## Forming a Compound and Spelling it Out

#### Gísli Rúnar Harðarson

### 1 Introduction

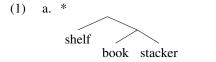
The goal of this paper is twofold. On the one hand, the goal is to derive the asymmetries in morphosyntactic behaviors of synthetic and primary compounds through differences in terms of their formation. On the other hand, the goal is to examine how the resulting structures may regulate the interactions between the elements therein in terms of morphophonology and contextual allomorphy.

The paper is organized as follows: in Section 2, I review some of the morphosyntactic differences between synthetic and primary compounds and outline a proposal for compound formation that accounts for the asymmetries observed. In Section 3, I provide an overview of the different ways elements within compounds interact. In Section 4, the paper is summarized and some outstanding questions are discussed.

# 2 Synthetic versus Primary Compounds

Synthetic and primary compounds differ in terms of various properties (see e.g., Marchand 1969, Roeper and Siegel 1978, Selkirk 1982).

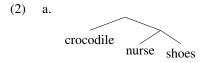
First, synthetic compounds are restricted in terms of the complexity of right-branching structures, i.e., even where the head of the compound corresponds to a ditransitive verb, only a single argument is possible.

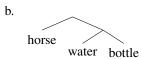


book shelf stacker

'someone who stacks books onto shelves' (Peter Smith, p.c.)

Primary compounds are, however, not subject to such restrictions:





Second, synthetic compounds always have a (roughly) synonymous phrasal counterpart, as is shown below.

(3) a. truck driver = driver of trucks
b. book stacker = stacker of books
c. head movement = movement of a head
d. candle stick maker = maker of candle sticks

Although is is certainly possible for primary compounds to have a roughly synonymous phrasal counterpart, it is not always so.

(4) a. nurse shoes = shoes of nurses
b. cream cheese bagel = a bagel with cream cheese

d. motherland  $\neq$  land of mothers

Third, the relationship between the head and the modifier is predictable and compositional with synthetic compounds (5), i.e., predicate—argument. That is, however not necessarily the case with primary compounds (6), where the relationship between the head and the modifier is much less predictable and the meaning is not necessarily compositional.

(5) a. truck driverb. book stacker

c. head movement

d. candle stick maker

(6) a. bird-dog

b. alligator shoes

c. red cap

d. cookbook

Finally, there appear to be selectional restrictions with respect to what elements can be combined in synthetic compounding, as is shown below. Such restrictions do not seem to apply to primary compounding.

(7) a. grim-acting

b. fast-mover

c. wage-earner

d. skin-grafting

e. expert-tested

(Roeper and Siegel 1978:207)

(8) a. \*grim-wanting

b. \*fast-finding

c. \*fool-looker

d. \*student-tested

e. \*doctor-grafting

Given the restrictions on composition and interpretation observed above, the question is how these asymmetries can be derived.

Under Lexicalism, the differences between synthetic and primary compounds were argued to stem from the two types of compounds being formed in different components of grammar (see e.g., Roeper and Siegel 1978, Fabb 1984, Roeper 1987, 1988). Primary compounds were argued to be formed in the lexicon, whereas synthetic compounds were argued to be formed in the syntax through incorporation of a complement.

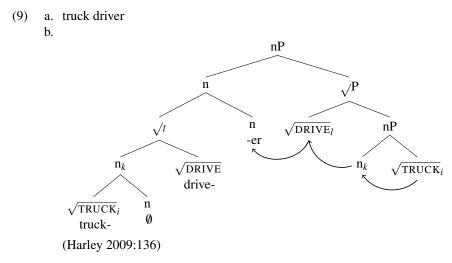
The benefit of such an account is that the asymmetries listed in the previous section follow. Synthetic compounds start off in a phrasal configuration and the relationship between the two elements is established as predicate—adjunct prior to compound formation. The meaning is hence (mostly) compositional and predictable. The restrictions on right-branching structure could then be due to the absence of the necessary structure to introduce additional arguments (e.g., Larson 1988, 1990, Harley 2002, Pylkkänen 2008).

The elements of a primary compound, however, being formed in the lexicon, are never in a structural configuration, hence the relationship between the two is not regulated by syntax and is much less predictable. It is not expected that there will necessarily be a synonymous phrasal counterpart to these compounds.

Under a non-lexicalist theory (e.g., Distributed Morphology, Halle and Marantz 1993), however, the two-component approach is not an option. There is only a single combinatorial engine. However, it can be translated into such a framework using existing syntactic mechanisms.

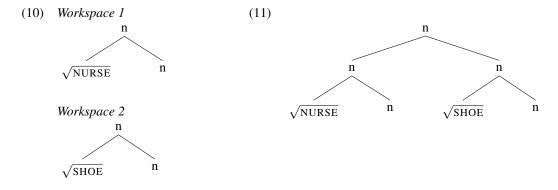
Both types of compounds must be formed in the syntax. They differ, however, in how they are put together. On the one hand, I argue that synthetic compounds are formed by incorporation, following Harley (2009). On the other hand, I argue that primary compounds are formed by merging the modifier directly to the head. Under this approach, the asymmetries between the two compounds can be derived as in the lexicalist approach discussed above. Synthetic compounds start in a phrasal configuration, whereas primary compounds do not.

Starting with synthetic compounds, I adopt Harley 2009 and assume the modifier in a synthetic compound adjoins to the head through subsequent head-movements.



The incorporated element is an argument (or first sister, cf. Roeper & Siegel 1978) of the root. Additional argument-introducing heads (e.g., Larson 1988, 1990, Harley 2002, Pylkkänen 2008) are absent from the structure. Hence failure to incorporate will result in a synonymous phrasal configuration. The structure establishes a predicate—argument relationship, hence both predictability of meaning and the selectional restrictions follow.

I assume that primary compounds formed by merging the modifier directly to the head, forming a complex head (cf. Chomsky 1970, Baker 1988, Lieber 1992, Roeper et al. 2002, Borer 2003). The two elements are formed in separate workspaces (10), and the resulting structures then merged together (11) (cf. Nunes and Uriagereka 2000, Piggott and Travis 2013).



The elements are never in a phrasal configuration, and hence the relationship between the elements is not limited to that of a head and a complement. The structure in (11) is not a result of syntactic/semantic selection and hence the selectional restrictions observed with synthetic compounds are not expected.

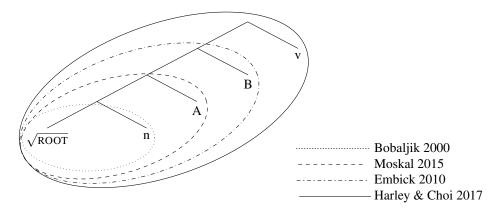
#### 3 Interactions

Two questions come up at this point: i) to what extent are interactions between the elements within a compound regulated by the structure, and ii) do we observe any asymmetries between synthetic and primary compounds in this respect? As will become clear below, the answer to (ii) appears to be—no; the answer to (i) is much less straight forward.

Under Nunes and Uriagereka 2000, Piggott and Travis 2013, it is expected that there will be asymmetry in terms of possible interactions. Modifiers in primary compounds would have to undergo spell-out prior to being merged with the head and hence expected to be islands in this respect. A cross-linguistic survey of various phenomena indicates however that the non-head elements in compounds are not islands.

In single-stem words, various locality domains have been proposed for contextual allomorphy, (12), and it has been argued that morphophonology is subject to the same locality restrictions (e.g., Embick 2010, Merchant 2015).

### (12) Domains of Contextual Allomorphy



Interactions between elements within compounds point, however, to a double dissociation between the domain of morphophonology and contextual allomorphy.

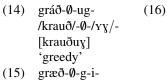
#### 3.1 Morphophonology

The i-umlaut in Icelandic is a set of morphologically triggered vowel alternations listed in (13).

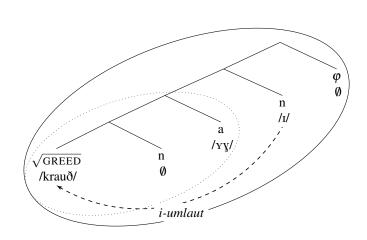
(13) Vowel alternation in i-umlaut (adapted from Árnason 2011:240)

$$\begin{array}{ccccc} /a/, /o/, /ee/ & \rightarrow & /\epsilon/\\ /au, /ou/ & \rightarrow & /ai/\\ /Y/, /o/ & \rightarrow & /i/\\ /u/, /ju/, /jou/ & \rightarrow & /i/\\ /eey/ & \rightarrow & /ei/\\ \end{array}$$

The i-umlaut only applies to certain morphemes in different contexts, and hence should be considered readjustment rules rather than automatic phonological processes (cf. Embick 2010:97ff; see Harðarson 2016 for further discussion of the i-umlaut). If that is the case, it is predicted that the i-umlaut is subject to the same phase-locality restrictions as contextual allomorphy. That does not seem to be the case as can be seen from the following examples.



(15) græð-Ø-g-i-/krauð/-Ø-/ʏy/-/ɪ/-[kraiðkɪ] 'greed/greediness'



In (15), the nominalizer -i triggers umlaut on the root across two category nodes. This would be ruled out by three out of the four domains in (12) and, although single-stem words seem to indicate that the domain spans the entire complex head, that does not seem to be the case when compounds are taken into consideration.

The u-umlaut shows similar behavior, but differs in that it applies throughout the word as long as there is a chain of potential undergoers (Anderson e.g., 1969, Orešnik e.g., 1977, i.a.).

(17) Vowel alternations in u-umlaut

a. 
$$/a/ \rightarrow /e/$$
  
b.  $/a/ \rightarrow /y/$  or  $/s/$ 

- (18) a. bak-ar-i bake-er-NOM.SG 'baker'
  - b. bök-ur-um bake-er-DAT.PL

(19) a. banan-i
banana-NOM.SG
'banana'
b. bönun-um
banana-DAT.PL

The u-umlaut never applies between two elements in a compound even when there is a chain of potential undergoers. This applies to both primary (20), and synthetic compounds (21).

(20) a. bakar-a#bönun-um baker-GEN#banana-DAT.PL 'a baker's banana' b. \*bökur-u#bunun-um

baker-GEN#banana-DAT.PL

(21) a. banana#bökur-um banana#baker-DAT.PL 'a banana baker'b. \*bönunu#bukur-um banana#banana-DAT.PL

Elements within compounds also appear to form domains for word stress. Starting with Icelandic, in single-stem words, primary stress falls on the leftmost syllable and secondary stress falls on every other subsequent syllable (22). This pattern breaks down in compounds, where secondary stress obligatorily falls on the leftmost syllable of each non-initial stem (23).

- (22) a. dr'ottnin-ˌar queen-GEN.SG
  - b. pr'ófess,or-Ø professor-NOM.SG
- (23) a. dr'ottnin-ar#mˌað-ur queen-GEN#man-NOM.SG 'the queen's consort'
  - b. pr'ofessor#b,indi-Ø professor#tie-NOM.SG 'a professor tie'

Russian typically only allows for a single stress regardless of the length of the word (e.g., Gouskova and Roon 2008).

(24) v'i-kristal-iz-ova-t<sup>j</sup>-s<sup>j</sup>a 'to crystalize' (adapted from Gouskova & Roon 2009:58)

In compounds, certain classes of stems can bear secondary stress as non-head elements, whereas main stress falls on the head of the compound (Gouskova and Roon 2008, Gouskova 2010).

(25) a. v<sub>er-o</sub>#ispoved'anije 'denomination' v'er-a 'faith' b. obor on-o#spos obnost defense capability obor on-a defense

(adapted from Gouskova & Roon 2009:59)

It is not the case, however, that elements within compounds never interact. There are various languages in which certain morphophonological processes can apply between two elements in compounds.

Vowel harmony is known to apply between two elements in a compound in various languages.

```
(26) Nez Perce (Adapted from Crook 1999:23–24)
```

(27) Nawuri (Adapted from Casali 2013:321)

```
    a. /ga/ + /tʃu/ + /toʊ/ → [gètʃútúú]
        NC 'water' 'throw' 'water throwing' (a funeral rite)
    b. /ɔ/ + /dɪ/ + /bojii/ + /pu/ → [òdìbójíípû]
        NC sleep break AGT 'gossiper'
```

(28) Chukchi (Adapted from Comrie 1981:245)

```
/pəlvəntə/ + /kupre/ + /n/ \rightarrow [pəlvəntəkopran] 
'metal' 'net' ABS.SG 'metal net'
```

There are also cases of morphophonological processes that apply specifically between two elements in a compound (Vogel 2010).

(29) Nasal assimilation in Marathi (adapted from Pandharipande 1997:563)

```
a. /b<sup>h</sup>agawat/ + /naːm/ → [b<sup>h</sup>agawannaːm]
'god' 'nam' 'god's name'
b. /ṣat/ + /maːs/ → [ṣanmaːs]
'six' 'month' 'six months'
```

(30) Final vowel lengthening in Hausa (adapted from McIntyre 2006:32)

```
a. /bì/ + /bango/ → [bì:bango]
'follow' wall 'leakage along the wall'
b. . /k'a:rè/ + /dangì/ → [k'à:rè:dangì]
'finish' 'relative' type of arrow poison
```

(31) Vowel deletion in Swedish (Josefsson and Platzack 2004:12)

```
a. flicka + skola → flickskola
'girl' school 'girls school'
b. loge + dans → logdans
'barn' 'dance' 'barn warming'
```

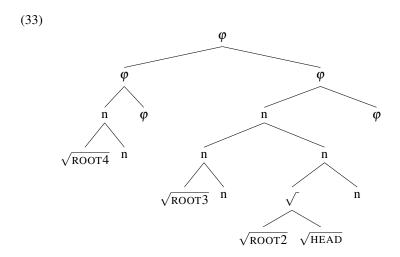
(32) Vowel change in Swedish (Josefsson and Platzack 2004:13)

```
    a. saga + bok → sagobok
    'story' 'book' 'storybook'
    b. gata + skylt → gatuskylt
    'street' 'sign' 'street sign'
```

There do not appear to be any asymmetries between synthetic and primary compounds with respect to these processes.

#### 3.1.1 Domain of Morphophonology

I follow various proposals in assuming that compounding takes place at different levels within the complex head and that modifiers must match the level where they are merged (e.g., Moskal 2015, Harðarson 2016, De Belder 2017). The structure of a primary compound would be along the lines of (33).



Possibilities for morphophonological interactions are determined by the position of the modifier within the complex head where uncategorized (e.g., Moskal 2015, Harðarson 2016, De Belder 2017). I furthermore assume that the domain of morphophonological interactions is defined by the highest functional morpheme in the extended projection of the root (Harðarson 2016), which is defined as follows.

### (34) Domain for morphophonological interactions

A domain for morphophonology is marked by the highest projection in the extended projection of the root.

- i. Morphophonological rules can apply to the exponents of two nodes, X and Y, if X and Y are within the same extended projection.
- ii. Morphophonological rules do not apply between two extended projections.
- (35) Extended projection within a complex head

 $\alpha$  is in the extended projection of a root R if:

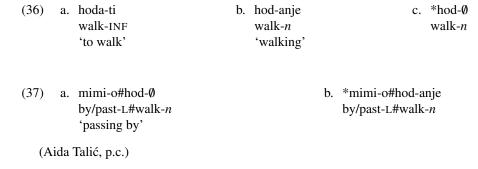
- i. the head of  $\alpha$  morphologically selects/subcategorizes R, or
- ii. the head of  $\alpha$  morphologically selects/subcategorizes  $\beta$ , where  $\beta$  is a head in the extended projection of R. (Harðarson 2016:23)

Under these definitions, uncategorized roots have no extended projection, hence do not form domains for morphophonology. Assuming that (25–27) are cases of root–root compounding (see Moskal 2015), these definitions allow for the interactions observed. Categorized roots, however, have an extended projection and hence form domains.

The compound specific processes in (28–29) point to the complex head serving as context for certain morphophonological processes specifically between two extended projections within the complex head but not within a single extended projection. This indicates that morphophonological operations apply at two levels within the complex head, in the vein of Lexical Phonology (e.g., Kiparsky 1982, Monahan 1982) and its descendants: i) within the extended projection of the root, and ii) between two extended projections within the complex head.

## 3.2 Contextual Allomorphy

The compound structure can furthermore serve as context for suppletive allomorphy. In Bosnian, compounding conditions null allomorph of a nominalizing suffix despite the linear intervention of the head root. Outside of compounds an overt nominalizing suffix is obligatory.



A similar pattern is attested in Russian (Jonathan Bobaljik, Ksenia Bogolomets, p.c.).

(38) a.  $vod-itel/*\emptyset$  b.  $ekskursa#vod-\emptyset$  lead-n tour#lead-n 'leader' 'tour leader'

The compound structure can also serve as context for allomorphy of the modifier. This is shown by the Bosnian example below, where the compound requires the plural form *ljud*- ('people') rather than the singular form *čovjek*- ('man').

(39) a. ljud.o#žder-0 b. \*čovjek.o#žder-0 man#devour-*n* 'man eater'

(Aida Talić, p.c.)

This indicates that the non-head element is not spelled out prior to attachment to the head of the compound. Note however, that the size of the non-head elements in these cases remains to be determined.

Assuming that Dutch linking morphemes are stem affixes (e.g., De Belder 2013), the selection of linking morphemes also shows contextual allomorphy where the selection of the linking morpheme can sometimes depend on the head of the compound.

(40) a. schaap-en#tong b. schaap-s#kooi c. schaap#herder sheep-L#tongue sheep-L#fold sheep#herder 'sheep's tongue' 'sheep fold' 'sheepherd'

(Krott et al. 2007:28)

Finally, Icelandic allows case and number marking on non-head elements in compounds and inflected modifiers in Icelandic are structurally peripheral to uninflected modifiers (Harðarson 2016).

(41) a. karl-a#hest#vagn
men-GEN#horse#wagon
b. [MAN<sub>infl</sub> [HORSE<sub>stem</sub> WAGON]]
'a horse-drawn carriage for men'
c. \*[[MAN<sub>infl</sub> HORSE<sub>stem</sub>] WAGON]
'a carriage drawn by male horses'
a. karl#hest-a#vagn
man#horse-GEN#wagon
b. \*[MAN<sub>stem</sub> [HORSE<sub>infl</sub> WAGON]]
'a horse carriage for men'
c. [[MAN<sub>stem</sub> HORSE<sub>infl</sub>] WAGON]
carriage drawn by male horses'

The class of strong neuter non-count nouns, however, cannot appear with an overt case marker but still appear structurally peripheral to inflected modifiers.

- (43) a. hör#[ vas-a#klútur ]
  flax<sub>STEM</sub># pocket-GEN#cloth
  'linen handkerchief'
  - b. silki#[ kodd-a#ver ]
    silk<sub>STEM</sub># pillow-GEN#case
    'silk pillow case'
  - c. tré#[ penn-a#standur ]
    tree<sub>STEM</sub># pen-GEN#stand
    'wooden pen stand'
- (44) a. \*hör-s#[ vas-a# klútur ] flax-GEN# pocket-GEN#cloth 'linen handkerchief'
  - b. \*silkis#[ kodd-a# ver ] silk-GEN# pillow-GEN#case 'silk pillow case'
  - c. \*tré-s#[ penn-a# standur ] tree-GEN# pen-GEN#stand 'wooden pen stand'

In the case of the stem *tré*- 'tree,' the base form is *trjá*-, which surfaces in dative and genitive plural, (45a) whereas the form *tré*- is conditioned by particular combinations of case and number. Both forms appear in compounds (45b,c).

(45) a. trjá-a b. tré#froskur c. trjá#froskur tree-GEN tree#frog tree#frog 'tree' 'wooden frog/tree frog' 'tree frog'

It appears then that the compound structure is conditioning a  $\emptyset$  allomorph of the genitive suffix. Contextual allomorphy then applies across boundaries that morphophonology in Icelandic cannot.

The data appear to be consistent with any of the four locality domains mentioned in (12). However, it is clear that the non-head elements cannot undergo spellout prior to formation of the compound.

#### 4 Conclusion

The asymmetries between synthetic and primary compounds can be derived from the difference in their formation, where synthetic compounds are formed through incorporation and primary compounds are formed by merging a modifier directly to the head after building the two elements in separate workspaces.

Although the elements of primary compounds are formed in separate workspaces, the modifiers are not spelled out prior to being merged with the head and can potentially be subject to contextual allomorphy. Furthermore, morphophonology applies at two levels: i) within the extended projection of the root, and ii) between two extended projections within the complex head.

The compound structure can serve as a condition for contextual allomorphy. Although questions remain with respect to the locality of contextual allomorphy within compounds, it is clear that the non-head element cannot have undergone spellout prior to the formation of the compound.

#### References

Anderson, Stephen R. 1969. An outline of the phonology of modern Icelandic vowels. *Foundations of Language* 5:53–72.

Baker, Mark. 1988. Incorporation: A Theory of Grammatical Function Changing. Doctoral dissertation, MIT. Bobaljik, Jonathan. 2000. The ins and outs of contextual allomorphy. In *University of Maryland Working Papers in Linguistics 10*, ed. K. Grohmann and C. Struijke, 35–71.

Borer, Hagit. 2003. Exo-skeletal vs. endo-skeletal explanations: syntactic projections and the lexicon. In *The Nature of Explanation in Linguistic Theory*, ed. J. Moore and M. Polinsky, 31–67. Stanford, CA.

Casali, Roderic F. 2013. [ATR] value asymmetries and underlying vowel inventory in Niger-Congo and Nilo-Saharan. *Linguistic Typology* 7:307–382.

Chomsky, Noam. 1970. Remarks on nominalization. In *Readings in Transformational grammar*, ed. R. Jacobs and P. Rosenbaum, 184–221. Oxford: Ginn & Company.

Comrie, Bernard. 1981. The Languages of the Soviet Union. Cambridge: Cambridge University Press.

Crook, Harold David. 1999. The Phonology and Morphology of Nez Perce Stress. Doctoral dissertation, University of California, Los Angeles.

De Belder, Marijke. 2013. Linking phonemes are class markers. In *Taalkunde in Nederland-dag TIN-dag*, 1–14.

De Belder, Marijke. 2017. The root and nothing but the root: primary compounds in Dutch. *Syntax* 20:138–169. Embick, David. 2010. *Localism Versus Globalism in Morphology and Phonology*. Cambridge, MA: MIT Press. Fabb, Nigel. 1984. Syntactic Affixation. Doctoral dissertation, MIT, Cambridge, MA.

Gouskova, Maria. 2010. The phonology of boundaries and secondary stress in Russian compounds. *The Linguistic Review* 17:387–448.

Gouskova, Maria, and Kevin Roon. 2008. Interface constraints and frequency in Russian compound stress. In Proceedings of FASL 17, ed. M. Babyonyshev J. Reich and D. Kavitskaya, 49–63.

Halle, Morris, and Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In *The view from Building 20: essays in linguistics in honor of Sylvain Bromberger*, ed. K. Hale and S. Kayser, 111–176. Cambridge, MA: MIT Press.

Harðarson, Gísli Rúnar. 2016. Peeling away the layers of the onion: on layers, inflection and domains in Icelandic compounds. *The Journal of Comparative Germanic Linguistics* 19:1–47.

Harley, Heidi. 2002. Possession and the double object construction. Linguistic Variation Yearbook 2:31-70.

Harley, Heidi. 2009. Compounding in Distributed Morphology. In *The Oxford Handbook of Compounding*, ed. Rochelle Lieber and Pavol Štekauer, 129–144. Oxford: Oxford University Press.

Josefsson, Gunlög, and Christer Platzack. 2004. Introduction. In *The Acquisition of Swedish Grammar*, ed. G. Josefsson, C. Platzack, and G. Håkonsson, 1–22. John Benjamins.

Kiparsky, Paul. 1982. Lexical Phonology and Morphology. In *Linguistics in the Morning Calm*, ed. I. Yang. Seoul: Hanshin.

Krott, Andrea, Robert Schreuder, R. Harald Baayen, and Wolfgang Dressler. 2007. Analogical effects on linking elements in German compound words. *Language and Cognitive Processes* 22:25–57.

Larson, Richard. 1988. On the Double Object Construction. Linguistic Inquiry 19:335-391.

Larson, Richard. 1990. Double Objects Revisited: Reply to Jackendoff. Linguistic Inquiry 21:589-632.

Lieber, Rochelle. 1992. Deconstructing Morphology. Word Formation in Syntactic Theory. Chicago: University Of Chicago Press.

Marchand, Hans. 1969. The Categories and Types of Present-Day English Word-Formation: A Diachronic Approach. München: CH Beck.

McIntyre, Joseph Anthony. 2006. Hausa Verbal Compounds. Doctoral dissertation, Leiden University, Leiden. Merchant, Jason. 2015. How Much Context Is Enough? Two Cases of Span-Conditioned Stem Allomorphy. *Linguistic Inquiry* 46:273–303.

Monahan, Karuvannur Puthanveettil. 1982. Lexical Phonology. Doctoral dissertation, MIT, Cambridge, MA. Moskal, Beata. 2015. Domains on the Border: Between Morphology and Phonology. Doctoral dissertation, University of Connecticut, Storrs.

Nunes, Jairo, and Juan Uriagereka. 2000. Cyclicity and extraction domains. Syntax 3:20-43.

Orešnik, Janez. 1977. Orešnik: Modern Icelandic u-umlaut from the descriptive point of view. *Gripla* 2:151–182

Pandharipande, Rajeshwari. 1997. Marathi. London: Routledge.

Piggott, Glyne, and Lisa Travis. 2013. Adjuncts within words and complex heads. In *Syntax and its Limits*, ed. R. Folli, C. Sevdali, and R. Truswell, 157–174. Oxford: Oxford University Press.

Pylkkänen, Liina. 2008. Introducing Arguments. Cambridge, MA: MIT Press.

Roeper. 1988. Compound syntax and head movement. Yearbook of Morphology 1:187-228.

Roeper, Thomas. 1987. Implicit arguments and the head-complement relation. *Linguistic Inquiry* 18:267–310.

Roeper, Thomas, and Muffy E. A. Siegel. 1978. A lexical transformation for verbal compounds. *Linguistic Inquiry* 9:199–260.

Roeper, Thomas, William Snyder, and Kazuko Hiramatz. 2002. Learnability in the Minimalist framework: Root compounds, merger, and the syntax-morphology interface. In *The Process of Language Acquisition Proceedings of the GALA Conference*, 25–35. Frankfurt.

Selkirk, Elisabeth. 1982. The Syntax of Words. Cambridge: MIT Press.

Vogel, Irene. 2010. The phonology of compounds. In Cross-Disciplinary Issues in Compounding, ed. S. Scalise and I. Vogel, 145–163. Amsterdam: John Benjamins Publishing Company.

Department of Linguistics University of Connecticut Storrs, CT 06269-1145 gisli.hardarson@uconn.edu