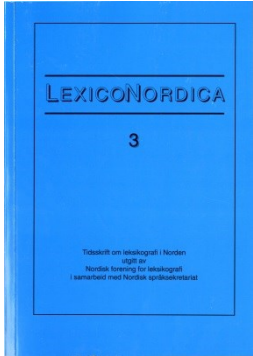


# LexicoNordica

Forfatter:	Sölve Ohlander [”The Big Four”. “Learners’ dictionaries” inför 2000-talet]	
Anmeldt værk:	<i>Cambridge International Dictionary of English</i> . Cambridge: Cambridge University, 1995. <i>Collins COBUILD English Dictionary</i> . HarperCollins, 1995. <i>Longman Dictionary of Contemporary English</i> . Longman, 1995. <i>Oxford Advanced Learner’s Dictionary of Current English</i> . Oxford: Oxford University, 1995.	
Kilde:	LexicoNordica 3, 1996, s. 257-279	
URL:	<a href="http://ojs.statsbiblioteket.dk/index.php/lexn/issue/archive">http://ojs.statsbiblioteket.dk/index.php/lexn/issue/archive</a>	

© 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.

*Sölve Ohlander*

## **"The Big Four". "Learners' dictionaries" inför 2000-talet**

*Cambridge International Dictionary of English* 1995. Editor-in-chief: Paul Procter. Cambridge University Press.

*Collins COBUILD English Dictionary*, Second ed. 1995. Editor in Chief: John Sinclair. HarperCollins.

*Longman Dictionary of Contemporary English*, Third ed. 1995. Director: Della Summers. Longman.

*Oxford Advanced Learner's Dictionary of Current English*, Fifth ed. 1995. Editor: Jonathan Crowther. Oxford University Press.

This review article discusses certain general tendencies and issues in connection with the "The Big Four", i.e. the 1995 editions of the four English learners' dictionaries now commonly regarded as dominant. One readily noticeable trend is that, on a fiercely competitive market, learners' dictionaries are becoming increasingly bulky, covering more and more words and providing ever new kinds of information. The 1995 dictionaries are all based on enormous computerized corpora of both written and spoken language. While this does not eliminate gaps in coverage, it does mean, among other things, that informal spoken English can be paid special attention to. On the basis of a learners' corpus, one of the dictionaries also provides contrastive information concerning English "false friends" in relation to a number of different languages, an interesting attempt, albeit poorly executed, to bridge the gap between monolingual and bilingual lexicography. Another novelty is the introduction of corpus-based frequency information in two of the dictionaries. With regard to definition, all four dictionaries now employ restricted defining vocabularies, whereas there is no agreement on definition format or style. Semantically, there is a clear tendency towards eliminating, in the presentation of word meanings, the distinction between homonymy and polysemy, a lexicographically convenient strategy which, however, may be pedagogically questionable. The article ends with some reflections on the possible limits, not least from a pedagogical point of view, of traditional, printed learners' dictionaries, especially in comparison with the potential of future electronic dictionaries.

### **Inledning**

1995 var ett stort år för engelska "learners' dictionaries". Som för att lägga sig i läge inför nästa årtusende utkom detta år de tre största i

branschen med nya, i flera avseenden kraftigt omarbetade upplagor: *Collins COBUILD English Dictionary (COBUILD)*, *Longman Dictionary of Contemporary English (LDOCE)*, *Oxford Advanced Learner's Dictionary of Current English (OALD)*. Samma år anmälde sig en fjärde intressent till medlemskap i denna grupp av väletablerade och arga konkurrenter på en lukrativ marknad – en uppkomling från Cambridge University Press: *Cambridge International Dictionary of English (CIDE)*. Den tidigare treklövern har förbytts i "The Big Four".

I den allt hårdare striden om marknadsandelar inom det enorma fält som engelska som andraspråk och främmande språk utgör ligger det i varje ny lexikonupplagas intresse att hävda sitt eget nyhetsvärde så otvetydigt som möjligt; blygsamhet är här en tvivelaktig dygd. Pionjären bland "learners' dictionaries", *OALD* (första upplagan 1948), omnämner sin senaste version som "This exciting new edition" (pärmens baksida); motsvarande karakteristik av den närmast föregående, fjärde upplagan från 1989 var "This comprehensively revised and greatly expanded new edition..." (pärmens baksida).<sup>1</sup> Därefter i senioritet följer *LDOCE* (första upplagan 1978), vars "third, completely new edition" (s. xi) "...reveals for the first time the richness of the English language" (pärmens baksida); andra upplagan av lexikonet (1987) varudeklarerades på sin tid som "A totally revised version..." (pärmens baksida). Nya *COBUILD* presenteras på följande sätt av John Sinclair: "This is a new book, a completely new edition..." (s. viii); om den första upplagan av *COBUILD* (1987) lät det så här (omslaget baksida): "The dictionary sets new standards in the presentation and layout of detailed information about the language."<sup>2</sup> Jämfört med självsäkerheten hos de redan etablerade lexikonen ter sig nykomlingen i gänget, *CIDE*, närmast blygsam: "The fresh approach that we have taken should appeal to all those who appreciate good lexicography..." (s. viii). Till grund för denna relativt återhållsamma formulering ligger måhända den omständigheten att *CIDE* är Cambridge University Press' första inbrytning bland "mainstream monolingual dictionaries" (s. viii).

Som kan utläsas av de ovannämnda citaten har det i accelererande grad gått inflation i uttryck för självuppskattning, med stark betoning

---

<sup>1</sup> Som en lätt pikant detalj kan man lägga märke till att den senaste upplagan av *OALD* presenterar sig som "The dictionary that really teaches English" (pärmens baksida), säkerligen en medveten travesti på *COBUILD*s välkända slogan: "Helping learners with real English."

<sup>2</sup> *COBUILD* har också förkortat sitt namn sedan första upplagan 1987 (*Collins COBUILD English Language Dictionary*), kanke för att markera att det verkligen rör sig om en helt ny upplaga. Detta framgår också av att samtliga 100 000 exemplar i lexikonet "are new for this edition" (pärmens baksida).

på varje ny lexikonupplagas omistlighet för användaren. Men trots allt finns det i 1995 års utgivning ett visst fog för den av förlagen eftertraktade uppmärksamheten. Skälet är att innovationerna i de nya lexikonerna till betydande del är av kvalitativ snarare än blott och bart kvantitativ natur. Det handlar således inte enbart om att lexikonerna innehåller fler och nyare uppslagsord än sina föregångare, utan lika mycket om vilket slags information som ges, hur den är strukturerad etc. Denna artikel avser inte att i detalj granska varje enskilt lexikon; vart och ett av dem har sina egna specialiteter. I stället kommer jag att uppehålla mig vid ett antal någorlunda klart urskiljbara tendenser av mer övergripande natur, särskilt sådana som ger sig till känna i fler än ett av lexikonerna. Detta innebär i huvudsak jämförelser på vissa punkter såväl mellan de olika lexikonerna från 1995 som mellan ny och tidigare upplaga av samma lexikon.<sup>3</sup> Även för enstaka lexikon karakteristiska "nischer" kan i en del fall vara värda att uppmärksamma; dagens nischer kan vara morgondagens – eller i alla fall nästa århundrades – trender.

### **Tätare utgivning, större omfång**

En skönjbar tendens, som kan kopplas till den allt hårdare konkurrensen men också till den snabba utvecklingen inom lexikografin, är att det i framtiden förmodligen kommer att bli ännu tätare mellan upplagorna av samma lexikon. Mellan tredje och fjärde upplagan av *OALD* dröjde det femton år (1974–1989), mellan fjärde och femte sex år. Första upplagan av *LDOCE* kom 1978, andra 1987, tredje 1995, det sista intervallet samma som mellan först och andra upplagan av *COBUILD*. Lexikonet som färskvara i en snabbt föränderlig språkvärld kommer säkert att betonas allt mer i marknadsföringen, gärna med argument hämtade från språkpedagogisk forskning kring inläring och undervisning av andraspråk och främmande språk.

Hand i hand med tendensen mot allt tätare utgivning av nya upplagor går en tydlig tendens mot alltmer omfattande inlärningsordböcker. En jämförelse mellan sidantal i olika upplagor ger en grov indikation:<sup>4</sup>

<sup>3</sup> Vid jämförelser med närmast föregående upplagor av lexikonerna kommer följande beteckningar att användas: *OALD* (89) = fjärde uppl. av *OALD*, 1989; *COBUILD* (87) = första uppl. av *COBUILD*, 1987; *LDOCE* (87) = andra uppl. av *LDOCE*, 1987. I vissa fall kan även beteckningar som *OALD* (95) etc. förekomma.

<sup>4</sup> Sifferuppgifterna avser endast den egentliga, alfabetiskt uppställda lexikondelen i böckerna, alltså inte olika "introductions", tabeller, listor eller andra slags appendix, där det finns stora skillnader mellan lexikonerna. Som jämförelse kan här

*LDOCE* (87) 1229 sidor – *LDOCE* (95) 1668 sidor, *COBUILD* (87) 1703 sidor – *COBUILD* (95) 1951 sidor. Det nya *CIDE* omfattar 1701 mycket tättryckta sidor. *OALD* utgör dock ett intressant undantag till den allmänna trenden: *OALD* (89) 1492 sidor – *OALD* (95) 1392 sidor (dock med något mindre stil). Även i fråga om sidstorlek avviker *OALD* från de övriga lexikonerna genom sitt mindre format. Huruvida *OALD* på detta sätt medvetet profilerar sig gentemot de andra, mer voluminösa lexikonerna är svårt att veta. Men frågan var gränsen går beträffande informationsmängd – praktiskt/pedagogiskt och även kommersiellt – rycker närmare för varje ny upplaga. Det rör sig inte enbart om antalet uppslagsord, definitioner och exempel, utan också om all den övriga information kring ords användning som inlärningslexikon under senare år visat allt större intresse för. Kanske håller vi på att närma oss den punkt där, för ordboksanvändaren, problemet inte längre är brist utan överflöd på information. Hela denna fråga hänger samman med en annan fråga, nämligen hur "learners' dictionaries" i praktiken används, ett område där vår kunskap är begränsad.<sup>5</sup>

### Fler och fler uppslagsord

En bidragande orsak, om än inte den enda eller kanske ens den främsta, till den förhärskande färdriktningen mot allt större lexikon är ambitionen att i varje ny upplaga spegla den snabba utvecklingen inom engelskans ordförråd. Så länge tillägg av nya uppslagsord och nya betydelser inte i motsvarande grad kompenseras genom eliminering av lika många föråldrade ord och betydelser kommer antalet uppslagsord och betydelser i "learners' dictionaries" ständigt att öka. Några sifferuppgifter kan illustrera detta.<sup>6</sup>

Trots relativt oförändrat sidantal omfattar den nya upplagan av *OALD* "63 000 references", att jämföra med "57 100 words and phrases" i *OALD* (89); den nya upplagan uppges innehålla "2 300 new

---

också nämnas att åttonde upplagan av *Concise Oxford Dictionary* (1990) omfattar 1428 sidor.

<sup>5</sup> På frågan om hur mycket vi egentligen vet om hur "learners' dictionaries" används ger Kernerman (1996:406) ett koncist svar: "Very little."

<sup>6</sup> Exakta jämförelser försvåras av bristen på en uniform terminologi vid lexikonens egna uppgifter om antal uppslagsord. Där ett lexikon talar om antalet "words and phrases" talar ett annat – eller en ny upplaga av samma lexikon – i stället om "references"; *Concise Oxford Dictionary* (1990) uppger "120,000 entries", varav "Over 20,000 entries new to this edition". Uppgifter om eliminerade uppslagsord eller betydelser saknas genomgående i nya upplagor.

words and meanings". *LDOCE* (87) redovisar "56,000 words and phrases", den nya upplagan "Over 80,000 words and phrases"; uppgift om antal nya ord eller betydelser saknas. *COBUILD* (87) anger "over 70,000 references", den nya upplagan "over 75,000 references", varav "More than 4,000 new words and meanings". *CIDE*, slutligen, uppger ett omfång av inte mindre än "100,000 words and phrases".<sup>7</sup>

Tendensen mot allt större täckning av ordförrådet, med ett kontinuerligt tillflöde av nya ord och betydelser, är således uppenbar. Det kan i detta sammanhang också vara av intresse att utröna huruvida det finns skillnader i täckning mellan lexikonen beträffande ett antal nyare, numera vanliga ord i engelskan (flera av dem känner vi också igen i dagens svenska). Först några ord som saknas i de tidigare upplagorna men som finns med i samtliga nya upplagor, inklusive *CIDE*: *bimbo*, *e-mail* (med stavningsvarianterna *email*, *E-mail*, *Email*), *multicultural*, *multimedia*, *wannabe*. Så några ord som tas upp av de flesta men inte alla de nya lexikonen. *Internet* och *PC* i betydelsen 'politically correct' är inkluderade i samtliga lexikon utom *OALD*; detsamma gäller substantivet *grunge* (välkänt bl.a. inom rockmusiken). Den i dag högfrekventa förkortningen *IT* redovisas i alla lexikonen utom i *COBUILD*, inte ens under uppslagsordet *information technology*.<sup>8</sup> Å andra sidan är *COBUILD* det enda lexikonet som har med det inte minst i MTV-sammanhang frekventa *unplugged*, medan *hypertext* förekommer i samtliga lexikon utom *LDOCE*. Ett mer vardagligt IT-ord som *home page* har däremot inte lyckats nästla sig in i något av lexikonen. Låt oss slutligen titta på det i informellt språk mycket vanliga substantivet *nerd*, sedan några år allt vanligare även i svenskan (*nörd*).<sup>9</sup> Ordet finns upptaget i samtliga nya lexikon utom *OALD* – något förvånande med tanke på att *nerd* är inkluderat i *Oxford Advanced Learner's Encyclopedic Dictionary* från 1992. Ordet saknas i såväl *OALD* (89) som *COBUILD* (87), till skillnad från *LDOCE* (87).

Som framgår av ovanstående exemplifiering företer de fyra lexikonen en ganska splittrad bild av en handfull välkända nyord i engelskan; ingen av ordböckerna redovisar samtliga ord. Detta ger ett kollektivt intryck av osystematisk täckning – olika för varje lexikon – med som det verkar tämligen slumpmässiga bortfall av ord som tveklöst hör hemma i ett inlärningslexikon för det sena nittio-talet. Några exotiska

<sup>7</sup> Jfr också *CIDE*, s. viii: "*CIDE* provides more entries and more examples than has been possible in other learners' dictionaries."

<sup>8</sup> Det är i sammanhanget värt att notera att *IT* finns med som uppslagsord redan i *LDOCE* (87), liksom även i *Concise Oxford Dictionary* (1990).

<sup>9</sup> Ordet saknas dock i *Nationalencyklopedins ordbok* (1996).

specialuttryck från begränsade fackområden är det ju inte fråga om. Det handlar om ord som man stöter på dagligen i vanlig tidningstext.

### Korpusbaserings

De ovan påtalade luckorna i ordtäckningen är i sig knappast något att uppröras över; oförklarliga luckor har alltid funnits i ordböcker. Ändå kan man tycka att t.ex. frånvaron av *nerd* i *OALD* eller *IT* i *COBUILD* är förvånande. Anledningen till denna förvåning är att sådana luckor kunnat uppstå trots att de aktuella lexikonerna samtliga är baserade på enorma material av samtida autentisk engelsk text, datoriserade mångmiljonkorpusar där representativitet beträffande olika texttyper i princip är garanterad. Korpusbaserings ligger också, direkt eller indirekt, till grund för många av de trender som är märkbara bland "learners' dictionaries".

Den första engelska inlärningsordbok som konsekvent baserades på en stor datoriserad korpus var *COBUILD* (87).<sup>10</sup> Sedan dess har tendensen varit klar. Inte bara *COBUILD* (95), utan även *OALD* (95), *LDOCE* (95) och *CIDE* är korpusbaserade. Detta framhålls också som ett försäljningsargument, ett slags kvalitetsgaranti för att urvalet av ord som behandlas i lexikonet är representativt och har objektiv täckning i det dagens språk, att upplysningar om betydelser, konstruktionssätt etc. är pålitliga, dvs. väl dokumenterade i autentisk text. Utvecklingen inom korpusområdet har som bekant varit mycket snabb och tendensen har varit entydig de senaste tio åren, mot större och större korpusar. Vad som för något decennium sedan var science fiction är i dag lexikografisk verklighet. Några sifferuppgifter, hämtade från de fyra lexikonerna (förord, inledningar etc.), kan belysa detta förhållande.

Till grund för *COBUILD* (87) låg en korpus på 20 miljoner ord löpande text, medan *COBUILD* (95) "has been written using the evidence of over *two hundred* million words – ten times the corpus made for the original dictionary" (s. viii). *COBUILD*s databas, "The Bank of English", motsvaras för *LDOCE* (95) av en korpus på drygt 130 miljoner ord. Tyndpunkten ligger i "British National Corpus" (100 miljoner ord), som även utgör basen för *OALD*s (95) korpus om c. 140 miljoner ord. *CIDE*, slutligen, bygger på "Cambridge Language Survey", ett material omfattande 100 miljoner ord. Samtliga korpusar innehåller såväl brittisk som amerikansk engelska.

---

<sup>10</sup> Se *COBUILD* (87), s. xv, och *COBUILD* (95), s. viii. Jfr även *LDOCE* (87), s. F8. Angående relationen mellan lexikon och korpusar, se Calzolari (1996).

Utvecklingen mot användning av gigantiska datoriserade korpusar inom lexikografin är således uppenbar och innebär ett revolutionerande framsteg jämfört med tidigare, mindre systematiska metoder att avtappa autentiskt språkbruk för kodifiering i ordböcker. Trots – eller kanske snarare på grund av – detta inställer sig frågan hur det då kan komma sig att samtliga fyra lexikon, om än i varierande grad, i sin korpusbaserade täckning av dagens ordförråd ändå uppvisar luckor av det slag som exemplifierats ovan.

I förordet till *LDOCE* (95) betonar Randolph Quirk att "readers have to be assured that the words they need to understand and use are included, and that such inclusion reflects up-to-date occurrence in material from a wide range of English-speaking countries and from sources dealing with a wide range of subject matter." Han tillägger: "The advent of computerised corpora enables us to achieve a greatly enhanced coverage." Som redan påpekats är sådana uppfattningar representativa för filosofin bakom samtliga här diskuterade lexikon. En stor datoriserad korpus skall bl.a. borga för att inte några i allmänspråket vanliga och användbara ord skall kunna "slip through the net". Som vi sett tycks det emellertid inte alltid förhålla sig på detta vis. Precis som i tidigare, icke korpusbaserade lexikon finns det svårförklarliga luckor. Vad de beror på i de enskilda fallen är inte lätt att ge något generellt svar på. Att t.ex. hänvisa till en ofrånkomlig eftersläpning vid intäckning av nya ord låter sig knappast göras, eftersom olika lexikon uppvisar olika luckor. En allmän slutsats kan dock dras: jättelika datoriserade korpusar av dagens modell må vara ett aldrig så kraftfullt arbetsredskap för moderna lexikografer, men lika lite som tidigare metoder vid ordboksproduktion tycks de kunna garantera heltäckning inom det för lexikonets syfte relevanta ordförrådet.

I sin "Introduction" (s. 9) till *COBUILD* (95) skriver John Sinclair: "Nowadays the problem is not finding the language..." Det ter sig då än mer gåtfullt hur ett ord som *IT* kan saknas i lexikonet. Att se en lösning i ännu större korpusar verkar inte rimligt. De som nu används förefaller mer än väl ha uppnått kritisk massa; de exemplifierade luckorna måste bero på något annat än otillräcklig korpusstorlek. Att uteslutande förlita sig på en korpus räcker tydligen inte. Det verkar i stället som om lexikografens intuition, baserad på egna språkliga erfarenheter och en upptränad observationsförmåga, måste balansera och i viss mån korrigera en aldrig så representativ korpus även vid själva urvalet av de ord som skall tas upp i lexikonet. Kanske är det i praktiken så att de stora korpusarna haft mer revolutionerande betydelse för andra saker i "learners' dictionaries" (grammatiska mönster, kollokationer, stilnivåer etc.) än för just urvalet av uppslagsord.



## Talspråk

Resonemanget i föregående avsnitt är inte avsett som ett argument mot korpusbaserade lexikon, utan som en påminnelse att en aldrig så stor korpus inte automatiskt undanröjer alla problem. Dagens korpusar ställer i själva verket större krav än någonsin på lexikograferna då det gäller att välja och vraka bland all den information som en stor datoriserad korpus ställer till förfogande. Vilka kriterier skall styra urvalet? Vilka typer av information och vilket slags språk skall redovisas i t.ex. en inlärningsordbok?

En klar tendens i dagens "learners' dictionaries" är att det informella talspråket fått större utrymme än tidigare. Delvis beror detta säkert på att den kommunikativt inriktade språkundervisning som anammats i allt fler länder sedan åttiotalet särskilt betonat vikten av att verkligen kunna tala det främmande språket. Det finns således en språkpedagogiskt motiverad lexikonmarknad för engelskt talspråk, även accentuerad av det stora utbud av informell talad engelska som finns tillgängligt genom t.ex. satellit- och kabel-TV. Men det finns också ett annat skäl till att talspråket hamnat i fokus på ett annat sätt än i tidigare inlärningslexikon, nämligen den ökade tillgången på stora talspråkskorpusar.

Samtliga de fyra ordböcker som här diskuteras anger att deras korpusar är uppbyggda av såväl skrivet språk som informellt talspråk. *COBUILD* anger storleken på sin talspråkskorpus, som är en del av "The Bank of English", till "about 15 million words", hämtade från "recordings of everyday casual conversation, meetings, interviews, and discussions" (s. xii). De tre övriga lexikonen ger ingen uppgift, men för *LDOCE* och *OALD* kan man räkna med en talspråkskorpus om åtminstone 10 miljoner ord, som ingår i "British National Corpus".<sup>11</sup> Enligt *LDOCE* (s. xi) handlar det om "natural speech, not radio or TV programmes or language that is in any way scripted."

Bland de fyra lexikonen är det framför allt *COBUILD* och *LDOCE* som särskilt markerar typiskt talspråkliga ord och uttryck, som en viktig stilistisk/sociolingvistisk upplysning.<sup>12</sup> I båda dessa ordböcker – till

<sup>11</sup> Uppgiften är hämtad från *Longman Language Activator* (1993: F8), som i likhet med *LDOCE* och *OALD* är baserad på bl.a. "British National Corpus".

<sup>12</sup> Jfr *LDOCE* (pärmens baksida): "The only dictionary to recognize the importance of spoken English..." – en karakteristik som *COBUILD* säkerligen inte skulle hålla med om. I sitt förord till *LDOCE* skriver Randolph Quirk om

skillnad från de två övriga – anges explicit att t.ex. adverbet *yeah* och frasen *you know* tillhör talspråket. De två övriga lexikonen nöjer sig med att mer traditionellt beteckna dessa uttryck som informella, en stilbeteckning som också används i såväl *LDOCE* som *COBUILD*.

Markören "spoken" i *LDOCE* och *COBUILD* kan i sig hälsas som ett välkommet tillskott till den gängse repertoaren av stil- och registerinformation i inlärningsordböcker, till stor del beroende på tillgången till stora talspråskorpusar.<sup>13</sup> Dock förefaller användningen av denna markör brista en hel del i konsekvens, inte bara – föga förvånande – mellan de båda lexikonen utan också inom ett och samma lexikon. I synnerhet tycks avgränsningen i förhållande till stilmarkören "informal" ha vållat problem. Ett uttryck som *no way!* räknas i *LDOCE* bland "spoken phrases"; i *COBUILD* betecknas det som "an informal expression". Detsamma gäller t.ex. *no such luck!*. Interjektionen *whoops* anges som "spoken" i *COBUILD*, men har ingen särskild stilmarkör i *LDOCE*. Den lika talspråkliga interjektionen *ouch* får däremot inte någon stilmarkör alls i *COBUILD*. Till skillnad från det jakande *yeah* (se ovan) etiketteras det nekande adverbet *nope* som "informal" i *LDOCE*, som "informal spoken" i *COBUILD*. Exempelen kunde mångfaldigas. I och för sig har kanske inte denna brist på ordning och reda så stor praktisk betydelse – det finns ju ett samband mellan t.ex. informellt språk och talspråk, även om de inte är synonyma begrepp. Men värdet av stil- och registerinformation, särskilt i en inlärningsordbok, minskar naturligtvis om läsaren inte riktigt kan lita på att aktuella markörer används konsekvent. Här finns det således, i båda lexikonen, utrymme för förbättringar.

### **Kontrastiv information: en ny trend?**

Det ligger i enspråkiga lexikons natur att de normalt inte ger kontrastivt differentierad information om ordförrådet, anpassad till vad man erfarenhetsmässigt vet är särskilda svårigheter för olika grupper av inlärare med olika förstaspråk. Exempelvis norska, finska och ryska inlärare har delvis olika problem beträffande engelskans (eller tyskans, franskans etc.) vokabulär, något som språklärare i regel känner väl till. Det är därför intressant att notera att två av de här aktuella lexikonen

---

talspråskorpusen att den "has been put to prominent use for the first time in this new edition [...]".

<sup>13</sup> *COBUILD*, till skillnad från *LDOCE*, använder också stilmarkören "written" vid ord som huvudsakligen förekommer i skriftspråket, t.ex. *anew*, där *LDOCE* i stället använder markören "literary".

uttryckligen säger sig ha utnyttjat var sin "learner corpus", dvs. en korpus bestående av "learner English". *LDOCE* anger (pärmens baksida) att man vid sidan av det övriga korpusmaterialet använt sig av "the 5 million word Longman Learner's Corpus". Något ytterligare utöver denna knapphändiga information får vi tyvärr inte reda på, t.ex. i vad mån denna korpus har påverkat lexikonets uppbyggnad.<sup>14</sup> *CIDE* å sin sida anger varken storlek eller några andra konkreta uppgifter om sin korpus. Men man talar om (s. viii) att "learner errors are codified for retrieval and analysis, This allows the learning difficulties of specific language groups to be targeted."

Ett konkret resultat av *CIDEs* intresse för språkspecifika, transferrelaterade problem vid ordinlärning är att lexikonet presenterar ett urval engelska "false friends" – dessa klassiska symptom på lexikal interferens – i förhållande till 16 olika språk.<sup>15</sup> Förutom japanska, koreanska och thai är den stora majoriteten av de kontrasterade språken europeiska. Danska, norska och svenska finns med, däremot inte isländska eller finska. Redovisningen av "false friends" sker på två sätt. Dels ges listor över engelska "false friends" i förhållande till vart och ett av de 16 språken, dels ges vid enskilda uppslagsord, som t.ex. *actual*, information om att ordet är en "falsk vän" i förhållande till bl.a. danska och tyska.

Urvalet "falska vänner" i begränsade listor kan alltid diskuteras. Att kräva fullständighet är omöjligt, bl.a. därför att förrådet inte är konstant. Nya medlemmar tillkommer ständigt, däribland en och annan falsk vän.<sup>16</sup> Vad man däremot kan kräva är att urvalet skall bygga på rimliga principer, så att de ord som tas upp verkligen har karaktären av "falska vänner". I detta avseende uppvisar *CIDE* klara brister. I lexikonet förklaras begreppet "false friend" (s. 502) på följande, tämligen okontroversiella sätt: "a word in a foreign language which looks or sounds similar to a word in your own language but does not have exactly the same meaning." En viktig del i denna definition är naturligtvis den semantiska aspekten ("not...exactly the same meaning"). "Falsk vänskap" kan sägas uppstå genom en konspiration mellan form

<sup>14</sup> Jfr *Longman Language Activator* (1993:F8): "The Longman Learners' Corpus of students' writing has given us insights into the problems shared by students from over 70 different countries [...]". Dock sägs ingenting om eventuella *skillnader* mellan grupper av inlärare med olika förstaspråk; jfr Ohlander 1995). För en mer ingående diskussion, se Granger (1996).

<sup>15</sup> Detta är också en av anledningarna till att lexikonet kallar sig "International"; jfr pärmens baksida, som dock felaktigt uppger antalet olika språk till 14 i stället för 16 (se s. 502).

<sup>16</sup> Jfr S. Ohlander (1991:258f.).

och betydelse i relationen mellan två språk. Enbart ljudlighet eller likartad stavning räcker i regel inte för att kvalificera ett ord, psykologiskt, som "äkta falsk vän". Vanligtvis finns där också en semantisk "frestelse", dvs. något slags betydelselikheter som gör bedrägeriet kontextuellt rimligt; ett minimikrav är samma ordklassstillhörighet.<sup>17</sup> Därför är alla överens om att engelska ord som *actual*, *announce* och *eventual* är "false friends" för inlärare med ett flertal olika förstaspråk, liksom *become* i förhållande till tyska *bekommen*. Dessa ord finns också, tillsammans med en hel del andra klassiker, inkluderade i *CIDEs* listor. Så långt är det lätt att samtycka, även om det också på denna fasta mark finns en del gåtfullheter och inkonsekvenser i form av omotiverade luckor i listorna för vissa språk jämfört med andra. Exempelvis saknas i den svenska listan just det välkända *actual*, medan det redovisas i t.ex. den danska och tyska. Ord som *flesh* och *stool* uppfyller också högt ställda krav på "falsk vänskap" i förhållande till svenskan men finns inte med.

Mer principiellt diskutabelt är att en mängd ord i de olika listorna enbart verkar ha tagits med på grund av ytlig likhet i uttal och/eller skrift, utan någon reell semantisk/kontextuell "frestelse". Några exempel från den svenska listan: *bucket* (jfr sv. *bukett*), *burst* (sv. *borsta*), *gang* (sv. *gång*), *glance* (sv. *glänsa*), *tax* (sv. *tax*; eng. *dachshund*), *under* (sv. *under*; eng. *wonder*). I så fall, varför inte lika gärna t.ex. *barn* (sv. *barn*), *glass* (sv. *glass*), *list* (sv. *list*; eng. *cunning*) *swim* (sv. *svimma*), *utter* (sv. *utter*; eng. *otter*) och en mängd andra liknande fall? Andra listor än den svenska ger rikhaltiga exempel på samma sak, dvs. godtyckligt valda, slumpvisa ytlikheter av den sort som i synnerhet "intermediate and advanced learners" uppfattar som roande snarare än problematiska, sådant som språkstudierande säkert i alla tider skämtat om mer än lurats av.

Ett annat problem är att en hel del ord i listorna för olika språk inte finns markerade som "false friends" vid enskilda uppslagsord (fr ovan). Detta gäller bl.a. fall som *announce*, *eventual*, *fabric* och *novel*, vilka finns redovisade i den svenska listan över "false friends" men som inte har någon motsvarande markering vid de enskilda uppslagsorden. Denna brist på intern konsekvens mellan listor och uppslagsord är besvärande.

Huvudintrycket av *CIDEs* presentation av "false friends" är att den skulle må väl av en rejäl uppstädning, både beträffande urval och kon-

---

<sup>17</sup> "Falsk vän" är inte något väldefinierat begrepp, särskilt inte beträffande "balansen" mellan fonologiska och semantiska likheter. Se Veisbergs (1996); jfr även Hill (1982:iv) och U. Ohlander (1991:68f.).

sekvens. Trots detta ger lexikonet en hel del nyttiga upplysningar inom området. Tanken i sig – att ge språkspecifik information i en enspråkig ordbok – är också intressant. Den kan ses som ett försök till inbrytning på traditionellt tvåspråkiga ordboksdömen. *CIDEs* "false friends" kan också uppfattas som ett lexikografiskt uttryck för det nyväckta intresset för kontrastiv lingvistik under de senaste tio åren, liksom för den därmed sammanhängande, urgamla tanken – numera åter teoretiskt respektabel – att olika förstaspråk på olika sätt påverkar inläringen av ett nytt språk.<sup>18</sup> Möjligen kan vi här se början till en ny trend bland engelska inlärningsordböcker.

### Frekvensuppgifter

Att veta om ett ord är vanligt eller ovanligt i ett visst textsammanhang är en viktig del av kunskapen i ett främmande språk. Som regel – om man bortser från språkets allra mest högfrekventa ord – är detta något som kräver lång tid och en myckenhet exponering för det främmande språket, i både tal och skrift. Här finns ytterligare ett område där användningen av stora datoriserade korpusar har möjliggjort mycket mer precisa upplysningar än som tidigare varit möjligt. Särskilt tydligt är detta i de nya upplagorna av *COBUILD* och *LDOCE*. Visserligen anger även *COBUILD* (87:xix) och *LDOCE* (87:F20), liksom *CIDE* (s. ix) och den nya upplagan av *OALD* (s. ix), att frekvens varit en bärande princip för ordningsföljden bland olika betydelser inom samma uppslagsord. Men det är först i och med 1995 års upplagor som *COBUILD* och *LDOCE* gått ett viktigt steg vidare: i båda lexikonerna ges nu även frekvensuppgifter för enskilda uppslagsord, om än på skilda sätt.<sup>19</sup>

*COBUILDs* ordfrekvenser är indelade i fem frekvensband. De olika banden innehåller fler och fler ord alltefter fallande frekvens (s. xiii): det första bandet innehåller c. 700 av språkets vanligaste ord (t.ex. *of*, *go*), det andra c. 1200 ord (t.ex. *bridge*, *obvious*), det tredje c. 1500 ord (t.ex. *gate*, *medicine*), det fjärde c. 3200 ord (t.ex. *artistic*, *puzzle*), det femte c. 8100 ord (t.ex. *denim*, *inconceivable*). Uppslagsord utan frekvensmarkering utgör en sjätte grupp som närmast representerar ett centralt passivt ordförråd, "words which you will probably read or hear rather than words which you will often need to use yourself" (s. xiii).

<sup>18</sup> Se Aijmer *et al.* (1996) och Granger (1996:46); jfr också Swan & Smith (1987).

<sup>19</sup> Jfr *COBUILD* (s. ix): "For the first time in a major dictionary, COBUILD gives information about the frequency of the headwords."

Frekvensuppgifterna i *COBUILD* har tveklöst klar pedagogisk relevans för språkinläraren/lexikonanvändaren, liksom för språkläraren, även om man inte utan vidare kan använda dem för att ställa samman engelskans "centrala ordförråd" i någon absolut mening. Att de också är en intressant innovation i "learners' dictionaries" är uppenbart.

Frekvensredovisningen i *LDOCE* har en delvis annan inriktning, med en klar orientering mot "the 3000 most frequent word in both spoken and written English" (s. xi). *LDOCE* arbetar till skillnad mot *COBUILD* med tre lika stora frekvensband. Varje grupp omfattar tusen ord, med särskilda markeringar för talat respektive skrivet språk. Ett ord som *action* har frekvensmarkörerna "S(poken)1" och "W(ritten)1", vilket innebär att ordet tillhör de tusen vanligaste orden i såväl talspråket som skriftspråket; på samma sätt har verbet *prefer* markörerna "S2" och "W2", dvs. ordet tillhör det andra tusentalet vanligaste ord i tal och skrift. Av särskilt intresse är de många fall där det finns en skillnad mellan tal- och skriftspråksfrekvenser. Substantivet *delay* är t.ex. markerat "S2" och "W3" (tredje vanligaste tusentalet) och är således vanligare i tal än i skrift. För en del uppslagsord finns frekvensmarkör endast för antingen tal eller skrift: *delay* som verb har enbart markören "S3", medan *demand* som verb enbart har "W2", till skillnad mot substantivet *demand*, som i stället har markörerna "S1" och "W1". Det särskilt i talspråket vanliga adverbet *hopefully* har mycket riktigt endast markören "S1".<sup>20</sup> Det går med andra ord att utläsa intressanta skillnader mellan talad och skriven engelska i *LDOCE*s frekvensmarkeringar för de tre tusen vanligaste orden.

För somliga uppslagsord i *LDOCE* ges också särskilda grafiska uppställningar över sådana skillnader. Vid verbet *bet* får man t.ex. reda på inte bara hur mycket vanligare detta verb är i talspråket än i skriftspråket, utan också att detta beror på ordets förekomst i en del mycket vanliga talspråksfraser som *I bet* och *You bet!*. Även andra typer av frekvensinformation ges i grafisk form vid vissa ord. Så ges vid verbet *receive* upplysningar om frekvensskillnaderna mellan *receive* och *get* i talspråk och skriftspråk; vid *trousers* finns frekvensinformation om användningen av *trousers* jämfört med *pants* i brittisk och amerikansk engelska, vid verbet *expect* om hur olika grammatiska mönster (*expect to do sth*, *expect sth/sb* etc.) frekvensmässigt skiljer sig åt. Allt detta är av givet intresse när det gäller att mer precist identifiera

---

<sup>20</sup> Däremot ges ej separata frekvensmarkörer för *hopefully* som satsadverb/ attityd-  
adverb jämfört med *hopefully* som sättsadverb; säkerligen beror dock den höga tal-  
språksfrekvensen nästan uteslutande på den första, av vissa talare ogillade första  
typen (*Hopefully we can solve the problem*).

autentiskt språkbruk för pedagogiska ändamål. Det är naturligtvis också, i likhet med *COBUILDs* frekvensangivelser, av lingvistiskt intresse, ett av många exempel på hur korpuslingvistisk metodik kan bidra till ökad kunskap om engelskan och andra språk.

Att vi i framtiden kommer att få se en utökad användning av pedagogiskt relevanta frekvensuppgifter i "learners' dictionaries" är ingen djärv gissning. Frågan är snarast hur mycket som är pedagogiskt lämpligt att pressa in i en inlärningsordbok och på vilka grunder urvalet bör ske.

### Definitioner och definitionsordlistor

I en inlärningsordbok är det särskilt viktigt att ordbetydelser definieras på ett sådant sätt och med sådana ord att inte orden i själva definitionen hindrar förståelse av uppslagsordets betydelse. Att i ett sådant lexikon förklara ett ord som *beaver* medelst definitionen "any large amphibious broad-tailed rodent of the genus *Castor*..." (*Concise Oxford Dictionary*) vore inte lämpligt. I stället har man alltid i "learners' dictionaries" strävat efter att använda så enkla definitionsord som möjligt. Den första upplagan av *LDOCE*, från 1978, introducerade en kontrollerad definitionsordlista på c. 2000 ord, varmed i princip betydelsen hos samtliga uppslagsord förklarades.<sup>21</sup> Även den senaste upplagan av lexikonet använder sig av denna "justly famous *LDOCE* defining vocabulary, now still further refined and improved" (Randolph Quirk, Preface, s. ix). I slutet av lexikonet finns, liksom i de två tidigare upplagorna, en lista över samtliga definitionsord, "carefully chosen to ensure that the definitions are clear and easy to understand, and that the words used in explanations are easier than the words defined" (s. B12).

Bruket av en begränsad uppsättning definitionsord är i och med 1995 års utgåvor ett genomgående drag i engelska inlärningsordböcker. Nykomlingen bland dem, *CIDE*, anger en definitionsordlista på "under 2000 words" (s. 1702), att jämföra med *LDOCEs* "about 2000 common words" (s. xvi). Pionjären *OALD* är å sin sida mån om att framhålla som en viktig nyhet i den senaste upplagan "the introduction of a 3500-word vocabulary within which definitions are written" (s. vi). Som synes är antalet definitionsord i *OALD* betydligt större än i *LDOCE* och *CIDE*.

---

<sup>21</sup> Till grund för denna definitionsordlista låg Michael Wests *A General Service List of English Words* (1953); se *LDOCE* (1978:F8). Jfr också Wests *An International Reader's Dictionary* (1965), där ordförklaringarna "are written within a vocabulary of 1,490 words held to be the commonest words in English or the words first learnt by foreigners" (s. vii).

Misstanken kan därmed infinna sig att *OALD*s definitionsvokabulär närmast är att betrakta som en måhända något modifierad inventering av de ord som de facto använts i föregående upplagas definitioner. Mot detta talar att långt ifrån alla definitioner i den senaste upplagan är direkta upprepningar från den närmast föregående. Ett exempel kan illustrera detta. Verbet *allocate* definieras i *OALD* (89) på följande sätt: "allot or assign sth...for a special purpose"; i den nya upplagan däremot: "to distribute sth officially...for a special purpose". Den nya definitionen kräver utan tvivel mindre ordkunskap för att kunna förstås. De mindre frekventa definitionsverben *allot* och *assign* har ersatts av det allmännare *distribute*, vilket till skillnad mot *allot* och *assign* också är inkluderat i *OALD*s definitionsordlista.<sup>22</sup>

Redan i sin första upplaga (1987) angav *COBUILD* att man använt sig av en noggrant kontrollerad definitionsvokabulär (omslagets baksida); några ytterligare upplysningar gavs inte. Även den nya upplagan (pärmens baksida) betonar användningen av en "carefully selected defining vocabulary". Till skillnad från de övriga lexikonen ger dock inte *COBUILD* någon separat lista över sina definitionsord utan talar enbart om "a natural defining vocabulary with most words in our definitions being amongst the 2,500 commonest words of English" (s. xviii).<sup>23</sup>

Trots skillnader i storlek m.m. mellan de fyra lexikonen är det uppenbart att den traditionella idén om ett begränsat och kontrollerat antal definitionsord inom spannet 2000–3500 ord numera är väl etablerad bland engelska inlärningslexikon. Man skulle kanske kunna tro att en sådan gemensam utgångspunkt skulle göra orddefinitioner tämligen likartade och förutsägbara. Detta är knappast fallet, vilket även en yttlig jämförelse mellan de olika ordböckerna ger vid handen. Som illustration kan vi kort återvända till verbet *allocate* och jämföra de olika definitionerna (för *OALD*, se ovan): "to give (something) to someone as their share of a total amount, for them to use in a particular way" (*CIDE*); "to decide officially that a particular amount of money, time etc or something such as a house or job etc should be used for a particular purpose" (*LDOCE*); "If one item or share of something is **allocated** to a particular person or for a particular purpose, it is given to

<sup>22</sup> Det kan noteras att varken *LDOCE* eller *CIDE* använder sig av verbet *distribute* som definitionsord. Enligt *LDOCE*s frekvensangivelser (jfr ovan) tillhör det inte heller de tre tusen vanligaste orden, vilket också stämmer överens med *COBUILD*s frekvensgruppering av ordet, för övrigt inom samma frekvensband som *assign*.

<sup>23</sup> Jfr *COBUILD Student's Dictionary* (1990): "[...] we have avoided the restrictions and awkwardness of a fixed defining vocabulary" (s. v). Detta lexikon innehåller också en lista över de 1860 definitionsord som används tio gånger eller mer.



that person or used for that purpose" (*COBUILD*). Även om det finns gemensamma ord i de olika definitionerna, så är det ändå skillnaderna mellan dem som är mest iögonfallande. Olikheterna beträffande ordval, längd och definitionsstil utgör ett värtaligt bevis på att enigheten bland lexikografer är begränsad i fråga om ett så centralt lexikografiskt område som definitioner. Detta i sin tur beror på svårigheten att på någorlunda fast, vetenskaplig grund etablera en allmänt omfattad och praktiskt användbar uppsättning kriterier på vad som utmärker en god definition t.ex. i en inlärningsordbok. Sådana kriterier måste bygga på omfattande forskning kring hur olika faktorer (ordval, längd, syntaktisk komplexitet etc.) påverkar begripligheten för den avsedda lexikonanvändaren.<sup>24</sup>

En central fråga i detta sammanhang är om alla uppslagsord i samma lexikon måste följa samma mall, oberoende av exempelvis uppslagsordets frekvens, ordklassstillhörighet, semantiska karaktär (abstrakt - konkret etc.), syntaktiska egenskaper, kollokationer m.m. Är det t.ex. maximalt användarvänligt att som *COBUILD* alltid definiera medelst en fullständig sats, oavsett uppslagsord? Samma fråga kan ställas beträffande det mer traditionella syntaktiska utbytbarhets-kriterium som de övriga lexikonerna ansluter sig till (jfr definitionerna av *allocate* ovan). Är det kanske i stället så att maximal användarvänlighet skulle kunna uppstå vid en mer eklektisk inställning till olika definitionsstilar inom samma lexikon?<sup>25</sup> Kan det finnas en konflikt mellan lexikografisk estetik och konsekvens å ena sidan och lexikonets rent praktiska syfte å den andra? Detta är inte rätt plats att mer ingående diskutera dessa frågor. Vi kan nöja oss med att konstatera att *COBUILD* fortsätter på den tidigare inslagna vägen med "full sentence definitions", en innovation som väckte diskussion när den först lanserades men som animerat försvaras av John Sinclair i den nya upplagan (s. xi). Denna definitionsstil är än så länge i minoritet bland engelska inlärlarlexikon. Ett begränsat inflytande från *COBUILD*s definitionsstil kan dock spåras i *CIDE* (jfr även not 25), nämligen i kombination med exempel på vissa bibetydelser, som t.ex. (under *pain*): "If you describe

<sup>24</sup> För diskussion kring definitioner och ordförklaringar, se Hanks (1987).

<sup>25</sup> Jfr *Longman Language Activator* (1993), där en sådan eklekticism faktiskt tillämpas; se Ohlander (1995). I *LDOCE* är det däremot ont om *COBUILD*-påverkade definitioner; ett av de fåtaliga exemplen är vid uppslagsordet *awry*, där frasen *go awry* definieras så här: "if something goes awry, it does not happen in the way that was planned", att jämföra med *go bananas*: "to become very angry or excited".

a person or a situation as **a pain in the neck** or **a pain**, you think that it or they are very annoying."<sup>26</sup>

### Semantisk differentiering: homonymi, polysemi m.m.

Ett välkänt problem för såväl lexikografer som lexikonanvändare utgör graford som karakteriseras av polysemi och/eller homonymi. Sådana ord tillhör i regel också den högfrekventa delen av ordförrådet, t.ex. starkt polysema verb som *run* och *take*, substantiv som *class* och *form* eller adjektiv som *dry* och *wrong*. Att sådana ord i ett språk som engelska dessutom ofta kan tillhöra mer än en ordklass gör inte saken lättare. Ofta har sådana ord presenterats som flera separata huvudord i stället för under ett enda huvudord. Variationen mellan lexikon har dock varit stor. Någon gemensam strategi inom detta notoriskt problematiska område – semantiskt/lexikologiskt och lexikografiskt – har inte funnits.

Den första upplagan av *COBUILD* introducerade ett för inlärningsordböcker radikalt grepp: såväl polysemi som homonymi sorterades in under ett och samma huvudord, även oavsett ordklassstillhörighet. Alla betydelser och användningar av *run* presenterades under samma uppslagsord med 67 underavdelningar (plus underavdelningar till många av dessa), med frekvens som en av ordningsprinciperna. Detsamma gällde klara fall av homonymi, som *bank* och *can*. I den nya upplagan av lexikonet medger John Sinclair (s. x) att denna "strict policy of 'one word, one entry' [...] occasionally irritated users." För att motverka detta har den nya upplagan infört s.k. "superheadwords", vilka delar upp vissa tidigare mycket långa "entries" i större underavdelningar. Exempelvis har *mean* i den nya upplagan tre ordklassbaserade "superheadwords" ("verb uses", "adjective uses" och "noun use"). Även mer renodlat semantiska grupperingar förekommer: *lie* är indelat i "position or situation" och "things that are not true", *fancy* i "wanting, liking, or thinking" och "elaborate or expensive"; *bank* och *can* har på liknande sätt numera vardera två "superheadwords". Införandet av "superheadwords" gör det onekligen lättare att hitta i *COBUILD* i de fall där de används. De kunde dock ha använts oftare. För exempelvis *run* är situationen i stort sett oförändrad, dvs. utan "superheadwords". Även många klara homonymifall står kvar under

---

<sup>26</sup> Jfr *COBUILD*: "If you think that a person, job, or situation is very annoying or irritating, you can say that they are **a pain** or **a pain in the neck**" (*pain*, 4).

samma huvudord: för *love* finns "tennisbetydelsen" ("fifteen–love") insprängd bland helt andra betydelser.

Även om förändringarna i *COBUILD* således inte är direkt överväldigande innebär "superheadwords" ändå en modifiering i riktning mot ett mer traditionellt, i det här fallet tydligare presentationsätt. Ökad användarvänlighet är också tanken bakom *LDOCEs* och *CIDEs* bruk av ett slags semantiska grovmarkörer för att göra det lättare att hitta den önskade betydelsen hos ett ord. Under verbet *love* anger *LDOCE* de tre första betydelserna med markörerna ROMANTIC ATTRACTION, CARE ABOUT och LIKE/ENJOY i fetstil före respektive definition, så att läsaren direkt skall ledas till rätt betydelse av uppslagsordet. Dessa "signposts" är avsedda som hjälp vid "entries with many definitions" (s. xvi). Detta är dock ej konsekvent genomfört. En sak är att de inte förekommer vid grammatiska småord som *but*, *of* eller *there*. Men de saknas också vid många lexikala uppslagsord med flera olika definitioner, t.ex. verbet *like* och substantivet *accident*. Logiken här är inte helt lätt att genomsåda. Annars fungerar dessa "signposts" väl som praktiska visuella hjälpmedel för lexikonanvändaren. Detsamma gäller även *LDOCEs* användning av introducerande, översiktliga "menus" över de huvudsakliga betydelsegrupperna vid vissa uppslagsord med ett stort antal betydelser och användningar, som *business*, *hot* och *run*.

I *CIDE* har den semantiska differentieringen i många fall gått ännu längre: "When a word has more than one meaning, [...] guide words help you to find the right one quickly" (s. x). Syftet med dessa "guide words" är ungefär detsamma som med *LDOCEs* "signposts". Det finns emellertid viktiga skillnader. I *LDOCE* används "signposts" för att tydliggöra olika betydelser inom *samma* huvudord med flera olika definitioner (jfr ovan). *CIDEs* "guide words" används i stället för att markera *olika* huvudord. Adjektivet *green* har i *LDOCE* bl.a. följande "signposts" under *samma* huvudord: COLOUR, WITHOUT EXPERIENCE, POLITICS; i princip samma semantiska begrepp (plus ytterligare några) används som "guide words" i *CIDE*, som visuella markörer för *olika* huvudord.<sup>27</sup> Någon systematisk parallelism av detta slag finns emellertid inte mellan de båda lexikonerna. Ett huvudord i *CIDE* kan innehålla ett antal betydelsenysanser (utan "guide words") som i *LDOCE* är tydligt markerade med var sin "signpost". Så är fallet

<sup>27</sup> Ett kuriosum i ett lingvistikhistoriskt perspektiv: ett av exemplen under *green* POLITICAL i *CIDE* är *green ideas*, selektionsregler är inte eviga utan är beroende av ords betydelseutveckling. Jfr dock Pinker (1994:93), där det i en diskussion kring Chomskys berömda exempel sägs att "the probability that *green* is followed by *ideas* [...] is surely zero."

med t.ex. verbet *kill*. Det motsatta förhållandet kan också gälla. I *LDOCE* är exempelvis samtliga betydelser av substantivet *depression* samlade under samma huvudord, men utan "signposts". I *CIDE* är situationen mer komplicerad, även för lexikonanvändaren. Det finns för det första två separata uppslagsord, markerade med WEATHER respektive HOLE. Vid dessa huvudord ges hänvisningar till två andra huvudord, nämligen verben *depress* CAUSE SADNESS och *depress* REDUCE. Under dessa huvudord återfinns sedan den psykologiska respektive den ekonomiska betydelsen hos *depression*. Den semantiska tanken bakom detta arrangemang – *depression* utspritt över fyra olika uppslagsord – må vara begriplig; praktiskt är det knappast. En annan olikhet mellan lexikonerna är att *CIDE* till skillnad från *LDOCE* (jfr ovan) även försöker särskilja betydelserna hos grammatiska funktionsord med hjälp av "guide words". I ett fall som *but*, där huvudbetydelserna är någorlunda klart åtskilda går det relativt bra: DIFFERENCE, EXCEPT, ONLY. I andra fall uppstår oundvikligen problem, t.ex. vid *on*, vars ofta överlappande betydelser och funktioner redovisas under 31 olika huvudord, markerade med "guide words" av de mest skiftande slag (ABOVE, CONNECTED, WRITING, PAIN, TO etc.). Uppgiften är svår, försöket heroiskt, resultatet inte helt övertygande.

Som framgått av ovanstående presentation finns det trots viktiga skillnader även grundläggande likheter mellan *LDOCE* och *CIDE*. Den viktigaste är kanske att båda lexikonerna, i olika grad och på olika sätt, i praktiken avskaffat skillnaden mellan homonymi och polysemi, denna klassiska lexikografiska knäckfråga. I den föregående upplagan av *LDOCE* markeras fall av klar homonymi, som t.ex. vid substantivet *bank*, genom olika uppslagsord, där vart och ett i sin tur kan inrymma polysemi. Den nya upplagan behandlar i stället de olika betydelserna av substantivet *bank* under samma huvudord, utan någon markerad skillnad mellan homonymi och polysemi.<sup>28</sup> I *CIDE* markeras visserligen traditionella homonymer i regel med var sitt huvudord, som i fallet *bank*, men detta gäller även många fall av vad som traditionellt skulle betraktas som klar polysemi (jfr exemplet ovan). *COBUILD*, å sin sida, gjorde ju redan i sin första upplaga rent hus med den intuitivt centrala men semantiskt och lexikografiskt ofta svårhanterade distinktionen mellan homonymi och polysemi. Införandet av "superheadwords" kan visserligen ses som ett slags reträtt men är i huvudsak en praktisk

---

<sup>28</sup> Däremot skiljer man fortfarande, genom olika huvudord, mellan *bank* som substantiv och *bank* som verb; jfr *COBUILD*, som visserligen genom olika "superheadwords" skiljer mellan "finance and storage" och "areas and masses" men som under den förra huvudbetydelsen tar upp *bank* som både substantiv och verb.

nödfallsåtgärd mot ett begränsat antal huvudord med särskilt många och svåröverskådliga underavdelningar. Endast *OALD* håller som en sista ropande röst fast vid tanken att trots alla problematiska fall försöka lexikografiskt markera skillnaden mellan homonymi och polysemi. Det skall bli intressant att se om kommande upplagor av lexikonet kan stå emot grupptricket.

Sammanfattningsvis kan man säga att engelska inlärningsordböcker tycks befinna sig i en process av fortlöpande desarmering av den besvärliga homonymi/polysemi-distinktionen, en utplattning av de semantiska hierarkierna om man så vill. Denna strategi, som kan ta sig olika uttryck, är bekväm för lexikografen. Huruvida den också är det för lexikonanvändaren är inte lika självklart. Skillnaden mellan homonymi och polysemi kan, trots mängder av gränsfall, betraktas som en psykologisk realitet. Att arrangera ordbetydelser på ett sådant sätt att gränsen mellan (synkront) besläktade och helt obesläktade betydelser blir oklar, eller helt suddas ut, kan knappast vara befrämjande för ordinläring, där sammanhang mellan ord och betydelser brukar framhävas som något väsentligt. Någon helt idealisk lösning på dilemmat – om det nu finns någon – är ännu inte i sikte. Problemet har alla möjligheter att förbli klassiskt.

### **Avslutande reflektioner**

På senare år har "learners' dictionaries" kommit att spänna över allt vidare fält, i takt med delvis nya perspektiv och betoningar inom såväl lingvistik som språkpedagogik. Sedan länge har inlärningsordböcker innehållit mängder av rent grammatisk information. Vi har under något tiotal år också kunnat bevittna en ökande fokusering på socio-lingvistiska, pragmatiska och textlingvistiska aspekter av språket; det samma gäller kollokationer. Hand i hand med denna kvalitativa utveckling går en kvantitativ: fler och fler uppslagsord, tjockare och tyngre lexikon. Konkurrenssituationen driver också fram något av en hetsjakt på nya "special features", innehållsligt såväl som i fråga om presentation, allt i syfte att förstärka den egna profilen på marknaden.

Jag har i denna artikel velat peka på ett antal någorlunda tydliga tendenser och därmed sammanhängande frågor i engelska "learners' dictionaries"; fler skulle ha kunnat nämnas. Som framgått finns det på viktiga punkter en samstämmighet mellan de olika lexikonen, även om skillnaderna mellan dem i många fall är lika iögonenfallande. Det är på sätt och vis fascinerande att det trots den kollektiva rörelsen, samsynen i stort om vad en inlärningsordbok bör innehålla, ändå är så lite som är självklart när det gäller realiseringen av målen, alltifrån vilka ord som

skall tas med till lexikonartiklarnas utformning i stort och smått. Frågorna är gemensamma, svaren ofta olika. Något säger detta onekligen om lexikografins och lexikografens situation i en ocean av valmöjligheter, en ocean som inte blivit mindre genom de datoriserade korpusarnas intåg.

Det råder ingen tvekan om att engelska inlärningsordböcker blivit allt bättre med åren. Det är lätt att hålla med Randolph Quirk i hans förord till *LDOCE* (s. 10): "I would have liked to have this dictionary when I was a student." Detta skulle kunna gälla samtliga de här diskuterade lexikonen, trots att inget av dem är invändningsfritt, ens på ett generellt plan.

Mot slutet av sin inledning till *COBUILD* (s. xi) säger John Sinclair: "Users expect more and more from their dictionaries [...]". Hur sant detta verkligen är kan man möjligen tvista om. Oomtvistligt är däremot att i alla fall engelska inlärningslexikon tycks drivas framåt av ambitionen att tillgodose vad man uppfattar som ständigt ökande användarbehov. Mycket gott kommer helt visst av detta. Men det kan också finnas en risk att ansamlingen av mer och mer – även mer och mer olika – information inom ett och samma, allt större lexikon leder till "information overload" för användaren. Om ett lexikon skall innehålla nästan allt riskerar lättillgängligheten och den pedagogiska tydligheten att gå förlorade. Det är inte säkert att all den korpusbaserade information som i princip är möjlig att inkludera också är pedagogiskt motiverad att ha med. Ett lexikon kan ändå aldrig bli en ersättning för egna språkliga erfarenheter, där inte minst textläsning – orden i verkligt autentisk kontext, snarare än i autentiska exempel – spelar en central roll vid uppbyggnaden av ett fungerande ordförråd.

Detta leder vidare till frågan om framtidens "learners' dictionaries". De fyra här diskuterade lexikonen utgör inte någon slutpunkt i utvecklingen; många problem återstår att lösa. Arbetet fortsätter – men i vilken riktning? Kanske börjar det bli dags att ställa frågan om vi i själva verket håller på att närma oss gränsen för den konventionella formen för inlärningslexikon, dvs. en bok. Även om det förmodligen alltid kommer att finnas behov av lexikon i bokform kommer kanske inlärningsordböckernas framtida utveckling huvudsakligen att äga rum inom andra medier och i andra format. Det handlar då inte främst om CD-ROM-versioner av redan existerande lexikon, snarare om helt andra sätt att utnyttja den nya teknikens potential, inte minst möjligheten för användaren att själv, efter egna behov, "skräddarsy" sitt

eget elektroniska lexikon med olika databaser som grund. Men vi är inte där än.<sup>29</sup>

För en mer överskådlig framtid kan vi nog trots allt vänta oss en fortsatt intensiv utveckling inom ramen för mer konventionella inlärningsordböcker, där några av de här diskuterade trenderna och frågorna säkerligen kommer att spela en central roll. Ett nytt millennium står för dörren: hur kommer 2000-talets upplagor av "The Big Four" att se ut? Startskottet har gått; jakten har börjat.

## Litteratur

- Aijmer, Karin, Altenberg, Bengt & Johansson, Mats (eds) 1996: *Languages in Contrast*. Lund University Press.
- Atkins, B. T. S. 1996: Bilingual dictionaries: past, present and future. I: Gellerstam *et al.* 1996, 515–546.
- Calzolari, Nicoletta 1996: Lexicon and corpus: a multi-faceted interaction. I: Gellerstam *et al.* 1996, 3–16.
- COBUILD 1987 = *Collins COBUILD English Language Dictionary*. First ed. Editor in Chief: John Sinclair. London/Glasgow: Collins.
- Collins COBUILD Student's Dictionary* 1990. Editor in Chief: John Sinclair. London/Glasgow: Collins.
- Concise Oxford Dictionary* 1990. Eighth ed. Oxford: The Clarendon Press.
- Gellerstam, M., Järborg, J., Malmgren, S.-G., Norén, K., Rogström, L. & Røjder Papmehl, C. (eds) 1996: *Euralex '96. Proceedings I–II*. Göteborg University, Department of Swedish.
- Granger, Sylviane 1996: From CA to CIA and back: an integrated approach to computerized bilingual and learner corpora. I: Aijmer *et al.* 1996, 37–51.
- Hanks, Patrick 1987: Definitions and explanations. I: Sinclair, John (ed), *Looking Up. An Account of the COBUILD Project in lexical computing*. Collins 1987, 116–136.
- Hill, Robert J. 1982: *A Dictionary of False Friends*. London: Macmillan.
- Kernerman, Lionel 1996: English learners' dictionaries: How much do we know about their use? I: Gellerstam *et al.* 1996, 405–414.
- LDOCE 1978 = *Longman Dictionary of Contemporary English*. First ed. Editor-in-Chief: Paul Procter. Longman.

---

<sup>29</sup> För en ingående diskussion om framtidens elektroniska lexikon, se Atkins (1996).

- LDOCE* 1987 = *Longman Dictionary of Contemporary English*. Second ed. Editorial Director: Della Summers. Longman.
- Longman Language Activator* 1993. Editorial Director: Della Summers. Longman.
- OALD* 1989 = *Oxford Advanced Learner's Dictionary of Current English*. Fourth ed. Chief Editor: A.P. Cowie. Oxford University Press.
- Ohlander, Sölve 1991: Efterskrift till U. Ohlander 1991, 242–295.
- Ohlander, Sölve 1995: Le mot juste? Några tankar kring "världens första produktionslexikon". I: *LexicoNordica* 2, 237–259.
- Ohlander, Urban 1991: *Språk som bro och barriär*. Stockholm: Carlssons.
- Oxford Advanced Learner's Encyclopedic Dictionary* 1992. Oxford University Press.
- Nationalencyklopedins ordbok*. Andra bandet. 1996. Vetenskaplig ledare: Sture Allén. Bra Böcker.
- Pinker, Steven 1994: *The Language Instinct*. Allen Lane, The Penguin Press.
- Swan, Michael & Smith, Bernard (eds) 1987: *Learner English. A teacher's guide to interference and other problems*. Cambridge University Press.
- Veisbergs, Andrejs 1996: False friends dictionaries: a tool for translators or learners or both. I: Gellerstam *et al.* 1996, 627–634.
- West, Michael 1953: *A General Service List of English words*. Longmans.
- West, Michael 1965: *An International Reader's Dictionary*. Longmans.