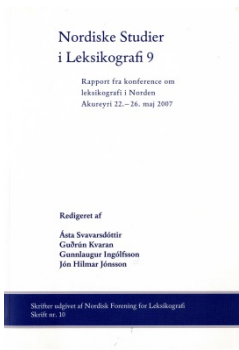


NORDISKE STUDIER I LEKSIKOGRAFI

Titel:	Modell för beskrivning av kollokationer i ett medicinskt lexikon (MedLex)	
Forfatter:	Emma Sköldberg og Maria Toporowska Gronostaj	
Kilde:	Nordiska Studier i Lexikografi 9, 2008, s. 433-446 Rapport fra Konference om leksikografi i Norden, Akureyri 22.-26. maj 2007	
URL:	http://ojs.statsbiblioteket.dk/index.php/nsil/issue/archive	

© Nordisk forening for leksikografi

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 Nordiske studier i leksikografi (1-5) 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.

Modell för beskrivning av kollokationer i ett medicinskt lexikon (MedLex)

The paper presents a pilot study of the treatment of collocations in an electronic medical lexicon, MedLex, currently under construction at the University of Gothenburg, Department of Swedish. The lexicon is descriptive and intended to support both reception and production. The study focuses on noun collocations, especially those which describe diseases and medical conditions. We present an analysis model based on a combination of semasiological and onomasiological premises. The model results in a classification into lexically conditioned collocations and type-related collocations. We also discuss how medically related collocations may be described in the MedLex lexicon.

1. Inledning

Dagens samhälle kännetecknas av en allt större konsumtion av hälso- och sjukvård. Människor förväntas vara aktiva hälsokonsumenter som tar del av all den medicinska information som är tillgänglig på Internet och i massmedierna. Dessutom ges patienter numera möjlighet att läsa sina egna patientjournaler. För att ta del av all denna hälsoinformation krävs olika grad av förkunskaper. Även i en konkret vårdssituation måste läkare, patienter och anhöriga förstå vad som sägs men också kunna uttrycka sig på ett nyanserat sätt för att undvika missförstånd. Kollokationer utgör en viktig del i en sådan kommunikation, men deras reception och produktion kan vålla problem för dem som är mindre bevandrade i det svenska språket. Av den anledningen utgör medicinska kollokationer ett intressant studieobjekt som bör uppmärksammas i högre grad i det lexikografiska sammanhanget.

Syftet med den här artikeln är att presentera en pilotstudie om hantering av kollokationer i ett framtida deskriptivt medicinskt elektroniskt lexikon, Medlex. Det är tänkt att vara polyfunktionellt och avsett för reception men i kanske ännu högre grad för produktion. Det ska kunna vara till nytta för både läkare och patienter med eller utan svenska som förstaspråk (jfr Bergenholtz & Tarp 1994: 16–25). Fokus i studien ligger på kollokationer med substantiv och de undersöks ur ett synkront och funktionellt perspektiv. Genom studien hoppas vi kunna bidra till en mer systematisk beskrivning av svenska medicinska kollokationer ur både ett lexikologiskt och lexikografiskt perspektiv. Tanken är att det blivande lexikonet ska vara ett slags brygga – och fylla en lucka – mellan allmänspråkliga

ordböcker och mer renodlade fackordböcker. Innehållet i lexikonet kan också komma till användning inom språkteknologi för bl.a. semi-automatisk igenkänning av kollokationer i medicinska texter.

Artikeln är disponerad på följande vis: inledningsvis redogör vi för vår syn på termen *kollokation* och därefter presenterar vi vårt material. Därpå diskuterar vi vår analysmodell och innehållet i det elektroniska lexikonet. Artikeln avslutas med en sammanfattning och några slutsatser.

2. Om termen *kollokation*

Litteraturen om kollokationer är omfattande och termen *kollokation* används med skiftande innehåll (se t.ex. Hausmann 2004, Svensén 2004 och Siepmann 2005 med anförd litteratur). De nämnda forskarna diskuterar bl.a. kollokationers grundläggande egenskaper inklusive den gradvisa övergången mellan kollokationer och angränsande kategorier som idiom och fria kombinationer. För att undvika oklarheter redogör vi nedan för vår syn på formella och semantiska drag hos prototypiska kollokationer. Vår syn är inspirerad av framför allt den strukturella och funktionella beskrivning av engelska kollokationer som presenteras i Bartsch (2004). Andra viktiga inspirationskällor är Hausmann (1985), Malmgren (2003), Heid (1994, 2004) och Martin (2006).

- En prototypisk kollokation är en återkommande kombination av två eller flera ord som har en intern semantisk och syntaktisk struktur och som uppvisar lexikalisk bindning, t.ex. *ta tempen* och *falla i koma*. Med det ovan sagda vill vi markera vårt väsentligen systemorienterade synsätt på kollokationer. Det står i kontrast till ett korpusinriktat perspektiv enligt vilket kollokationer betraktas som ett statistiskt fenomen, beskrivna i termer av bigram och trigram, t.ex. *läkaren och* och *patienter som lider* (jfr Sinclair 1991; Kjellmer 1994).
- En prototypisk kollokation kan delas upp i en bas och en kollokator. Basen betraktas som ett slags semantiskt huvudord och kollokatorn som ett biord. I kollokationen *ta tempen* är *tempen* bas och det delexikaliserade verbet *ta* kollokator. Man kan här tala om en form av direktionalitet inom ordförbindelsen i och med att basen *tempen* styr valet av kollokatorn *ta* och inte tvärtom. Det finns dock kollokationer där direktionaliteten är svårare att fastställa och det är t.ex. när verbet har en mer pregnant betydelse. Exempel är *ordinera vila* och *ordinera läkemedel*.
- En prototypisk kollokation är vanligen genomskinlig. Det medför att den ofta är relativt lätt att tolka. Men graden av genomskinlighet kan variera, och den påverkas av flera olika faktorer, t.ex. de ingående ordens eventuella polysemi. I många fall disambigueras polysema ord när de ingår i en kollokation, t.ex. *förband* i *lägga förband* (jfr användning av substantivet

i medicinsk respektive militär domän). Något som också påverkar genomskinligheten är den eventuella metaforiken hos kollokationens led, t.ex. *tung i tung andning* och *djup i djup medvetlöshet*.

- Många prototypiska kollokationer uppvisar samma typ av relation mellan bas och kollokator. De oftast förekommande relationerna har beskrivits i termer av lexikala funktioner (Mel'čuk 1998). Exempelvis uttrycker den lexikala funktionen **Magn** ett gradförhållande, och den kommer till uttryck i kollokationer som *hög feber*, *djup depression* och *svår huvudvärk*.
- En prototypisk kollokation rymmer ofta implicit kulturell eller encyklopedisk information. Tillgång till den implicita informationen är en förutsättning för att man ska kunna tolka och använda kollokationerna på ett korrekt sätt. Till exempel har kollokationen *fri abort* en innebörd som är samhällsrelaterad. För att tolka *medicinsk* respektive *kirurgisk abort* krävs det däremot fackkunskaper. Andra exempel, som kan vålla både receptions- och produktionsproblem, är medicinska kollokationer med attributen *positiv* eller *negativ*, t.ex. *positivt/negativt graviditetstest*.

Vidare brukar kollokationer klassificeras utifrån deras inre syntaktiska struktur och klassificeringar av svenska, engelska, tyska respektive spanska kollokationer görs av bl.a. Svensén (2004), Bartsch (2004), Heid & Gouws (2006) samt Koike (2001). Nedan återges några återkommande syntaktiska klasser inom det medicinska området:

ADJ + SUBST	<i>öppet sår, molande värk</i>
V + SUBST	<i>ställa diagnos, lägga om ett sår</i>
V + PP	<i>hoppa ur led, svara på medicinen</i>
V + ADV	<i>andas djupt, må illa</i>
ADV + ADJ	<i>djupt medvetlös, gravt synskadad</i>
SUBST + V	<i>febern stiger, blodtrycket sjunker</i>

Ett tydligt drag är också att det bland kollokationerna förekommer parallella lexikala strukturer. I princip samma semantiska innehåll kan, genom morfologisk variation, realiseras med olika syntaktiska strukturer. Ett exempel är den redan nämnda kollokationen *djupt medvetlös* (ADV + ADJ), som är starkt besläktad med *djup medvetlöshet* (ADJ + SUBST). Ett annat fall är *blodtrycket sjunker* (SUBST + V) och *sjunkande blodtryck* (ADJ + SUBST). Vad beträffar morfologin uppträder vissa kollokationer ofta i en särskild grammatisk form. Det gäller t.ex. *tillstöta komplikationer*, vars bas främst förekommer i pluralis.

Ett annat fenomen, som uppmärksammats av bl.a. Heid (2004: 732), är att kollokationer kan kombineras med varandra och bilda mer komplexa flerordsenheter. Exempelvis kan kollokationerna *starkt beroendeframkallande* och *beroendeframkallande medel* fogas samman till *starkt beroendeframkallande medel*.

3. Material

Som utgångspunkt för vår studie har vi valt cirka 50 allmänspråkliga ettordsenheter som är representativa för ämnesområdet medicin. Det rör sig t.ex. om beteckningar på sjukdomar, ingrepp, behandlingar och symtom. Med tanke på det begränsade utrymmet här koncentrerar vi oss främst på ord som betecknar sjukdomar och sjukdomstillstånd. De aktuella orden upptas i allmänspråkliga svenska ordböcker, t.ex. NEO (1995–96), och fackspråkliga medicinska lexikon, exempelvis Lindskogs *Medicinsk terminologi* (2004).

För att finna relevanta kollokationer med de utvalda orden har vi använt både allmänspråkligt och fackspråkligt korpusmaterial. De allmänspråkliga texterna ingår i Språkbanken (på cirka 100 miljoner ord) vid Institutionen för svenska språket, Göteborgs universitet (<<http://spraakbanken.gu.se/>>). De medicinska texterna är hämtade ur Medlex korpus vid samma institution. Medlex korpus, som ännu inte är tillgänglig för allmänheten, innehåller drygt 14 miljoner ord och består av medicinska texter hämtade från bl.a. Läkartidningen och Socialstyrelsens hemsida. Våra data har också jämförts med motsvarande sökningar på Google (<<http://www.google.se>>). Som verktyg har vi dels använt allmänt tillgängliga konkordansprogram, dels ett kollokationsprogram, utarbetat av Christian Sjögreen vid Institutionen för svenska språket, Göteborgs universitet.

Fokus i vår studie ligger således på kollokationer. Däremot utesluter vi för tillfället etablerade medicinska flerordstermer och eponymer som t.ex. *gula fläck-en* och *Downs syndrom*. Det kan dock noteras att det enligt Manning & Schütze (1999: 141) föreligger en överlappning mellan begreppet kollokation och fackspråkliga flerordstermer.

4. Analysmodellen

Vår analysmodell bygger på en kombination av ett semasiologiskt och ett onomasiologiskt synsätt. Det semasiologiska perspektivet har diskuterats mer utförligt i avsnitt 2 ovan. Hädanefter uppmärksammas det onomasiologiska perspektivet och mer onomasiologiskt grundade kategorier. Inom det medicinska området finns det flera subdomäner: ANATOMI, SJUKDOM, BEHANDLING etc. Varje subdomän förutsätter en uppsättning av specifika, och sinsemellan ganska olika, scenarier som kan beskrivas med vissa onomasiologiska parametrar. Till exempel kan subdomänen SJUKDOM analyseras utifrån parametrar som Insjuknande, Tillstånd, Grad, Förlopp, Typ och Tillfrisknande. För att illustrera tankegången anges här några typiska kollokationer med *depression* som bas:

- Insjuknande: *drabbas av (en) depression, gå in i en depression*
- Tillstånd: *ha en depression, lida av depression*
- Grad: *lätt(are) depression, djup depression,*
- Förlopp: *akut depression, kronisk depression*

- Typ: *årstidsbunden depression, bipolär depression*
- Tillfrisknande: *komma ur en depression, ta sig ur en depression*

För att kunna upptäcka generella mönster har vi valt att jämföra *depression* med tre andra sjukdomar och sjukdomstillstånd, närmare bestämt *cancer*, *diabetes* och *förkylning*, som inbördes är mycket olika, vad gäller t.ex. symtom och allvar. I samband med beteckningarna aktualiseras ett antal typiska parametrar, men listan är inte uttömmande. I samtliga exempel nedan utgör sjukdomarna bas i de enskilda kollokationerna.

Parameter	<i>depression</i>	<i>cancer</i>	<i>diabetes</i>	<i>förkylning</i>
Insjuknande	<i>få en d.</i> <i>drabbas av (en) d.</i> <i>gå in i en d.</i>	<i>få c.</i> <i>drabbas av c.</i> <i>utveckla c.</i>	<i>få d.</i> <i>drabbas av d.</i> <i>utveckla d.</i>	<i>få en f.</i> <i>drabbas av en f.</i> <i>dra på sig en f.</i> <i>ådra sig en f.</i>
Tillstånd	<i>ha d.</i> <i>lida av (en) d.</i>	<i>ha c.</i> <i>lida av c</i>	<i>ha d.</i> <i>lida av d.</i>	<i>ha en f.</i>
Grad	<i>djup d.</i> <i>svår d.</i> <i>lätt(are) d.</i> <i>mild d.</i>	<i>långt framskriden c.</i>	<i>svår d.</i>	<i>elak f.</i> <i>kraftig f.</i> <i>lätt f.</i>
Förlopp	<i>kronisk d.</i> <i>långdragen d.</i> <i>akut d.</i>	<i>kronisk c.</i>	<i>långvarig d.</i>	<i>långvarig f.</i> <i>långdragen f.</i> <i>envis f.</i>
Typ	<i>årstidsbunden d.</i> <i>bipolär d.</i>	<i>ärfilig c.</i> <i>obotlig c.</i> <i>aggressiv c.</i>	<i>ärfilig d.</i>	
Tillfrisknande	<i>komma ur en d.</i> <i>ta sig ur en d.</i>	<i>botad c.</i>		

Tabell 1: Parameterbaserad klassificering av kollokationer med substantiven *depression*, *cancer*, *diabetes* och *förkylning*

Av exemplen i tabell 1 framgår att det finns kollokationer som är typiska för endast en av sjukdomarna men som inte används i samband med de andra. Man kan t.ex. *gå in i en depression*, men inte **gå in i en cancer/diabetes/förkylning*. Man talar om en *djup depression* men inte om **djup cancer/diabetes/förkylning*. Andra kollokationer med något begränsad användning är *dra på sig i dra på sig en förkylning* och *obotlig i obotlig cancer*.

Vidare finns det kollokationer som uppträder med några av de aktuella substantiven. Till exempel finns det flera förekomster av *lida av (en) depression/cancer/diabetes* i materialet, men däremot inte *lida av (en) förkylning*. Detta faktum torde kunna förklaras med att den senare åkomman inte är fullt så allvarlig som de andra. Med andra ord är det språkanvändarnas kunskaper om faktiska förhållanden som spelar en viktig roll för deras val av adekvata kollokationer.

Slutligen finns det kollokatorer som är aktuella i samband med samtliga sjukdomar ovan. Det gäller t.ex. *få* i *få (en) depression/cancer/diabetes/förkylning* och *drabbas av* i *drabbas av (en) depression/cancer/diabetes/förkylning*. En annan återkommande kollokator är *ha*. Dessa kollokatorer förekommer troligtvis med alla typer av sjukdomar, och utgör därför underlag för ett återkommande mönster med begreppet SJUKDOM i basen.

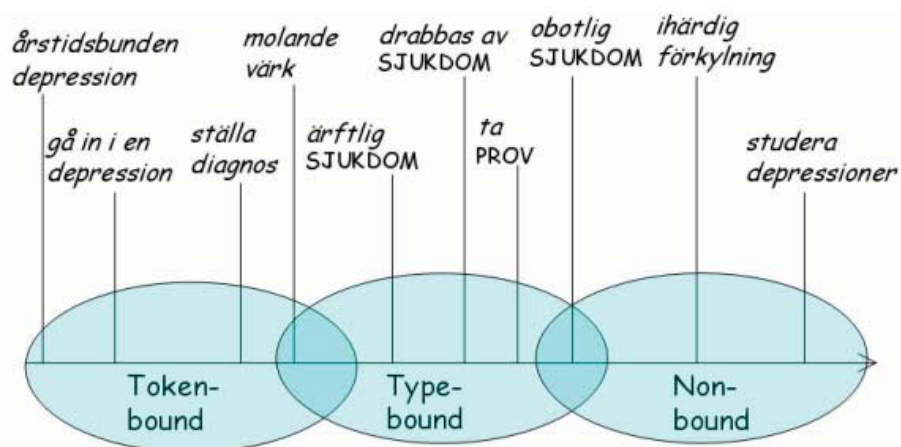
De förhållanden som illustreras i tabell 1 kan relateras till ett arbete av Willy Martin (2006). I sin artikel argumenterar han för en ramsemantisk modell för beskrivning av lexikonenheter i elektroniska ordböcker. Som illustration använder Martin exempel hämtade från det semantiska fältet Musikinstrument. Modellen utgår från en uppsättning parametrar, som är relevanta för det aktuella semantiska fältet, samt deras språkliga realisationer. De postulerade parametrarna utgör också ett viktigt stöd för en systematisk och uttömmande beskrivning av kollokationer.

Martin (2006: 289) förespråkar en uppdelning i *token-bound* och *type-bound collocations* och relaterar dem till *non-bound/free combinations*. Skillnaden mellan kollokationstyperna och de fria kombinationerna ligger i en fallande grad av bindning mellan de ingående leden. För fria kombinationer är bindningen svagast eftersom dessa ordförbindelser uppvisar minst begränsningar beträffande konstituenternas semantiska kombinatorik. Av Martins modell framgår det dock tydligt att gränserna mellan kategorierna inte är särskilt skarpa och att de går in i varandra.

Med utgångspunkt i Martin (2006) skiljer vi mellan *lexikalt betingade kollokationer* (token-bound) och *typrelaterade kollokationer* (type-bound). De lexikalt betingade kollokationerna uppvisar ett starkt beroende mellan leden och tillåter endast begränsade synonymiska byten vad gäller ingående konstituent, t.ex. *årstidsbunden/årstidsrelaterad depression*. Kännetecknande för typrelaterade kollokationer är att deras kollokatorer återkommer i ett antal olika kollokationer och att deras bas i de flesta fall kan generaliseras och uttryckas med hjälp av en kategori, t.ex. SJUKDOM. Möjligheten att byta ut basen mot en hypo- eller hyperonym tyder på att en typbunden relation finns inbäddad i kollokationen, t.ex. *ärfilig SJUKDOM* och *drabbas av SJUKDOM*. De lexikalt betingade kollokationerna uppvisar således högre grad av bindning mellan leden än de typrelaterade.

Förhållandet mellan lexikalt betingade och typrelaterade kollokationer samt fria ordkombinationer kan illustreras med hjälp av figur 1.

Hittills har vi diskuterat den modell som vi använder för att analysera de medicinska kollokationerna. Nedan följer en mer ingående presentation av hur de beskrivs i själva lexikonet.



Figur 1: Ordkombinationer ordnade efter grad av bindning (efter Martin 2006)

5. Kollokationer i MedLex ordbok

Den elektroniska MedLex-ordboken kommer att bestå av två länkade moduler. Den ena är en allmänspråklig lexikonmodul och den andra är en kollokationsmodul. Den senare ska ses som ett komplement till den första, men tanken är att modulerna även ska kunna användas oberoende av varandra.

Artiklarna kommer att vara uppbyggda kring ord som utgör bas i kollokationerna. Dessa ord betraktas som primära ingångar i Medlex. Som sekundära ingångar fungerar kollokatorer inklusive relevanta kollokationer. Som exempel kan *ta pulsen* anföras. Substantivet *puls* är primär ingång till båda modulerna. Däremot är *ta* och *ta pulsen* sekundära ingångar. Man kan undra varför vi kommer att lista kollokatorer som t.ex. *ta* som enskild ingång. Skälet är att man ska kunna navigera i lexikonet utifrån alla ord som förekommer som bas eller kollokator i det medicinska ordförrådet utan att på förhand behöva veta kollokationens exakta form. Om man söker på hela kollokationen, får man direkt information om hur just den används (se vidare avsnitt 5.2 nedan).

I det följande ger vi en översikt över typer av information som ska inkluderas i lexikonet. Däremot gör vi inte anspråk på att visa hur informationen ska visualiseras i lexikonet.

5.1. Allmänspråklig lexikonmodul

I artiklarna i den allmänspråkliga lexikonmodulen ges information om uttal (i hörbart format), ordklassinformation och uppgifter om böjning samt betydelse i form av betydelseindikator, definition och definitionstillägg. Med en betydelseindikator menas en koncis formulering som speglar det huvudsakliga innehållet i ett visst betydelsemoment.

Det finns flera syften med att inkludera betydelseindikatorer. För det första

blir ordets övergripande begreppskategori synlig. För det andra vägleds användaren snabbt i valet av betydelsemoment. Detta gäller särskilt för polysema medicinska ord, t.ex. *cancer*, som får två betydelseindikatorer: SJUKDOM respektive TUMÖR. För det tredje möjliggör betydelseindikatorerna att man kan göra tesauristiska sökningar. Ordboksanvändaren kan exempelvis få fram en lista över alla ingångar som har betydelseindikatorn SJUKDOM.

Artiklarna ska också innehålla semantiskt relaterade ord (som synonymer, antonymer, hyperonymer och hyponymer) samt information om kollokationer. I artiklarna presenteras även några aktuella sammansättningar, enklare redigerade språkexempel och information om morfologiskt relaterade ord som tillhör samma eller andra ordklasser. Dessutom ges en länk till motsvarande ingång i kollokationsmodulen. Mer konkret kan det, utifrån substantivet *depression*, se ut på följande sätt:

uppslagsord: *depression*

uttal: (ljud)

ordklass: substantiv

böjning: *depression, depressionen, depressioner, depressionerna* m.fl.

betydelseindikator: SJUKDOM

definition: sjukligt tillstånd med svår (långvarig) nedstämdhet

definitionstillägg: ofta omfattande ångest, handlingsförlamning och livströtthet

sammansättningar: *utmattningsdepression, förlossningsdepression* m.fl.

semantiskt relaterade ord: motsats *mani*

kollokationer:

ADJ + *depression*: *djup, svår; lätt(are) | akut, långdragen, kronisk | årstidsbunden*

V + *depression*: *drabbas av (en), gå in i en | ha, lida av | ta sig ur en*

exempel: *hon lider av en djup depression som orsakar sömnlöshet*

morfologiskt relaterade ord: v. *deppa*; adj. *deprimerad, deprimerande, deppig, depressiv*

länk: till kollokationsmodul

Som synes listas kollokationerna först utifrån grammatisk struktur. I nästa steg indelas de med lodstreck efter aktuella parametrar: Insjuknande, Tillstånd, Grad etc. (jfr OCDSE 2002). Semikolon avgränsar semantiska undergrupper inom respektive parameter. Denna parameterbaserade indelning förefaller i många fall sammanfalla med en klassificering utifrån typer av lexikala funktioner, presenterade av Mel'čuk (se avsnitt 2 ovan).

5.2 Kollokationsmodul

Den andra modulen, kollokationsmodulen, kan som sagt användas antingen självständigt eller som ett komplement till den allmänspråkliga modulen. I kollokationsmodulen ligger tyngdpunkten på en mer uttömmande beskrivning av kollokationernas strukturella, typologiska och pragmatiska egenskaper. Denna

information kommer att göras åtkomlig för användaren dels i form av löpande artikeltext, dels genom ett utökat utbud av sökmöjligheter. I den löpande artikeltexten är den alfabetiska ordningen default. Det utökade utbudet av sökmöjligheter innebär främst tesauristiska sökningar utifrån betydelseindikatorer samt parametrar. Tack vare dessa sökmöjligheter skapas nya genvägar till innehållet i lexikonet, och mycket av den implicita informationen i den allmänspråkliga modulen blir explicit. Nedan ges en översikt över mikrostrukturen i artiklarna och sökmöjligheterna i kollokationsmodulen.

I artikelns mikrostruktur inkluderas följande delmoment: uppslagsord, ordklass, betydelseindikator (enbart för kollokationers basord) och definition. Informationen om uppslagsordets ordklass och betydelseindikator fyller främst en disambiguerande funktion och signalerar vilken betydelse som aktualiseras i artikeln. (Det kan tilläggas att en sådan hantering av ingångar förespråkas i många elektroniska ordböcker, exempelvis CALD 2003, eftersom hanterandet effektiviserar åtkomsten av information.) Efter den inledande delen följer en mer uttömmande lista över kollokationer. Vid listningen av kollokationerna följer vi samma format som används i den allmänspråkliga modulen, dvs. kollokationerna sorteras efter grammatisk struktur och relevanta parametrar. Det ska också finnas en länk till motsvarande ingång i den allmänspråkliga modulen. De ovan nämnda momenten i mikrostrukturen kan exemplifieras på följande sätt utifrån substantivet *depression*:

uppslagsord: depression

ordklass: substantiv

betydelseindikator: SJUKDOM

definition: sjukligt tillstånd med svår (långvarig) nedstämdhet

kollokationer:

ADJ + *depression*: *djup, svår, allvarlig; mild, lätt(are) | akut, långdragen, kronisk | årstidsbunden, årstidsrelaterad | bipolär | postnatal | egentlig*

V + *depression*: *få, drabbas av, gå in i en | ha, lida av | leva med | behandla, medicinera mot | ta sig ur en, komma ur en | bota*

länk: till allmänspråklig modul

I kollokationsmodulen är varje enskild kollokation försedd med en länk till mer specifik information om just den aktuella kollokationen. Om man klickar på länken, får man explicit information om bakomliggande parameter, kollokationens relativa frekvens och bruklighet (med hänsyn till ämnesområde och stilnivå). Dessutom ges autentiska exempel hämtade ur Medlex korpus samt ett kommentarfält med annan relevant information om kollokationen ifråga. Där förtydligas kollokationens innebörd och kombinatoriska egenskaper. Kommentarfältet kan också innehålla uppgifter av encyklopedisk eller pragmatisk karaktär som kan vara nödvändiga vid reception och produktion av ordförbindelsen. Om man t.ex.

väljer att aktivera länkarna till kollokationerna *svår depression* respektive *årstidsbunden depression*, får man ta del av följande uppgifter:

svår depression

parameter: Grad

frekvens: **** (= mycket frekvent)

bruklighet: allmänspråk

exempel: *Vid en medelsvår eller svår depression väljer läkaren ofta att kombinera samtalsterapi med en medicinsk behandling.*

kommentar: Indelning i *lätt*, *måttlig* och *svår depression* bygger på en skattningsskala utifrån förekommande symtom. Det är betydelsefullt att gradera depressionen eftersom patientens förutsättningar att bli fullt frisk ökar om behandlingen anpassas efter depressionens svårighetsgrad.

årstidsbunden depression

parameter: Typ

frekvens: ** (= vanligt)

bruklighet: allmänspråk

exempel: *Ljusbehandling är ett behandlingsalternativ vid årstidsbunden depression*

kommentar: De flesta patienter med årstidsrelaterad depression har en s.k. vinterdepression, dvs. de blir dåliga på hösten och tillfrisknar när våren kommer. Sommardepression verkar inte vara så vanligt i Sverige.

Som redan antytts kompletteras den deskriptiva delen i den löpande artikeltexten med ett utökat utbud av sökmöjligheter. Följande tillval vore önskvärda:

- **Enkel tesauristisk sökning** med betydelseindikator eller parameter som sökkriterium. I lexikonet kan man t.ex. söka på alla ingångar som har betydelseindikatorn SJUKDOM och få fram en lista med *cancer, depression, diabetes* etc. Söker man på parametern Grad, får man träffarna *svår, lätt, framskriden* etc.
- **Kombinerad tesauristisk sökning** där en betydelseindikator och/eller parameter kombineras med ett konkret ord. Man kan t.ex. vilja veta vilka sjukdomar som tar *akut* som attribut. En lista med t.ex. följande kollokationer skulle presenteras: *akut depression, akut pankreatit* och *akut lymfatisk leukemi*. Vidare kan man vilja veta hur en specifik parameter, exempelvis Grad, kan uttryckas när det gäller ett visst ord, t.ex. *depression*. Sökningen skulle ge följande resultat: *djup depression, svår depression, lätt(are) depression* etc.
- **Mönstersökning** efter typer av semantiska mönster som finns i lexikonet. En sådan sökning lämpar sig om man t.ex. vill veta vilka parametrar som är aktuella vid SJUKDOM. Här följer några exempel på sådana mönster: Förlopp + SJUKDOM, Grad + SJUKDOM, Insjuknande + SJUKDOM och Tillfrisknande + SJUKDOM.

- **Mönsterbaserad sökning** syftar till att finna alla belägg i hela lexikonet som har ett visst mönster. (För att hitta befintliga mönster används den mönstersökning som beskrivs strax ovan.) Om man väljer mönstret Insjuknande + SJUKDOM, får man resultatet *få en depression, drabbas av en depression, gå in i en depression, få cancer, drabbas av cancer, utveckla cancer* etc.

Med hjälp av dessa tillvalsökningar fås såväl lexikal som metaspråklig information om kollokationer, deras strukturer och typer. Denna information kan förhoppningsvis stödja lexikonanvändaren i jakten på en lämplig kollokator till en viss bas. Den kan också tjäna som ett verktyg för lexikografer som vill kontrollera systematiken och konsekvensen i material som de arbetar med.

6. Sammanfattning och slutsatser

I den här artikeln presenterar vi en pilotstudie om hantering av kollokationer i ett medicinskt elektroniskt lexikon, MedLex. Lexikonet är under uppbyggnad vid Institutionen för svenska språket, Göteborgs universitet. Det är deskriptivt och avsett att stödja såväl reception som produktion. Fokus i studien ligger på kollokationer med substantiv, framför allt sådana som beskriver sjukdomar och sjukdomstillstånd. Den analysmodell som presenteras ovan bygger på en kombination av semasiologiska och onomasiologiska utgångspunkter. Modellen resulterar i en indelning i dels lexikalt betingade, dels typrelaterade kollokationer. I artikeln diskuteras också hur de medicinskt relaterade kollokationerna kan beskrivas i det konkreta elektroniska lexikonet.

Vår pilotstudie har bekräftat vikten av att anlägga ett helhetsperspektiv, dvs. att inkludera såväl en semasiologisk som onomasiologisk dimension vid analys- och struktureringsarbetet. Vi hoppas att våra iakttagelser vad gäller kollokationer kring sjukdomar och sjukdomstillstånd också är relevanta vid beskrivning av kollokationer i andra medicinska sub-domäner. Studien visar även att det föreligger långtgående korrelationer mellan metodologin och det elektroniska formatet som skapar nya och mycket raffinerade sök- och åtkomstmöjligheter. Med andra ord utgör en nyanserad lexikologisk och encyklopedisk beskrivning av dessa enheter en nödvändig förutsättning för att skapa ett nytt elektroniskt lexikon över medicinska kollokationer. I dessa framtida elektroniska lexikon bör det också finnas möjligheter att på ett smidigt sätt komplettera det lexikala perspektivet med ett encyklopediskt genom att förse användaren med ett antal relevanta länkar till andra elektroniska källor.

Referenser

Ordböcker

- CALD = *Cambridge Advanced Learner's Dictionary* 2003. Cambridge: Cambridge University Press.
- Medicinsk terminologi* 2004. Utarbetad av Bengt Lindskog. Stockholm: Norstedts Akademiska Förlag.
- NEO = *Nationalencyklopedins ordbok* 1995–96. Utarbetad vid Språkdata, Göteborgs universitet. Höganäs: Bra Böcker.
- OCDSE = *Oxford Collocations dictionary for students of English* 2002. Oxford: Oxford University Press.

Övrig litteratur

- Bartsch, Sabine 2004: *Structural and Functional Properties of Collocations in English. A corpus study of lexical and pragmatic constraints on lexical co-occurrence*. Tübingen: Gunter Narr Verlag.
- Bergenholtz, Henning & Sven Tarp 1994: *Manual i fagleksikografi. Udarbejdelse af fagordbøger – Problemer og løsningsforslag*. Herning: systime.
- Hausmann, Franz Josef 1985: Kollokationen im deutschen Wörterbuch. Ein Beitrag zur Theorie des lexikographischen Beispiels. I: Bergenholtz, H. & J. Mugdan (eds.), *Lexikographie und Grammatik. Akten des Essener Kolloquiums zur Grammatik im Wörterbuch 28. –30.6. 1984*. Tübingen: Niemeyer, 226–257.
- Hausmann, Franz Josef 2004: Was sind eigentlich Kollokationen? I: Steyer, K. (Hrsg.), *Wortverbindungen – mehr oder weniger fest*. (Institut für Deutsche Sprache, Jahrbuch 2003.) Berlin/New York: Walter de Gruyter, 309–334.
- Heid, Ulrich 1994: On Ways Words Work Together. Topics in Lexical Combinatorics. I: Martin, W., W. Meijs, M. Moerland, E. ten Pas, P. van Sterkenburg, & P. Vossen (eds.), *Euralex 1994 Proceedings. Papers Submitted to the 6th EURALEX International Congress on Lexicography in Amsterdam, The Netherlands*. Amsterdam, 226–257.
- Heid, Ulrich 2004: On the presentation of collocations in monolingual dictionaries. I: Williams, G. & S. Vessier (eds.), *Proceedings of the Eleventh EURALEX International Congress, Lorient, France July 6–10, 2004*. Lorient: UBS, 729–738.
- Heid, Ulrich & Rufus H. Gouws 2006: A Model for a Multifunctional Dictionary of Collocations. I: Corino, E., Marelllo C. & C. Onesti (eds.), *Proceedings of the XII Euralex International Congress, Torino, Italia, September 6–9, 2006*. Alessandria, 979–988.
- Kjellmer, Göran 1994: Introduction. I: *A Dictionary of English Collocations. Based on the Brown Corpus*. Utarbetad av Göran Kjellmer. Oxford: Clarendon Press, ix–xliii.
- Koike, Kazumi 2001: *Colocaciones léxicas en el español actual: estudio formal y léxico-semántico*. Alcalá de Henares: Universidad de Alcalá.
- Malmgren, Sven-Göran 2003: Begå eller ta självmord? Om svenska kollokationer och deras förändringsbenägenhet. I: *Språk & Stil. Tidskrift för svensk språkforskning* 13 Ny följd 2003, 123–168.
- Manning, Christopher & Hinrich Schütze 1999: *Foundations of Statistical Natural Language Processing*. MIT Press.

- Martin, Willy 2006: Frame-based Lexicons and Making of Dictionaries. I: Corino, E., Marengo C. & C. Onesti (eds.), *Proceedings of the XII Euralex International Congress, Torino, Italia, September 6–9, 2006*. Alessandria, 281–293.
- Mel'čuk, Igor 1998: Collocations and Lexical Functions. I: Cowie, A.P., (ed.), *Phraseology: Theory, Analysis, and Applications*. Oxford: Clarendon Press, 23–53.
- Siepmann, Dirk 2005: Collocation, Colligation and Encoding Dictionaries. Part I. Lexicological Aspects. I: *International Journal of Lexicography* 18:4, 2005, 409–443.
- Sinclair, John 1991: *Corpus, Concordance, Collocations*. Oxford: Oxford University Press.
- Svensén, Bo 2004: *Handbok i lexikografi. Ordböcker och ordboksarbete i teori och praktik*. (2 uppl.). Stockholm: Norstedts Akademiska Förlag.

Internet

- Google <<http://www.google.se>>, hämtat 2007-08-28.
- Språkbanken <<http://spraakbanken.gu.se/>>, 2007-08-28.

Emma Sköldberg
FD, f. 1968
Institutionen för svenska språket
Göteborgs universitet
Box 200
SE-405 30 Göteborg
emma.skoeldberg@svenska.gu.se

Maria Toporowska Gronostaj
FD, f. 1945
Institutionen för svenska språket
Göteborgs universitet
Box 200
SE-405 30 Göteborg
maria.gronostaj@svenska.gu.se