

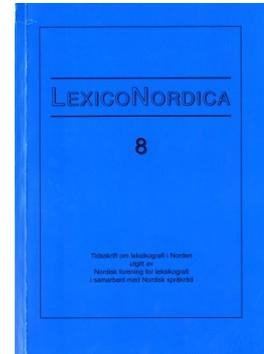
# LexicoNordica

Titel: Verbal Syntax in an Electronic Bilingual Icelandic Dictionary: A Preliminary Study

Forfatter: Kristín Bjarnadóttir

Kilde: LexicoNordica 8, 2001, s. 5-23

URL: <http://ojs.statsbiblioteket.dk/index.php/lexn/issue/archive>



© LexicoNordica og forfatterne

## Betingelser for brug af denne artikel

Denne artikel er omfattet af ophavsretsloven, og der må citeres fra den. Følgende betingelser skal dog være opfyldt:

- Citatet skal være i overensstemmelse med „god skik“
- Der må kun citeres „i det omfang, som betinges af formålet“
- Ophavsmanden til teksten skal krediteres, og kilden skal angives, jf. ovenstående bibliografiske oplysninger.

## Søgbarhed

Artiklerne i de ældre LexicoNordica (1-16) er skannet og OCR-behandlet. OCR står for 'optical character recognition' og kan ved tegngenkendelse konvertere et billede til tekst. Dermed kan man søge i teksten. Imidlertid kan der opstå fejl i tegngenkendelsen, og når man søger på fx navne, skal man være forberedt på at søgningen ikke er 100 % pålidelig.

Kristín Bjarnadóttir

## Verbal Syntax in an Electronic Bilingual Icelandic Dictionary: A Preliminary Study

Denne artikkelen handler om verbbeskrivelsen i en elektronisk utgave av den enspråklige islandske standardordboken, *Íslensk orðabók* (2000), og hvordan den kan utnyttes i tospråklig sammenheng. Hovedforskjellen mellom en trykt og en elektronisk ordbok ligger i de ulike presentasjonsmåtene, i et bokformat er det tekstens omfang som bestemmer utformingen av beskrivelsen men i en elektronisk framstilling må det tas hensyn til hvor mye tekst det kan vises på skjermen. Teksten må deles opp i passe store enheter. For å få det til i den elektroniske utgaven av *Íslensk orðabók* føres verbale konstruksjoner opp som sublemmaer under de enkelte verbene. Dette betyr at den syntagmatiske beskrivelse blir mer utførlig og systematisk enn i de eldre utgaver av ordboken. I artikkelen blir verdien av denne framstillingen vurdert, spesielt med hensyn til en tospråklig ordboksbeskrivelse.

### 1. Introduction

The 3rd edition of *Íslensk orðabók* (hereafter *ÍO*), the only extant monolingual Icelandic dictionary, was published on disc in November 2000 by Edda hf.<sup>1</sup> As a printed book this dictionary is immensely popular and very influential; when one hears the words “það stendur í orðabókinni” (“it says so in the dictionary”), this is the dictionary being referred to. This self-same dictionary has also been criticized on various grounds, notably for its lack of syntactic information, as by Eiríkur Rögnvaldsson (1998) and by Jón Hilmar Jónsson (1985).

The changes made from the printed dictionary to the electronic one are in one sense very slight, i.e., in the actual text of the entries, as this is not really a complete revision of the 1983 edition as far as vocabulary, explanations, etc., are concerned. In another sense the changes are quite substantial, i.e. in the format of the entries, as is made inevitable by the difference in media, and, in case of the verbs especially, in the addition of syntactic information. The topic of this paper is a tentative appraisal of how the changes made to this monolingual dictionary could be utilized for other projects in process at Edda, such as in the production of bilingual Icelandic dictionaries. The case in point is an Icelandic-Danish Dictionary being edited by Halldóra Jónsdóttir.

<sup>1</sup> The 1st edition of *ÍO* was published in 1963, the 2nd in 1983; the editor in chief of both was Arni Böðvarsson. The editor in chief of the 3rd edition is Mördur Árnason. The author of this paper was on the editorial team for the 3rd edition from 1997. Sections of the dictionary are accessible for viewing at the website <http://ord.is>.

The changes discussed here are those made to the format of verbs and the idea of using this format from the monolingual dictionary as a base for the bilingual one stems from the very simple reason that no other suitable lexicographic description of Icelandic verbs is available at the moment.<sup>2</sup> The need is acute; bilingual Icelandic dictionaries (i.e. dictionaries where Icelandic is the source language) are either too small or too old – or both.

## 2. From book to disc

The changes made from book to disc in the monolingual dictionary are the base of the speculations that follow on the presentation of the verbs in a bilingual one. These changes are made necessary by the change in medium. These reflect the limitations of the printed page, on the one hand, and the limitations of the computer screen, on the other hand.

### 2.1 *The printed page: The problem of compactness*

The limitation of the printed page is a familiar one to all lexicographers, i.e., the problem of compactness. In the 1983 edition of *ÍO*, all available methods are used to compress the text into as compact a format as possible. The dictionary contains 85.000 entries on 1.259 pages of fine print, the difference in fonts is minimal, and examples are kept to the barest minimum, quite often to the user's grief, according to reviewers (cf. Jónsson 1985).

A list of a few compounds, as shown in figures 1 and 2, demonstrates the difference in layout in print and on the screen, showing the difference in access to the entries, although this is, strictly speaking, a digression from the main topic of this paper.

---

<sup>2</sup> There are approximately 8.500 verbs in the 3rd edition of *IO*. The restructuring of the verbs for the computerized version was the responsibility of the author of this paper.



These examples show the difference in presentation between the two editions, and they pinpoint the problem of space on the printed page, i.e., the problem of density.

## 2.2 The problem of space on the screen: The over-large entry

The screen poses a completely different problem from the printed page, which shows up in the over-large entry. If the problem of the printed book is condensing the text enough to make the dictionary easy to handle, the problem of the screen is making the text clear enough not to overtax the readers' eyes. To that purpose, the text is distributed on the screen, using linebreaks, indentation, colours and fonts as needed. The ideal entry fits one screen or less, as in the noun *bein* in figure 3, to give an idea of the amount of information easily handled in this manner. The size of the window can be varied; this is a simulation for print containing exactly one screenful from a 17 inch screen.

<p><b>bein</b> -s, -HK</p> <p><b>1</b> líffræði/læknisfræði</p> <ul style="list-style-type: none"> <li>sterkasta gerð af stoðvef, hluti af beinagrind</li> </ul> <p><i>beinin í mannslíkamanum eru um 220</i></p> <p><b>2</b></p> <ul style="list-style-type: none"> <li>fótleggur</li> </ul> <p><b>sitja flötum beinum</b></p> <p><b>3</b></p> <p>í ýmsum orðasamböndum</p> <p><b>bera beinin, bera bein sín deyja</b>  <b>bíta bein fyrir sig</b> sjaldgæft eiga auðvelt með að svara fyrir sig  <b>hafa bein (beinin) til e-s</b> vera nægilega sterkur til e-s  <b>hafa bein í nefi (nefinu)</b> vera ákveðinn, láta ekki hlut sinn  <b>hafa mikið bein í hendi</b> vera völdugur  <b>hvíla lúin bein</b> hvílast, einkum eftir göngu eða áreynslu</p> <p>...</p> <p><b>4</b> sjómennska</p> <ul style="list-style-type: none"> <li>leggur notaður sem sökkvibúnaður á neti, beinatygill, leggjateinn</li> </ul> <p><b>5</b></p> <p>um fisk, aflabrogð</p> <p><b>fá ekki bein (úr sjó)</b> veiða ekki neitt</p> <p><b>6</b></p> <ul style="list-style-type: none"> <li>bitlingur</li> </ul> <p><b>7</b></p> <p>hnjóðsyrði um mann</p> <p><b>bannsett beinið</b></p>
---

FIGURE 3. The entry for *bein* (ÍO 2000)

The problem of space arises when the entry gets to be too long to fit one screen, especially in the really long entries, such as those for some common verbs. The verb *taka* ('take') for example, stretches to almost 3½ pages in the 1983 edition of *ÍO*. Browsing through the same entry on the screen by using 'Page Down' gives 20 screenfuls. Keeping track of information that way is exceedingly difficult.

These long entries therefore need to be broken up into manageable parts, and for the verbs that is done by making sub-entries of verbal phrases (particles and prepositional phrases), as well as of some verbal forms, such as the mediopassive, present and past participles, etc. These forms are shown in an index on the right of the screen in a separate window, and the items there appear on the screen at a click. A simulation of this can be seen in figure 4, i.e., the first screenful from the verb *vera* ('to be').

<p><sup>3</sup><b>vera</b> var, vorum/vórum, verið (nt. er, (1. p. fornt/úrelt em) vð. sé, (einnig veri sem „óskháttur“) vð. þt. væri; í elstu textum er -s- í nt. og þt.et.: vesa, vas o.s.frv.) s</p> <p>1</p> <p>* eiga sér stað, vera til, gerast, vara, haldast það var eitt sinn orustan var á öðrum degi vikunnar er ekki svo? er það ekki rétt?</p> <p>2</p> <p>* dveljast vera e-s staðar</p> <p>* dveljast, gista um nætursakir, vera staddur lofa e-m að vera hýsa e-n yfir nótt hann er á fundi vera af jólin dveljast um jólin vera um nótt gista um nætursakir fari það og veri þ.e. fari það grákollótt</p> <p>3</p> <p>(til að tengja við frumlag setningar það sem sagt er um það (tengja sagnfyllingu við frumlag)) hún er stór hún er góð hún er vondur Jón er bóndi hann er sjómaður það væri gaman</p> <p>4</p> <p>(í ýmsum merkingum) hún er stærri ef nokkuð er þ.e. ef þær eru þá ekki jafnar, ef e-u munar það er ekki margra ekki eru margir færir um það er á meðan er þ.e. það varir meðan það stendur yfir, þ.e. sagt um e-ð sem búast má við að standi skemur en æskilegt er það varð (hlaut) að vera það gat ekki verið óðrúfisi það verður svo að vera ekki er unnt að breyta því það er það sem er þ.e. það er nú meinið hver á að vera hann? hver á að vera tilekinn aðili? ((leik, t.d. sá sem eltir í eltingarleik) láta sem ekki (ekkert) sé láta sem ekkert sérstakt sé um að vera</p>	<p>vera 3</p> <p>1</p> <p>2</p> <p>3</p> <p>4</p> <p>5</p> <p>6</p> <p>7</p> <p>vera + að</p> <p>vera + af</p> <p>vera + á</p> <p>vera + á milli</p> <p>vera + á móti</p> <p>vera + á við</p> <p>vera + án</p> <p>vera + eftir</p> <p>vera + frá</p> <p>vera + fyrir</p> <p>vera + fyrir framan</p> <p>vera + hjá</p> <p>vera + inn í</p> <p>vera + inni</p> <p>vera + í</p> <p>vera + með</p> <p>vera + milli</p> <p>vera + móti</p> <p>vera + nærri</p> <p>vera + saman</p> <p>vera + til</p> <p>vera + til í</p> <p>vera + til með</p> <p>vera + um</p> <p>vera + undan</p> <p>vera + undir</p> <p>vera + upp á</p> <p>vera + upp með</p> <p>vera + upp úr</p> <p>vera + uppi</p> <p>...</p>
--	---

FIGURE 4. Beginning of *vera* (ÍO 2000)

By clicking on the items in the list on the right of the screen, the user gets to the appropriate part. Two such subentries with added English transcriptions from the list in figure 4 are found in figure 5, *vera + saman* and *vera + til með*:

<b>vera + saman</b> <b>vera saman</b> 1 vera samvistum 2 (um karl og konu) vera í nánum kynnum, eiga nán mök 3 vera samherjar í spili eða leik	'be together' 'be lovers' 'be on the same team or side'
<b>vera + til með</b> <b>vera til með e-ð</b> 1 vera tilbúinn með e-ð 2 vera tilleiðanlegur (að gera e-ð), geta átt það til	'have something ready or prepared' 'be prepared to do something', 'be apt to do something'

FIGURE 5. *vera + saman* and *vera + til með*

An explanation of the use of the '+' is in order here. The term 'verbal phrase' is very loosely interpreted because the average user is not going to distinguish between particle verbs, verbs that take obligatory prepositional phrases, etc.<sup>3</sup> The word-order in the structure can also be varied, and the '+' is only used as a heading to indicate subordinate constructions where both verb and particle (or preposition, adverbial, etc.) occur, without any information on the actual structure. These are, in the jargon, 'search strings'. The actual verbal phrases are listed under the main heading containing the '+' as subordinate constructions. On the screen colours are used for further differentiation between these two; the heading with the '+' is in black boldface, whereas the actual verbal phrases are in bright blue boldface on the screen.

Setting up verbal phrases as subentries, and thus, in a way, recognizing their status as lexical items or lexemes in their own right, inevitably leads to a fairly massive reorganization of the entries from the form used in the two older editions of the dictionary. One of the reasons for this is that the treatment of such material in the printed dictionary is by no means uniform. Verbal phrases are sometimes listed in the body of the text, i.e., under the appropriate semantic description, but they can also appear in a separate list at the end of the entry, ordered alphabetically by particles and prepositions. In some cases the same phrase even appears in both parts of the entry, sometimes with two different explanations. If phrasal verbs are to be treated

<sup>3</sup> In Icelandic, all these can be combined under the term *smáorð* (literally 'little word').

as sub-entries in a format where the perspective is limited by the size of the screen, a uniform approach has to be maintained. This does, of course, call for guidelines on what is to be regarded as a verbal phrase. In the 3rd edition of *ÍO*, the term is interpreted widely, even so far as to include some adverbials, the idea being to try to imagine what the average user might search for under the dual heading of verb and particle/preposition.

### 2.3 Differentiation within the subentry

The subordinate constructions are listed under the main heading described above (i.e., *vera + saman*) in different type (bright blue boldface on the screen), using the traditional short forms to indicate arguments as used in dead examples in the printed version of the dictionary. The term *dead example* is used in a wide sense here to include the appropriate short forms for subjects as well as objects, to cope with non-nominative or impersonal subjects, as in *taka + fyrir* in figure 6:<sup>4</sup>

<b>taka + fyrir</b>	
<b>taka fyrir e-ð</b>	
1 láta e-ð ekki viðgangast lengur, afnema, hindra e-ð	'stop sth. from happening, rescind sth.'
2 taka [þvert] fyrir e-ð (þver)neita e-u, fortaka e-ð	'refuse (to do) sth.'
3 taka [greiðslu] fyrir e-ð þiggja greiðslu fyrir e-ð	'accept payment for sth.'
<b>taka e-n fyrir</b>	
1 taka e-n til meðferðar, setjast að e-m (á e-n)	'harrass sby.'
2 lögfræði/félagsfræði taka e-n til yfirheyrslu	'call on sby. (for interrogation, etc.)'
<b>taka e-ð fyrir</b>	
1 taka e-ð á dagskrá (t.d. mál fyrir rétti)	
2 fara að vinna að e-u, leggja e-ð fyrir sig	'start working on sth., make a career of sth.'
<i>hún tók lögfræðina fyrir</i>	
<b>taka e-ð fyrir sig</b> hefjast handa við e-ð	'begin to do sth.'
<i>taka fyrir sig verk</i>	
<b>e-ð tekur fyrir e-ð</b> e-ð skyggir á e-ð	'sth. obstructs the view of sth.'
<i>sandrokið tók fyrir sólu</i>	
<b>það tekur fyrir e-ð</b> ÖP e-u lýkur	(impersonal) 'sth. comes to an end'
<i>þá tók fyrir ásóknina þá hætti ásóknin</i>	
<i>þá tók fyrir sjóróðra þá hættu veiðar</i> (vegna ótiðar)	

FIGURE 6. *taka + fyrir*

<sup>4</sup> The short forms used are the ones traditionally used in Icelandic dictionaries and books on grammar: *e-ð* (nom./acc. -animate), *e-u* (dat. -animate), *e-s* (gen. ±animate), *e-r* (nom. +animate), *e-n* (acc. +animate), *e-m* (dat. +animate), etc.

As evident in the example, the dead examples from the 1983 edition have been elevated to a form of sub-sub-entries; they are given the status of proper headings or lexical items on the screen, i.e., they are 'heads' or, in primitive computer-jargon, these constructions within each verbal phrase have 'search value'. In actual fact, these former dead examples are not confined to the infinitive, and they do include short forms for subject as well as for object, as can be seen at the bottom of figure 6.

### 3. Constructions as 'heads' (or lexical items)

The change from dead examples to 'heads' can be seen to represent a move away from the traditional dictionary headword where structural information is usually not a part of the headword itself, but shown in examples. The move is towards a fully specified lexeme, perhaps in the sense of Rochelle Lieber (1992) or Di Sciullo & Williams (1987). The 'head', representing a complete (or discrete) unit meriting an explanation, examples, and whatever serves to make the user familiar with it, is then a unit in the lexicon in the sense of generative grammar, or a 'listeme', to use Di Sciullo & Williams' term. Full specification implies giving the predicate argument structure for each lexeme, as Lieber does, and for Icelandic the intricacies of case-marking make that a very important part of the description. One could even argue that a lexeme does not exist without such information, i.e., that words do not exist in a vacuum.

Bringing this idea to its logical conclusion would perhaps entail a description on par with the one used at *Orðabók Háskólans* for verbs in the prospectus to a dictionary of verbs (Svavarsdóttir et al. 1993), where the microstructure of the entries is based on syntactic principles. A diagram of the structure used there is shown at the end of this paper for comparison, but a description based on this kind of structure is very complicated and probably not at all suited to the needs of the ordinary user, but then it was never intended to be.

To make a long story short, the conclusion in the work on the new edition of the *ÍO* was to include syntax as expedient. The use of syntax in the headings is primarily intended to help the user find what he is looking for on the screen, and in this he is aided by the colours on the screen. A by-product is trying not to lead the user astray by falsifying data, e.g. by implied non-existent constructions such as finite forms of verbs in senses only found with the past participle or in impersonal constructions, or by giving the active voice of verbs only used in the middle voice (mediopassive).

The guiding principle was trying to be practical and user-friendly, keeping in mind that the older versions of *ÍO* have been criticized for lack of information on usage, sometimes to such an extent that the assumption seems to be that the user is thoroughly familiar with all aspects of a word except for its meaning. In places one gets the feeling that all syntactic information is regarded as redundant. The verbs *aftra* and *hindra* show this clearly; their meaning is similar, but *hindra* takes an accusative object (*hindra e-ð*, *hindra e-n í e-u*) and *aftra* a dative one (*aftra e-m frá e-u*, *aftra e-u*). Neither verb contains any syntactic information for the modern language in the 1983 edition, as seen in figure 7:

**aftra, -aði** S hindra; † a. sér að e-u hika við e-ð.  
**hindra, -aði** S 1 koma í veg fyrir, tálma; tefja, slóra. 2 †  
 hika, fresta. 3 LH PT *hindruð kona* ∴ ófrísk.

FIGURE 7. The entries for *aftra* and *hindra* (ÍO 1983)

In the computerized versions the common constructions are stated, as can be seen in figure 8. The verb *henda* provides good examples of constructions as 'heads'; the verb takes a dative or an accusative object, *henda e-u* (dat., cf. section 1B2), *henda e-ð* (acc., cf. section 3). The subject in these sections is unspecified, i.e., either animate or inanimate, and thus not shown. In the sense in section 4, the subject is always inanimate (*e-ð*, nom., reflexive or not), or impersonal (*e-n*, acc., +animate); see next page.

This verb is specially selected here to show different senses coinciding with difference in syntax, although sections 1 and 2 could be combined, both on grounds of syntax and sense.<sup>5</sup>

The minimal requirements in the case of Icelandic verbs is sufficient information on case-marking, both of objects and subjects. Sometimes semantic restrictions on either subject or object are also useful (e.g. ±animate). This, however, is not followed strictly, as nominative subjects (unspecified for ±animate) are assumed to be the norm, shown by the traditional description with infinitives, as in sections 1B3 in *henda* in figure 8. In the last part of that verb the short forms for subjects are also shown, *e-ð* for an inanimate subject (nominative) and *e-n* for an accusative subject; see figure 9.

<sup>5</sup> The meaning in section 2 is very common, but strangely enough it was missing in both older editions.

<b>henda henti s</b>	
<b>1</b>	
<b>henda e-u</b> kasta e-u <i>hann henti stönginni hátt í loft</i> <i>hentu til mín hamrinum</i>	[Obj.=dat. –animate] 'throw sth.' 'he threw the rod ...into the air' 'throw me the hammer'
<b>2</b>	
<b>henda e-u</b> fleygja e-u, setja e-ð í ruslaflát  <i>ég er búinn að henda gömlu skónum</i> <i>ekki henda ryggjóinu á götuna</i>	[Obj.=dat. –animate] 'throw sth. away, put sth. in the garbage' 'I've thrown away the old shoes' 'don't throw the chewing gum on the street'
<b>3</b>	
<b>henda e-ð</b>	[Obj.=acc. –animate]
1 grípa e-ð	'catch sth.'
<b>henda e-ð á lofti</b> grípa e-ð og notfæra sér það <i>henda frétt á lofti</i>	'make use of sth.' (idiom)
2 öðlast e-ð	'obtain sth.'
<b>henda e-n / e-ð</b> ná e-m / e-u	[Obj.=acc. ±animate] 'catch sby./sth.'
<b>hendir seinn hvatan</b> seinfær maður nær oft hinum fljóta <b>henda sund (spor)</b> byrja að ganga óstuddur	
3 <u>fornt/úrelt</u> nema, læra e-ð <i>henda e-ð af bókum</i>	'learn sth.' 'learn sth. from books'
<b>4</b>	
<b>e-ð hendir, e-ð hendir sig</b> e-ð kemur fyrir, ber við, gerist <i>þetta hendir öðru hverju</i> <i>nú henti það sig að hann varð of</i> <i>seinn</i>	[Subj.=nom. –animate] 'sth. happens' 'this happens now and then'
<b>e-n hendir e-ð</b> ÓP e-r verður fyrir e-u	[Subj.=acc. +animate] 'sth. happens to sby.'

FIGURE 8. The entry for *henda* (ÍO 2000)

1–2	<b>henda e-u</b>	dative object, –animate
3.1–3	<b>henda e-ð</b>	accusative object, –animate
3.2	<b>henda e-n/e-ð</b>	accusative object, ±animate <sup>1</sup>
4	<b>e-ð hendir, e-ð hendir sig</b>	nominative subject, –animate
	<b>e-n hendir e-ð</b>	accusative subject, +animate

FIGURE 9. Arguments with the verb *henda* (ÍO 2000)

The details for the arguments of verbs to be expected in the headings are mainly the case of objects and oblique subjects, ±animate (when distinctive), 'dummy' (impersonal) subjects, reflexion, and retroactivity.

The difficulty lies in deciding how much of that information is necessary and what is redundant. The borderline for that is not the

same for the average user as for the syntactician who would probably interpret this as giving full information on lexical case-marking, for instance, and treating all structural case-marking as a default. This, however, would probably not be seen by the general public to be a very user-friendly way of doing things as most people are very often not aware of grammatical regularity. The guideline for the changes made in the description of the verbs in the 3rd edition of the *ÍO* was to make the material as accessible and clear to the reader as possible in the very short time allotted to the job, using the framework described here to rearrange material from the older editions. It must be stressed that this is not a new dictionary being made from scratch, so to speak.

The computerized version of *ÍO* is in many ways a pilot project, and very many questions are still unresolved. To take the format of the verbs as an example, these are not very standardized. There are a few possibilities of layout for each and every aspect and these are used as deemed appropriate for each individual verb.

#### 4. From a monolingual to a bilingual dictionary

Bilingual dictionaries published in Iceland are primarily made to suit Icelanders, i.e., they are monodirectional, but dictionaries published in other countries tend to be more varied. All of them have had to do bidirectional duty. The needs of Icelanders and foreign learners are of course not the same, but of necessity the two must be addressed at once.

The question is then, how much of a problem that is. By taking advantage of the new medium, the mass of the whole work is no longer a problem, only the disposition of the information on the screen. The point is that this may be the means to serve the dual purpose of monodirectional and bidirectional dictionaries, and that going towards full specification of 'listemes' (or subentries) is equally important to both sets of users.

For a monolingual Icelandic dictionary, I maintain that the constructions are needed in order to find what one is looking for, as the meaning is construction specific. In a bilingual dictionary, it automatically follows that the equivalent constructions in the target language are given if the constructions are set up as 'heads' in the source language. In both cases, the same dictionary would therefore ideally serve both as a passive and an active dictionary. This point is well illustrated by impersonal verbs, as in the following examples.

#### 4.1 Examples of impersonal verbs

The impersonal verbs proved to be quite a problem in the switch from book to disc for the new edition of *Íslensk orðabók*, as the aim was providing minimal syntactic information wherever it was missing in the book. This in fact proved to be impossible at times as the data was sometimes unavailable for rare or obsolete words. There is, however, substantial material on impersonal constructions in Icelandic, notably the work of Jóhannes Gísli Jónsson and Joan Maling, some of which they were kind enough to provide in manuscript when these verbs were being edited. The Institute of Lexicography's archives and corpus were then used to fill in the gaps as possible.

The facts on the impersonal verbs in the 1983 edition are these: Two types of constructions are marked with 'ÓP' (for *ópersónulegt* 'impersonal'), i.e., constructions with oblique subjects (in the accusative, dative or genitive), and constructions with a 'dummy' subject, *það* or *hann*.<sup>6</sup> The information on those verbs in the 1983 edition is very limited, and sometimes there was absolutely no syntactic information apart from a bare 'ÓP', i.e., no examples of any kind, and thus no indication of case-marking.

The verb *fatra* in figure 10 is one of those verbs:<sup>7</sup>

*fatra*, -aði S ÓP *fata*, skjátlást; MM *fatrast fara* í flækju; tefjast, fara í handaskólum.

FIGURE 10. *The entry for fatra (ÍO 1983)*

A rough transcription of *fatra* would be 'to be wrong', but such complete lack of information on argument structure can make it very difficult to determine what is happening to whom or what – i.e., the argument structure can be so vague that the meaning becomes unclear. This probably poses even greater problems in a bilingual dictionary

<sup>6</sup> *Hann* is only used for verbs describing the weather.

<sup>7</sup> Possible constructions with oblique subject are *e-n fatrar*, *e-m fatrar*, *e-s fatrar* (the genitive is unlikely, the genitive subjects are very rare); the 'dummy' subject construction *það fatrar* + AdvP with argument for person, e.g., *?það fatrar fyrir e-m* is remotely possible in the modern language, (cf. *það skeði fyrir honum*, which makes anyone over a certain age cringe), but as the verb is exceedingly rare, that possibility is very remote. A comparable construction in middle voice *það fatrast fyrir e-m* is shown in Fritznér: *fatrast mjök fyrir honum* (Mork., probably the same example as found in the Institute's corpus, in *Odds þáttur Ófeigssonar*). The explanations in the 1983 edition of *ÍO* (and in Asgeir Blöndal Magnússon's *Íslensk orðsifjabók* (1989)) are probably based on these two citations, the different meanings for the middle voice being the result of uncertainty implied by the use of the noun *fatur* in the same text rather than actual attestations.

for foreigners, who are probably not even familiar with the possible structures.

In the new edition, the active voice of this verb has an animate, dative subject, whereas the middle voice has a 'dummy' one as in figure 11:

<b>fatra</b> -aði S ÓP e-m fatrar e-m fatar, skjáttlast	'sby. makes a mistake, sby. is mistaken'
<b>fatrast + fyrir</b> það fatrast fyrir e-m ÓP e-ð fer í flækju hjá e-m, e-r tefst (við e-ð), e-ð fer í handaskolum hjá e-m	'sth. gets tangled up for sby., sby. is delayed (by sth.)'

FIGURE 11. *The entry for fatra (ÍO 2000, update)*

The new version shows all and only the constructions attested in the archives and corpus at the Institute of Lexicography. The data is therefore not falsified.

For differences in meaning, the importance of the differentiation between constructions can be seen in the verb *batna*, which can take either a nominative or a dative subject. The meaning changes between these constructions; with a nominative subject it means 'improve', whereas with a dative subject it means 'recover from an illness' (+animate subject only). In the 1983 version one explanation serves both; the impersonal construction is only shown as an example in figure 12:

**batna**, -aði S verða betri, skána: # *hagur hans batnaði*; ÓP: *honum batnar (veikin)*.

FIGURE 12. *The entry for batna (ÍO 1983)*

Native speakers have no problem with the meaning here; they know it anyway as this is a very common verb. For learners of the language the new version is more helpful, and the difference between a singer whose singing gets better and a singer whose laryngitis gets better is clear, as seen in figure 13. The distinctive meaning in the sentences *söngvarinn batnaði með hverjum tónleikum* and *söngvaranum batnaði hálsbólgan* is no less clear when they are abbreviated to *söngvarinn batnaði* (nominative, and he sings better), and *söngvaranum batnaði* (dative, and he is no longer sick):

<b>batna -aði s</b>	
<b>1</b>	
<b>e-ð/e-r batnar</b> e-ð/e-r verður betra/betri, e-ð/e-r skánar	'sth./sby. improves, sth./sby. gets better'
<i>hagur hans batnaði</i>	'his finances improved'
<i>ástandið batnaði</i>	'the situation improved'
<i>söngvarinn batnaði með hverjum tónleikum</i>	'the singer got better at each recital'
<b>2</b>	
<b>e-m batnar</b> ÓP e-r nær sér (af veikindum)	'sby. recovers (from an illness), sby. gets better ...'
<i>söngvaranum batnaði hálsbólgan</i>	'the singer recovered from his laryngitis'

FIGURE 13. *The entry for batna (ÍÓ 2000)*

There are regular patterns of theta-roles and case marking in some classes of impersonal verbs (e.g. dative for +animate subject (experiencer) and nominative for –animate subject/object (theme)), but the impression is that the average native speaker is rather muddled by the grammatical structures as such, even if his usage conforms to the norms. This confusion is one of the justifications for having constructions as 'heads' in the monolingual dictionary. For users who are unfamiliar with the structures themselves (e.g. foreigners) and for users who need to know the corresponding structure in another language (e.g. translators), being presented with a construction instead of a single headword is bound to be more helpful as it is better guidance to usage.

A case in point is that an argument can be a subject in one construction and an object in an equivalent one, as can be seen in the entry for *gagna* (cf. figure 14). This can be true in paraphrases in the same language, as in sense 1 of *e-m gagnast e-ð*, at the end of figure 14; see next page.

The same kind of switching of argument can also be seen in translations (or transcriptions), as in the same place in the entry for *gagna* in figure 15 (corresponding arguments marked for clarity); see next page.

<p><b>gagna</b> -aði S  e-ð gagnar → e-ð gagnast  <i>himnan gagnar til þess að festa garnirnar við önnur líffæri</i></p> <p><b>e-r gagnar e-m</b>  1 → e-r gagnast e-m 1  <i>hvermig má sá mér gagna sem er sjálfum sér verstur</i>  <i>hver maður gagnar sinni fösturjörð</i>  2 → e-r gagnast e-m 2  <i>bolarnir gagna kínum</i></p> <p><b>e-m gagnar e-ð ÓP</b> → e-m gagnast e-ð  <i>það gagnar mér ekki</i></p> <p><b>gagnast</b> MM  e-ð gagnast, e-ð gagnar e-ð verður að gagni  <i>hér gagnast ekki veiðarfæri í sjó</i>  <i>orð þeirra gagnast til að afsanna þetta</i>  <i>vel gagnist staðbundið verði þér (ykkur) að góðu</i></p> <p><b>e-r gagnast e-m</b>  1 e-r verður að liði, kemur að gagni  <i>Jóhannes gagnast okkur vel í hreingeringunni</i>  2 e-r kelfir, fyljar, lembir e-n  <i>nautið gagnaðist kínni</i>  3 e-r hefur kynmök við e-n  <i>hann mátti eigi gagnast konum</i></p> <p><b>e-m gagnast e-ð, e-m gagnar e-ð ÓP</b>  1 e-ð verður e-m að gagni  <i>honum gagnaðist lyfið ágætlega</i>  2 e-r fær frið til e-s  <i>okkur gagnast ekki að sofa</i></p>
---

FIGURE 14. The entry for *gagna* (ÍO 2000)

<p>e-m, gagnast e-ð, e-m, gagnar e-ð, ÓP  1 e-ð, verður e-m, að gagni</p> <p><i>honum, gagnaðist lyfið, ágætlega</i></p>	<p>'sth.<sub>j</sub> works for sby.<sub>i</sub>, sth.<sub>j</sub> is useful for sby.<sub>i</sub>, sth.<sub>j</sub> serves sby.<sub>i</sub>'s, purpose'  'the medicine<sub>j</sub> did its job for him<sub>i</sub>.'</p>
--	---

FIGURE 15. Last part of *gagna* with transcriptions

As the entry for *gagna* in figure 14 shows, the actual descriptions of individual verbs can be quite complex when it comes to giving all and only the constructions the ordinary user can expect to find in more or less ordinary texts. In comparison the entry for *gagna* in the first two editions of *Íslensk orðabók* is very simple, and the same can be said for the entries for the verb in some bilingual Icelandic dictionaries shown in figure 16 in the appendix. A common denominator for these entries

is that constructions are only shown in examples, and the choice of examples seems to be rather the sketchy, only showing a part of the range of constructions. Complexity on par with the entry shown in figure 15 may not be everyone's idea of clarity, but the choice seems to be between the simplicity of the older dictionaries and the detailed description made possible by the new format. The task at hand at the moment is finding out what kind of description is suitable for the new Icelandic-Danish dictionary in preparation at Edda.

## 5. Conclusion

The decision to give 'lexeme'/'listeme' status to constructions and verb phrases by giving them 'search value' should ideally be a reflection of the mental lexicon, providing the proper argument structure. The constructions could then be used as 'heads' more or less in the same way as they must appear in our mental lexicon. Sometimes full specification is not needed, as the words do more or less enter into all regular constructions. In this way, leaving the arguments unspecified implies that anything goes. Very often, however, usage of a certain word is limited to a subset of the possibilities of its word-class, in some cases even to the point of one or two constructions for a verb, as in *fatra* in figure 11.

The new format, the screen, provides the possibility of showing this, and at the same time the use of constructions as heads is a way of splitting up unmanageable chunks of information into pieces small enough for the screen. In the end, the reactions of the users will show whether this is a move in the right direction, but as the disc is fairly new that feedback is still to come.

## Bibliography

- A *Prospectus to a Dictionary of Icelandic Verbs*. 1993: [Sýnihefti sagnorðabókar.] English Version of the Introduction. Rannsóknar- og fræðslurít 5. Reykjavík: Orðabók Háskólans.
- Árnason, Mördur (ed.) 2000: *Íslensk orðabók. Tölvuútgáfa*. Reykjavík: Edda hf.
- Blöndal, Sigfús (ed.) 1920–1924: *Íslensk-dönsk orðabók*. Reykjavík.
- Blöndal Magnússon, Ásgeir 1989: *Íslensk orðsifjabók*. Reykjavík: Orðabók Háskólans.
- Böðvarsson, Árni (ed.) 1963: *Íslensk orðabók handa skólum og almenningi*. Reykjavík: Bókaútgáfa Menningarsjóðs.

- Böðvarsson, Árni (ed.) 1983: *Íslensk orðabók handa skólum og almenningi*. Önnur útgáfa aukin og bætt. Reykjavík: Bókaútgáfa Menningarsjóðs.
- Di Sciullo, Anna Maria, & Edwin Williams 1987: *On the Definition of Word*. Cambridge, Mass.: The MIT Press.
- Fritzner, Johan 1883–1896: *Ordbog over det gamle norske sprog I–III*. Omarbeidet, forøget og forbedret Udgave. Kristiania.
- Halldórsson, Björn 1992: *Orðabók. Íslensk-latnesk-dönsk*. Orðfræðirit fyrri alda II. Reykjavík: Orðabók Háskólans. [Ed. 1992 Jón Aðalsteinn Jónsson; 1st ed. 1814.]
- Hólmarsson, Sverrir, Christopher Sanders & John Tucker 1989: *Íslensk-ensk orðabók*. Reykjavík: Iðunn.
- Jón Ólafsson úr Grunnnavík: *Dictionarii tentamen Islandici*. AM 433 fol. [Manuscript at the Arnarnaganean Institute, Stofnun Árna Magnússonar á Íslandi.]
- Jónsson, Jóhannes Gísli 1999: *List of predicates that take a quirky subject in Icelandic*. Manuscript. Háskóli Íslands.
- Jónsson, Jóhannes Gísli 2000: *Not so quirky: On lexical (subject) case in Icelandic*. Manuscript. Háskóli Íslands.
- Jónsson, Jón Hilmar 1985: *Íslensk orðabók handa skólum og almenningi*. Ritstjóri: Árni Böðvarsson. Önnur útgáfa aukin og bætt. [Review.] *Íslenskt mál* 7:188–207.
- Leijström, Gunnar, Jón Magnússon, Sven B. F. Jansson 1994: *Isländsk-Svensk Ordbok*. Sjätte upplagan. Rabén Prisma.
- Lieber, Rochelle 1992: *Deconstructing Morphology*. Word Formation in Syntactic Theory. Chicago: The University of Chicago Press.
- Maling, Joan & Jóhannes Gísli Jónsson 2000: *Það rignir þágufalli á Íslandi: Icelandic Verbs with Dative Objects*. Manuscript. Brandeis University & Háskóli Íslands.
- Rögnvaldsson, Eiríkur 1998: Málfræði í orðabókum. Hvernig og til hvers? *Orð og tunga* 4:25–33.
- Svavarsdóttir, Ásta, Guðrún Kvaran, Jón Hilmar Jónsson & Kristín Bjarnadóttir 1993: *Sýnihefti sagnorðabókar*. Rannsóknar- og fræðslurit 3. Reykjavík: Orðabók Háskólans.
- Widding, Ole, Haraldur Magnússon & Preben Meulengracht Sørensen 1976: *Íslensk-dönsk orðabók*. Reykjavík: Ísafoldarprentsmiðja hf.

## Appendix

**gagna**, -aði S ÓP 1 verða að gagni: það gagnar (mér) ekki: ☉ vel gagnist verði (ykkur) að góðu. 2 fá frið til e-s: okkur gagnast ekki að sofa. 3 notast af e-u, kelfa, fylja, lemba: nautið gagnaðist kúnni. (ÍO 1983)

**gagna** v. Eg *gagna*, gagnadi, at *gagna*. ... frugi sum, utilitatem adfero. Þat *gagnar* eigi, non est usui. Þat *gagnar*, expedit. §Mjer *gagnast* þat ei, non possum id propter turbulentiam et inquietudinem ab aliis adlatam efficere, non id mihi facere expedit. at *gagnast* manni ad vitni, pro teste alicui utilis esse. Mier *gagnast* ei at vera at því, non dicet mihi id allaborare. (Orðabók Jóns Ólafssonar úr Grunnavík 1734–1779)

**Gagnaz** (at *gagnaz*), *prodesse*, gavne, være til Nytte, **Hönum gagnaz ecki at skrifa fyrir hávada**, ob *garritum facultatem scribendi amittit*, han har, formedelst Støj, ikke Lejlighed til at skrive. **At gagnaz einum**, *auxilio*, v. *commodo alicui esse*, være En til Nytte, gjøre Bistand. **Það gagnaz ecki**, *frustratur*, det hjælper ikke. (Björn Halldórsson 1992 [1814])

**Gagnar** (at *gagna*), *prodesse*, nytte, gavne. **Þaðagnar ecki**, *non prodest*, det nytter ikke noget. (Björn Halldórsson 1992 [1814])

**gagna** (a) [gag-na] v. 1. v. impers. med dat. (*verða að gagni*) gavne, nytte: það *gagnar* (mjer) ekki, det nytter ikke noget (for mig). — 2. refl. *gagnast*. a. gavne, være til Nytte: nær er at landseti *gagnist* húsbónda sínum til sjóróðra, enn öðrum útifrú (LFR. VII. 154). — b. spec. om Tyre og Vædere: *tarfurinn gagnadist kúnni* (Breiðd.). — c. faa Lov til: því *gagnadist ekki að liggja í gröfnni*, det kunde ikke faa Lov til at ligge stille i Graven (JÁþj. I. 306); *mikill helv. hávadi er í ykkur ytra, það liggur við, að okkur gagnist ekki að sofa* (MelBr. 85).

(Sígfús Blöndal 1920–1924)

**gagna** -aði v *upers*. gavne, nytte; g. e-m være en til hjælp *refl.*: *gagnast* gavne, være til nytte; bedække (om tyr, væder). (Widding et al. 1976)

**gagna** v *impers* (-aði): hverjum ~ar það? who will that help? það ~ar ekki it is no use (Sverrir Hólmarsson et al. 1989)

**gagna** -aði *opers* v *gagna*, vara til nytta, tjäna till (g. e-m) (Leijström et al. 1994)

FIGURE 16. Sample entries from older dictionaries for *gagna*

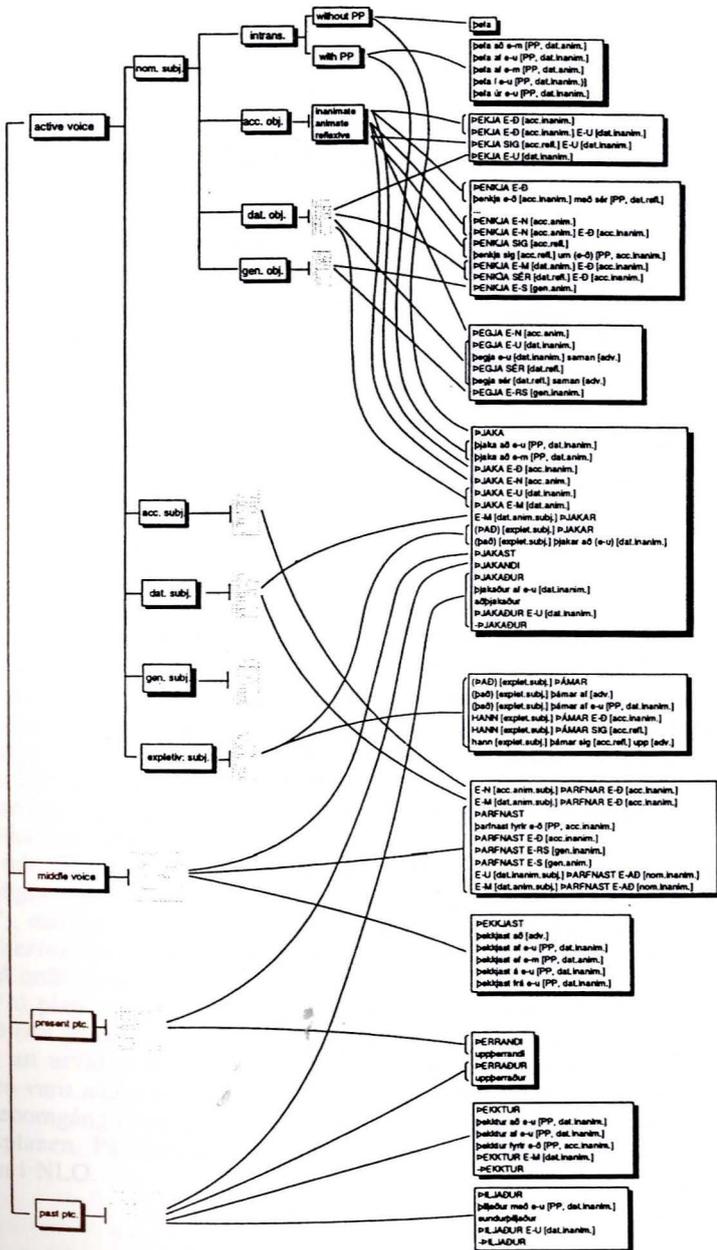


FIGURE 17. The structure used in Sýnihefti sagnorðabókar

LeuonaNordica 8 - 2011