

Kristín Bjarnadóttir

# Málfræði í orðabókum

## 1 Inngangur

Í orðabókum er að finna lýsingu einstakra orða eða annarra þeirra einda sem mynda uppfletteiningar í verkinu, hvort heldur það er hefðbundin prentuð orðabók eða rafræn orðabók af einhverju tagi. Þær upplýsingar sem orðabókarlýsing byggist á og kemur til skila eru af ýmsu tagi en efni þessarar greinar er það hvernig málfræðiupplýsingar eru settar fram fyrir orðabókarnotendur, bæði að formi og efni.

## 2 Málfræði, orðasafnið og orðabækur

Málfræði er lýsing á málkerfinu, sett fram á ýmsan hátt eftir tíma og skólum. Viðfangsefni málfræðinga er að setja fram reglur og lögmál um það hvernig kerfið í heild eða einstakir hlutar þess virka; það má því segja að málfræðingar fáist við málið „ofan frá og niður“. Orðabækur lýsa kerfinu hins vegar „neðan frá og upp“ ef svo má segja, þ.e. út frá einstökum orðum eða flettum.

Bæði fyrirbærin, málfræði og orðabækur, eru flókin en hvort á sinn hátt. Orðabækur byggjast á málfræðilegri hugsun en orðabókarfræði er aðeins að hluta til málfræðileg. Miðlun efnisins sjálfs og framsetning þess er þar fullgildur þáttur. Þessi munur sést af þeirri flokkun sem gjarnan er viðhöfð þegar rætt er um þessar greinar.

Málfræði skiptist í fjölmargar undirgreinar, sbr. skilgreiningu Höskuldar Þráinssonar í *Handbók um málfræði* (1995:89):

MÁLFRÆÐI er samheiti yfir allar þær fræðigreinar sem lýsa eðli mannlegs máls og skýra það, svo sem hljóðfræði, beygingarfræði, orðmyndunarfræði, setningafræði, merkningarfræði.

Þarna má bæta við fjölmörgum greinum og reyndar má telja orðabókarfræði til málfræði líka.

Orðabækur eru af ýmsu tagi, en skilgreiningu á því fyrirbæri sem íslenskum orðabókarnotendum er tamast að hugsa til þegar minnst er á orðabók er að finna í dæmigerðustu íslensku orðabókinni:

**orðabók** KVK 1 bók með safni orða (í stafrófsröð) úr e-u máli og (venjulega) skýringum á merkingu (notkun) þeirra á sama eða öðru máli > *íslensk orðabók / færeysk-íslensk orðabók / Orðabók Sigfúsar Blöndals / stafsetningarorðabók*. 2 stofnun þar sem unnið er að samningu slíkrar bókar > *Orðabók Háskóla Íslands*

(*Íslensk orðabók* 2002)

Tæknibreytingar í orðabókargerð eru slíkar að orðabók er ekki lengur endilega bók, stafrófið er ekki eini hugsanlegi leitaraðgangurinn að efninu og merkingarlýsing er aðeins einn af þeim þáttum sem orðabókarlýsing nær til. Orðabækur eru af ýmsum gerðum og nægir þar að nefna almennar móðurmálsbækur (eins og *Íslensk orðabók* sem vísað er til hér á undan), tvímálaorðabækur, stafsetningarorðabækur, orðsifjabækur, samheitabækur, orðtengslabækur, þesárusa, málnotkunarbækur, iðorðabækur o.s.frv. Allar þessar orðabækur eiga það þó sameiginlegt að útgangspunkturinn er orðið sjálft eða flettan.

Í generatífri málfræði er það orðasafnið (orðasafn hugans) sem geymir allar upplýsingar um einstök orð og önnur fyrirbæri sem þar eru sambærileg við orðabókarflettur. Þar er kominn tengiliður málfræði sem er lýsing á málkerfinu og orðabókar sem er upplýsingaveita um einstakar flettur. Í stuttu máli má segja að í orðasafninu séu les ásamt öllu öðru sem ekki er fyrirsegjanlegt samkvæmt þeim reglum og lögmálum sem sett eru fram um málkerfið í málfræði, með öllum sínum undirgreinum. Í ljósi þessa mætti ætla að í góðri orðabók eigi að vera úrval lesa með nauðsynlegum upplýsingum um þá málfræði sem ekki er fyrirsegjanleg miðað við málfræðikunnáttu notandans, þ.e. sú málfræði sem hann er ekki sjálfbjarga með. Framsetning þessara upplýsinga hlýtur að mótast af gerð orðabókarinnar og notandahópnum

sem hún er ætluð, auk miðilsins sem hún birtist í, þar sem möguleikarnir í prentaðri bók og rafrænni (til dæmis) eru mjög mismunandi. Hér á eftir verður sjónum aðallega beint að almennri íslenskri orðabók, enda er nokkur hefð á framsetningu málfræðiupplýsinga þar.

### 3 Hvaða málfræði á að vera í orðabók?

Með málfræðiupplýsingum í orðabók er átt við kerfisbundna framsetningu tiltekinna málfræðilegra þátta sem á sér fastan stað í flettugreinunum. Gera þarf greinarmun á málfræði sem notuð er við skipulag á flettugreinum í orðabók og málfræði sem beinlínis er birt notandnum sem slík, skýrt afmörkuð innan flettugreinanna. Þá þarf einnig að gera greinarmun á upplýsingum sem leynast í flettugreinunum, t.d. í dæmum, útskýringum, og jafnvel í leiðbeiningum um málnotkun, og skipulega framsettum málfræðiatríðum. Skipulag flettugreina í orðabókum Jóns Hilmars Jónssonar (1994, 2002) ákvarðast t.d. algjörlega af málfræðilegum þáttum, af setningargerð í *Orðastað* og af merkingarlegum þáttum í *Orðaheimi*.<sup>1</sup>

Þegar aðrar íslenskar orðabækur eru skoðaðar kemur í ljós að málfræðiupplýsingar eru oftast fólgnar í orðflokksmerkingu flettna, ásamt vísbendingum um beygingu orða sem yfirleitt eru settar fram með því að gefa endingar kennifalla nafnorða eða kennimynda sagna, og heilar beygingarmyndir þar sem stofnbrigði krefjast þess. Þessi háttur er t.d. hafður á í *Íslenskri orðabók* (2002) þar sem beygingar grunnorða eru sýndar á þennan hátt:<sup>2</sup>

**hundur** KK -s, -ar  
**labba** S -aði, -að  
**fjörður** KK fjarðar, firðir

<sup>1</sup>Í *Stóru orðabókinni* (2005) hafa bæði þessi verk verið felld saman í eina heild. Í þessum verkum brýtur Jón Hilmar blað í íslenskri orðabókargerð; hefðin er sú að málfræðiupplýsingar séu fremur af skornum skammti og ekki endilega settar fram á kerfisbundinn hátt, nema þá helst upplýsingar um beygingu.

<sup>2</sup>Kenniföll og kennimyndir eru þær beygingarmyndir nafnorða og sagna sem taldar eru nægja til að gefa vísbendingu um beygingarflokk. Kenniföll nafnorða eru nefnifall og eignarfall eintölu og nefnifall fleirtölu, t.d. *köttur, kattar, kettir; nál, nálar, nálar*. Kennimyndir sterkra sagna eru fjórar: nafnháttur, þátíð framsöguháttar í 1. persónu eintölu og fleirtölu, og lýsingarháttur þátíðar, t.d. *ganga, gekk, göngum, gengið*. Kennimyndir veikra sagna eru þrjár: nafnháttur, þátíð framsöguháttar í 1. persónu eintölu og lýsingarháttur þátíðar, t.d. *æpa, æpti, æpt*.

Málfræðiupplýsingar af öðrum toga eru yfirleitt ekki settar fram með kerfisbundnum hætti í íslenskum orðabókum, þótt dálitlar upplýsingar um setningargerð sé einnig að finna í *Íslenskri orðabók*. Ekki er samt gerð tilraun til að sýna fallstjórn sagna og setja rökliði þeirra fram á kerfisbundinn hátt fyrr en í 3. útgáfu.<sup>3</sup>

Minna fer fyrir annars konar málfræðilegum upplýsingum í íslenskum orðabókum. Hljóðkerfisfræði er afskipt grein í íslenskri orðabókargerð og kerfisbundnar upplýsingar um framburð er yfirleitt ekki að finna, nema í Blöndalorsorðabók þar sem hljóðritun fylgir uppfletti-orðum. Ekki er gerð grein fyrir orðmyndun á skipulagðan hátt nema í *Íslenskri orðsifjabók* (Ásgeir Blöndal Magnússon 1989), þótt stöku aðskeyti séu uppflettiorð í *Íslenskri orðabók*. Loks eru merkingarlegir þættir ekki settir fram á skipulagðan hátt þó að merkingarskýringar séu uppistaðan í verkum á borð við *Íslenska orðabók*; reyndar er þessi þáttur stundum svo allsráðandi að svo virðist sem gert sé ráð fyrir að notandinn viti bókstaflega allt um orðið sem hann leitar að, nema hvað það merkir. Oft er engar aðrar upplýsingar um uppflettiorðið að finna en merkingarskýringuna eina og notandinn er stundum engu nær um það hvernig orðið er notað.

Notendahópur íslenskra orðabóka er lítill og þess vegna þarf almenn íslensk móðurmálsorðabók að gegna mörgum hlutverkum í einu. Í hana leita allir, bæði skólafólk, innlent og erlent, almenningur og fræðimenn. Hún þarf því helst að vera nothæf bæði til skilnings og skrifta, þ.e. sem aktív og passív orðabók í senn. Lágmarkskröfur til almennrar íslenskrar orðabókar eru því talsvert miklar:

Til að orðabók komi að fullum notum þarf hún að innihalda nægilegar upplýsingar til að notandanum sé kleift að nota flettiorðin á réttan hátt. Því verður að spyrja: Hvaða þættir þurfa að koma fram í orðabókarflettu; og hvernig á að setja þá fram?

(Eiríkur Rögnvaldsson 1998:25)

Grein Eiríks sem þarna er vitnað til er byggð á erindi sem flutt var á málþingi um almenna íslenska orðabók sem haldið var í tilefni af því að endurskoðun á *Íslenskri orðabók* vegna 3. útgáfu var á mótunarstigi. Í greininni setur Eiríkur fram hugmyndir um það hvernig bæta

<sup>3</sup>Með *röklið* er hér átt við skyldubundna nafnliði sem fylgja sögn, t.d. í stöðu frumlags, andlags eða í forsetningarlið.

má ýmsa ágalla í framsetningu málfræðiupplýsinga í bókinni. Hann bendir á að beygingarlýsinguna í bókinni megi laga með því að samræma efnið betur og bæta þurfi inn bráðnauðsynlegum upplýsingum um grunnatriði í sambandi við setningarstöðu orða. Ábendingar Eiríks voru hafðar að leiðarljósi við endurskoðun *Íslenskrar orðabókar*, að svo miklu leyti sem unnt var vegna tímaskorts.<sup>4</sup>

Þótt stefnt sé að því takmarki að setja beyginga- og setningagerðarupplýsingar fram á kerfisbundinn hátt er ekki sjálfgefið hvaða leið á að fara. Þar þarf að athuga hvað gera má ráð fyrir að notandinn kunni og hvað komi honum að mestu gagni.

## 4 Málfræði í orðabókum: Þrjár höfuðleiðir

Í handbók um orðabókarfræði (Hausmann o.fl. 1989:518) setur Joachim Mugdan fram þrjár höfuðleiðir í framsetningu beygingarupplýsinga í almennum móðurmálsorðabókum. Hann tekur fram að valið milli þessara aðferða markist af málkunnáttu notendahópsins en orðabókarmaðurinn velur sér einnig leið eftir hugmyndum sínum um sambandið á milli málfræðibóka og orðabóka og verkaskiptingu þar á milli. Hér verða þessar hugmyndir Mugdans skoðaðar í ljósi íslenskra gagna. Fyrst eru beygingaratriði skoðuð en síðan er vikið að upplýsingum um setningargerð en áhugavert er að skoða hvort hægt er að beita sömu aðferðafræði þar.

### 4.1 Beygingarlýsing í orðabókum

Leiðir Mugdans eru miðaðar við framsetningu beygingarupplýsinga í almennri móðurmálsorðabók þar sem hægt er að gera ráð fyrir tiltölulega einsleitum hópi notenda. Leiðirnar þrjár eru þessar:

1. Beygingarlýsing á heima í málfræðibókum.
2. Lýsa þarf óreglulegri beygingu og undantekningum. Gert er ráð fyrir að notandinn kunni reglulegar beygingar eða geti fundið þær í málfræðibókum.
3. Orðabókin ein á að duga til að finna beygingu orða.

<sup>4</sup>Sjá formála 3. útgáfu *Íslenskrar orðabókar* og grein mína um endurskoðun á sagnlýsingunni í verkinu (Kristín Bjarnadóttir 2001).

Notendahópur *Íslenskrar orðabókar* er ekki af sama tagi og hópurinn sem Mugdan miðar við; bókin þarf að þjóna mun breiðari hópi, eins og fram kemur hér að ofan, enda er hún eina íslenska móðurmálsorðabókin sem gefin hefur verið út.

Hér á eftir eru þessar leiðir skoðaðar í ljósi íslenskra gagna. Tekið skal fram að grundvallaratriðin sem talin eru hér að ofan eru ættuð frá Mugdan; eftirfarandi staðfærsla er hins vegar mín eigin.

#### 4.1.1 Beygingarlýsing á heima í málfræðibókum

Jafnvel þótt henta þætti að hafa klára verkaskiptingu milli málfræðibóka og orðabóka þannig að orðabókarnotandinn þyrfti að sækja sér upplýsingar um beygingar í málfræðibækur væri ekki hægt að komast hjá því að merkja flettur þannig að notandinn kæmist á leiðarenda og fyndi orðinu stað í réttum beygingarflokki. Til þess að tryggt sé að uppflattimynd sé einkvæm þarf upplýsingar um orðflokk og kyn nafnorða, t.d. í orðunum *reiði* (hk og kvk) og *leiði* (hk og kvk). Þetta dugar ekki einu sinni í öllum tilfellum þar sem dálítið er til af orðum þar sem beygingarflokkarnir sjálfir skilja á milli orða, t.d. í samhljóma veikum og sterkum sögnum (*brenna*), og í nafnorðum eins og *ró* þar sem beyging er mismunandi eftir merkingu og kenniföllin ein duga til að skilja á milli orða. Í merkingunni 'kyrrð, friður' er *ró* eintöluorð og eignarfall eintölu er *róar/rór* en *ró* í merkingunni 'stálplata . . . með gati' er líka til í fleirtölu og kenniföllin eru *róar* (ef.et.) og *rær* (nf.ft).

Lágmarksupplýsingar um beygingar í íslensku til þess að hægt sé að **finna** upplýsingarnar í málfræðibókum eru því samkvæmt þessu annað tveggja:

1. Kennimyndir og kenniföll (ásamt orðflokki og kyni), a.m.k. í einsrituðum flettiorðum.
2. Tákn (fyrir beygingarflokka) sem vísa í málfræðibók.

Íslenska hefðin er að sýna kenniföll og kennimyndir, eins og gert er í *Íslenskri orðabók*, en dæmi eru um síðari aðferðina líka. Í *Rússnesk-íslenskri orðabók* eftir Helga Haraldsson (1996) er t.d. notað flókið tákna-kerfi, sem vísar í ítarlega beygingarlýsingu aftast í bókinni. Framsetningin er að því leyti óvenjuleg að beygingarupplýsingar fylgja báðum málunum, bæði rússnesku og íslensku.

Báðar þessar leiðir hafa galla. Táknakerfið gerir talsverðar kröfur til notandans vegna þess hve ógagnsætt það er og ósennilegt er að notendum verði það verulega tamt nema með mjög mikilli og stöðugri notkun. Táknakerfið vísar samt nákvæmlega á réttan stað, ef notandinn hefur í sér einurð til að fylgja því eftir og málfræðin sjálf er sett fram á nægilega ítarlegan máta. Kenniföll og kennimyndir gefa í verunni heldur meiri upplýsingar en ætla mætti ef allar málfræðiupplýsingar eiga heima í málfræðibók en án þessara upplýsinga er stundum ekki hægt að greina samhljóða orð í sundur í les. Samt duga kenniföll og kennimyndir nafnorða og sagna ekki alltaf til þess að hægt sé að ganga að allri beygingu uppflettiorðs vísri í málfræðibókum. Kenniföll og kennimyndir eru ekki einu beygingarmyndir orða sem eru ófyrirsegjanlegar út frá stofngerð og í málfræðibókum er yfirleitt aðeins gefið yfirlit um beygingarflokka, með mismörgum dæmum. Tæmandi er lýsingin aldrei og notandinn er því stundum engu nær um þá beygingarmynd sem hann vill fá. Nægir þar að nefna þágufall eintölu í sterkum karlkynsnafnorðum (með eða án *-i*) og *-na* eða *-a* í eignarfalli fleirtölu af veikum kvenkynsnafnorðum. Loks má geta þess að lítið stoðar að fara í málfræðibók til að finna ítarlega beygingarlýsingu íslenskra lýsingarorða; hana er hvergi að finna á bók.<sup>5</sup>

#### 4.1.2 Ófyrirsegjanleg beyging birt í orðabókinni

Ein leið til að nýta málkunnáttu notandans til að ákvarða hvaða beygingarupplýsingar eiga heima í orðabók er að gera ráð fyrir því að stærstu beygingarflokkar séu sjálfgildi í lýsingunni og upplýsingar um þá þurfi ekki að birta. Stærstu flokkarnir eru jafnframt opnir og ómarkaðir beygingarflokkar, sbr. greiningu Ástu Svavarsdóttur í bókinni *Beygingakerfi nafnorða í nútímaíslensku* (1993), t.d. veikar sagnir sem fá endinguna *-aði* í þátíð, veik beyging karlkyns- og kvenkynsnafnorða o.s.frv. Útkoman verður þá þessi:

<sup>5</sup>Samkvæmt reynslunni við vinnu við *Beygingarlýsingu íslensks nútímamáls* (sjá vefsíðu Orðabókar Háskólans: [www.lexis.hi.is](http://www.lexis.hi.is)) þarf fimm beygingarmyndir til þess að skipta lýsingarorðum upp í beygingarflokka svo að ótvírætt sé. Þessar orðmyndir ættu því að geta gegnt sama hlutverki og kenniföll nafnorða og kennimyndir sagna. Orðmyndirnar eru nefnifall eintölu í öllum kynjum í sterkri beygingu, ásamt miðstigi og efastigi (kk.nf.et.sb.), t.d. *góður, góð, gott, betri, bestur; heitur, heit, heitt, heitari, heitastur; stuttur, stutt, stutt, styttri, stýstur; háður, háð, háð, háðari, háðastur*. Þessar upplýsingar hafa ekki verið settar fram á skipulegan hátt í orðabók og hér á eftir verða nafnorð og sagnir látin duga sem dæmi um beygingarupplýsingar í orðabókum.

1. Regluleg beyging er ekki birt:  
Beygingin er fyrirsegjanleg út frá orðflokki, kyni og stofngerð, t.d.

**hest | ur** KK, **penn | i** KK, **bjór |** KK, **mó | r** KK  
**rústa** S

2. Allar aðrar beygingar eru birtar í orðabókinni (á einhvern hátt), t.d.

**gest | ur (-s, -ir)** KK  
**fjörð | ur (fjarðar, firðir, þgftet. firði)** KK  
**keyra (keyrði keyrt)** S

Þarna er notendum ætluð talsverð þekking á beygingarkerfinu og ekki er víst að allir notendur geri sér ljóst hvaða merkingu það hefur þegar engar beygingarupplýsingar fylgja orði. Íslenskir notendur orðabókar þar sem þessi leið er farin eru reyndar ekki mjög líklegir til að koma að tómunum kofunum; það er fremur ósennilegt að þeir þurfi að fletta upp kenniföllum eða kennimyndum orða úr opnum (reglulegum) beygingarflokkum. Ófyrirsegjanlegar myndir þyrfti samt alltaf að birta, þ.m.t. þágufall eintölu af sterkum karlkynsorðum og nefnifall fleirtölu af veikum kvenkynsorðum. Þar virðast engar reglur duga til að ákvarða hver beygingin er.<sup>6</sup>

Þessi leið byggir m.ö.o. á því að notendur viti hvaða merkingu það hefur ef enginn beygingarlýsing fylgir orði. Málfræðiupplýsingarnar sem fylgja þessum orðum eru þá aðeins orðflokkur, kyn nafnorða og táknun á stofngerð, t.d. með lóðréttu striki á undan nefnifallsendingu í karlkynsnafnorðum.

#### 4.1.3 Orðabókin á að duga ...

Þriðja leið Mugdans í framsetningu beygingarupplýsinga í orðabókum er að gera kröfu um að orðabókin ein dugi til þess að notandinn fái fullnægjandi upplýsingar um beygingu orðs, án þess að gert sé ráð fyrir mikilli málkunnáttu. Í íslensku væri hægt að gera þetta á tvo vegu:

<sup>6</sup>Þetta kom skýrt í ljós við vinnu við *Beygingarlýsingu íslensks nútímamáls* á Orðabók Háskólans. Þar eru settar fram tilgátur um þessar beygingarmyndir orða, byggðar á talsverðum rannsóknnum, en rækileg úttekt á þessu er ekki til.



1. Með því að birta kenniföll og kennimyndir af öllum orðum, auk óreglulegra mynda
2. Með því að birta heil beygingardæmi . . .

Fyrri liðurinn hér er mjög svipaður síðari lið í næsta kafla hér á undan, þ.e. að birta allar kennimyndir og kenniföll og annað efni sem þarf til að gefa nægjanlegar upplýsingar um ófyrirsegjanlegar myndir. Munurinn er þó sá að hér eru birtar upplýsingar um alla beygingarflokka og ekki treyst á málkunnáttu notandans. Hin leiðin er að birta beygingardæmi í heild en sú leið hefur til skamms tíma ekki verið fær í orðabókum vegna plássleysis.

#### 4.1.4 . . . en fyrir hvern á orðabókin að duga?

Valið milli þeirra kosta sem hér hafa verið settir fram er e.t.v. ekki lengur jafnmikið lykiltríði í orðabókargerð og verið hefur. Orðabókarhefðin hefur að verulegu leyti mótast af þeirri nauðsyn að spara pláss eftir fremsta megni, stundum svo að það verður til verulegra vandræða fyrir notendur. Notkun á táknakerfi í framsetningu á beygingarupplýsingum sparar meira pláss en nokkur önnur leið, en hefur líka galla eins og Helgi Haraldsson bendir á í formála *Rússnesk-íslenskrar orðabókar* sem nefnd er sem dæmi um verk þar sem beygingarupplýsingar eru gefnar sem tilvísanir í sjálfstæða málfræði:

Margvíslegum brögðum er beitt til að spara rými, og kann þetta í fyrstunni að vefjast fyrir sumum.

(Helgi Haraldsson 2002:ii)

Víst er að það reynir á þolinmæðina að læra á táknakerfi eins og það sem Helgi notar, en hann hefur það til síns máls að rússneskumælandi fólk mun vera vant að nota táknakerfi af þessu tagi í orðabókum og vissulega sparar hann með þessu umtalsvert rými í bókinni sem er mjög stór samt.

Jafnvægislistin sem þarf til þess að koma öllum nauðsynlegum upplýsingum um beygingu orða fyrir í orðabók án þess að umfangið fari úr böndunum er mikil kúnst, sérstaklega þegar gera þarf ráð fyrir jafnfjölbreyttum notendahópi og við útgáfu almennrar íslenskrar orðabókar. Í rafrænni orðabók, hvort heldur er á diskum eða á vefnum, eru möguleikarnir allt aðrir. Þar er jafnvel hægt að koma efninu

Þannig fyrir að notandinn sjálfur geti ákveðið hvers konar upplýsingar henta honum best og sótt sér þær að vild. Möguleikarnir eru þá þeir sem taldir eru upp hér á undan, allt frá því að sleppa öllum beygingarupplýsingum öðrum en þeim sem nægja til að uppflettimynd sé einkvæm, yfir í að sýna heil beygingardæmi þar sem allar beygingarmyndir og öll afbrigði eru sýnd. Þar gefst jafnvel kostur á leiðbeiningum til notandans, um málfar, merkingarmun beygingarmynda o.s.frv. Skipting Mugdans er þá úr sögunni en í staðinn koma viðfangsefni fólgin í því að finna framsetningu sem er nægilega kerfisbundin og samt nægilega sveigjanleg til þess að þola breytingar á birtingarformi eftir því hvort notandinn er skólanemi, fræðimaður eða jafnvel tölva.

## 4.2 Setningafræði í orðabókum

Orðabókarnotendur á Íslandi ætlast til að finna upplýsingar um beygingu orða í orðabókinni sinni; það er okkar hefð. Upplýsingar um setningargerð eru mun gloppóttari í íslenskum orðabókum, eins og sjá má af gátlista í grein Eiríks Rögnvaldssonar (1998:32) sem einnig er vitnað í hér að framan (sjá 3. hluta). Á listanum eru helstu atriði sem Eiríkur taldi ástæðu til að endurskoða í *Íslenskri orðabók* vegna 3. útgáfu:

### Gátlisti ER við endurskoðun *Íslenskrar orðabókar*:

- Fjöldi rökliða (áhrifsgildi) sagna þarf að koma skýrt fram.
- Hvaða rökliðir eru skyldubundnir og hverjum má sleppa.
- Hvort og hvenær forsetningarliðir geta komið í stað andlaga.
- Fallstjórn sagna þarf að taka sérstaklega fram.
- Miðmynd þarf mjög oft að vera sjálfstætt flettiorð.
- Lýsingarháttum þarf að gera mun hærra undir höfði.
- Skerpa þarf notkun táknunarinnar ÓP.
- Takmarkanir á setningarstöðu orða verða að koma fram.

Meðal þeirra atriða sem *Íslensk orðabók* hafði einmitt helst verið gagnrýnd fyrir voru gloppóttar og ófullnægjandi upplýsingar um setningargerð, eins og Eiríkur bendir á. Í þessum lista eru þau atriði sem venjulegur málnotandi þarf að kunna skil á til þess að nota orðaforðann í bókinni, ekki bara til skilnings heldur líka til málnotkunar.

Talsverðar breytingar voru gerðar á framsetningu setningargerð-

arupplýsinga í 3. útgáfu *Íslenskrar orðabókar* og var reynt að samræma eftir bestu getu og bæta við efni eins og tími gafst til (sjá Kristín Bjarnadóttir 2001). Í ljósi þeirra breytinga er áhugavert að skoða höfuðleiðir Mugdans í framsetningu beygingarupplýsinga og hvort skipting af svipuðu tagi varpar ljósi á vandamál í framsetningu setningargerðarupplýsinga í íslenskum orðabókum.

#### 4.2.1 Setningafræði á heima í málfræðibókum

Það að vísa orðabókarnotendum algerlega á setningafræðibækur í leit að setningargerðarupplýsingum einstakra orða er einmitt stefnan sem Eiríkur Rögnvaldsson gagnrýnir með framsetningu gátlistans síns. Sem dæmi um þessa framsetningu má nefna ótölulegan fjölda sagna í 1. og 2. útgáfu *Íslenskrar orðabókar* (1963, 1983) þar sem engir rökliðir eru sýndir með sögnum, hvorki sem formlega skilgreindur setningafræðihluti né í dæmum. Af sama toga eru sagnir sem aðeins eru merktar með ÓP án þess að tiltekið sé hvort um aukafallsfrumlag eða gervifrumlag er að ræða.<sup>7</sup>

Bækur um setningafræði eru, ef eitthvað er, enn verri heimildir um málfræði einstakra orða en bækur um beygingarfræði. Í setningafræðibókum er sjónarhornið alltaf „ofan frá og niður“ og fjallað er um hverja setningargerð fyrir sig. Þess vegna er útilokað að finna tæmandi úttekt á rökliðum einstakra sagna og öðru efni af sama toga. Í þessu kemur fram eðlismunur á umfjöllun um beygingu og umfjöllun um setningargerð í málfræðibókum. Beygingardæmi eru venjulega skoðuð sem ein heild, þótt vissulega sé líka fjallað um einstök föll. Sögnum er hins vegar yfirleitt ekki skipt upp í setningargerðarflokka sem jafnast á við beygingarflokkana heldur er fjallað um hverja setningargerð fyrir sig. Það er því fremur lítil von um árangur við leit í málfræðibók þegar orðabókin gefur ekki fullnægjandi upplýsingar um rökliði sagna.

#### 4.2.2 Ófyrirsegjanlegar setningargerðir birtar í orðabókinni

Hér er hugmyndin sú að sleppa því að lýsa öllu sem er fyrirsegjanlegt út frá meginreglum íslenskrar setningargerðar og sleppa með því

<sup>7</sup>Sjá umfjöllun um sögnina *að fatra* í grein minni í *Orði og tungu* 5 (Kristín Bjarnadóttir 2001:91).

öllum upplýsingum um setningarfræðileg sjálfgildi. Þar er t.d. átt við að ekki þurfi að sýna formgerðarföll, en orðasafnsbundin föll þurfi að sýna.

Ef þessi leið væri notuð til hins ítrasta þyrfti ekki að gefa upplýsingar um frumlag í nefnifalli með sögnum, og eftir kenningunni ætti líka að vera óhætt að sleppa upplýsingum um andlag í þolfalli, þar sem þetta eru formgerðarföll. Þessi leið er fær með nefnifallið; hún var t.d. notuð í *Sýnihefti sagnorðabókar* (Ásta Svavarsdóttir o.fl. 1993) þar sem nafnháttur er notaður til að sýna setningargerðir þar sem frumlagið er sjálfgefið:

**þinga:** ...*halda fund e. ráðstefnu*

Í *Sýniheftinu* var hins vegar tekinn sá kostur að sýna þolfallsandlag með skammstöfun fyrir fallmynd rökliðarins þótt segi megi að fallið sé sjálfgildi þar sem gera þurfti greinarmun á áhrifssögn með þolfallsandlagi og áhrifslausri sögn. Þágufallsandlag þurfti svo að sýna, bæði sem orðasafnsfall og sem röklið með áhrifssögn:

**þeyta:** *hræra e. slá hratt*

**þeyta e-ð:** *hræra eða slá e-ð hratt ...*

**þeyta e-u:** *kasta e-u; losa sig við e-ð*

**þeyta e-m ...:** *kasta e-m, hrinda e-m*

Ekki er raunhæft að gera ráð fyrir að almennir notendur þekki muninn á formgerðarfalli og orðasafnsfalli og munurinn á áhrifssögn og áhrifslausri sögn gerir að engu möguleikann á því að sleppa því að sýna röklið sem er sjálfgilt fall, a.m.k. með þeirri aðferð sem hér er höfð til að sýna rökliði.

Í *Sýniheftinu* er einn merkingarþáttur í rökliðum sagna sýndur um leið og fallstjórnin, þ.e. +/-lifandi, eins og sjá má í dæmunum hér á undan. Sá háttur var tekinn upp að nokkru leyti við endurskoðun *Íslenskrar orðabókar* (2002), þannig að frumlagi er sleppt ef það er í nefnifalli og vísar til lifandi veru. Nokkur misbrestur er þó á því að svona upplýsingar séu fullkomlega samræmdar. Þarna er samt vísir að þeirri aðferð sem hér er verið að skoða og vert er að athuga hvernig upplýsingar af þessu tagi nýtast almennum notendum.

### 4.2.3 Og orðabókin á að duga ...

Samkvæmt þessari skipan á orðabókin ein að duga til að finna mögulega setningarstöðu orða. Til þess að þessu skilyrði sé fullnægt þarf að sjá til þess að a.m.k. öll atriðin á gátlista Eiríks sem vitnað er til hér að ofan komi fram. Orðtengslabækur Jóns Hilmars Jónssonar, *Orðastaðurur* (1994) og *Stóra orðabókin* (2005), leysa þennan vanda að verulegu leyti en almenn íslensk orðabók með merkingarskýringum þar sem á þessu er tekið að fullu er ekki til. Til glöggvunar fylgir hér tilbúið sýnishorn með lágmarksupplýsingum um alla rökliði þar sem notuð er sama framsetning og í *Sýniheftinu* en það var haft til fyrirmyndar þegar sagnlýsingin í *Íslenskri orðabók* var endurskoðuð:

**e-r þeytir:** *e-r hrærir e. slær hratt*

**e-r þeytir e-ð:** *e-r hrærir eða slær e-ð hratt ...*

**e-r þeytir e-u:** *e-r kastar e-u; e-r losar sig við e-ð*

**e-r þeytir e-m ...:** *e-r kastar e-m, e-r hrindir e-m*

Til þess að setja fram svona lýsingu svo að vel fari þyrfti líka að gera grein fyrir skyldubundnum atviksorðum, forsetningarliðum o.þ.h.

Höfuðleiðir Mugdans við lýsingu á beygingu orða í orðabókum bregður ljósi á vandamálin í framsetningu setningargerðarupplýsinga. Þar hefur glíman við plássleysið orðið til þess að málfræðihluti orðabókargreinanna er stórlega vanræktur og lausnin er ekki fólgin í málfræðibókum. Vandamál orðabókarmannsins er einnig sá að ekki er alltaf auðvelt að finna nauðsynleg gögn til þess að setja fram setningarlegar upplýsingar þannig að vel fari og gagn sé að.

## 5 Framsetning málfræðiupplýsinga í orðabókum

Mismunurinn á upplýsingum um beygingar og setningargerð í íslenskum orðabókum er að hluta fólgin í framsetningunni. Upplýsingar um beygingar eru alltaf settar fram á formlegan hátt og þær eiga sér fastan stað í flettugreininni. Setningargerðarupplýsingar birtast hins vegar ýmist sem svokölluð setningargerðardæmi, t.d. *gefa e-m e-ð*, eða sem eiginleg dæmi, t.d. *skrifa skýrt og rétt* (sjá Kristín Bjarnadóttir 2001 þar sem gerður er greinarmunur á eiginlegum dæmum, setningargerðardæmum og setningargerðarhausum sem hafa sjálfstætt uppflöttigildi). Oftar en ekki er setningargerðarlýsingin glöppótt

og ófullnægjandi, þ.e. hún nægir engan veginn til þess að orðabókarnotandinn geti notað orðið með stuðningi af orðabókarlýsingunni. Til þess að setningargerðarupplýsingar komi að sama gagni og beygingarupplýsingarnar þarf að setja þær fram á jafnkerfisbundinn hátt. Sjálf framsetningin skiptir einnig máli og tengist nokkuð aðferðum Mugdans sem hér hafa verið skoðaðar. Þar skiptir notendahópurinn meginmáli og viðhorf orðabókarmannsins til notandans.

Framsetning málfræðiatríða í orðabókum er einkum með þrennu móti. Í fyrsta lagi eru notaðir kótar eða tákn sem vísa í skýringar í leiðbeiningum í orðabókinni; í öðru lagi eru notaðar skammstafanir fyrir málfræðihugtök sem ætlast er til að notandinn hafi á valdi sínu; og í þriðja lagi eru notaðar skammstafanir fyrir orð sem gegna því hlutverki að sýna sama efni og felst í málfræðihugtökunum, t.d. *e-u* sem stytting á *einhverju* sem stendur í stað *þgf* eða *DAT*. Fjórða leiðin, og sennilega sú notendavænasta, er ekki algeng en hún er að sýna dæmigerða orðnotkun sem vísbendingu um setningargerð, eins og Jón Hilmar Jónsson gerir í bókum sínum. Með þessu móti fær notandinn mjög miklar upplýsingar um málnotkun, bæði setningarlegar og merkingarlegar, og ætti að vera fullfær um að nota orð sem svona eru sýnd.

Þessar leiðir gera mjög mismiklar kröfur til notandans, rétt eins og orðabókartexti sem byggður er á aðferðum Mugdans, og svigrúm fyrir málfræðiupplýsingarnar sem hægt er að koma fyrir er líka mjög misjafnt:

### Hve abstrakt má framsetningin vera?

Kótar, skammstafanir eða mannamál?

<i>Rússnesk-íslensk orðabók</i>	hneppa 2T
<i>Nimb</i> (2004)	formentlig_1 Dd1_PP
<i>tilbúið dæmi</i>	gefa DAT ACC
<i>Íslensk orðabók</i> (2002)	gefa e-m e-ð
<i>Orðastaður</i>	gefa <honum, henni><eitthvað; hlut, dýrgrip, peninga> (<fyrir greiðann, að launum, í þakklætis- skyni, að skilnaði>); ...

Tvö dæmi skera sig úr vegna þess hve miklar upplýsingar eru í þeim fólgnar, þ.e. dæmið frá Sanni Nimb sem sýnir eina gerð af setningu-

arstöðu danska atviksorðsins *formentlig* og dæmið úr *Orðastað*. Framsetningin er samt með gjörólíkum hætti; framsetningin í *Orðastað* miðar að því að gera notandanum lífið sem léttast og þar er að auki rúm fyrir blæbrigði í merkingu sem birtist í orðavali fyrir rökliði með sögninni sem þarna er sýnd. Dæmið frá Sanni Nimb er úr doktorsritgerð hennar um setningarstöðu atviksorða í dönsku (Nimb 2004:230) en í því verki er sett fram formleg greining á setningarstöðu af mikilli nákvæmni, eins og sést í flettugreininni fyrir orðið í heild:

formentlig\_1 Dd1\_PP

*Formentlig for første gang herhjemme, blev en reklamespot sendt ud i æteren til ære.*

formentlig\_2 Dd1\_PP

*Formentlig et andet forhold gør, at en række østeuropæiske mærker også trygt kan stå på gaden.*

formentlig\_3 Dd1FTe\_S

*Vi har formentlig kun set toppen af isbjerget af konsekvenser endnu.*

Þetta dæmi er sýnt hér til þess að vekja athygli á tvennu, þó að efnið í þessari grein sé annars bundið við íslenska orðabókargerð. Í fyrsta lagi mótast framsetningin og greiningin af því að lýsingin er hluti af tungutækniverkefni þar sem fengist er við öll svið setningafræðinnar og er efnið ætlað í gagnabanka sem síðar má vinna orðabókarflettur upp úr. Í öðru lagi nægir þetta dæmi til að benda á að verkefnin í orðabókarlýsingu fyrir íslensku eru mörg hver óunnin; í *Íslenskri orðabók* (2002) vottar til dæmis ekki fyrir kerfisbundinni lýsingu á setningarstöðu atviksorða.

Greining af því tagi sem Sanni Nimb setur fram er sennilega ekki sérlega vel til þess fallin að létta venjulegum orðabókarnotanda lífið; til þess er hún of tæknileg. Hún er hins vegar vísbending um það sem hægt er að gera með því að nota stór textasöfn með nákvæmri greiningu. Framsetning hverrar orðabókar er svo eftir sem áður á valdi orðabókarmannsins sem þarf að taka fullt tillit til notenda sinna og nýta allar þær aðferðir sem nú standa til boða. Orðabækur eru ekki lengur í því hafti sem prentuð bók skapar; nú er veigamesti þátturinn í allri framsetningu ekki að spara pláss heldur að ganga frá öllum upplýsingum á kerfisbundinn hátt.

## Heimildir

- Ásgeir Blöndal Magnússon. 1989. *Íslensk orðsifjabók*. Reykjavík: Orðabók Háskólans.
- Ásta Svavarsdóttir. 1993. *Beygingakerfi nafnorða í nútímaíslensku*. Reykjavík: Málvísindastofnun Háskóla Íslands.
- Ásta Svavarsdóttir, Guðrún Kvaran, Jón Hilmar Jónsson og Kristín Bjarnadóttir. 1993. *Sýnihefti sagnorðabókar*. Rannsóknar- og fræðslurít 3. Reykjavík: Orðabók Háskólans.
- Eiríkur Rögnvaldsson. 1998. Málfræði í íslenskri orðabók: Hvernig og til hvers? *Orð og tunga* 4:25–32.
- Hausmann, Franz Josef, Oskar Reichmann, Herbert Ernst Wiegand, Ladislav Zgusta (eds.). 1989. *Dictionaries. An International Encyclopedia of Lexicography*. Berlin: Walter de Gruyter.
- Helgi Haraldsson. 1996. *Rússnesk-íslensk orðabók*. Reykjavík: Nesútgáfan.
- Höskuldur Þráinsson. 1995. *Handbók í málfræði*. Reykjavík: Námsgagnastofnun.
- Íslensk orðabók handa skólum og almenningi*. 1963. Ritstjóri: Árni Böðvarsson. Reykjavík: Bókaútgáfa Menningarsjóðs. [1. útg.]
- Íslensk orðabók handa skólum og almenningi*. 1983. Önnur útgáfa aukin og bætt. Ritstjóri: Árni Böðvarsson. Reykjavík: Bókaútgáfa Menningarsjóðs.
- Íslensk orðabók*. 2002. Þriðja útgáfa aukin og endurbætt. Ritstjóri: Mördur Árnason. Reykjavík: Edda hf.
- Jón Hilmar Jónsson. 1994. *Orðastaður. Orðabók um íslenska málnotkun*. Reykjavík: Mál og menning.
- Jón Hilmar Jónsson. 2002. *Orðaheimur. Íslensk hugtakaorðabók*. Reykjavík: JPV útgáfa.
- Jón Hilmar Jónsson. 2005. *Stóra orðabókin um íslenska málnotkun*. Reykjavík: JPV útgáfa.
- Kristín Bjarnadóttir. Um sagnlýsinguna í 3. útgáfu Íslenskrar orðabókar. *Orð og tunga* 5:87–114.
- Mugdan, Joachim. 1989. Information on Inflectional Morphology in the General Monolingual Dictionary. Í: Hausmann o.fl. 1989, bls. 518–525.
- Nimb, Sanni. 2004. *Danske adverbier mellem leksikon og syntaks*. Doktorsritgerð við Kaupmannahafnarháskóla.



Sigfús Blöndal. 1920–1924. *Íslensk-dönsk orðabók*. Reykjavík. [Blöndals-orðabók.]

Sjá líka *Beygingarlýsingu íslensks nútímamáls* á vefsetri Orðabókar Háskólans: [lexis.hi.is](http://lexis.hi.is)

## Abstract

The topic of the paper is the scope and presentation of grammatical information in Icelandic dictionaries. Grammar (or linguistics) is the foundation on which dictionaries rest, but there is nevertheless a fundamental difference between a grammar and any kind of dictionary. The topic of a grammatical description is the structure of the language in question, focusing on the whole, i.e. top-down. The structure of a dictionary is based on individual entries, and the focus is bottom-up. Taking information on inflectional morphology and the transitivity and case assignment of verbs as examples, three different methods of division of labour between grammar and dictionary are discussed. The conclusion is that the dictionary needs to provide its own stand-alone description of grammatical features, with sufficient detail to provide for the user's needs, because of the difference in focus between a dictionary and a grammar.

## Lykilorð:

málfræði í íslenskum orðabókum  
beyging í íslenskum orðabókum  
rökliðir sagna í íslenskum orðabókum

## Keywords:

grammatical information in Icelandic dictionaries  
inflectional morphology in Icelandic dictionaries  
transitivity and case assignment in Icelandic dictionaries

*Kristín Bjarnadóttir*  
*Orðabók Háskólans*  
*Neshaga 16*  
*Reykjavík*  
*kristinb@lexis.hi.is*