second group, PCIIs, vary widely, from noun phrases to full finite sentences. The one aspect all the PCs here have in common is that they contain full agreement and case assignment, with the exception of some PCs with foreign non-heads, as “*step-by-step*” *bók* “‘step by step’ book’. The Icelandic PCs do, for example, exhibit full genitive agreement in the non-heads of PCs such as *tveggjamannaherbergi* ‘room for two’, which is not the case in the German counterpart, *Zweibettzimmer*, according to Pafel (2015) (cf. §3.1.3).

The classification of the more complex PCIIs is in line with Carola Trips’ analysis of English PCs, and all the PCII constructions here have a counterpart in her analysis of English. Due to scarcity of the Icelandic data, some of the English PC constructions are not found, despite seeming quite plausible. This could be the result of problems with spelling, which makes searching difficult. The only type of Icelandic PCII that seems to be unusual, in comparison with the English data, are PCs with adjectival heads, as in *klukkan-tíu-á-laugardegi-skemmtilegur* “‘ten o’clock on a Saturday’ amusing’ (cf. §3.2.2), and *utanríkisþólitískur* ‘outside state political’ (cf. §3.1.5).

As seen in §2, inflected word forms are also a feature of very many non-phrasal compounds in Icelandic. It is maintained here that non-head genitives and stems cannot be assumed to belong to specified layers of the morphology. Semantically, attempts to analyse them in this way result in bracketing paradoxes, as in *hörvasa-klútur* ‘linen pocket cloth’, i.e., [N.STEM [N.GEN.PL N]] ‘linen handkerchief’ (the correct meaning) and [[N.STEM N]GEN.PL N] ‘cloth in a linen pocket’. Splitting the analysis of morphological structure and semantics also seems counter-intuitive in the light of the fact that the choice of stems or genitives is to a large degree determined by the inflectional class of words. This has a bearing on the PCs. The traditional PCs (PCIs) always contain a genitive, either as a head of a non-head noun phrase, or as a part of a prepositional phrase. If there is no fun-

damental difference between stem and genitive non-heads in non-phrasal compounds, where do the genitive phrasal non-heads fit in? This question remains unresolved.

Although the more complex PCs are sometimes deemed outlandish, some PCs are totally acceptable on all levels to native speakers of Icelandic. This applies to the structures classified as PCIs in this study, i.e., the traditional ones. Words of this type are acceptable to the degree that very few people actually realize that they are structurally different from simpler structures, as in *milliþinganefnd* ‘between sessions committee’ [PP + N] and *efnahagsnefnd* ‘financial matters committee’ [[N.GEN N].GEN N]. Lexicalization does not seem to be a crucial point either, as most of the structures seem to be freely available for new formations, including nonce formations.

Now to the *gaman-væri-að-kaflinn* ‘the “would-be-fun-to-do” section’, i.e., thoughts on future work. This study is focussed on data. This data, however, is quite limited and should be expanded. Work on a large corpus of Modern Icelandic has just started at The Árni Magnússon Institute of Icelandic Studies, and that should produce extensive data, which will be analysed with a new version of our automatic compound splitter producing binary trees with full analysis of all constituents (cf. Daðason & Bjarnadóttir 2014). As Icelandic word formation exhibits interesting interaction between syntax and morphology, presenting this data seems a very worthwhile undertaking.

Acknowledgements

I would like to thank all the participants in the workshop on phrasal compounds in Mannheim on June 26th 2015, and especially the organizers Carola Trips and Jaklin Kornfilt for a most inspiring event. Thanks are also due to my colleagues, Jóhannes Bjarni Sigtryggsson, Jón Friðrik Daðason, Jón Hilmar Jónsson and Kristján Rúnarsson.

Abbreviations

ACC	accusative	ART	article
ADJ	adjective	DAT	dative
AdjP	adjectival phrase	DEF	definite
ADV	adverb	DMII	The Database of Modern Icelandic Inflection
AdvP	adverbial phrase		

FEM	feminine	ONP	The Dictionary of Old Norse Prose
GEN	genitive	PC	phrasal compound
IMP	imperative	PERS	person
INDEF	indefinite	PL	plural
INF	infinitive	POSS	possessive pronoun
MASC	masculine	PP	prepositional phrase
N	noun	PREP	preposition
NEUT	neuter	SG	singular
NOM	nominative	SUP	superior degree
NP	noun phrase	V	verb
NUM	numeral		

References

- Árnason, Kristján. 2011. *The phonology of Icelandic and Faroese*. Oxford: Oxford University Press.
- Bjarnadóttir, Kristín. 1995. Lexicalization and the selection of compounds for a Bilingual Icelandic dictionary base. In *Nordiske studier i leksikografi*, vol. 3, 255–263.
- Bjarnadóttir, Kristín. 2012. The database of modern Icelandic inflection. In *LREC 2012 Proceedings: Proceedings of “Language Technology for Normalization of Less-Resourced Languages”*, SaLTMiL 8 – AfLaT, 13–18.
- Bjarnadóttir, Kristín. 1996[2005]. *Afleiðsla og samsetning í generatífri málfræði og greining á íslenskum gögnum*. Reykjavík.
- Bjarnadóttir, Kristín. 2002. *A Short Description of Icelandic Compounds*. The Website of The Institute of Lexicography. <http://notendur.hi.is/~kristinb/comp-short.pdf>, accessed 2002-04-01.
- Booij, Geert E. 1994. Against split morphology. *Yearbook of Morphology* 1993. 27–50.
- Botha, Rudolf P. 1981. A base rule theory of Afrikaans synthetic compounds. In Michael Moortgat, Harry van der Hulst & Teun Hoekstra (eds.), *The scope of lexical rules*, 1–77. Dordrecht: Foris.
- Daðason, Jón Friðrik & Kristín Bjarnadóttir. 2014. Utilizing constituent structure for compound analysis. *LREC 2014 Proceedings*. 1637–1641.
- Hallsteinsdóttir, E., T. Eckart, C. Biemann, U. Quasthoff & M Richter. 2007. *Íslenskur Orðasjóður – Building a Large Icelandic Corpus*. *Proceedings of NODALIDA-07, Tartu, Estonia*.

- Indriðason, Þorsteinn G. 1999. Um eignarfallssamsetningar og aðrar samsetningar í íslensku. *Íslenskt mál* 21. 107–150.
- Indriðason, Þorsteinn G. 2014. Fallbeygðir fyrri liðir og tvær kenningar um orðhlutafræði. *Íslenskt mál* 36. 9–30.
- Indriðason, Þorsteinn G. 1994. *Regluvirkni í orðasafni og utan þess. Um lexíkalska hljóðkerfisfræði íslensku*. Reykjavík.
- Jónsson, Baldur. 1984. Samsett orð með samsetta liði. Fáeinar athuganir. 158–174.
- Lieber, Rochelle. 1989. On percolation. In Geert Booij & Jaap van Marle (eds.), *Yearbook of morphology 2*, 95–138. Heidelberg: Springer.
- Meibauer, Jörg. 2007. How marginal are phrasal compounds? Generalized insertion, expressivity, and I/Q-interaction. *Morphology* 17. 233–259.
- Pafel, Jürgen. 2015. Phrasal compounds are compatible with Lexical Integrity. *Language Typology and Universals* 68. 263–280.
- Perlmutter, David. 1988. The split morphology hypothesis: Evidence from Yiddish. In Michael Hammond & Michael Noonan (eds.), *Theoretical morphology*. 77–99. San Diego: Academic Press.
- Rask, Rasmus Christian. 1811. *Vejledningen til det islandske eller gamle nordiske sprog*. Copenhagen: Schuboth.
- Rögnvaldsson, Eiríkur. 1986. *Íslensk orðhlutafræði: Kennslukver handa nemendum á háskólastigi*. Reykjavík: Málvísindastofnun Háskóla Íslands.
- Snædal, Magnús. 1992. Hve langt má orðið vera? *Íslenskt mál* 14. 173–207.
- Svensén, Bo. 1993. *Practical lexicography. Principles and methods of Dictionary-Making*. Oxford: Oxford University Press.
- Thráinsson, Höskuldur. 2007. *The syntax of Icelandic* (Cambridge Syntax Guides). Cambridge: Cambridge University Press.
- Thráinsson, Höskuldur, Hjalmar P. Petersen, Jógvan í Lon Jacobsen & Zakaris Svabo Hansen. 2004. *Faroese. An overview and reference grammar*. Tórshavn: Føroya Fróðskaparfelag.
- Trips, Carola. 2016. An analysis of phrasal compounds in the model of parallel architecture. In Pius ten Hacken (ed.), *The semantics of compounding*, 153–177. Cambridge: Cambridge University Press.

Websites

- The Database of Modern Icelandic Inflection (Beygingarlýsing íslensks nútímamáls). <http://bin.arnastofnun.is>
- Íslenskur orðasjóður (Wortschatz Universität Leipzig). http://wortschatz.uni-leipzig.de/ws_isl/
- The Dictionary of Old Norse Prose. <http://www.onp.ku.dk>