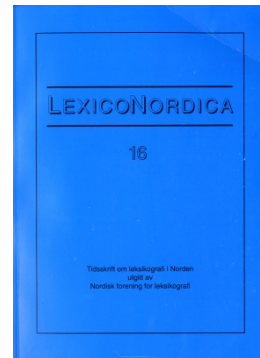


LexicoNordica

Titel: Ordforbindelser: Grunnelementer i ordboken?
Forfatter: Jón Hilmar Jónsson
Kilde: LexicoNordica 16, 2009, s. 161-180
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.

Jón Hilmar Jónsson

Ordforbindelser: Grunnelementer i ordboken?

The article describes the construction of a lexicographic description where different types of phraseological units have a wider and more active role than is usually the case in traditional dictionaries. It relates how the description of Icelandic phrasemes and collocations has been assembled in stages, starting with a printed combinatory dictionary, going towards an electronic dictionary description. A large collection of phrasemes and collocations has been used to identify and analyse lexical and semantic relations, and on the basis of such an analysis to classify the vocabulary according to semantic features.

Moreover, the phrasemes have been widely lemmatised on a level with single words which makes it possible to pinpoint their semantic relations to single-word lemmas. Additionally, as the multi-word lemmas have a standard presentation form, the lemma strings can be grammatically tagged and categorised by syntactic constructions.

A lexicographic model of this sort is based on the onomasiological standpoint. The lemmas are generally unambiguous and the description combines different roles that have usually been fulfilled by different types of dictionaries.

1. Innledning

1.1. Ordforbindelser som elementer i allmenne ordbøker

De fleste ordbøker, i hvert fall de tradisjonelle allmenne ordbøker, forventes å ha et betydelig innslag av ordforbindelser av ulike slag som leksikalske elementer underordnet de enheter som utgjør selve lemmalisten til ordboken.¹ Det er derimot mindre alminnelig at ordforbindelser opptrer i selve lemmalisten og dermed blir direkte tilgjengelige på linje med de ettordslemmaer som dominerer i tekstbildet. Dette betyr at det bilde ordboken gir av de ulike typer ordforbindelser, henger nøye sammen med beskrivelsen av de enkelte ord de er knyttet til (se bl.a. Burger 1989, Moon 2007). Dette forholdet har uunngåelig og på mange måter en begrensende effekt på ordforbindelsenes andel og omfang i ordboksteksten. Det er også en alminnelig erfaring at det er vanskelig å lage seg et bilde av den rolle ordforbindelser spiller i oppbyggingen og utformingen av en ordbok og en ordbokstekst, og vurdere deres bidrag til den

¹ Termen **ordforbindelse** er her brukt i vid forstand og dekker forskjellige typer ordkombinasjoner, i samsvar med den definisjon og forklaring termen får i NFL.

leksikografiske beskrivelse. Når det gjelder begrensningen av omfanget, må det ut fra leksikografens synsvinkel blant annet tas hensyn til en indre balanse i de enkelte artikler, der en klar betydningsinndeling som regel er det overordnede målet. Tilknytningen til de enkelte delbetydninger vil også være uklar i mange tilfeller så det kan bli vanskelig å avgjøre hvor en konkret ordforbindelse med en særpreget semantisk karakter skal plasseres innenfor artikkelen. Dertil kommer at fremstillingen av ordforbindelsene sjelden er standardisert i den grad at den viser en klar skilnad mellom ulike typer, og de skiller seg heller ikke alltid tydelig fra regelrette eksempler som har til funksjon å illustrere bruken av enkelte ord eller delbetydninger.

I mange allmenne ordbøker kan beskrivelsen og behandlingen av ordforbindelser dermed ikke anses å utgjøre en selvstendig del av oppbyggingen og utformingen av ordboksteksten.

Deres inkludering kan også være nokså tilfeldig. Den forutsetter for det første at ordforbindelsene inneholder enkeltord som (gjerne med hensyn til ulike kriterier) anses å være aktuelle som lemmaer. Det må også avgjøres hvilket lemma blant flere potensielle lemmaer skal prioriteres, ut fra de brukerbehov ordboken skal dekke og i samsvar med allmenne leksikografiske prinsipper. Det betyr at ordforbindelsene sjelden blir synlige i ulike kontekster, og det forventes at brukerne er klar over de prinsipper som best fører til et effektivt oppslag. Det bør også nevnes at den uttrykksmessige variasjon som karakteriserer mange ordforbindelser gjerne blir sterkt begrenset.

Men samtidig som beskrivelsen av ordforbindelser i de konkrete ordbøker blir underkastet sterke begrensninger må det framheves at bak de fleste ordbøker skjuler det seg et omfattende kilde- og tekstmateriale der ordforbindelser av ulike slag bidrar med vitnesbyrd om de leksikalske, semantiske og syntagmatiske egenskaper som blir trukket fram i beskrivelsen av de enkelte ord og på den måte setter sitt preg på ordboksteksten.

De leksikalske enheter som dekkes av den allmenne termen **ordforbindelse** har et ulikt forhold til de enheter de er sammensatt av, noe som også preger deres opptreden i en konkret leksikografisk sammenheng. På den ene side kan de fungere som naturlige elementer i beskrivelsen av lemmaet, der de bl.a. kan vise ordets valens eller understreke og fordype en bestemt definisjon eller forklaring. Hertil hører primært kollokasjoner og forskjellige typer konstruksjoner. I den andre enden av skalaen har vi faste ordforbindelser og idiomer der en sterk konvensjonalisering av uttrykket og en selvstendig semantisk karakter er typiske egenskaper.

1.2. I skyggen av det semasiologiske aspektet

På bakgrunn av dette er det forståelig at ordforbindelser ofte forekommer som ganske statiske og isolerte enheter i ordboksteksten, ikke minst i de mest allmenne ordbøker der en betydningsbeskrivelse med forklaringer og definisjoner står i fokus. Overfor en ordbok av den type vil det alltid være lett å påpeke svakheter og mangler, der ordforbindelser av ulike slag enten saknes, eller man kan forestille seg at de ville ha hatt en belysende og kompletterende funksjon med hensyn til en gitt informasjon.

Men uansett hvor mye plass det gis til fraser og ordforbindelser i en allmenn ordbok, må man innse at en ordbok av den type der betydningsbeskrivelse av ettordslemmaer blir prioritert, alltid vil gi fraseologien en underordnet rolle. Grunnen er at det ensidig tas hensyn til det semasiologiske aspektet der uttrykkssiden av de leksikalske enhetene bestemmer sammensetningen og struktureringen av lemmabestanden, og den sentrale oppgave består i å skille imellom og definere de ulike betydninger lemmaene opptrer i. Det overskygger det onomasiologiske aspektet, der innholdssiden av de leksikalske enhetene står i fokus for beskrivelsen av lemmaene. En slik innholdsbasert beskrivelse tilhører etter tradisjonen spesialordbøkernes oppgavefelt. Hvis en allmenn leksikografisk beskrivelse skal kunne skildre hvordan de ulike ordforbindelser fungerer som språklige elementer og hvordan de preger språkb Bruken, både generelt og i forhold til de enkelte ord de er sammensatt av, trengs det en ny innfallsvinkel der det tas mer hensyn til det onomasiologiske aspektet. Det vil for det første bety at ordforbindelsene i større grad kan opptre som selvstendige lemmaer, men det vil også ha betydning for beskrivelsen av enkeltordene og deres innbyrdes relasjoner, som resultat av en systematisk kartlegging av ordforbindelsenes paradigmatisk og syntagmatiske egenskaper.

2. Ordforbindelser som elementer i fraseologiske ordbøker

For å få til en sterkere fokusering på ordforbindelser og deres karakteristiske egenskaper enn den allmenne ordboken tillater, kan de beskrives innenfor rammen av ulike typer fraseologiske ordbøker (se Schemann 1989, se også Farø 2007:950 f.). Her dreier det seg om ulike formål og motivasjoner som da også setter sitt preg på ordbøkernes innhold og fremstilling. For leksikografen kan det simpelthen være viktig å få plass til et større antall fraser og ordforbindelser for dermed å gi en bredere

dekning av fraseologien. I andre tilfeller gjelder oppmerksomheten bestemte typer fraser og ordforbindelser ut fra en typologisk klassifisering. Generelt får man regne med at framstillingen og struktureringen av ordboksteksten her i større grad blir styrt av reelle brukerbehov enn det er tilfellet i en gjennomsnittlig allmenn ordbok.

Men så lenge lemmastrukturen bestemmes av de samme prinsipper som tradisjonelt gjelder for allmenne ordbøker og ordforbindelsene blir underordnet ettordslemmaer, vil beskrivelsen være preget av de samme begrensninger. Man får regne med at de ordforbindelser ordboken inneholder, blir valgt ut fra ulike kriterier, dels som selvstendige leksikalske enheter som da også krever en selvstendig beskrivelse, i andre tilfeller vil hensikten være å gi et sammenhengende bilde av den fraseologien som karakteriserer bruken av enkelte lemnaer. Inn i dette bildet kommer leksikografens avgjørelse om hvilket lemma en gitt ordforbindelse skal tilknyttes. Særlig når det gjelder kollokasjoner, tilsier tradisjonen at det bør tas logiske hensyn med tanke på ordbokens genuine formål og aktuelle brukerbehov. Det betyr f.eks. at substantiver, som basis i kollokasjoner, generelt skal ha prioritet som lemnaer overfor de adjektiver de står sammen med. Tilsvarende prinsipper blir framhevet for plasseringen av idiomer og andre faste ordforbindelser. Dermed vil de enkelte ordklasser få ulik vekt innenfor lemmalisten, og det gis en ufullstendig innsikt i hvordan ordforbindelser er knyttet til hverandre leksikalskt, syntaktisk og semantisk.

Man kan trygt fastslå at de tradisjonelle fraseologiske ordbøkene, uansett hvor solide og innholdsrike de kan være, har måttet lide for en sterk strukturell påvirkning fra de ordbøker der ordet som en avgrenset enhet står i fokus, slik at den eneste adgangen til den leksikografiske informasjon er gjennom en alfabetisert liste av ettordslemmaer. Denne påvirkningen og de restriksjoner den innebærer, er naturligvis opprinnelig forbundet med de begrensninger papirordboken har pålagt leksikografien. Og den lever forståelig nok videre i mange av de leksikografiske produkter som nå formidles i elektronisk form.

Men selv om ordbøker av denne typen ikke kan gi en fyllestgjørende dekning av ordforbindelser og belyse deres funksjon som sentrale leksikalske enheter, utgjør de et viktig bidrag til en mer omfattende og dynamisk ordboksbeskrivelse der fraseologien står i fokus og får prege struktureringen og fremstillingen av hele den leksikografiske beskrivelse.

3. Ordforbindelser i fokus for islandsk leksikografi

Fra sent i 80-årene har undertegnede arbeidet med en leksikografisk beskrivelse av islandske ordforbindelser og fraseologi. Dette arbeidet har foregått i etapper og er dels blitt konkretisert i utgivelsen av tre fraseologiske ordbøker. I det siste har oppmerksomheten i økende grad blitt rettet mot å utvikle en ordboksmodell der ordforbindelser utgjør basen i materialet og fungerer som grunnlag for kartleggingen av de semantiske, syntagmatiske og leksikalske relasjoner lemmaene, der også flerordsenheter er inkludert, opptrer i.

I det følgende kommer jeg til å redegjøre for noen hovedtrekk i utformingen av denne ordboksmodellen der ordforbindelser blir tildelt en sentral rolle og er blitt frigjort fra det å være strukturelt underordnet andre leksikalske enheter.

3.1. *To fraseologiske ordbøker*

Det første bidraget til en sammenhengende beskrivelse av islandsk fraseologi besto i utgivelsen av ordboken *Orðastaður* i 1994. Denne ordboken kan dels defineres som en kollokasjons- og konstruksjonsordbok, men den inneholder dessuten en morfologisk komponent med en semantisk sortert liste over sammensetninger til de enkelte oppslagsord. Struktureringen er ellers relativt enkel, lemmalisten består av omtrent 11.000 ettordslemmer som til sammen dekker bortimot 45.000 ordforbindelser og fraser.

Orðastaður ble gjenutgitt i en utvidet og revidert utgave i 2001, men i 2002 utkom ordboken *Orðaheimur*, som kan betegnes som en fraseologisk begrepsordbok. Her er struktureringen litt mer sammensatt og avansert, lemmalisten består av drøyt 800 begrepsoverskrifter i alfabetisk rekkefølge, som til sammen dekker omtrent 33.000 ulike fraser. I tillegg kommer et alfabetisert register over samtlige innholdsord i frasene der hvert lemma dekker de fraser det forekommer i. Alfabetisk sortering av frasene under hvert nøkkelord gir her en direkte tilgang til de enkelte frasene, der det henvises til vedkommende lemma eller lemmer i begrepslisten. Når det gjelder mikrostrukturen i disse ordbøkene, inneles den for en stor del ut fra semantiske faktorer, spesielt i *Orðaheimur*.

3.2. Ordbaserte vs. begrepsbaserte lemmaer

Når det gjelder lemmastrukturen tas det i *Orðaheimur* et første skritt til å frigjøre fraser og ordforbindelser fra å være underordnet ettordslemmaer. Her er hensikten å føre sammen semantisk beslektede fraser, særlig med hensyn til de brukerbehov som gjelder produksjon, slik at brukeren kan vurdere å velge mellom ulike uttrykksformer (på samme måte som i en synonymordbok). I det tilfellet vil brukeren ikke nøye seg med konkrete ord som lemmaer, for et ordbasert oppslag må føre videre til de uttrykk og fraser som tilhører vedkommende begrepsfelt. De enkelte begrepsfelt forsynes i sin tur med overskrifter eller etiketter som belyser deres semantiske innhold og karakter. Det betyr at ordboken inneholder to kombinerte alfabetisk ordnete lemmalister, en ordbasert liste over samtlige innholdsord i frasene som dekker alle fraser der vedkommende ord forekommer som konstant ledd, og en begrepsbasert liste over de overskrifter som dekker semantisk beslektede fraser. Ved enkelte fraser i den første listen henvises det til den begrepsoverskrift (eller -overskrifter) frasen er underordnet.

3.3. Standardisert framstilling, fast sorteringsregel

For å sikre målbevisste og effektive oppslag i den ordbaserte lemma-listen, kreves det en klar standardisering av frasene og en sorteringsregel som bevirker at en gitt frase lett kan lokaliseres innenfor en lengre liste. Ved hjelp av standardiseringen kan det trekkes fram viktige syntaktiske egenskaper, spesielt i verbfraser, der det bl.a. kan skilles mellom animate og inanimate subjekt, og nominale og adverbiale ledd får gjenspeile verbets syntaktiske og semantiske valens. Foruten å gi hver enkelt frase en klar plassering får brukeren ved hjelp av en fast sorteringsregel en samlet oversikt over leksikalsk og formelt sammenhengende fraser. Også dette kommer klarest til uttrykk blant verbene, noe som gir verblemmaene atskillig større verdi med hensyn til fraseologien enn de vanligvis har i en tradisjonell ordbok. Her viser det seg hvor nyttig det er å ha tilgang til frasene ut fra ulike innholdsord, istedenfor å måtte slå opp under ord av en bestemt ordklasse.

3.4. *Selvstendig fraseologisk register*

Genereringen av den ordbaserte lemmalisten i *Orðaheimur* ga et klart vitnesbyrd om at en slik liste også kan ha en selvstendig verdi uten på forhånd å være knyttet til en helhetlig semantisk klassifisering av frasene. Denne oppfatningen lå bak avgjørelsen om å bygge opp et omfattende fraseologisk register over fraser og ordforbindelser hentet fra sitatsamlingen til Orðabók Háskólans (Islands universitets leksikografiske institutt, senere den leksikografiske avdelingen til Árni Magnússon instituttet for islandske studier). Her gjelder de samme prinsipper for standardisering og sortering av materialet som ble innledet ved utarbeidelsen av *Orðaheimur*, men registeret inneholder langt flere fraser og nøkkelord (ordbaserte lemmer), ialt omkring 135.000 fraser knyttet til bortimot 60.000 nøkkelord (se <http://www.lexis.hi.is/osamb/osamb.pl>). Det er også formelt og innholdsmessig mer variert. Her er rene kollokasjoner også inkludert, og frasene blir valgt og registrert uten at de på forhånd skal tilhøre en lukket mengde semantiske felt.

3.5. *Kombinert fraseologisk ordbok*

Den tredje ordboken i denne familien med fraseologiske ordbøker kom i 2005 under tittelen *Stóra orðabókin um íslenska málnotkun*, og er i grunnen en kombinert utgave av de to andre som i tillegg til papirutgaven også ble utgitt som Cd-rom. Etableringen av det fraseologiske registeret utgjorde en viktig inspirasjon i denne sammenheng, for det viste at det også er nyttig å avdekke sammenhengen mellom leksikalsk og semantisk beslektede kollokasjoner innenfor rammen av et slikt register. Med tanke på det var det ikke noe i veien for å utvide og komplettere den ordbaserte lemmalisten i *Orðaheimur* ved også å dekke samtlige fraser i *Orðastaður* (der hovedparten består av kollokasjoner og konstruksjoner). De egentlige lemmalister til de to ordbøker (ordbasert i *Orðastaður*, men begrepsbasert i *Orðaheimur*) kunne på en enkel måte slås sammen i en alfabetisert lemmaliste med en grafisk skilnad mellom ord- og begrepslemmer. Dette framgår av figur 1 (se appendiks), som viser to tekstavsnitt fra *Stóra orðabókin um íslenska málnotkun* der ordet *ágreiningur* 'uenighet, meningsforskjell' forekommer i tre lemmafunksjoner, i hovedlisten som ordbasert (**ágreiningur**) og begrepsbasert (**ÁGREININGUR**) lemma og som ordbasert lemma i registerdelen, der det henvises til vedkommende artikkel i hoveddelen.

Denne sammenslåingen av de to fraseologiske ordbøkene var på mange måter et avgjørende tiltak som la grunnlaget for det videre arbeidet og sikret overgangen fra den trykte til den elektroniske ordboken.

Inkluderingen av kollokasjoner og andre ordforbindelser fra *Orða-staður* i den ordbaserte lemmalisten betyr at de ikke lenger ligger skjult i store artikler med en ofte uoversiktlig mikrostruktur, og sorteringen fremkaller en interessant sammenlikning av formelt beslektede fraser. I tillegg bygges det opp en ordliste med enkeltord som kompletterer den primære lemmalista. Dermed skapes det viktige forutsetninger for å tildele ordforbindelsene en ny funksjon som aktive faktorer i oppbyggingen av den leksikografiske beskrivelsen.

4. Oppbyggingen av *Íslenskt orðanet*

En enkel, men ganske grunnleggende fordel ved overgangen fra papirordboken til det elektroniske mediet består i at ordboksteksten lett kan tillegges nye fraser uten begrensninger, og uten at det på noen måte forstyrrer brukerens tilgang til beskrivelsen eller enkelte deler av den. Dette understreker ytterligere at ordforbindelsene på denne måten har fått en selvstendig status, noe som gjør det aktuelt å vurdere forholdet mellom de enheter vi definerer som lemmaer og adresser på den ene side, og de enheter som opptrer som angivelser til disse lemmaene på den andre. Sammenlikning med papirordbøker viser at det absolutte skillet mellom makrostrukturelle og mikrostrukturelle enheter er blitt opphevet. Den ordbaserte lemmalisten gir en verdifull innsikt i de innbyrdes relasjoner mellom ordforbindelsene, noe som må være relevant for en systematisk kartlegging av de ulike relasjoner i en større sammenheng. Søking i materialet kan foregå på ulike plan, og det kan sorteres på ulike måter og ut fra ulike kriterier.

På bakgrunn av dette stilles det to fundamentale spørsmål:

- Kan ordforbindelsene i enda høyere grad opptre og fungere som selvstendige leksikalske enheter i leksikografisk sammenheng?
- Kan ordforbindelsene tildeles en større funksjon enn hittil som aktive faktorer i oppbyggingen og utformingen av en allmenn leksikografisk beskrivelse?

Disse overveielsene ga motivasjonen til å bygge opp en ordboksmodell i elektronisk form der hele det fraseologiske materialet (i *Stóra orðabókin um íslenska málnotkun* og det fraseologiske registeret) blir utnyttet som grunnlag for en leksikografisk beskrivelse der fraser og ordforbindelser utgjør den aktive kjernen. Det utviklet seg til utformingen av et leksikografisk prosjekt under navnet *Íslenskt orðanet* (Islandsk ordnett), som fortsatt pågår i samarbeid mellom undertegnede og Þórdís Úlfarsdóttir. Her omfatter den ordbaserte lemmalisten i første omgang samtlige nøkkelord fra de to hovedkildene pluss alle sammensetninger i den morfologiske komponenten til *Orðastaður*, til sammen bortimot 200.000 lemmar. Til denne lemmalista knyttet det seg et liknende antall fraser og ordforbindelser av ulike slag.

4.1. Ordforbindelser avdekker semantiske relasjoner

Utgangspunktet for ordnettet var å kartlegge og markere semantiske relasjoner innenfor det islandske ordforrådet ut fra de syntagmatiske relasjoner som ordforbindelser (og da spesielt kollokasjonene) avslører. Ved å linke hver enkelt frase til alle innholdsord den inneholder, gis det blant kollokasjonene en klar oversikt over semantisk beslektede kollokatorer under de enkelte lemmaer. Her kan det bl.a. foretas sammenlikning mellom lemmaer som viser hvor mange og hvilke nøkkelord de er linket til, hvilke av disse nøkkelordene de er linket til begge to, og hvilke bare er linket til det ene av dem. Figur 2 viser en slik sammenlikning mellom lemmaene *löngun* 'lengsel' og *þrá* 'sterk lengsel' der det fremgår at 26 nøkkelord er linket til begge ordene, mens 42 hhv. 46 nøkkelord bare er linket til det ene av dem. En slik sammenlikning innenfor et større felt vil gi viktig informasjon om semantisk nærhet og ulik nærhetsgrad. Nært beslektede ord kan også føres sammen innenfor en liste med kollokasjoner ut fra leksikografens vurdering og slås sammen med ordgrupper under andre lemmaer gjennom felles ord, for dermed å bygge opp et større sammenhengende felt.

Semantiske relasjoner er ikke begrenset til enkeltord, de er også synlige blant konstruksjoner og idiomer, enten de er relatert til hverandre eller til enkeltord. Også her kan slike relasjoner komme til uttrykk gjennom felles leksikalske enheter. Her har verbene nok en gang en sentral funksjon ved å føre sammen en stor mengde kollokasjoner, konstruksjoner og idiomer, og dermed legge grunnlaget for en nærmere sammenkobling og gruppering.

Men verbene som ettordslemmaer kan ikke uten videre inngå i de semantiske relasjoner de viser i de ulike flerordsenheter, f.eks. i idiomer med en klar egenbetydning. Her må det tas hensyn til at hele argumentstrukturen og den syntaktiske variasjon i mange forbindelser og konstruksjoner har en avgjørende virkning på betydningen. For å kunne aktivisere verbene som lemmaer i ordnettet, må de i de fleste tilfeller omstruktureres som flerordslemmaer i samsvar med uttrykksformen i vedkommende forbindelse(r). Dette gir lemmalisten et helt nytt preg og fører til en kolossal økning i antall verblemmaer der en stor mengde verbale ordforbindelser og konstruksjoner blir inkludert og inngår i semantiske relasjoner med andre lemmaer.

4.2. Flerordslemmaer, entydiggjøring av lemmaene

Lemmatiseringen av verbene som flerordslemmaer bygger på den tidligere standardisering av ordforbindelsene som på visse punkter vil måtte presiseres ytterligere. Dette medfører at de allerfleste verblemmaene vil være semantisk entydige. Denne entydigheten er i grunn forutsetningen for en problemfri kartlegging av de semantiske relasjoner mellom lemmaene. Det forholdet får samtidig en dyptgående virkning for hele ordboksmodellen, for dermed blir det onomasiologiske aspektet satt i fokus. Istedenfor at ordforbindelsene må tilpasse seg som underordnede enheter en beskrivelse av (mest) flertydige enkeltord opptrer, de som selvstendige og i grunnen entydige lemmaer. Dermed er det ordforbindelsene (som flerordslemmaer) som gir tonen og krever en gjennomgående entydiggjøring av hele lemmabestanden, også blant enkeltordene. Resultatet er en lemmaliste med en blanding av ettords- og flerordslemmaer der en eventuell flertydighet blir markert i form av homonymindeks og en identifiserende merknad i lemmalisten.

Flerordslemmaer er ikke bare aktuelle for verbene, de har også stor betydning for lemmatiseringen av forskjellige adverbialer som i stor grad er semantisk relatert til andre leksikalske enheter. Blant adjektivene, der en klar entydiggjøring er problematisk, kan lemmaformen tillegges et entydiggjørende element i form av et prototypisk substantiv (basis). Også nominaler med en klar egenbetydning vil være aktuelle som flerordslemmaer på linje med substantiviske ettordslemmaer.

4.3. Grammatisk tagging av flerordslemmaer

Standardisering av flerordslemmaene har ikke bare en kontrolleffekt med hensyn til sammenhengen mellom semantisk beslektede lemmaer, den åpner også veien for en ytterligere markering av lemmaformene som består i å foreta en grammatisk tagging av samtlige flerordslemmaer. En slik operasjon gir en helt ny innsikt i de paradigmatisk relasjoner blant flerordslemmaene, den synliggjør karakteristiske strukturtrekk til enkelte verb, lemmaene kan sorteres ut fra grammatisk mønstre, og det kan fokuseres på enkelte mønstre og deres utbredelse. Den bidrar også til å styrke ordforbindelsen selvstendighet ved å tildele dem en grammatisk markering som belyser deres funksjon og karakter og kan sammenliknes med den tradisjonelle ordklassemarkering for ettordslemmaer. Figur 3 inneholder en grammatisk sortert liste med verblemmaer som innledes med verbet *borða* 'spise' og viser forekomsten av et bestemt mønster innenfor lista i tilknytning til andre verb.

I ordnettet spiller den grammatisk tagging en viktig rolle ved bearbeidelsen av dataene, ikke minst i forbindelse med den semantisk klassifisering av lemmaene, der det viser seg at strukturelle likheter gjerne gjenspeiler semantisk likheter (jf. Braasch 2008).

Denne innfallsvinkelen til en systematisk og selvstendig beskrivelse av ordforbindelser kan vel sies å stå i kontrast til den tradisjonelle tilnæringsmåte i allmenne ordbøker der også verblemmaene blir uttrykt som ettordsenheter. Men det betyr ikke at en tradisjonell ordboksbeskrivelse ikke kan dra nytte av den kartlegging og markering dataene gjennomgår i denne modellen. Foruten den strukturelle standardisering som fremkaller karakteristisk konstruksjoner og belyser verbets valens, vil den semantisk klassifisering være til stor nytte ved betydningsbeskrivelsen, både av enkelte verb og de verbale ordforbindelser som krever forklaringer og definisjoner. Innenfor ordnettet kan slike betydningsforklaringer anses å være en naturlig fortsettelse av den strukturelle og semantisk kartleggingen.

4.4. Semantisk klassifisering av lemmabestanden

De semantisk relasjoner ordnettet avslører, berører lemmaene på ulik måte, dels dreier det seg om synonymi og antonymi, i andre tilfeller består relasjonen mellom over- og underbegrep. I atter andre tilfeller har vi med å gjøre mer løse relasjoner eller felles tilhørighet til et semantisk felt (begrep). Det er en naturlig framgangsmåte å markere først de fast-

ere relasjoner og legge den markeringen til grunn for en videre kartlegging og klassifisering. Men her må det tas et visst hensyn til ulike forhold i de enkelte ordklasser. Blant verbene (som flerordslemmaer) vil innbyrdes relasjoner f.eks. ikke gjenspeiles i relaterte nøkkelord i samme grad som blant mange substantiver. Med tanke på det ble de verbale lemmaene tildelt begrepetiketter, på liknende måte som ordforbindelsene i *Orðaheimur*. Denne begrepsmessige klassifiseringen har vært en av de mest tidkrevende tiltak i oppbyggingen av ordnettet. I motsetning til *Orðaheimur*, der begrepsoverskriftene utgjør selve lemmalisten, fungerer de i det elektroniske mediet utelukkende som semantiske markører i bakgrunnen. I denne forbindelse har den grammatiske taggingen også hatt en viktig funksjon ved å avdekke og skjerpe sammenhengen mellom form og betydning.

Foruten en klassifisering som styres og kontrolleres av leksikografen, byr ordnettet på å foreta mekanisk klassifisering av ulike slag, som dels kan utnyttes som praktiske operasjoner under bearbeidelsen av materialet men som også kan ha en egenverdi og bli synliggjort overfor brukerne. Det gjelder f.eks. kumulative grupperinger av parataktiske uttrykk som i mange tilfeller gir en interessant innsikt i et semantisk og konnotativt sammenhengende ordforråd.

4.5. Flerdimensjonell ordboksmodell

Den ordboksmodell jeg her har beskrevet, forener i grunn ulike funksjoner som ut fra tradisjonell leksikografisk terminologi blir dekket av forskjellige typer ordbøker. Men samtidig er den fleksibel slik at leksikografen får bestemme sammensetningen og omfanget av den leksikografiske beskrivelsen som blir synliggjort overfor de allmenne brukerne. Lemmaene, så vel ettords- som flerordslemmaer og deres innbyrdes relasjoner, utgjør hovedkjernen og formidler informasjon som tradisjonelt dels tilhører begreps- og synonymordbøker, men som i tillegg dekker viktige funksjoner fraseologiske ordbøker skal oppfylle. Men bak denne beskrivelsen ligger det et mer omfattende strukturert fraseologisk materiale som i større eller mindre grad også kan bli gjort tilgjengelig og knyttet til den begrepsorienterte kjernen.

Entydiggjøring av lemmaene og lemmatisering av flerordsenheter har til følge at betydningsforklaringer og definisjoner ikke lenger er i blikkfeltet. Det spørres allikevel hvilken status betydningsforklaringer kan ha i en slik modell. Som tidligere nevnt tildeles de en enkel differensierende funksjon overfor homonymer i lemmalisten. Det er naturlig å

tenke seg at enkelte lemmaer kan tillegges forklaringer når det anses å være viktig for brukerne, f.eks. metaforiske uttrykk blant flerordslemmaer. Men leksikalske og syntaktiske elementer i selve lemmaene har i mange tilfeller en betydningsforklarende funksjon, foruten at de relasjoner lemmaene står i til hverandre indirekte vil gi uttrykk for det semantiske innholdet, noe som forminsker behovet for formelle definisjoner og forklaringer. Hvis vi sammenlikner dette forholdet med det som gjelder for semasiologiske ordbøker, er det klart at materialet til ordnettet ikke vil vitne om alle de delbetydninger ordene blir tildelt i en ordbok av den type. Men det uttrykker til gjengjeld et interessant skille mellom det ordforrådet og de ordforbindelser som er mest aktive i språkbruken, og det mindre aktive og mer sjeldne ordtilfanget. Når det gjelder å avgrense en sentral kjerne av ordforrådet, f.eks. i forbindelse med konkrete ordboksprosjekter, vil opplysninger om ordenes forekomst i ordforbindelser av ulike slag kunne brukes som en interessant målestokk.

5. Grunnelementer i ordboken

I samsvar med oppbyggingen av ordnettet er det en naturlig framgangsmåte at de datamengder i grunndatabasen som anses å være ferdig markert og analysert etter hvert blir gjort tilgjengelig for allmennheten på nettet. Websiden www.ordanet.is tjener her dette formålet, og slik den ser ut nå, er innholdet konsentrert om de semantiske relasjoner mellom lemmaene. Det understreker at en ordboksbeskrivelse av denne type ikke er en avgrenset størrelse. Den må tvert imot kontinuerlig kunne utvides med nytt materiale, der nye ordforbindelser på den ene side styrker og kompletterer de beskrivelser allerede er blitt foretatt og på den andre side tilføyer nye lemmaer.

Overskriften til dette foredraget uttrykker på en forsiktig måte det synet at de elementer som i leksikografisk terminologi blir forent under termen **ordforbindelse**, burde tildeles en mer fremtredende posisjon leksikografisk teori og i utarbeidelsen av konkrete leksikografiske produkter enn tradisjonen tilsier. Det henger ikke bare sammen med at ordbøkene i konkrete tilfeller ikke klarer å dekke mangfoldigheten og bredden i fraseologien. Det går mer på den plassering ordforbindelsene har hatt i skyggen av de enkelte ord ordbøkene beskriver. Hvis de med leksikografenes hjelp kan tre ut av den skyggen og beskrives på deres egne premisser, vil det også kunne påvirke vår innstilling til beskri-

velsen av de enkelte ord. I den sammenheng reiser seg igjen det spørsmålet som stilles i overskriften.

Litteratur

- Braasch, Anna 2008: Verber med fælles syntaktiske konstruktioner og semantisk slægtskab. I: Ásta Svavarsdóttir et al. (red.), *Nordiske Studier i Leksikografi 9. Rapport fra konference om leksikografi i Norden, Akureyri 22.–26. maj 2007*. (Skrifter udgivet af Nordisk Forening for Leksikografi 10), Reykjavík, 43–59.
- Burger, Harald 1989: Phraseologismen im allgemeinen einsprachigen Wörterbuch. I: *Wörterbücher. Dictionaries. Dictionnaires. Ein internationales Handbuch zur Lexikographie*. Hausmann, F.J. et al. (red.), 593–599. Berlin/New York: de Gruyter.
- Farø, Ken 2007: Danish Phraseology. I: *Phraseologie/Phraseology. Ein internationales Handbuch der zeitgenössischen Forschung*. Burger, H. et al. (red.), 949–958. Berlin/New York: de Gruyter.
- Jónsson, Jón Hilmar 2008: Lemmatisering og tilgangsstruktur: refleksjoner på bakgrunn av *Íslenskt orðanet*. I: Ásta Svavarsdóttir et al. (red.), *Nordiske Studier i Leksikografi 9. Rapport fra konference om leksikografi i Norden, Akureyri 22.–26. maj 2007*. (Skrifter udgivet af Nordisk Forening for Leksikografi 10), Reykjavík, 269–281.
- Moon, Rosamund 2007: Phraseology in general monolingual dictionaries. I: *Phraseologie/Phraseology. Ein internationales Handbuch der zeitgenössischen Forschung*. Burger, H. et al. (red.), 909–918. Berlin/New York: de Gruyter.
- NFL = *Nordisk leksikografisk ordbok* 1997. Red.: Henning Bergenholtz, Ilse Cantell, Ruth Vatvedt Fjeld, Dag Gundersen, Jón Hilmar Jónsson og Bo Svensén. Oslo: Nordisk forening for leksikografi/Nordisk språksekretariat.
- Orðastaður = Jón Hilmar Jónsson 1994: *Orðastaður. Orðabók um íslenska málnotkun*. Reykjavík: Mál og menning. (2. utgave, utvidet og revidert 2001. Reykjavík: JPV-útgáfa.)
- Orðaheimur = Jón Hilmar Jónsson 2002: *Orðaheimur. Íslensk hugtaka-orðabók með orða- og orðasambandaskrá*. Reykjavík: JPV-útgáfa.
- Schemann, Hans 1989: Das phraseologische Wörterbuch. I: *Wörterbücher. Dictionaries. Dictionnaires. Ein internationales Handbuch zur Lexikographie*. Hausmann, F.J. et al. (red.), 1019–1032. Berlin/New York: de Gruyter.

Stóra orðabókin um íslenska málnotkun = Jón Hilmar Jónsson 2005:
Stóra orðabókin um íslenska málnotkun (en Cd-utgave følger med).
Reykjavík: JPV-útgáfa.

Nettsteder

Íslenskt orðanet, <http://www.ordanet.is>
Fraseologisk register til Orðabók Háskólans,
[https:// www.lexis.hi.is/osamb/osamb.pl](https://www.lexis.hi.is/osamb/osamb.pl)

Jón Hilmar Jónsson
Forskningsprofessor
Árni Magnússon instituttet for islandske studier / Islands universitet
Neshagi 16
IS-107 Reykjavík
Island
jhj@hi.is

Appendiks: figur 1–3

FIGUR 1. Tekstavsniitt fra *Stóra orðabókin um íslenska málnotkun*. Ordet *ágreiningur* 'uenighet, meningsforskjell' forekommer i tre lemmafunksjoner, i hovedlisten som ordbasert (**ágreiningur**) og begrepsbasert (**ÁGREININGUR**) lemma og som ordbasert lemma i registerdelen, der det henvises til vedkommende artikkel i hoveddelen.

1a

<p>ágreiningur no kk</p> <p>(það er) ágreiningur um <þetta, málið, ákvörðun hans, stefnu flokksins>/út af <þessu> <i>það er ágreiningur um það hvort eigninni hefur verið ráðstafað löglega</i></p> <p>það er ágreiningur við <hann, hana> (út af <þessu>) það er ágreiningur á milli <þeirra> það er ágreiningur með <þeim> ágreiningur kemur upp ágreiningur rís (út af <formsatriðum>) ágreiningur er uppi ágreiningurinn magnast gera (ekki/engan) ágreining (um <málið>) <þetta, málið> veldur ágreiningi (<reyna að>) jafna ágreininginn (<það tókst að>) leysa ágreininginn (<reyna að, það tókst að>) eyða ágreiningnum ágreiningurinn er úr sögunni ákvæði <djúpstæður, alvarlegur> ágreiningur</p>

1b

ÁGREININGUR

vera ósammála <honum, henni> (um <þetta>)
 <þeir, þær, þau> eru ósammála (um <þetta; margt>)
 <þeir, þær, þau> eru ekki á sama máli
 vera annarrar skoðunar (en <hann, hún>)
 skoðanir <þeirra> fara ekki saman
 <fundarmenn> eru ekki á eitt sáttir (um <þetta>)
 vera á öndverðum meiði við <hann, hana>
 standa á öndverðum meiði við <hann, hana>
 <þeir, þær, þau > eru á öndverðum meiði (um <þetta;
 málið>)
 <þeir, þær, þau> standa á öndverðum meiði
 <þá, þær, þau> greinir á um <það hvar húsið eigi að
 standa; margt>
 <þá, þær, þau> skilur á um <það hver sé undirrót
 vandans>

1c

alvarlegur ágreiningur	alvarlegur	lo
ágreiningur er uppi	ágreiningur	no kk
ágreiningur kemur upp	ágreiningur	no kk
ágreiningur rís	ágreiningur	no kk
ágreiningur um <þetta; málið, ákvörðun hans, ste	ágreiningur	no kk
ágreiningurinn er úr sögunni	ágreiningur	no kk
ágreiningurinn magnast	ágreiningur	no kk
djúpstæður ágreiningur	djúpstæður	lo
eyða ágreiningnum	ágreiningur	no kk
gera ágreining	ágreiningur	no kk
gera ágreining	ÁGREININGUR	
gera ekki/engan ágreining	AFSTADA	
jafna ágreininginn	ágreiningur	no kk
leysa ágreininginn	ágreiningur	no kk
það er ágreiningur á milli <þeirra>	ágreiningur	no kk
það kemur/kom upp ágreiningur <í nefndinni>	ÁGREININGUR	

FIGUR 2. Databasen til *Íslenskt orðanet*. Sammenlikning mellom lemmaene *löngun* 'lengsel' og *þrá* 'sterk lengsel' der det fremgår at 26 nøkkelord er linket til begge ordene, mens 42 hhv. 46 nøkkelord bare er linket til det ene av dem. Ulike ikoner markerer de forskjellige relasjonstyper.

þrá 46	bæði 26	löngun 42
H eftir- forl	○ ómótstæðilegur lo ○	í- forl H
H heim- forl	○	blíður lo H
H hug- forl	○ óstjórnlegur lo ○ ○	dulvitaður lo ○
H út- forl	○ sterkur lo ○ ○	fullur lo H
○ brennandi lo	○ sterkur [tilfinning] lo	inngróinn lo ○
○ dulinn lo	no ○ ○	kenndrænn lo ○
○ eldheitur lo	H athöfn no kvk H	lostugur lo ○
ot hræddur lo	H ○ ást no kvk ○	mannkærlegur lo ○
○ leyndur lo	H ferð no kvk H	óstöðvandi lo ○
○ ljúfsár lo	Ⓡ fíkn no kvk T	vanmáttugur lo ○
○ logandi lo	H framfarir no kvk flt H	auga no hvk H
H áttahagar no kk flt	H frami no kk H	augu no hvk flt H
○ ○ brjóst no hvk	H framkvæmd no kvk H	ásókn no kvk ○
H draumur no kk	H frelsi no hvk H	bardagi no kk H
H eilífð no kvk	H fróðleikur no kk H	brennivín no hvk H ○
H eining no kvk	Ⓡ fýsn no kvk Ⓡ ○	dráp no hvk H
H far no hvk	Ⓡ girnd no kvk Ⓡ ○	forvitni no kvk ○
H fegurð no kvk	H lestur no kk H	gróði no kk H
H flug no hvk	H líf no hvk H	hefnd no kvk H
○ friður no kk	Ⓡ lyst no kvk Ⓡ	hvíld no kvk ○
H frægð no kvk	H mennt no kvk H	-leysi no hvk H
H fullkomleiki no kk	H menntun no kvk H	matarlyst no kvk ○
H fullkonnun no kvk	H sjálfræði no hvk H	metorð no hvk flt H
○ hamingja no kvk	H starf no hvk H	nám no hvk H
○ hjarta no hvk	H völd no hvk flt H	næði no hvk ○

FIGUR 3. Databasen til *Íslenskt orðanet*. En grammatisk sortert liste med verblemmaer som innledes med verbet *borða* 'spise', knyttet til en tilsvarende taggliste. Det vises hvordan én bestemt tagg innenfor lista forener ulike verblemmaer.

3a

5	<u>MARK</u>	so fn-r-a lo-a	<u>borða sig fullsaddan so</u>
6	<u>MARK</u>	so fn-r-a lo-a	<u>borða sig pakksaddan so</u>
7	<u>MARK</u>	so fn-r-a lo-a	<u>borða sig saddan so</u>
8	<u>MARK</u>	so fs no-a	<u>borða undir spreng so</u>
9	<u>MARK</u>	so fs no-dg	<u>borða með guðsgófflunum so</u>
10	<u>MARK</u>	so fs no-g <ao>	<u>borða til miðdags <þar> so</u>
11	<u>MARK</u>	so lo-a	<u>borða þurrt so</u>
12	<u>MARK</u>	so no-a	<u>borða morgunmat so</u>
13	<u>MARK</u>	so no-a	<u>borða morgunverð so</u>
14	<u>MARK</u>	so no-a fn-a	<u>borða fylli sína so</u>
15	<u>MARK</u>	so no-a fn-a	<u>borða nægju sína so</u>
16	<u>MARK</u>	so <no-a>	<u>borða <kjöt, fisk, matinn> so</u>
17	<u>MARK</u>	so <no-a>	<u>borða <morgunmat, kvöldmat> so</u>

3b

4	<u>MARK</u>	so fn-r-a lo-a	<u>borða sig saddan so</u>
13	<u>MARK</u>	so fn-r-a lo-a	<u>drekka sig fullan so</u>
14	<u>MARK</u>	so fn-r-a lo-a	<u>drekka sig ofurólvi so</u>
15	<u>MARK</u>	so fn-r-a lo-a	<u>drekka sig ófæran so</u>
16	<u>MARK</u>	so fn-r-a lo-a	<u>drekka sig svindrukkinn so</u>
17	<u>MARK</u>	so fn-r-a lo-a	<u>drekka sig svinfullan so</u>
18	<u>MARK</u>	so fn-r-a lo-a	<u>drekka sig ölóðan so</u>
19	<u>MARK</u>	so fn-r-a lo-a	<u>éta sig máttlausan so</u>
20	<u>MARK</u>	so fn-r-a lo-a	<u>éta sig saddan so</u>
21	<u>MARK</u>	so fn-r-a lo-a	<u>fá sig fullreyndan so</u>
22	<u>MARK</u>	so fn-r-a lo-a	<u>ganga sig skólausan so</u>
23	<u>MARK</u>	so fn-r-a lo-a	<u>gera sig athlæqilegan so</u>
24	<u>MARK</u>	so fn-r-a lo-a	<u>gera sig blíðan so</u>
25	<u>MARK</u>	so fn-r-a lo-a	<u>gera sig breiðan so</u>
26	<u>MARK</u>	so fn-r-a lo-a	<u>gera sig digran so</u>