



Kristín Bjarnadóttir

**Ef-og-pá-kannski-hlutir:
Um setningarliði í samsettum orðum**

29. Rask-ráðstefnan um íslenskt mál og almenna málfræði
31. janúar 2015



Viðfangsefnið

Tvær megingerðir samsettra orða með fyrri hluta sem ekki eru stök orð eða les (samsett eða ósamsett):

1. Rótgróin orð: **innanríkisráðherra, áttablaðabrot**
2. Tækifærisorð: **á-tali-sónn, ef-og-pá-kannski-hlutir**

Gagnasafnið er u.þ.b. þúsund orð, aðallega úr *Beygingarlýsingu íslensks nútímamáls og Íslenskum orðasjóði*, með dálitlum viðbótum af Google.



Yfirlit

- Inngangur: Er orð stafastrengur á milli bila?
- Kenningarnar:
 - Eiga setningarliðir sér einhvern stað inni í orðum?
No Phrase Constraint (Botha 1981)
 - Sér setningarhluti málkerfisins inn í orð?
- Gagnasafnið:
 - Venjulegt ritmál (BÍN)
 - Efni af vefsíðum (Íslenskur orðsjóður)
- Greining, flokkun og dæmi
- Niðurstaða



Stafastrengur á milli bila?

allt eða ekkert dæmi	3 bil
“allt eða ekkert” dæmi	3 bil, gæsalappir
„allt eða ekkert“ aðferð	3 bil, gæsalappir
‘allt eða ekkert’ týpa	3 bil, gæsalappir
allt-eða-ekkert hugsunarháttur	2 bönd, 1 bil
„munn-við-munn“ aðferð	2 bönd, 1 bil, gæsir
munn-við-munn-öndun	3 bönd
millipinganefnd	skrifað í einu orði
áttablaðabrot	skrifað í einu orði
átta gata tryllitæki	2 bil
einnar-hæðar-skýjakljúfur	2 bönd



Hvað er sérstakt við setningarliðasamsetningar?

- Venjulegar reglur um orðmyndun ná ekki yfir þær
 - Hvorki fastar né lausar samsetningar
 - Ekki bandstafur
 - Ekki samtenging **tveggja** lesa (orða)
- Stundum notaðar í mjög óformlegu máli (og texta)
 - Vegna stafsetningarinnar er erfitt að safna dæmum
 - Lítil vitneskja um virkni og tíðni
- Óljós mörk samsetts orðs og orða í setningu
- Lítil umfjöllun um íslensku (nema væntanleg grein Þorsteins G. Indriðasonar um frásasamsetningar).



Talsverð gróska í rannsóknum erlendis

Since compounds are an interface phenomenon of the different components of grammar they serve well as a testing ground for the validity of various models of grammatical theory.

[Phrasal compounds] are also different in being structurally more complex consisting of both a phrase and a lexeme, exploring their nature, we believe, is even more interesting and insightful.

(Carola Trips & Jaklin Kornfilt, óútg.)

Enska: Texti í gæsalöppum, bil, nafnorð
(sbr. Carola Trips 2007)





No Phrase Constraint (Botha 1981)

Lexical rules do not apply to syntactic phrases to form morphologically complex words.

This hypothesis implies that syntactic rules are excluded not only from the class of word formation rules, but also from the class of rules that form the bases to which word-formation rules may apply.

(Botha 1984:137)

**Orðmyndunarreglur eiga ekki að hafa
aðgang að setningarreglum og öfugt.**



Gögnin

- **Beygingarlýsing íslensks nútímamáls/gögn úr KB 1996:**
 - 250 þúsund samsett orð með orðhlutagreiningu
 - 600 setningarliðasamsetningar
 - Í BÍN er ekki mikið um orðaforða úr mjög óformlegu máli. Öll orð eru án orðabila.
- **Íslenskur orðasjóður** (textasafn af vefnum):
 - 600-700 milljónir lesmálsorð
 - 3 þúsund línur úr leit: “*_*_* “
(stafastrengur m. 2 bandstrikum og 1 bili)
 - 200 setningarliðasamsetningakandídatar
 - Í Ío er talsvert mikið af mjög óformlegu máli
- Leitað var í Google til að finna stafsetningarafbrigði.



BÍN: Greining I

ORD:innanborðs_vél

[

[

innan |innan |fs

|borðs |borð |no.hk.et.ef

]<FI>

+

[vél |vél |no.kvk.et]<no.kk>

]<no.kk>

|HEIM:Ío

Ío = *Íslensk orðabók* (2000)



BÍN: Greining II

ORD:sjöundadags_aðventisti

[

[

sjöunda |sjöundi |lo.num.kk.et.ef

|dags |dagur |no.kk.et.ef

]<NI>

+

[aðventisti |aðventisti |no.kk.et

]<no.kk>

|HEIM:norr

norr = *Norræna verkefnið* (Íslenskur orðabókarstofn)



BÍN: Flokkun I

Forsetningarliðir: Algengasta gerðin

[[fs + no.ef]_{<FI>} + no]_{<no>}

innanflokksátök

[[fs + no.ef]_{<FI>} + ao]_{<ao>}

utanlandsfrá

[[fs + no.ef]_{<FI>} + lo]_{<lo>}

utanríkispólitískur

[[fs + no.pf]_{<FI>} + lo/lh.nt]_{<lo>}

íhöndfarandi

[[fs + abfn.pgf]_{<FI>} + lo/lh.tp]_{<lo>}

úrsérgenginn

Nafnliðir í eignarfalli: Nokkuð algengt

[[lo + no]_{<NI.ef>} + no]_{<no>}

heilsdagsvinna

[[to_{<ef>} + no]_{<NI.ef>} + no]_{<no>}

tveggjaatkvæðaorð

einsatkvæðisorð

[[to_{<ób>} + no]_{<NI.ef>} + no]_{<no>}

fimmliðaháttur

[[to_{<raðt.>} + no]_{<NI.ef>} + no]_{<no>}

sjöundadagsaðventisti

[[fn_{<ef>} + no]_{<NI.ef>} + no]_{<no>}

engramannaland



BÍN: Flokkun II

Annað:

Atviksliðir	utanaðlærdómur	34 orð
Nafnorð+lýsingarorð	Helgamagrastræti	1
Lýsingarliður	allraheilagramessa	1
Hliðskipaðir liðir:	tuttuguogfimmeyringur	?
Boðháttur:	haltukjaftibrjóstykur	1

Til athugunar:

tveggjarúðugluggi, sexrúðugluggi (sbr. sexraðabygg)

margrastjörnukoníak

millistríðsár (sbr. millipilsaþytur, millilandaferð)

Sautjándajúnítorg

Annars er sambeyging í setningarliðunum nánast alltaf „eðlileg“, b.e. eins og í setningum.



Íslenskur orðasjóður: 3.000 línur

Strengur með 2 bandstöfum, bil, nafnorð

```
19 gratarlækur-stekkjarmoar-djupidalur breyting
19 kvenföt-feminin-fashion verslun-feminin-fashion
18 2010-09-04hljóð og
18 ever-so-red skrifaði:
18 líkami-fyrrir-lífið áskorun
17 prentari-skanni-fjölföldun-fax prenttækni:
17 hofsjökull-akureyri-vonarskarð <
17 p-2-p og
17 2010-08-23hljóð og
16 dg-2gb-bk supertalent
16 mso-pagination:widow-orphan; font-family:"calibri","sans-serif";
16 font-family:"calibri","sans-serif"; mso-ascii-font-family:calibri;
16 kerlingafjöll-setrið-hofsjökull <
16 emily-rocker-chick og
16 mso-ascii-font-family:calibri; mso-ascii-theme-font:minor-latin;
16 ship-o-hoj ship-o-hoj
16 ki-adi-mundi var
15 norðausturhálandið-mars-2005 <
15 aix-en-provence í
15 guðs-leik-húsum bókaumfjallanir
15 hólaskógur-sultarfit-16 ókt
15 nýtt-stáss-eldklukka hálsmen
```



Íslenskur orðasjóður I Sömu orðgerðir og í BÍN

Fyrri hlutar: Lýsingarorð og nafnorð í eignarfalli

allra-spila-klúbbur

einnar-hæðar-skýjakljúfur

einnar-mínútu-myndbanda-vinnustofa

fjögra-potta-kúturinn

þriggja-non-eu leikmanna-reglan

Fyrri hlutar: Forsetningarliður

af-því-bara rök

eftir-vinnu-teiti (ekki 'teiti í eftirvinnu')

inn-í-eyra tæki

niður-brekku fíling



Íslenskur orðasjóður II

Ekki í BÍN-dæmunum*

Nafnliðir, nefnifall:

bland-í-poka efni

búð-í-búð verslun

fyrirtækja-til-neytenda markaður

hland-í-sundlaug dæmið

kortér-í-þrjú-náungi

líkami-fyrir-lífið aðferð*

líkaminn-sem-söluvara pakki

maður-á-mann aðferð**

málgrein-á-dag tólið

munni-við-hlíf aðferðin

munni-við-munni aðferðin

mynd-á-viku mynd

mynd-og-ljóð-sýning

næring-fyrir-lífið áætlun

orð-fyrir-orð þýðing

poki-í-öskju kerfi

rör-í-rör kerfi

skref-fyrir-skref leiðbeiningar

sól-yfir-vegi vinur

*áskorun, áætlun, bók, pakki, prógrammið, vefur, æfing, æfingaráætlun

**staða, vörn



Íslenskur orðasjóður III

Erlendir fyrri hlutar, íslenskir síðari hlutar (hausar):

board-a-match sveitakeppni

box-to-box miðjumaður

fish-and-chips réttur

food-and-fun-hádegisverður

game-of-the-year verðlaun

champ-to-champ-nod kallarnir

coast-to-coast skautahlaup

hand-in-hand viðskipti

peer-to-peer tækni

pointe-à-pitre umhverfi

point-in-time afritun

side-by-side fjórhjól

step-by-step bók

wake-on-lan fíðus

wheel-to-wheel barátta

Hreinni útlensku er sleppt, t.d. *foot-and-mouth disease*



Íslenskur orðasjóður IV

- Neitun:** ekki-svo-dapurlegi-dagurinn
ekki-svo-fjarlæg framtíð
- Sagnir:** gettu-betur-lið (sbr. haltukjaftibrjóstsýkur)
gerðu-það-sjálfur tónlist
hefði-átt-að regla
spurt-og-svarað-dæmi
ekkert-um-það-segjandi von
lemja-vondu-kalla stórmynd
- Lýsingarorð +:** fúll-á-móti-flokkurinn
- Spurning:** hvers-vegna-ekki samfélagið
- ...
- næstum-því-vinur



Íslenskur orðasjóður V

„mögulega-kannski-hver-veit guð“ eru þetta orð ...

ég-rek-þau-í-gegnum-þig horn

ég-verð-að-vita-hvað-gerist-næst-bók

ég-þoli-ekki-illar-tölvur-klúbburinn

ruglaðu-mig-í-hausnum mynd

hið-eina-sanna-opna-þjóðar-grunnnet

ohh-ég-vildi-að-ég-gæti-kúkað kúkur

ódýrari-en-í-frakklandi nettenging

'skrifaðu-í-gestabókina-eða-ég-kýli-þig-í-andlitið' dagurinn

stóri-verðtryggði-laxár pakkinn

svona-er-nú-það drottningin

mögulega-kannski-hver-veit guð

un-be-fucking-leaveable laugardagur



Niðurstaða

- Sumt í muninum á gögnunum skýrist af mismunandi forsendum í dæmasöfnuninni: BÍN = stafastrengur á milli bila, Ío = “*_*_* “
- Samsetningar í Ío geta verið mun frjálssari í formi og lengri en í BÍN
- Mun meira virðist vera um einnota samsetningar í Ío en tíðnina þarf að skoða betur.
- Staðfesta setningarliða- og frásasamsetningarnar þá kenningu að orðmyndunin sé ekki blind á setningargerð og öfugt?
 - Fall, tala og kyn í setningarliðasamsetningunum er (yfirleitt alltaf) eins og í setningarliðunum frjálsum. →



Í íslensku er full sambeyging í setningarliðasamsetningunum: Góðrarvonarhöfði

Sbr. við kenningar um þýsku: **Phrasal compounds:**

The AN [lo+no] constituent cannot be a true syntactic phrase: it does not exhibit the properties a corresponding well-formed syntactic phrase would have to process Firstly, their parts do not agree as the parts of the corresponding phrase would do. ...

Zweibettzimmer ('double bedroom')

zwei Betten ('two beds') *zwei Bett

Jurgen Pafel. 2015. Phrasal compounds are compatible with lexical integrity. Trips, Carola, & Jaklin Kornfilt. Handrit, án titils, 29.1.2014.



Nafnorð með innri beygingu

Í Beygingarlýsingu íslensks nútímamáls (BÍN) eru nú **1.035** nafnorð með beygða fyrri hluta:

- Stóri-Botn, Litlibær, Gulikambur, Stóribjörn
- æðstitemplar, hæstiréttur, langatöng, miklihvellur
- Árnibjörn, Annamaría

Að langmestu leyti eru þetta örnefni.

Örnefni í BÍN* eru 23.047

Samsett nafnorð í BÍN eru 218.222

*þ.m.t. götuheiti o.þ.h.

BÍN 22.1.2015



Takk fyrir áheyrnina

Kristín Bjarnadóttir
Stofnun Árna Magnússonar í íslenskum fræðum

kristinb@hi.is



Heimildir

- [Beygingarlýsing íslensks nútímamáls](#). Kristín Bjarnadóttir (ritstj.) Án ártals.
- Botha, Rudolf P. 1981. A base rule theory of Afrikaans' synthetic compounding. Moortgat, Michael, Harry van der Hulst & Teun Hoekstra (ritstj.). *The scope of lexical rules*. Foris, Dordrecht, bls. 1-77.
- Botha, Rudolf P. 1984. *Morphological Mechanisms: Lexicalist Analyses of Synthetic Compounding*. Pergamon, Oxford.
- *Íslensk orðabók*. Mörður Árnason (ritstj.) 2000. Edda hf., Reykjavík. [Rafræn útgáfa].
- [Íslenskur orðasjóður](#). (án árt.). Universitát Leipzig.
- Kristín Bjarnadóttir. 2005. [Afleiðsla og samsetning](#) í generatívrí málfræði og greining á íslenskum gögnum. Rannsóknar- og fræðslurit 7. Orðabók Háskólans.
- Kristín Bjarnadóttir. 1998. [Norræna verkefnið](#). Skýrsla um íslenskan orðabókarstofn.
- Pafel, Jurgen. Væntanl. 2015. Phrasal compounds are compatible with lexical integrity. Trips, Carola, & Jaklin Kornfilt. (ritstj.) Handrit, án titils. [29.1.2014].
- Trips, Carola & Jaklin Kornfilt. Væntanl. 2015. Handrit, án titils. Greinasafn um samsett orð með setningarliðum. [Workshop on prasal compounds, Mannheim, June 2014.]
- Trips, Carola. 2012. Empirical and theoretical aspects of phrasal compounds: against the “syntax explains it all” attitude. Ralli, Angela, Geert Booij, Sergio Scalise & Athanasios Karasimos (eds). [On-line Proceedings of the Eighth Mediterranean Morphology Meeting](#) (MMM8), 322-346.