

Göteborgsstudier i nordisk språkvetenskap 27

# Ordboksanvändning på nätet

En undersökning av användningen av  
*Lexins svenska lexikon*

*(English Summary)*

av Ann-Kristin Hult



GÖTEBORGS UNIVERSITET

TITLE: Dictionary use on the internet. Empirical studies of the use and users of the *Swedish Lexin dictionary*

SWEDISH TITLE: Ordboksanvändning på nätet. En undersökning av användningen av *Lexins svenska lexikon*

LANGUAGE: Swedish (English summary)

AUTHOR: Ann-Kristin Hult

## Abstract

In this thesis a study of the use and users of the Swedish online, monolingual, learner's dictionary *Lexins svenska lexikon*, LSL, is presented. The target group of LSL is learners at a relatively advanced beginner's level. In 2008–2011, the LSL underwent a comprehensive revision and the study is based on data from before (LSL3) and after (LSL4) the revision. The thesis has a threefold aim, namely to 1) investigate the use of LSL, 2) find out who the users of LSL are, and 3) take a closer look at individual informant's search behaviour. This, in turn, forms the basis for three sub studies. A general question to be answered is whether the LSL is actually used by the dictionary's target group and if so, how?

In sub study 1 the use of LSL is examined by means of a log file analysis which produces a general view of the users' most frequent look-ups. The data from before (i.e. 2008) and after (i.e. 2011) the revision are strikingly different. For example, the 2011 data contains more central and common Swedish words than the 2008 data. The users of LSL4 appear, to a greater extent, to belong to the target group of LSL, than the users of LSL3.

In sub study 2 a web questionnaire survey is used to find out information about the users of LSL which results in an important, though generalized, user profile. Evidently, both LSL3 and LSL4 are used by the target group, but, to judge from the two sets of questionnaire data, to a greater extent in the updated version of LSL, i.e. LSL4.

Lastly, in sub study 3 log file data and questionnaire data are triangulated which made it possible to observe more closely individual target group informants and their dictionary activities. This qualitative study results in yet another picture of the use of LSL.

To sum up, combining methods and triangulating data seem to have resulted in a more complete and balanced picture of the use of LSL.

KEYWORDS: electronic lexicography, dictionary use, monolingual dictionary, learner's dictionary, log file analysis, web questionnaire, data triangulation

© Ann-Kristin Hult

DISTRIBUTION: Institutionen för svenska språket  
Box 200  
405 30 Göteborg

OMSLAGSBILD: Sven Lindström

FOTO PORTRÄTT: Niklas Sjöholm

ISSN: 1652-3105

ISBN: 978-91-87850-62-2

LÄNK TILL E-PUBLICERING: <http://hdl.handle.net/2077/49441>

SÄTTNING: Sven Lindström

TRYCKNING: Ineko AB

# Innehåll

Tabeller .....	I
Figurer .....	V
1. Inledning.....	1
1.1 Syfte och frågeställningar .....	2
1.2 Disposition.....	5
2. Allmän teoretisk bakgrund .....	7
2.1 Lexikografi – från hantverk till vetenskap .....	7
2.2 Lingvistikorienterad lexikografisk teori .....	9
2.3 Autonom lexikografisk teori.....	12
2.3.1 Teori utifrån ordbokens struktur .....	12
2.3.2 Speciella teoretiska inriktningar med fokus på tvåspråkig lexikografi.....	14
2.3.3 Funktionsläran .....	17
2.4 Diskussion.....	19
3. Ordinläring, inlärningsordböcker och Lexin .....	23
3.1 Ordinläring och ordböcker .....	24
3.1.1 Ordinläring ur ett andraspråksperspektiv .....	24
3.1.2 Ordböcker ur ett andraspråksperspektiv .....	32
3.1.2.1 Lemmaselektion .....	34
3.1.2.2 Uttal.....	37
3.1.2.3 Morfologi: böjning och ordbildning.....	38
3.1.2.4 Betydelsebeskrivning.....	39

3.1.2.5	Språkprov .....	41
3.1.2.6	Konstruktioner .....	43
3.1.2.7	Bruk och stil .....	44
3.1.2.8	Illustrationer .....	45
3.2	Lexin – inlärningsordböcker för invandrare .....	47
3.2.1	Projektet Lexikon för invandrare.....	47
3.2.2	<i>Lexins svenska lexikon 3, LSL3</i> .....	51
3.2.3	<i>Lexins svenska lexikon 4, LSL4</i> .....	56
3.2.4	Mer om Lexins internetversion.....	59
3.2.4.1	Webbgränssnittet .....	59
3.2.4.2	Rättstavningsfunktionen i Lexin .....	62
3.2.4.3	Användningen av lexikonen .....	63
4.	Användarforskning: metoder och resultat .....	67
4.1	Möjliga tillvägagångssätt .....	67
4.2	Valda tillvägagångssätt .....	68
4.2.1	Surveyundersökning genom frågeformulär .....	68
4.2.2	Surveyundersökning genom observation .....	71
4.2.3	Metodkombination .....	73
4.2.4	Urval – sannolikhet och representativitet.....	75
4.3	Resultat av tidigare användarstudier .....	76
4.3.1	Några traditionella enkätstudier .....	78
4.3.2	Några traditionella observationsstudier .....	82
4.3.3	Studier av ordböcker i elektroniskt format .....	85
4.3.3.1	Webbenkäter.....	87
4.3.3.2	Loggfilsanalyser.....	91
4.3.4	Sammanfattning.....	97
5.	Metod och material .....	99
5.1	Metod i delstudie 1: Loggfilsanalys.....	100
5.1.1	Insamlingsförfarande .....	100
5.1.2	Metodologiska överväganden .....	103
5.2	Metod i delstudie 2: Webbaserat frågeformulär .....	104
5.2.1	Insamlingsförfarande .....	104
5.2.2	Metodologiska överväganden .....	105
5.3	Metod i delstudie 3: Datatriangulering .....	107
5.3.1	Insamlingsförfarande .....	107
5.3.2	Metodologiska överväganden .....	108
5.4	Material i delstudie 1: Loggfilerna.....	109
5.4.1	Lexikonuppslagningar.....	109
5.4.2	Navigeringsval .....	114
5.5	Material i delstudie 2: Frågeformulären.....	116
5.5.1	Beskrivning av frågeformulären.....	117
5.5.2	Redovisning av bortfall .....	126
5.6	Material i delstudie 3: Enskilda informanternas aktiviteter .....	129
5.7	Etiska aspekter .....	131
5.8	Inför resultatredovisningen .....	132

6. Delstudie I: Användningen av LSL .....	133
6.1 Loggfiler av typ a: Lexikonuppslagningar .....	133
6.1.1 Träffar 2008 och 2011.....	136
6.1.1.1 Söksträngar.....	136
6.1.1.2 Mest sökta lemman.....	142
6.1.2 Korrigerade uppslagningar 2008 och 2011.....	148
6.1.2.1 Automatkorrigering.....	148
6.1.2.2 Korrigering med förslag.....	151
6.1.3 Icke-träffar 2008 och 2011.....	155
6.1.4 Flerordsuppslagningar.....	157
6.2 Loggfiler av typ b: Navigeringsval.....	159
6.3 Sammanfattning.....	161
7. Delstudie II: Användarna av LSL .....	163
7.1 Bakgrundsinformation.....	163
7.1.1 Födelseår .....	164
7.1.2 Kön .....	165
7.1.3 Högsta avslutade utbildning.....	166
7.1.3.1 Utbildning i Sverige.....	166
7.1.3.2 Utbildning i hemlandet.....	169
7.1.4 Sysselsättning.....	170
7.1.5 Sammanfattning.....	172
7.2 Informanternas utsagor om användningen av LSL .....	173
7.2.1 Sökfrequens .....	173
7.2.2 Användningssituation .....	174
7.2.3 Mest eftersökta informationskategorier .....	177
7.2.4 Instruktioner till Lexin.....	180
7.2.5 Hjälpmedel vid sidan av Lexin .....	181
7.2.6 Sammanfattning.....	182
7.3 Användare med annat modersmål än svenska .....	183
7.3.1 Ankomst till Sverige.....	183
7.3.2 Sfi och/eller annan kurs i svenska språket .....	184
7.3.3 Modersmål .....	185
7.3.4 Bästa språk.....	189
7.3.5 Bruk av svenska och svenskkunskaper .....	189
7.3.6 Sammanfattning.....	190
7.4 Kommentarer om LSL.....	191
7.4.1 Typer av kommentarer .....	192
7.4.2 Informanternas kommentarer om Lemma.....	195
7.4.3 Informanternas kommentarer om Betydelse.....	196
7.4.4 Informanternas kommentarer om Exempel.....	197
7.4.5 Sammanfattning.....	198
8. Delstudie III: Målgruppsanvändning .....	201
8.1 Beskrivning av målgruppen.....	201
8.2 Informant 1: Mervan .....	203
8.2.1 Mervans enkätsvar .....	203
8.2.2 Mervans aktiviteter i LSL.....	204

8.2.2.1 Datum och tidpunkt för användningen.....	205
8.2.2.2 Typer av aktiviteter.....	206
8.2.2.3 Aktiviteternas förlopp .....	207
8.2.2.4 Kommentar .....	207
8.3 Informant 2: Elira.....	208
8.3.1 Eliras enkätsvar .....	208
8.3.2 Eliras aktiviteter i LSL.....	210
8.3.2.1 Datum och tidpunkt för användningen.....	212
8.3.2.2 Typer av aktiviteter.....	212
8.3.2.3 Aktiviteternas förlopp .....	213
8.3.2.4 Kommentar .....	216
8.4 Informant 3: Luciana.....	217
8.4.1 Lucianas enkätsvar .....	217
8.4.2 Lucianas aktiviteter i LSL.....	218
8.4.2.1 Datum och tidpunkt för användningen.....	219
8.4.2.2 Typer av aktiviteter.....	219
8.4.2.3 Aktiviteternas förlopp .....	220
8.4.2.4 Kommentar .....	221
8.5 Informant 4: Kazim .....	221
8.5.1 Kazims enkätsvar .....	221
8.5.2 Kazims aktiviteter i LSL.....	222
8.5.2.1 Datum och tidpunkt för användningen.....	223
8.5.2.2 Typer av aktiviteter.....	224
8.5.2.3 Aktiviteternas förlopp .....	224
8.5.2.4 Kommentar .....	226
8.6 Informant 5: Farah .....	227
8.6.1 Farahs enkätsvar.....	227
8.6.2 Farahs aktiviteter i LSL .....	228
8.6.2.1 Datum och tidpunkt för användningen.....	229
8.6.2.2 Typer av aktiviteter.....	229
8.6.2.3 Aktiviteternas förlopp .....	229
8.6.2.4 Kommentar .....	231
8.7 Sammanfattning.....	232
9. Sammanfattning och framåtblick.....	235
9.1 Användningen av LSL .....	235
9.2 Användarna av LSL.....	238
9.3 Målgruppsanvändning på individnivå.....	239
9.4 Framtidens LSL och vidare forskning.....	241
Efterord.....	245
Summary.....	247
Referenser.....	259
Bilaga 1. Frågeformuläret från 2007.....	277
Bilaga 2. Frågeformuläret från 2011.....	279

# Tabeller

TABELL 3:1.	<i>Ordboksinformation och vad det innebär att kunna ett ord. (Efter Nation 2013:428).</i> . . . . .	33
TABELL 3:2.	<i>Tillgängliga och planerade Lexin-lexikon (2016-07-22)</i> . . . . .	49
TABELL 3:3.	<i>Informationskategorier i LSL (efter Gellerstam 1999:7)</i> . . . . .	52
TABELL 3:4.	<i>Antal uppslagningar i Lexins svenska lexikon</i> . . . . .	65
TABELL 5:1.	<i>Antal uppslagningar per dag i LSL under datainsamlingsperioderna 2008 och 2011</i> . . . . .	109
TABELL 5:2.	<i>Uppslagningar 2008, måndag 6 oktober mellan kl. 10.00.00 och 10.00.59</i> . . . . .	111
TABELL 5:3.	<i>Uppslagningar 2011, måndag 3 oktober mellan kl. 10.01.00 och 10.01.59</i> . . . . .	113
TABELL 5:4.	<i>Antal klickningar som görs i rutan framför respektive informationskategori</i> . . . . .	115
TABELL 5:5.	<i>Antal uteblivna svar på enskilda frågor i enkätmaterial från 2007 och 2011</i> . . . . .	127
TABELL 5:6.	<i>Aktiviteter knutna till informanten bakom enkät-id 2011:287.</i> . . . . .	130
TABELL 5:7.	<i>Delstudie 1–3: Strategi, metod och material</i> . . . . .	132
TABELL 6:1.	<i>Antal uppslagningar per dag under datainsamlingsperioderna 2008 och 2011</i> . . . . .	134
TABELL 6:2.	<i>Frekvenser för uppslagningar i materialet från 2008 och 2011 fördelade på träff, automatkorrigering (A), korrigering med förslag (F) och icke-träff.</i> . . . . .	135
TABELL 6:3.	<i>De 50 mest frekventa träffarna i 2008 och 2011 års material</i> . . . . .	137
TABELL 6:4.	<i>Frekvenser i gp08 och gp11 för de 25 vanligaste söksträngarna i 2008 och 2011 års material</i> . . . . .	141

TABELL 6:5.	<i>De 50 mest sökta lemmarna som ger träff i 2008 och 2011 års loggfils-material . . . . .</i>	143
TABELL 6:6.	<i>Antal lemmar, betydelsemoment (lexem) och uttryck i LSL4 för de 25 mest frekventa lemmarna i 2008 och 2011 års material. . . . .</i>	145
TABELL 6:7.	<i>De 25 mest frekventa lemmarna från 2008 och 2011 års material och deras placering i Kelly-listan . . . . .</i>	147
TABELL 6:8.	<i>De 25 vanligaste uppslagningarna som leder till automatkorrigering 2008 och 2011, i frekvensordning . . . . .</i>	149
TABELL 6:9.	<i>De 25 vanligaste uppslagningarna som leder till korrigering med förslag 2008 . . . . .</i>	151
TABELL 6:10.	<i>De 25 vanligaste uppslagningarna som leder till korrigering med förslag 2011 . . . . .</i>	153
TABELL 6:11.	<i>De 25 vanligaste icke-träffarna under insamlingsperioderna 2008 och 2011 . . . . .</i>	156
TABELL 6:12.	<i>Antal klickningar som görs i rutan framför respektive informationskategori . . . . .</i>	160
TABELL 7:1.	<i>Informanter per åldersgrupp i Studie<sup>11</sup> och Studie<sup>07</sup>, i absoluta och relativa tal . . . . .</i>	164
TABELL 7:2.	<i>Informanternas sysselsättning i Studie<sup>11</sup>, i frekvens och procent . . . . .</i>	170
TABELL 7:3.	<i>Typ av kommentar, frekvens och exempel, Studie<sup>11</sup>. . . . .</i>	171
TABELL 7:4.	<i>Sysselsättning Studie<sup>11</sup> och Studie<sup>07</sup>, i procent . . . . .</i>	171
TABELL 7:5.	<i>Studie<sup>11</sup>: Informanternas uppgifter om användningssituation . . . . .</i>	175
TABELL 7:6.	<i>Informanternas uppgifter om mest eftersökta informationskategorier i Studie<sup>11</sup> . . . . .</i>	177
TABELL 7:7.	<i>Informanternas uppgifter om mest eftersökta informationskategorier i Studie<sup>07</sup> . . . . .</i>	179
TABELL 7:8.	<i>Studie<sup>11</sup>: Användning av andra språkliga hjälpmedel vid sidan av LSL. . . . .</i>	181
TABELL 7:9.	<i>Antal informanter per Lexin-språk och övriga språk i Studie<sup>11</sup> respektive Studie<sup>07</sup> . . . . .</i>	186
TABELL 7:10.	<i>Antal uppslagningar i de tvåspråkiga Lexin-lexikonerna september 2011 och oktober 2007, efter frekvens . . . . .</i>	188
TABELL 7:11.	<i>Kommentarernas antal, längd (i antal ord räknat) samt medianlängd. . . . .</i>	191
TABELL 7:12.	<i>Absoluta och relativa frekvenser för respektive typ av kommentar. . . . .</i>	192
TABELL 7:13.	<i>Exempel på kommentarer av varje typ i Studie<sup>07</sup> och Studie<sup>11</sup>. . . . .</i>	193
TABELL 8:1.	<i>Aktiviteter knutna till Mervan, enkät-id 2011:287. . . . .</i>	205
TABELL 8:2.	<i>Typ av aktivitet och dess frekvens. . . . .</i>	206
TABELL 8:3.	<i>Aktiviteter knutna till Elira, enkät-id 2011:635. . . . .</i>	210
TABELL 8:4.	<i>Typ av aktivitet och dess frekvens. . . . .</i>	213
TABELL 8:5.	<i>Aktiviteter knutna till Luciana, enkät-id 2011:839 . . . . .</i>	218
TABELL 8:6.	<i>Typ av aktivitet och dess frekvens. . . . .</i>	219
TABELL 8:7.	<i>Aktiviteter knutna till Kazim, enkät-id 2011:72 . . . . .</i>	222



TABELL 8:8.	<i>Typ av aktivitet och dess frekvens</i> . . . . .	224
TABELL 8:9.	<i>Aktiviteter knutna till Farah, enkät-id 2011:434</i> . . . . .	228
TABELL 8:10.	<i>Typ av aktivitet och dess frekvens</i> . . . . .	229



## Figurer

FIGUR 3:1.	<i>Scenisk illustration över lemmat vardagsrum. Lexin/Bildteman.</i>	46
FIGUR 3:2.	<i>Uppslagsordet potatis och dess illustration (NoKSO 2001).</i>	47
FIGUR 3:3.	<i>Artikeln väntar i LSL3 och vänta i NoKSO, i komprimerad form.</i>	53
FIGUR 3:4.	<i>Artikeln väntar i LSL3, i komprimerad form</i>	56
FIGUR 3:5.	<i>Artikeln väntar, väntan och väntar sig i LSL4, i komprimerad form.</i>	57
FIGUR 3:6.	<i>Webbgränssnitt före uppdateringen: uppslagsord nyckel</i>	60
FIGUR 3:7.	<i>Webbgränssnitt efter uppdateringen 2011: startsida.</i>	60
FIGUR 3:8.	<i>Det senaste gränssnittet: uppslagsordet nyckel. Hämtad 2016-07-22.</i>	61
FIGUR 3:9.	<i>Resultat av uppslagning på söksträngen cyklar i LSL4. Hämtat 2016-11-18.</i>	62
FIGUR 4:1.	<i>Exempel på uppslagsordet nyckel. Hämtad 2007-08-28.</i>	94
FIGUR 5:1.	<i>Exempel på lexikonuppslagning som genererar loggfiler av typ a. Hämtad 2016-05-09.</i>	100
FIGUR 5:2.	<i>Typer av användaraktiviteter som genererar loggfiler av typ b. Hämtad 2016-05-09.</i>	101
FIGUR 5:3.	<i>Popup-fönster som visades vid klickning på Inställningar vid undersökningstillfället. Hämtad 2016-08-09</i>	101
FIGUR 5:4.	<i>Komplett artikel för fot i LSL4. Hämtad 2016-05-25.</i>	102
FIGUR 5:5.	<i>Manipulerad artikel för fot i LSL4. Hämtad 2016-05-25.</i>	102
FIGUR 5:6.	<i>Resultatet av söksträngen habitat. Hämtad 2016-05-18</i>	112
FIGUR 5:7.	<i>Typer av användaraktiviteter som genererar loggfiler av typ b. Hämtad 2016-05-09.</i>	114
FIGUR 5:8.	<i>BILD SVENSKA från uppslagsordet rulltrappa (pil infogad i efterhand). Hämtad 2016-05-09</i>	114

FIGUR 5:9.	<i>VISA FILM från uppslagsordet trappa. Hämtad 2016-05-09 . . . . .</i>	115
FIGUR 7:1.	<i>Studie<sup>11</sup>: Högsta avslutade utbildning i Sverige. . . . .</i>	166
FIGUR 7:2.	<i>Studie<sup>11</sup>: Högsta avslutade utbildning i hemlandet . . . . .</i>	169
FIGUR 7:3.	<i>Studie<sup>11</sup>: Informanternas uppgifter om hur ofta LSL används . . . . .</i>	173
FIGUR 7:4.	<i>Studie<sup>11</sup>: Informanternas ankomstår till Sverige . . . . .</i>	183
FIGUR 8:1.	<i>Rullista med sökbara ordformer som visas allteftersom en söksträng skrivs in i sökfältet: exemplet fragment/fragmentarisk. Hämtad 2016-08-10 . . . . .</i>	225
FIGUR 8:2.	<i>Komplett artikel för bagatell i LSL4. Hämtad 2016-08-10 . . . . .</i>	226
FIGUR 8:3.	<i>Manipulerad artikel för bagatell i LSL4. Hämtad 2016-08-10 . . . . .</i>	227
FIGUR 8:4.	<i>Nomenklatorisk illustration till rubriken "Hygien" . . . . .</i>	231

# 1. Inledning

I denna avhandling undersöks användningen av internetordboken *Lexins svenska lexikon* (hädanefter oftast LSL). Det är en del av projektet Lexin och utgörs av en serie lexikon innehållande den enspråkiga svenska ordboken LSL samt ett tjugotal tvåspråkiga lexikon. Lexin-ordböckerna har funnits tillgängliga på internet sedan mitten av 1990-talet. Förlagan utgörs emellertid av ordböcker i tryckt format. I såväl de tryckta verken som internetversionerna ingår ett bildtema innehållande omkring 1 800 illustrationer. I internetversionen illustreras dessutom över 700 verb med hjälp av korta videofilmer. Lexin-ordböckerna riktar sig främst till invandrare som nyligen anlänt till Sverige och som befinner sig på nybörjarnivå i sin svenskinläring, men ordböckerna är också tänkta att tillfredsställa de behov av förbättrade kunskaper i svenska som finns hos invandrare som har vistats i Sverige under en längre tid (Beijer & Fre Woldu 1997:8). Medan de tvåspråkiga lexikonen riktar sig till *absoluta nybörjare* är LSL främst till för de inlärare som är *relativt avancerade nybörjare*. De tvåspråkiga lexikonen finns på större minoritetsspråk<sup>1</sup> i Sverige där utbudet av tvåspråkiga ordböcker på marknaden är starkt begränsat. De största Lexin-ordböckerna, däribland LSL, har ca 30 000 uppslagsord.

Fram till mitten av 1900-talet genomfördes få undersökningar av användare och användning av ordböcker. Alltsedan början på 1960-talet har emellertid ordböksanvändning, internationellt sett, beforskats i allt större utsträckning.

---

1 Termen *minoritetsspråk* används här i den allmänna betydelsen 'språk som i ett flerspråkigt samhälle används av en språklig minoritet' (jfr *Nationalencyklopedin*, minoritetsspråk).

På 1990-talet ökade studiernas antal i rask takt. Vad gäller svenska förhållanden saknades länge större studier inom detta forskningsområde (Svensén 2004:534). Gemensamt för många av de internationella användarundersökningarna är att informanterna är universitetsstuderande och har engelska som främmande språk, och vanligen också studerar engelska. Studier om mindre kvalificerade informanter råder det däremot en brist på (Hulstijn & Atkins 1998:10, se även översikten hos Welker 2006).

Inom användarforskningen<sup>2</sup> har naturligen intresset alltmer riktats mot användningen av elektroniska ordböcker. En elektronisk ordbok är en ordbok vars innehåll existerar i digital form och som görs tillgängligt genom olika typer av elektroniska format såsom internet och mobilapplikationer. Internetordböcker utgjordes inledningsvis av digitala versioner av redan existerande tryckta ordböcker. Numera utarbetas ofta ordböcker direkt för det elektroniska formatet och efter den senaste teknikens möjligheter. Den information som användaren är ute efter är dock densamma, men sättet att nå denna ett annat.

Metodikerna inom just internetforskningen handlar dels om att anpassa traditionella metoder till det nya mediet och finna nya sätt att använda dem på, dels om att utveckla nya metodologiska verktyg. Det elektroniska formatet möjliggör framför allt undersökning av faktisk (spontan) ordboksanvändning. Uppslagningar och andra aktiviteter på en webbsida kan registreras och sparas i s.k. loggfiler. Man kan på så sätt få tillgång till stora material, till skillnad från vid traditionella användarstudier av tryckta ordböcker där informanterna många gånger är färre och där brukarsituationen vanligtvis är mer eller mindre konstruerad. I jämförelse med traditionella arbetsätt är också den kraftigt minskade tidsåtgången vid datainsamling slående.

## 1.1 Syfte och frågeställningar

Lexin är alltså ursprungligen en serie tryckta ordböcker – till försäljning – som riktar sig till en mycket specifik målgrupp. Så småningom digitaliseras ordböckerna och görs fritt tillgängliga på internet. De blir därmed åtkomliga för en potentiellt mycket bredare användargrupp. Internetformatet ger möjlighet att få fram autentiska uppgifter om hur Lexin faktiskt används. Det är intressant oavsett om användarna tillhör målgruppen eller inte.

---

2 Termen *användarforskning* inkluderar här forskning om såväl användare som användning av ordböcker. Motsvarande gäller för termen *användarstudier*.

Undersökningen av LSL genomförs genom en kombination av metoder – loggfilsanalys och webbaserat frågeformulär.<sup>3</sup> LSL genomgick en omfattande revidering hösten 2008 till våren 2011. Materialet för undersökningen utgörs av loggfilsdata från 2008 och 2011 och enkätdata från 2007 och 2011, dvs. före och efter revideringen av ordboken. Materialet från 2011 är inte bara kvantitativt större utan innehåller också fler typer av information. Därför står 2011-materialet i fokus genom hela avhandlingen.

Syftet med undersökningen är trefaldigt, närmare bestämt att 1) undersöka användningen av LSL, 2) granska vilka det är som använder LSL samt 3) närstudera enskilda informanternas sökbeteende. De tre delsyftena bildar i sin tur utgångspunkt för tre delstudier.

Det första delsyftet är alltså att undersöka *användningen* av LSL. Den studeras med hjälp av loggfiler av två typer som dokumenterar dels lexikonuppslagningar (typ a), dels så kallade navigeringsval (klickningar) på webbsidan (typ b). Den senare typen återfinns endast i 2011-materialet. I delstudie 1 behandlas frågor som:

- Vilka uppslagningar är mest frekventa när det gäller söksträngar som ger träff i LSL? Vad kännetecknar dem?
- Vilka söksträngar ger inte träff i LSL? Vad är utmärkande för dem?
- Vilka likheter och skillnader uppvisar de två loggfilsmaterialen beträffande de två första frågorna? Verkar revideringen av LSL ha haft någon effekt på användningen av LSL?
- Vilka navigeringsval utnyttjas av användarna?

Det andra delsyftet är att undersöka *vilka det är som använder* LSL. Sammantaget har 1 158 informanter besvarat frågor om sig själva och sin användning av LSL. Delstudie 2 önskar besvara frågor som:

- Vilka lexikografiska behov ger informanterna uttryck för att de har?
- Används LSL av den åsyftade målgruppen?
- Vilka likheter och skillnader uppvisar de två uppsättningarna av enkätmaterial? Finns det olikheter som kan knytas till revideringen av LSL?

Slutligen är det tredje delsyftet att närstudera *enskilda informanternas sökbeteende*. Här ingår endast material från 2011. Ett övergripande mål med delstudie 3 är att pröva metoden att länka samman de två typerna av data, dvs. loggfilerna

---

3 Inte minst tack vare goda kontakter med Christian Mattsson, tidigare projektledare för Lexin, och Viggo Kann, webbansvarig för internetversionen av Lexin, fanns bra förutsättningar att rent praktiskt genomföra en undersökning av detta slag.

och enkäterna. Fokus läggs på de informanter som räknas till Lexins målgrupp. Undersökningen inriktar sig sedan på ett litet urval informanter och deras användning av LSL. I delstudie 3 behandlas följande frågeställningar:

- Vilka likheter och skillnader uppvisar de enskilda informanternas sökbeteende?
- Överensstämmer informanternas lexikografiska självbild med hur de faktiskt använder LSL?
- Kan den aktuella metoden generera relevanta och tillämpliga resultat i anknytning till praktiskt lexikografiskt arbete?

Det finns flera omständigheter som gör just LSL till ett lämpligt undersökningsobjekt. Lexin är ett unikt och omfattande projekt som har inspirerat liknande projekt i andra nordiska länder och som även har väckt internationell uppmärksamhet (Beijer & Fre Woldu 1997, Gellerstam 1999; Holmboe 1999, Bjørneset & Svavarsdóttir 2005, Hovdenak 2008; Pálfi & Tarp 2009; se även Hult et al. 2010). Projektet i sig är väldokumenterat men däremot har inte användningen av Lexin tidigare undersökts i någon större omfattning. Lexins internetversion är dessutom en frekvent använd webbsida och har så varit alltsedan 1995 då den lanserades. Under oktober månad 2015 utfördes t.ex. totalt nästan 37 miljoner lexikonuppslagningar på webbsidan. Drygt 1,6 miljoner uppslagningar gjordes i LSL vilket motsvarar omkring 4 %.

LSL har vidare utarbetats av professionella ordboksredaktörer (se vidare Toporowska Gronostaj 2005:90–91). Det innebär att innehållet presenteras på ett för målgruppen anpassat och genomtänkt sätt vilket gör det intressant att undersöka i vilken utsträckning och på vilket sätt ordboken används av just målgruppen.

Under 2008–2011 genomfördes som sagt en omfattande revidering av LSL och det faktum att jag har tillgång till data från både före och efter uppdateringen gör det möjligt att göra jämförelser av användningen av de två olika versionerna av LSL. När jag i det följande skriver LSL3 avses den icke uppdaterade versionen av ordboken. Den uppdaterade versionen kallas LSL4. Ges ingen uppgift om vilken version det rör sig om avses LSL i allmänhet, eller så framgår det av sammanhanget vilken version som åsyftas.

Ett övergripande mål med denna undersökning kan därmed sägas vara att ge ett relevant empiriskt bidrag till användarforskningen, dvs. den teoretiska lexikografien. Min förhoppning är också att undersökningen ska bidra med tillämpbar kunskap till den praktiska lexikografien.



## 1.2 Disposition

Avhandlingen omfattar nio kapitel. I det inledande kapitlet har undersökningens syfte och frågeställningar introducerats. I kapitel 2 presenteras avhandlingens teoretiska bakgrund. Kapitel 3 innehåller en allmän introduktion om ordinläring och ordböcker ur ett andraspråksperspektiv. Där beskrivs också projektet Lexin, uppdateringen av LSL samt internetversionen av densamma. Kapitel 4 redovisar metoder inom användarforskningen och resultat av tidigare studier. Kapitel 5 redogör för undersökningens metod och material. Därefter avhandlas undersökningens resultat i tre kapitel, 6–8. I kapitel 6 redogörs för användningen av LSL i form av lexikonuppslagningar och övriga aktiviteter på webbsidan. I det andra resultatkapitlet, kapitel 7, redogörs för informanternas enkätsvar. Kapitel 8 redovisar resultatet av länkningen av loggfilsdata och enkätdata utifrån några näranalyser. Kapitel 9 ger en sammanfattning av resultaten och avslutas med en framåtblick.



## 2. Allmän teoretisk bakgrund

I detta kapitel presenteras kort lexikografins vetenskapliga utveckling samt användarens alltmer framträdande roll inom denna.

### 2.1 Lexikografi – från hantverk till vetenskap

Lexikografen har två huvudriktningar – en praktisk och en teoretisk. Den praktiska delen rör själva utarbetandet av ordböcker. Den teoretiska delen, metalexikografen, inbegriper ”undersökning och utveckling av teorier rörande ordböckers tillkomst, egenskaper, ändamål och användning” (Svensén 2004:3). Längre likställdes lexikografi med just det praktiska hantverket och som sådan har den en månghundraårig tradition. En lexikograf var en person som skrev ordböcker – med största sannolikhet ett tillfredsställande arbete, dock icke utan utmaningar:

There can be no doubt that lexicography is a very difficult sphere of linguistic activity. Many lexicographers have given vent to their feelings in this respect. Perhaps the most colourful of these opinions based on a lexicographer's long experience is that of J. J. Scaliger (16th – 17th cent.) who says in fine Latin verses that the worst criminals should neither be executed nor sentenced to forced labour, but should be condemned to compile dictionaries, because all the tortures are included in this work. (Zgusta 1971:15.)

Ordböckerna var i regel en angelägenhet för förlagen, med undantag för historiska ordböcker som oftast utarbetades av akademiska institutioner – etymologisk forskning förutsatte språkhistoriska teorier och metoder. En historisk ordbok var en empirisk realisering av lexikologisk forskning. Inom historisk lexikografi har praktik och vetenskap på så sätt gått hand i hand under lång tid (Hannedóttir 1998, Gouws 2005) och fram till omkring 1970-talet var detta den huvudsakliga förbindelselänken mellan lexikografen och språkvetenskapen.

I den mån man över huvud taget tänkte teoretiskt inom den synkrona lexikografen ansågs lexikografi som helhet länge tillhöra allmän språkvetenskap, dvs. lingvistik, antingen som tillämpad sådan eller som en del av lexikologin. I slutet av 1960-talet börjar emellertid det praktiska lexikografiska arbetet med samtidsordböcker dokumenteras, systematiseras och, framför allt, diskuteras i vetenskapliga verk och artiklar. Relevanta teoretiska perspektiv hämtades naturligen från det lingvistiska och lexikologiska området. I takt med utvecklingen av en ordbokstypologi etableras lexikografi alltmer som en vetenskaplig disciplin, framför allt i Frankrike, dåvarande Tjeckoslovakien (Prag) och Tyskland (se vidare Gouws 2005). När Zgusta 1971 ger ut *Manual of lexicography*, ett verk skrivet på engelska och sålunda tillgängligt för en större och internationell publik, tar den metalexikografiska utvecklingen fart på riktigt.<sup>4</sup>

Inom loppet av ett decennium uppstår i Europa en metalexikografi som inte uteslutande handlar om historisk lexikografi eller om enskilda ordböcker och deras specifika problem, utan om teoribildning och metodologi (Gouws 2005:159). Zgusta visar att inte bara den historiska utan också den synkrona lingvistik har betydelse för lexikografen. Hans idéer har vidareutvecklats av ledande lexikografer, bl.a. i England (se avsnitt 2.2).

Lexikografens alltmer vetenskapliga karaktär bekräftas också av det faktum att det anordnas akademiska sammankomster i ämnet, t.ex. ett kollokvium över lexikografi i Köpenhamn 1976. En av de första lexikografiska konferenserna hålls i Exeter, England 1983. Den första nordiska konferensen genomförs 1991 och då bildas också Nordiska föreningen för lexikografi (NFL). Vidare utarbetas en lexikografisk terminologi och i slutet av 1980-talet publiceras en internationell handbok i lexikografi i tre band: *Wörterbücher/Dictionaries/Dictionnaires* (1989–1991), ett samlingsverk med huvudredaktörerna Franz Josef Hausmann, Oskar Reichmann, Herbert Ernst Wiegand och Ladislav Zgusta. Utgivningen av *Nordisk leksikografisk ordbok* (NLO 1997) är en del i utvecklingen av en lexikografisk terminologi i Norden. Som förlaga används den internationella handboken liksom Bo Svenséns *Handbok i lexik-*

---

4 Termen *metalexikografi* tillskrivs dock inte Zgusta utan en fransk lexikograf vid namn Terence R. Wooldridge (Gouws 2005:159).

ografi (1987). Därtill startas tidskrifter inriktade på lexikografi (*Lexicographica* 1984–, *International Journal of Lexicography* 1988–, *LexicoNordica* 1994–).

I samband med denna etablering av metalexikografi som vetenskaplig disciplin kom också ordboksanvändaren (med sina olika behov) i blickfånget. Användarforskningen växer sig sedan starkt framför allt under 1990-talet och framåt (se vidare kapitel 4). Dessutom leder införandet av ny teknik och datoriserade system till att det praktiska lexikografiska arbetet ges en tydligare struktur och, inte minst, att man kan utföra mer eller mindre avancerade sökningar i stora textsamlingar, alltså korpusar. Tillsammans bidrar detta till att lexikografin uppnår en alltmer självständig status. Lexikografins ökade vetenskaplighet innebär dock inte med automatik att det praktiska ordboksarbetet har någon självklar plats inom den akademiska sfären. Lexikografiskt arbete bedrivs oftast med hjälp av privat finansiering och/eller av kommersiella aktörer. Målet med praktiskt ordboksarbete är i regel en produkt som ska tillgodose kundernas behov och sälja (bra). Både de praktiska och de teoretiska lexikografiska insatserna görs dock i regel av lingvistiskt skolade akademiker vid det universitet som de är knutna till.

I det följande presenteras två grundläggande perspektiv på lexikografisk teori, en lingvistikorienterad inriktning och en autonom inriktning som ser lexikografin som självständig i förhållande till andra vetenskapsgrenar.

## 2.2 Lingvistikorienterad lexikografisk teori

Zgusta (1971) representerar en lingvistiskt orienterad lexikografisk inriktning med fokus på synkrona ordböcker. Enligt Gouws är verket ”one of the strongest influences in establishing lexicography as a theoretical discipline” (2012:458). Zgusta lägger stort fokus på grundläggande lingvistisk teori. Bokens första fyra kapitel avhandlar lexikal betydelse, ordbildning, ordkombinationer och språkvarieteter. Det första kapitlet upptar så mycket som en tredjedel av boken. De resterande fyra kapitlen inriktar sig på den lexikografiska produkten och fokuserar på ordbokstypologi, den enspråkiga ordboken, den tvåspråkiga ordboken och till sist planering och organisering av lexikografiskt arbete. Zgusta betonar lingvistikens roll som teoretisk utgångspunkt i formulerandet av lexikografiska problem. En ordboksartikel ska baseras på en grundlig lingvistisk analys. Zgusta lyfter också fram vikten av att beskriva språket som det används – med all dess variation. Sålunda anläggs ett mer deskriptivt än ett preskriptivt förhållningssätt.

Antalet teoretiskt inriktade lexikografiska artiklar ökar i kölvattnet av Zgusta (1971). Alltfler ordböcker uppvisar också uppenbara förbättringar

när det gäller det språkvetenskapliga innehållet (Gouws 2005:162). Inom framför allt brittisk lexikografi är Zgustas idéer fortfarande i hög grad aktuella. B. T. Sue Atkins och Michael Rundell är två ledande lexikografer i England, bland flera, som verkar i Zgustas anda. De anser inte att det finns ett behov av specifikt lexikografiska teorier utan att teoretisk lexikografi tar sin utgångspunkt i närliggande discipliner såsom lingvistik och informationsteknologi. I sin handbok *Oxford Guide to Practical Lexicography* (2008) fastställer de att "we do not believe that such a thing exists", dvs. en oberoende lexikografisk teori (Atkins & Rundell 2008:4). Också i Frankrike återfinns denna orientering. Béjoint (2010:381) fastslår t.ex.: "I simply do not believe that there exists a theory of lexicography, and I very much doubt that there can be one". Atkins & Rundell konstaterar också att "lexicographers are applied linguists" (2008:130). Oberoendet av specifikt lexikografiska teorier innebär dock inte att lexikografernas arbete är teorilöst. Snarare understryker författarna att de i allra högsta grad är behjälpta av teoretiska kunskaper:

To do the job well, we need to feel confident that our reliance on intuition and subjective judgments is kept to a minimum; that our approach is systematic, internally-consistent, and driven by what the language data is telling us; and the description we end up with is compatible with our observation of usage, and is complete, with nothing of importance omitted (either by accident or design). Given these needs, it would be foolish not to embrace any theories that may help us do the job [...]. (Rundell 2012:47.)

De relevanta teorierna återfinns inom det enorma lingvistiska forskningsfältet, framför allt lexikal semantik, kognitiv teori, pragmatik och korpuslingvistik. Framstående forskare inom dessa fält, som Atkins & Rundell ofta refererar till, är Charles Fillmore, Igor Mel'chuk, John Sinclair, Jurij Apresjan, George Lakoff och Douglas Biber. Deras fokus ligger på språket i sig och inte på lexikografiska spørsmål eller lösningar på dessa, men "they show us different ways of looking at language, which we can take and adopt to our needs" (Atkins & Rundell 2008:4). Sinclairs banbrytande arbeten inom korpuslingvistik bidrog t.ex. till att lexikografer kunde få statistiska underlag till sin kunskaps- och erfarenhetsbaserade intuition om ordens betydelse och användning (se Moon 2007). Att detta har varit ett framgångsrecept för den praktiska lexikografins vidkommande står utom allt tvivel:

[I]n the last three decades English-speaking (specifically British) lexicography – despite being seen as light on theory – has been among the most innovative areas of activity in the field, producing dictionaries which have broken new ground, been critically and commercially successful, and been widely imitated. (Rundell 2012:48.)

Därmed inte sagt att lexikografer alltid har omedelbar praktisk nytta av den stora lingvistiska teoribildningen. Det handlar också i hög grad om att läsa in sig på ämnet av "consciousness-raising value" om hur språk är strukturerat och fungerar. Rundell nämner som exempel John Searles *Speech acts* (1969), Eleanor Roschs arbeten om prototyp teorin (1977 m.fl.), George Lakoff & Mark Johnsons *Metaphors we live by* (1980) och D. Alan Cruses *Lexical semantics* (1986). Och ännu, menar Rundell (2012:83), har inte lexikografin dragit tillräcklig nytta av kunskaper inom ämnesområden som korpusdesign, andraspråksinläring, inlärningskorpusar och den senaste utvecklingen inom maskinell översättningsteknik.

Trots tilltron till det omfattande lingvistiska forskningsfältet så är det den praktiska användningen av slutprodukten, dvs. ordboken, som ger ett kvitto på hur väl lexikografin lyckats i sitt arbete. Det konstaterar redan Samuel Johnson i sin skrift om planen för sin berömda engelska ordbok:

The value of a work must be estimated by its use; it is not enough that a dictionary delights the critick, unless, at the same time, it instructs the learner; as it is to little purpose that an engine amuses the philosopher by the subtilty of its mechanism, if it requires so much knowledge in its application as to be of no advantage to the common workman. (Johnson 1747.)

Citatet visar att lexikografer i alla tider har haft användaren i åtanke i sitt praktiska arbete. Johnson betonar att det är användningen av ordboken och inte huruvida ordboken uppskattas av de lärda kritikerna som räknas. Längre fram i historien ser vi hur lingvister, tillika lexikografer, lyfter fram behovet av att användaranpassa ordböckers lingvistiska innehåll och det sätt varpå detta innehåll presenteras (t.ex. hos Householder 1962:279 och Hornby 1965:4). Zgusta (1971) pekar på att lexikografin måste ta hänsyn till ordboksanvändarens förutsättningar och behov: "[T]he general user must always be taken into consideration" (1971:17). Det är alltså den vanlige användaren, användaren i gemen, som åberopas.

Användaren spelar numera en avgörande roll inom den lingvistikorienterade lexikografin i såväl teoretiska som praktiska sammanhang. Atkins & Rundell anser att lexikografer, i allt som rör en ordboks innehåll och utseende, måste ta hänsyn till "who the users will be and what they will use the dictionary for" (2008:5). Användaren har en central roll vid planeringsstadiet av ett ordboksprojekt och det finns två sätt att närma sig denna, antingen genom så kallad användarprofilering eller genom användarforskning. Användarprofilering ska karakterisera den typiska användaren och dennas syfte med användningen samt vilka lingvistiska och lexikografiska förkunskaper hen besitter (Atkins & Rundell 2008:28–30). Användarforskning i sin tur avser alla de metoder som tas i anspråk för att ta reda på vad människor gör när de konsulterar sina

ordböcker. Används t.ex. ordboken för det åsyftade ändamålet? Täcker den upp användarnas behov av lexikografiska data? Uttolkar användaren innehållet såsom det är tänkt att uttolkas?

I slutet av 1990-talet genomför Atkins tillsammans med Krista Varantola en omfattande undersökning av ordboksanvändningen hos studenter med engelska som främmande språk (se Atkins & Varantola 1998a). Framstående forskare som Paul Bogaards och Hilary Nesi har under de senaste decennierna också bidragit med otaliga användarundersökningar vars resultat utgjort viktiga vetenskapliga underlag i arbetet med att utforma ordböcker på ett så användarvänligt sätt som möjligt (se t.ex. Bogaards 1990, 1991, 1998, Bogaards & van der Kloot 2001, Nesi 1996, 2000a,b, Nesi & Meara 1994, Nesi & Haill 2002). Användarfokuset kommer sedan att spela en central roll även i de autonoma lexikografiska teorier som utvecklas under 1980-talet och framåt (se vidare följande avsnitt, 2.3).

## 2.3 Autonom lexikografisk teori

Rundell med flera understryker som synes kopplingen mellan lexikografi och lingvistik. Men som redan framgått betraktar andra forskare lexikografin som autonom i förhållande till övriga vetenskaper. Från 1970-talet och framåt etableras den inriktning inom metalexikografen som betonar lexikografins självständighet gentemot omgivande discipliner, i synnerhet lingvistik. Teoretiska bidrag inom denna inriktning presenteras i det följande.

### 2.3.1 Teori utifrån ordbokens struktur

En av de framstående forskare från framför allt Tyskland och Danmark som under de senaste decennierna har önskat markera lexikografins självständighet gentemot lingvistik är Herbert Ernst Wiegand.

Wiegands tongivande artikel från 1977, "Nachdenken über Wörterbücher – Aktuelle Probleme", verkar föranledd av ett behov att ifrågasätta den samtida uppfattningen av vad en ordbok är, vilket språkligt innehåll den bör ha och vilken struktur som är den bästa för en potentiell ordboksanvändare. Vad gäller enspråkiga ordböcker hade dessa frågor inte varit föremål för lika mycket uppmärksamhet som vid tvåspråkiga, i synnerhet inte i fråga om användaren och hans behov. Nu gör denna användare ett definitivt inträde i den teoretiska lexikografen och följande flitigt citerade rader brukar få markera starten för den ökade uppmärksamheten:



Die Lexikographie muß nähere Bekanntschaft mit dem vorausgesetzten Benutzer suchen; er sollte nicht länger der bekannte Unbekannte bleiben. (Wiegand 1977:62.)

Lexikografen måste stifta närmare bekantskap med den förmodade användaren; han bör inte längre förbli den kända okända. (Min översättning.)

Wiegand lyfter inte minst fram behovet av lexikografisk empiri. Det saknas empiriskt grundade svar på frågor som: I vilka typer av situationer används enspråkiga ordböcker? När använder invandrare enspråkiga ordböcker? Används enspråkiga ordböcker till att söka svar på både språkfrågor och sakfrågor? (Wiegand 1977:61). Genom analyser av ordboksartiklar utformar Wiegand förslag på hur artiklarna strukturellt kan göras mer användaranpassade och användarvänliga, dels på makronivå, dvs. rörande lemmalistan, dels på mikronivå, dvs. rörande de lexikografiska data som ingår i själva ordboksartiklarna. Beroende på ordbokens funktion, som *receptionsordbok* (för att få hjälp att förstå) eller *produktionsordbok* (för att få hjälp att tala/skriva), ställs skilda strukturella krav vad gäller makro- och mikrostruktur.

Åren som följer utvecklar Wiegand ett utkast till en generell lexikografisk teori vilken bland annat inbegriper en taxonomi över lexikografiteorin. För den teoretiska lexikografins framväxt innebär det en ny milstolpe. Wiegand omarbetar sedan sina teoretiska perspektiv och 1998 utkommer första delen av tvåbandsverket *Wörterbuchforschung*. Denna del innehåller för det första en kartläggning av innebörden av termerna *ordboksforskning*, *metalexikografi* och *språklexikografi* och hur de förhåller sig till varandra. För det andra fokuserar Wiegand på forskning om ordboksanvändning. Han ger först en översikt över användarstudier från mitten av 1990-talet. Därefter sätter Wiegand in forskning om ordboksanvändning i ett aktionsteoriskt ramverk (Wiegand 1998:268–567). Detta går i huvudsak ut på att utforska användarens alla typer av potentiella ageranden i en användningssituation. Slutligen behandlas metodologiska perspektiv på forskning om ordboksanvändning. I detta första delband återfinns också en mycket rik terminologi för beskrivning av ordboksstruktur.

Wiegand önskar etablera ordboksforskning som en vetenskaplig disciplin och i förlängningen även få den att uppnå status av officiell akademisk disciplin. I slutet av 1990-talet anser Wiegand att ordboksforskningen ännu inte uppnått denna nivå utan snarare går att beskriva som ett vetenskapligt forskningsfält. Ordboksforskning har som centralt objekt *språklexikografi* (Wiegand 1998:63). Specifika forskningsobjekt kan vara tvåspråkig ordboksforskning eller användningen av engelska enspråkiga ordböcker (Smit 2001:300). Språklexikografi utgör hos Wiegand i sin tur en av tre underkategorier till den övergripande termen *metalexikografi*. De två övriga är *Sachlexikographie* (facklexikografi) och

*Allbuchlexikographie* (encyklopedier) (jfr avsnitt 2.3.3 om funktionslärares syn på språklexikografi). Fokuset på ordbokens struktur innebär en tydligare åtskillnad mellan lingvistik och lexikografi. Wiegand förnekar emellertid inte lingvistikens betydelse för lexikografien utan accepterar kunskapsmässiga och metodmässiga inflytanden av olika fält både inom och utom det lingvistiska fältet, i synnerhet lexikologi. Språkvetenskap bör alltså enligt Wiegand ses som ett – av flera – forskningsfält som har ett inflytande på lexikografien.

Ordbokens struktur å ena sidan och användarens behov, kunskaper och användningsbeteende å andra sidan är de två ben som en lexikografisk teori ska stå på. Fokuset på ordbokens struktur får också till följd att layoutfrågor undersöks vetenskapligt. En god layout förutsätter tydliga typografiska indikationer i syfte att optimera användarens förutsättningar att komma åt ordbokens innehåll. I linje med detta utvecklar Wiegand ett lexikografiskt metaspråk, kommentarspråk, i syfte att praktiskt förmedla information om ords semantiska och hierarkiska relationer. Kommentarspråket kan förmedlas med så kallade struktur-symboler. Hyperonymi- och hyponymiförhållanden markeras t.ex. med pilar.

Wiegand utgår vidare från att ordboken har två huvudsakliga funktioner, tillika två potentiella ordboksanvändningssituationer: att lösa (1) *konfliktrelaterade* lexikografiska problem eller (2) *icke-konfliktrelaterade* problem. De konfliktrelaterade lexikografiska problemen är av kommunikativ art och utgörs av problem av textreceptions-, textproduktions- eller översättningsunderstödjande slag. De icke-konfliktrelaterade lexikografiska problemen är icke-kommunikativa och rör behov av uppgifter i studiesyfte m.m. Det är sedan utifrån användarens typ av behov som ordbokens funktion bestäms. Den praktiska lexikografins uppgift är att, med hänsyn till användarens behov, utveckla språkliga uppslagsverk av alla sorter, där strukturen är så organiserad att användare kan komma åt det språkliga innehållet så bekvämt som möjligt. Istället för att utgå från den språkliga aktiviteten är det ordboksanvändarens behov som sätts i förgrunden.

I Norden ser man t.ex. tydliga spår av Wiegands tankar om ordboksstruktur i Svenséns *Handbok i lexikografi* (2004 [1987]), och terminologin i *Nordisk leksikografisk ordbok* (NLO, 1997) är i mångt och mycket wiegandsk.

### 2.3.2 Speciella teoretiska inriktningar med fokus på tvåspråkig lexikografi

Parallellt med Wiegands arbeten presenteras från 1970-talet och framåt andra modeller över ordbokstypologi och användarsituationer. Hausmann (1977) ställer upp en ordbokstypologi utifrån vilken funktion ordböckerna ska uppfylla. Han skiljer i första steget på användarens behov av lärande och använ-

dares behov av konsultation. Lärandeordböcker (i svensk tradition vanligen inlärningsordböcker) används för att utöka ordförrådet och behärsknin- gen av det, och syftet är att metodiskt studera ett främmande språk. Konsultationsordböcker används när ordboksanvändaren behöver lösa spe- cifika språkliga problem. Denna användarsituation delas upp i läsande och skrivande. Läsande delas i sin tur in i reception och översättning till moders- målet (L2–L1). De ordböcker som ska användas för att förstå text delar han slutligen in i ordböcker som har modersmålet som objektspråk och de som har ett främmande språk som objektspråk. Skrivande delas in i fritt skrivande och översättning till främmandespråket. Ordböcker för skrivande delar Hausmann sedan in i ordböcker för fri produktion och ordböcker för översättning till det främmande språket (L1–L2).

Två andra inflytelserika artiklar är Kromann et al. (1984a,b)<sup>5</sup> som re- dogör för ett helhetsperspektiv på lexikografi, liksom Hausmann med fokus på översättning och tvåspråkiga ordböcker. Kromann et al. sätter brukaren i centrum i sin så kallade aktiv/passiv-modell (produktion/reception; jfr Nation 2013:46–58 om produktiv och receptiv behärskning av ett ordförråd) på så sätt att de tar hänsyn även till användarens modersmål. I en av författarnas teser för tvåspråkig lexikografi fastställs att lemmaurvalet och redigeringsar- betet på såväl makro- som mikrostruktur ska styras av användarens kompetens och behov. Även tillhörigheten till ordbokstyp ska ha stor betydelse för lem- maurvalet. Indelningen i aktiva och passiva ordböcker ställer sedan skilda krav på utformning av och innehåll i de enskilda ordboksartiklarna. Det innebär att det för varje språkpar behövs fyra ordböcker för att tillfredsställa användarnas olika behov:

active dictionary: native -> foreign language  
Danish-English: for Danish users  
English-Danish: for English users  
passive dictionary: foreign -> native language  
English-Danish: for Danish users  
Danish-English: for English users  
(Kromann et al. 1984a:208.)

En dansk-engelsk ordbok för produktion (L1–L2) skiljer sig alltså så mycket från en dansk-engelsk ordbok för reception (L2–L1) att de två ändamålen inte kan kombineras i en och samma ordbok. Omvänt, en engelsk-dansk recep-

---

5 Kromann et al. publicerar först en kortare konferensartikel på engelska där aktiv/passiv- modellen lanseras. Senare samma år utkommer en utförlig beskrivning av modellen, men då på tyska.

tionsordbok (L2–L1) tillfredsställer inte kraven på en engelsk-dansk ordbok för produktion (L1–L2). För att ta ett konkret exempel: den svensk-ryska delen av Norstedts ryska ordbok har ett begränsat lemmaantal på 40 000 ord. Många svenska genomskinliga sammansättningar tas av den anledningen inte upp som lemman. Urvalet räcker för ryssar och deras receptionsbehov eftersom de i de flesta fall kan sluta sig till vad de genomskinliga sammansättningarna betyder. För svenskar som använder den svensk-ryska ordboken för att översätta svenska texter till ryska, dvs. har produktionsbehov, är urvalet däremot inte tillräckligt. De kan inte sluta sig till vad en genomskinlig svensk sammansättning heter på ryska (se Ljunggren & Malmgren 2007).

Sådana tankar om att ordböckers innehåll bör anpassas efter både användarens modersmål och efter deras typer av behov, receptiva eller produktiva, är emellertid inte nya. Intressant nog hänvisar Kromann et al. till en så tidig lexikograf som Ščerba som redan i mitten av 1930-talet lade fram ett förslag till en optimal lösning, ”an ideal one for enabling us to meet well-defined user needs” vad gäller ordböckers typologi (Kromann et al. 1984a:209). Brukarperspektivet är heller inte nytt utan har sitt ursprung i Wiegands (1977) resonemang om att användarens bruk av enspråkiga ordböcker för reception och produktion ställer olika krav på ordboksartiklarnas mikrostruktur. Kroman et al. tillämpar alltså detta på tvåspråkiga ordböcker.

En teoretisk modell som i praktiken förutsätter att man har fyra ordböcker till hands är å andra sidan inte vidare användarvänlig, syftet till trots. Av ekonomiska skäl kan aktiv/passiv-modellen också vara besvärlig att realisera, i synnerhet när det gäller tryckta ordböcker. Med tanke på att ordboksförlagens största konsumentgrupp är studenter, dvs. i regel personer med begränsade ekonomiska resurser, har förlagen svårt att se varför de ska avsätta pengar till att utveckla ordböcker som svårigen kan säljas. Dessutom ska ordboksanvändarna övertygas om att de har behov av, och således bör investera i, fyra ordböcker istället för en eller två. Förlagen har av förklarliga skäl alltså valt att satsa på multifunktionella ordböcker. I regel innebär det att man lägger störst energi på receptions- och produktionsbehov hos det mindre språkets användare av det större språket. Tvåspråkiga ordböcker mellan svenska och engelska är alltså framför allt anpassade för svenskars behov av reception och produktion på engelska. Bidirektionella ordböcker, dvs. ordböcker som i lika hög grad uppfyller både aktiv och passiv funktion, är ovanliga. Som en föredömlig bidirektionell ordbok har Valeriy Berkovs rysk-norska ordbok från 1987 lyfts fram (se Liberman 1990, Berkov 1996, Ljunggren & Malmgren 2007).

I den elektroniska ordbokseran vi nu lever i är förutsättningarna i mångt och mycket annorlunda, och bättre, och det finns tekniska (och kanske ekonomiskt försvarbara) möjligheter att åter fråga sig om aktiv/passiv-modellen är praktiskt realiserbar.

### 2.3.3 Funktionsläran

Wiegand vill med sin generella lexikografiska teori etablera lexikografi som en självständig akademisk disciplin. Han avfärdar dock inte betydelsen av närliggande discipliner, framför allt språkvetenskapen. I början av 1990-talet inleds emellertid arbetet på en teori som ännu kraftigare understryker lexikografins autonomi från lingvistik. Inom ramen för Århuskolan i Danmark och med Sven Tarp och Henning Bergenholtz i spetsen utformas *funktionsläran*, en teori om lexikografiska funktioner. Den utgår från de äldre idéer som just har beskrivits i 2.3.2. Den största källan till inspiration får emellertid Tarp och Bergenholtz från Wiegands arbeten, speciellt tankarna om en mer användaranpassad presentation av innehållet i en ordbok. Tarp och Bergenholtz utformar sin teori med större fokus på enspråkiga ordböcker och i centrum för teorin placeras användaren. Dessutom tonas skillnaden mellan språk- och saklexikografi ned.

I Tarp (2006, på engelska 2008) beskrivs funktionslärans bakgrund, födelse och utveckling. Tarps (2008b:81) definition av en lexikografisk funktion lyder: ”the satisfaction of the specific types of lexicographically relevant need that may arise in a specific type of potential user in a specific type of extra-lexicographical situation”. Ordbokens lexikografiska funktion definieras alltså utifrån användarens behov och inte utifrån vilken typ av språklig aktivitet som ska utföras. I fokus för funktionsteorin står de potentiella användarna (till skillnad från faktiska användare av kött och blod, jfr nedan) och de sociala situationer vari de ingår. Där uppstår lexikografiska behov. En potentiell ordboksanvändare är en person som har ett lexikografiskt relevant behov – t.ex. läser en bok och stöter på ett ord som hen inte förstår – men som ännu inte beslutat sig för att försöka lösa sitt problem med hjälp av en ordbok. Hen kan lika gärna fatta beslutet att fråga någon om råd, att googla ordet eller helt enkelt ignorera problemet och fortsätta läsa. Ordböcker är bruksföremål som ska tillfredsställa dessa, som Tarp kallar det, mänskliga behov. Tarp menar att den potentiella användaren av en ordbok hitintills endast har haft en blygsam roll i den lexikografiska teoribildningen.

Det finns två typer av användarsituationer och dessa ska avgöra ordbokens funktion. De (1) kommunikativa användarsituationerna uppstår i behovssituationer av reception, produktion och översättning men också vid korrektur och rättning. Här har alltså Tarp utvidgat de traditionella användarsituationerna med ytterligare två. De (2) kognitiva användarsituationerna uppstår när man är i behov av och har en önskan om (mer) kunskap om något. Uppdelningen har uppenbara likheter med Wiegands konfliktrelaterade och icke-konfliktrelaterade behov.

Det finns således många sätt att tillfredsställa lexikografiskt relevanta behov (varav ett är att vända sig till en ordbok). Det är alltså i dessa sociala situationer, som Tarp kallar extra-lexikografiska användarsituationer, som de verkliga lexikografiska behoven finns och som man i första hand måste få kunskap om. Dessa behov är funktionsrelaterade, medan den faktiska brukarens behov är bruksrelaterade. De senare är varken relevanta eller lika intressanta. Tarp åtskiljer vidare faktisk användare från potentiell användare respektive faktisk ordbokskonsultation från extra-lexikografisk situation. Användarfokuset till trots är Tarp inte särskilt intresserad av vad användarstudier kan avslöja om faktiska användare och de faktiska ordboksanvändningssituationer de befinner sig i. Användarstudier ger endast svar på subjektiva, användningsrelaterade behov, dvs. de behov användarna själva anser sig ha. Enligt Tarp är detta helt irrelevant och oanvändbar information. Det är de objektiva behoven, behov som uppstår hos en potentiell användare i en potentiell användarsituation, som ger relevant information till lexikografen i dennas arbete att utforma en användaranpassad ordbok. Eftersom användaren sällan är medveten om sitt eget bästa, dvs. sina objektiva behov, är det lexikografens ansvar att uttolka dessa och sedan tillfredsställa dem i ordboken.

Från att teorimässigt har stått varandra ganska nära har Wiegand och Tarp/Bergenholtz sedermera alltmer divergerat. Enighet råder gällande vissa grundläggande lexikografiska antaganden: (1) lexikografen är ett autonomt vetenskapligt forskningsfält och inte en gren av lingvistik och (2) ordböcker utgör lexikografins forskningsområde och de ska ses som bruksvaror som framställs för ett bestämt ändamål, nämligen att tillfredsställa mänskliga behov. I flera avseenden skiljer sig dock funktionsläran och Wiegands generella lexikografiska teori åt. Tarp/Bergenholtzs huvudsakliga invändning mot Wiegands lexikografiska teori är dess lingvistiska premisser. Funktionsläran har över huvud taget inga beröringspunkter med lingvistik, i varje fall inte närmare än med andra vetenskaper. Det är inte den språkliga aktiviteten utan den potentiella användaren som står i fokus. Brukaren använder det referensverk som finns tillgängligt och baserar inte sitt val på lingvistisk grund och därför är det inte meningsfullt att skilja på språk- och saklexikografi. Denna distinktion verkar fortfarande vara central för Wiegand som ju endast behandlar språklexikografi. Dessutom inför han en tredje underkategori inom metalexikografen, nämligen encyklopedisk lexikografi. Trots detta säger han sig sätta användarens behov före den språkliga aktiviteten.

Det är viktigt att understryka att Tarp och Bergenholtz båda har en bakgrund som facklexikografer. Det är då inte så förvånande att de har utvecklat en generell lexikografisk teori där språk och sak inte skiljs åt. I en fackordbok om genteknologi t.ex. är genteknologin långt viktigare än språkvetenskapen. En annan skillnad mellan Wiegand och Tarp/Bergenholtz är att Wiegand

kartlägger användarens behov i närmare anslutning till handlingen att använda en ordbok medan Tarp/Bergenholtz säger sig utgå från en användarsituation där beslutet att använda sig av just en ordbok för att lösa sitt lexikografiskt relevanta behov ännu inte har tagits. Det råder vidare oenighet kring hur centrala lexikografiska termer ska definieras (jfr Wiegand 2001 och Bergenholtz & Tarp 2003).

Förutom en beskrivning av funktionsläran som en generell lexikografisk teori ger Tarp (2008b) exempel på hur en generell teori för inlärningsordböcker bör utformas. Den följs dessutom upp av en specifik teori för inlärningsordböcker. I LSL, som ju är en inlärningsordbok, kan man se spår av funktions-teorin (se vidare kapitel 4). Den har t.ex. också satt avtryck i senaste upplagan av *Svenska Akademiens ordlista*, SAOL 14 (2015:IX, XV, XXXIII).

Funktionsteorin har avnämare världen över, t.ex. i Brasilien (Herbert Andreas Welker), Spanien (Pedro Fuertes Olivera), Indonesien (Deny Kwary) och Sydafrika (Rufus Gouws).

## 2.4 Diskussion

Frågan huruvida den praktiska lexikografin är i behov av specifikt lexikografiska teorier eller om andra vetenskaper, framför allt lingvistik, kan bidra med tillräcklig och relevant vetenskaplighet är ingalunda slutdiskuterad. Funktionsteorins teoretiska anspråk har ifrågasatts från flera håll. Den huvudsakliga kritiken handlar om avsaknaden av empiriska data (Farø & Gottlieb 2007:194, Piotrowski 2009:485, Bogaards 2010:316). En mer positiv recension är Lew (2008) som diskuterar funktionsteorin i förhållande till inlärningsordböcker och, med vissa reservationer, betraktar Tarp (2006, 2008b) som ett mycket viktigt bidrag i utvecklingen av en lexikografisk teori. Rundell (2012) ifrågasätter på ett mer generellt plan den konkreta nyttan med specifikt lexikografiska teorier. I sitt praktiska lexikografiska arbete finner Rundell ingen hjälp när han söker sig till teoretiska skrifter av lexikografer som Ščerba, Wiegand, Tarp och Bergenholtz. Vad han finner är ibland intressanta diskussioner och iakttagelser, men sällan något nytt, och inte något som kan kallas lexikografisk teori. Det han återkommande stöter på är en tendens att problematisera, att se svårigheter och komplexitet där det inte existerar: "The emphasis is less on 'how should we approach the task of defining?' [...] and more on 'what name should we give to this or that element in the entry?'" (Rundell 2012:54). Wiegand och hans efterföljare verkar befinna sig i ett parallellt universum fortsätter Rundell, i vilket människor som inte är involverade i praktiskt ordboksarbete utformar teoretiska modeller om hur man framställer ord-

böcker. Rundell lyfter också fram vad funktionsteorin *inte* behandlar, nämligen vilket lingvistiskt material som ligger till grund för alla ordböcker och hur detta har tagits fram och analyserats.

Är då lexikografi att likställa med tillämpad lingvistik eller är det en vetenskap i sin egen rätt? Inom framför allt brittisk lexikografi ser man som sagt inte poängen med att i lexikografiska sammanhang så strikt särskilja lexikografi från lingvistik. För anhängare av den motsatta uppfattningen, dvs. att specifika lexikografiska teorier existerar och behövs, med företrädare som Wiegand, Bergenholtz och Tarp, kan denna till synes nonchalanta inställning måhända provocera. De pläderar för att lexikografin i egenskap av självständig disciplin är i behov av specifika lexikografiska teorier. Enligt dessa autonoma lexikografiska teorier bestäms en ordboks funktion på basis av en profil av potentiella typer av användare och typer av användarsituationer. Men är det inte just detta som seriöst praktiserande lexikografer av idag gör, alldeles oavsett om de är anhängare av funktionsläran eller inte? Lorentzen & Trap-Jensen kallar detta arbetssätt helt enkelt för ”sund fornuft” (2006:200). Den grundläggande motsättningen i hur man ser på språklexikografi och facklexikografi har säkert sin förklaring i att för Tarp och Århuskolan har den huvudsakliga lexikografiska verksamheten länge ägnats åt facklexikografi, medan Rundell och hans kollegor inom brittisk lexikografi haft mycket lite att göra med just facklexikografi.

Under de senaste årtiondena har ordbokens funktion och det användarorienterade perspektivet kommit att prioriteras inom både praktisk och teoretisk lexikografi. Vi kan här se hur den praktiska och teoretiska lexikografin drar åt samma håll. För den mer praktiskt orienterade brittiska traditionen märks det inte minst i den redan omnämnda *The Oxford Guide to Practical Lexicography* (Atkins & Rundell 2008). Tarp har i senare skrifter också börjat anta ett mer avslappnat förhållningssätt till lexikografiskt närliggande discipliner: Lexikografi är ”a discipline particularly characterized by its big interdisciplinary vocation and its cooperative and integrating nature” (Tarp 2010:461). Man får annars lätt uppfattningen att man som ny medlem i det lexikografiska samfundet måste välja sida. Även Gouws (2005) öppnar upp för ett mer ömsesidigt och pragmatiskt förhållningssätt när det gäller den metalexikografiska teoribildningen.

Alle können voneinander lernen. Dies wird die lexikographische Praxis ebenso wie die metalexikographische Theorie und die Wörterbuchforschung fördern. (Gouws 2005:172.).

Alla kan lära av varandra. Det främjar såväl den praktiska lexikografin som den metalexikografiska teorin och ordboksforskningen. (Min översättning.)



Och, som t.ex. påpekas av lexikografen och användarforskaren Bogaards, ”dependence on other sciences does not deny an independent status of the field of lexicography” (Bogaards 2010:318).

Enligt min mening råder det ingen tvekan om att grundläggande lingvistik är själva utgångspunkten för all (språk)lexikografi, t.ex. vad gäller Lexinordböckerna. Den lingvistikorienterade lexikografiska teorin har på så sätt stor relevans för föreliggande undersökning. Det hindrar däremot inte att även tankarna bakom de mer autonoma lexikografiska teorierna som har beskrivits här är av stor betydelse. Det inbegriper t.ex. specifik lexikografisk kompetens om ordbokstypologi och en enhetlig lexikografisk terminologi att luta sig mot. Därjämte naturligtvis teoretiska perspektiv som utgår ifrån användaren och bruket av ordboken.



### 3. Ordinläring, inlärningsordböcker och Lexin

I föregående kapitel används frekvent ordet *ordbok*, dock utan att vidare förklaras. *Nordisk leksikografisk ordbok* (NLO) ger följande definition:

[L]eksikografisk oppslagsverk som inneholder et utvalg av et språks eller en språkvarietetets ord, orddeler og ordkombinasjoner, enten med ordforklaringer på samme språk eller med ekvivalenter på ett eller flere andre språk. (NLO 1997:197.)

Ordböcker kan sedan grupperas efter olika kriterier (se NLO 1997:21–24, jfr Svensén 2004:26–47 och Fjeld & Vikør 2008:134–140). Vi har till att börja med allmänordböcker. De kan i sin tur t.ex. indelas efter förhållandet till världen. Man brukar då avse huruvida de bara ger information om språket (språkordbok), om världen (sakordbok) eller både om språket och om världen (encyklopedisk ordbok eller allordbok). Indelningen känns igen från kapitel 2. Ordböcker kan också kategoriseras efter sitt förhållande till tiden (synkron eller diakron, historisk eller samtida), förhållande till den språkliga normen (deskriptiv eller normativ), eller på grundval av antalet språk (ensspråkig, tvåspråkig, flerspråkig ordbok). Ordböcker kan vidare grupperas efter användning, funktion eller konstruktion. Här skiljer NLO t.ex. mellan konsultationsordböcker och ordböcker för läsning (jfr Hausmann 1977) och mellan receptionsordböcker och produktionsordböcker. I förbindelse med funktion kan ordböcker klassificeras efter om de är mono-, bi- eller polyfunktionella och efter vem den tilltänkte användaren är, exempelvis modersmålstalare eller främman-

despråkstalare. Det finns dessutom ordböcker som avpassas för specifika ändamål, t.ex. inlärningsordböcker. NLO använder termen *pedagogisk ordbok* och definitionen lyder: ”ordbok som er särskilt utformat för visse brukargrupper som hjälpemiddel ved språkinnlæring” (NLO 1997:214). Vid sidan av allmänordböcker finns specialordböcker. Några exempel på sådana är: konstruktionsordbok, synonymordbok, idiomordbok, förkortningsordbok, rättstavningsordbok, fackordbok, slangordbok och frekvensordbok.

När det gäller elektroniska ordböcker kan de också delas in efter sina olika egenskaper. Några viktiga distinktioner är: digitaliserade vs ursprungligt elektroniska ordböcker, offline- vs onlineordböcker, gratis online-ordböcker vs avgiftsbelagda online-ordböcker och statiska vs dynamiska ordböcker (Bergenholtz & Malmgren 2005:3).

Den ordbok vars användning undersöks i föreliggande avhandling, *Lexins svenska lexikon*, är som redan framgått en inlärningsordbok för relativt avancerade nybörjare. Den är dessutom enspråkig, synkron och i första hand en receptionsordbok. LSL är vidare ytterst en språkordbok men den har också inslag av sakinformation och encyklopedisk information (se vidare avsnitt 3.2.2). Vad gäller formatet är LSL en fritt tillgänglig statisk online-ordbok. Den är ursprungligen en digitaliserad ordbok och har i takt med den tekniska utvecklingen utökats med allt fler sökfunktioner. Nesi (2012:363–364) kategoriserar vidare Lexin som en experimentell elektronisk ordbok mot bakgrund av dess fokus på att möta de lexikografiska behoven hos immigranter i Sverige. Hon understryker att den här typen av ordbok är utvecklad för forskning och utbildning snarare än för kommersiell vinning.

Ordböcker är eminenta hjälpmedel för den som vill lära sig (något om) ord. Vid ordinlärning på ett annat språk än modersmålet utgör ordböcker särskilt ovärderliga verktyg. I det här kapitlet riktas först uppmärksamheten mot ordinlärning och inlärningsordböcker ur ett andraspråksperspektiv. Därefter presenteras projektet Lexin, den svenska ordbasen samt den senaste uppdateringen av densamma, dvs. *Lexins svenska lexikon 4*.

## 3.1 Ordinlärning och ordböcker

### 3.1.1 Ordinlärning ur ett andraspråksperspektiv

Andraspråksinlärning och andraspråksanvändning etablerade sig som ett internationellt forskningsområde i slutet av 1960-talet. Det är numera ett mångfacetterat vetenskapsfält med många inriktningar. En inriktning handlar om

förhållandet mellan inläraren och (inlärningen av) andraspråket och vad som är ”karakteristiskt för andraspråksanvändning på olika stadier av språkbehärskning och i givna situationer” (Hyltenstam & Lindberg 2013:8). En annan orientering har ett språkpolitiskt intresse som t.ex. kan handla om resursfördelning, attityder och ideologier i förhållande till andraspråksinlärare och andraspråksanvändare. En tredje inriktning har ett pedagogiskt och didaktiskt fokus där forskare är intresserade av hur andraspråksinlärning kan stödjas genom olika typer av undervisning samt ”relationen mellan undervisning, språkutveckling, identitetsutveckling och kunskapsutveckling hos flerspråkiga elever i allmänhet” (Hyltenstam & Lindberg 2013:8). Etableringen som internationellt forskningsområde har befästs i omfattande referensverk som t.ex. Doughty & Long (2003), Ritchie & Batia (2009) och Gass & Mackey (2011).

Andraspråksforskningen i Sverige går tillbaka till början av 1970-talet och växer fram i nära kontakt med den internationella forskningen. I likhet med vad som är fallet med den internationella forskningen stimuleras det vetenskapliga arbetet initialt särskilt av de behov av kunskap som finns inom undervisningssektorn.

Inom andraspråksforskningen skiljer man mellan *förstaspråk* och *andraspråk*. Ett förstaspråk är det språk man lär sig först, och ett andraspråk är ”vilket språk som helst som tillägnats *efter* det att förstaspråket etablerats eller börjat etableras” (Abrahamsson 2009:13). Termerna syftar sålunda på den ordningsföljd i vilken en individ exponerats för språken, inte på ”behärskningsgraden eller ens på hur mycket eller i vilken ordning en talare identifierar sig med sina olika språk” (a.a s. 14). Förstaspråk likställs här med *modersmål*. Det behöver dock inte betyda att detta är individens bästa språk (jfr avsnitt 7.3.4). Det är t.ex. i immigrationssammanhang vanligt att i synnerhet barn till slut behärskar andraspråket bättre än förstaspråket, då andraspråket är majoritetsspråket i det samhälle de växer upp i. I inlärningssammanhang brukar andraspråket även kallas *målspråk*, dvs. det språkliga system som är målet för inlärningen. *Målspråk* är en välkänd beteckning också inom den lexikografiska terminologin och då i betydelsen det språk som källspråkets lexikaliska enheter översätts till i en två- eller flerspråkig ordbok (NLO 1997:187). Den version av målspråket som en inlärare har under inlärningsperioden kallas för *interimspråk* eller *inlärarespråk* och är enklare, mindre stabil och mer föränderlig än målspråket (se vidare Hammarberg 2013:29–32 och där angivna referenser).

En viktig distinktion inom andraspråksforskningen är också den mellan *andraspråk* och *främmandespråk*. Ett andraspråk lärs in i den miljö där det används och där det utgör huvudsakligt kommunikationsspråk, medan ett främmandespråk lärs in i en miljö där språket *inte* används som huvudsakligt kommunikationsspråk (Abrahamsson 2009:15). Den långsiktiga målsättningen för en andraspråksinlärare är att ”uppnå mycket goda språkfärdigheter med

ett stort och varierat ordförråd” men det finns också ett kortsiktigt behov av att snarast möjligt kunna använda språket för att kommunicera med sin omgivning (Enström 1996:11). En främmandespråksinlärare har inte detta omedelbara behov av kommunikation.

Språkinläring sker i regel på två sätt, dels metodiskt och strukturerat i ett formellt undervisningssammanhang (*språkinläring*), dels spontant i mer eller mindre naturliga omgivningar (*språktillägnande*). För en infödd talare är språkinläringen en relativt autonom process. Att lära sig ett andra språk utöver modersmålet kräver mycket mer aktiva inlärningsstrategier. Språkutvecklingen i ett andraspråk kan vidare påverkas av faktorer som ”ålder, inflytande från andra språk, individuella och sociolingvistiska skillnader samt motivation och attityder” (Enström 1996:10, se även Boyd & Fraurud 2010). Det kan medföra att olika individer lär sig olika mycket och med olika hastighet.

Att lära sig ett språk handlar till stor del om att lära sig ord. Språkinläring omfattar emellertid både inläring av lexikonet och inläring av regler (fonologiska, morfologiska, syntaktiska m.m.) för hur orden ska hanteras. Det är främst i undervisningssammanhang och i synnerhet i undervisning av främmande språk som dessa två delar brukar skiljas åt. Vid utvecklingen av modersmålet är inläringen av lexikon och grammatik en sammanflätad process. Inlärningsprocessen för ett andraspråk är dels tudelad (inom ramen för den formella undervisningen), dels sammanflätad (i målspråksmiljön). I det följande fokuserar jag på inläring av ord.

En språkbrukares kunskaper om ords egenskaper och användning brukar kallas *lexikal kompetens* (Enström 1996:49, 2013b:181). Utmärkande för en språkbrukare på eller nära modersmålsnivå är att hen kan många ord (kvantitet), vet mycket om de enskilda orden (kvalitet) och kan använda dem i kommunikation (Henriksen 1995:13). Stroud (1979) specificerar vad denna kompetens innefattar, dvs. en språkbrukares kunskap om orden i sitt modersmål (jfr Järborg 2007:81–86):

- (1) Ordens fonologiska uppbyggnad, dvs. deras ljudbilder.
- (2) Ordens morfologiska uppbyggnad, dvs. ordbildnings- och ordböjningsegenskaper.
- (3) Ordens kategoritillhörighet, dvs. vilken ordklass ett givet ord hör till.
- (4) Ordens syntaktiska egenskaper, dvs. vilka krav på frasens eller satsens övriga delar ett givet ord ställer.

- (5) Ordens betydelsemässiga egenskaper. Denna punkt innefattar:
- (a) Ordens semantiska uppbyggnad, dvs. vilka betydelsekomponenter som sammanfattas i ett givet ord.
  - (b) Betydelserelationer mellan ord, t.ex. vilka motsatstyper som förekommer och vad det innebär att ett ords betydelse är överordnad ett annat ords betydelse.
  - (c) Betydelserelationer inom ord, dvs. hur de betydelser som ett och samma ord kan ha är relaterade till varandra.
  - (d) Ordens bibetydelser.
  - (e) Ordens selektionsrestriktioner, dvs. vilka betydelser de ord får ha som ett givet ord kan kombineras med.

(Stroud 1979:174–175)

Lexikal kompetens omfattar såväl receptiv/passiv som produktiv/aktiv behärskning av ett ordförråd (jfr Kromann et al. 1984a och aktiv/passiv-modellen ur ett lexikografiskt perspektiv, avsnitt 2.3.2). Den förra kompetensen har med ordförståelse att göra och den senare med ordinläring och ordanvändning. Nation (2013:49) tar fasta på denna distinktion i sin uppställning av den lexikala kompetensens beståndsdelar. Han presenterar en modell som visar hur den lexikala kompetensen inkluderar både form, betydelse och användning som alla i sin tur kan relateras till såväl receptiv som produktiv kunskap (se även Lindberg 2007:42–43). I anslutning till form så är t.ex. receptiv kunskap det att känna igen ett ord när man hör det. Att veta hur ordet uttalas är däremot produktiv kunskap. I receptiva sammanhang utgår man från ordets form för att hitta fram till dess betydelse. I produktiva sammanhang är det den givna betydelsen som leder fram till ett lämpligt ord (Enström 1996:52).

Det receptiva ordförrådet är alltid större än det produktiva. Den naturliga förklaringen är att det produktiva ordförrådet förutsätter djupare kunskap om ord, vilket tar längre tid att uppnå. Ordinlärningsprocessen kan enligt Lindberg (2007:39) beskrivas som ”ett kontinuum från enbart igenkänning [...] till kunskap om alla aspekter av ett ords betydelse och användning i egen produktion” (se vidare Melka 1997). Måhända är partiell kunskap om ords betydelse och användning ”normen för många av de ord som inte tillhör de allra mest frekventa”, i synnerhet för andraspråksinlärare (Enström 2013b:176, se även Laufer 1997b:141–142).

När det talas om storleken på en individs ordförråd är det följaktligen skillnad på det receptiva ordförrådets storlek och det produktiva. I det här sammanhanget fokuserar vi på det förra. Laufer (1997a:24) hävdar att för engelskans del täcker ett passivt ordförråd på 5 000 lexikala enheter in i genomsnitt 90–95 % av orden i en text, vilken som helst. Men frågan är om det räcker för

att förstå innebörden av de resterande 5–10 % av orden. Nation (2013:206–207) anger en något högre andel ord, 98 %, som bör vara känd för att man ska kunna gissa sig till de okända orden och således kunna ta till sig hela innehållet i en given text. Det innebär att texten får innehålla högst ett okänt ord på 50 kända ord för att gissningsstrategierna ska fungera optimalt.

I det här sammanhanget aktualiseras frågan om så kallade ”svåra ord”. Det är enligt Enström ett faktum att vissa ord är svårare att lära sig och att komma ihåg än andra. Svårigheten kan ha flera förklaringar.<sup>6</sup> I Frick & Malmström (1983:16–20) urskiljs t.ex. följande typer av orsaker hos svenska modersmålstalare:

- 1) feltolkningar styrda av ljudlikhet (*egalt* förväxlas med *legalt*, *myt* förväxlas med *muta*)
- 2) förväxling med ord som har motsatt betydelse (t.ex. *disharmoni/harmoni*, *emigrant/immigrant*)
- 3) annan förväxling med ord som berör samma område (*komfort/lyx*, *modifiera/förklara*)
- 4) svårigheter att hålla isär olika betydelser av ett och samma ord (*deklarera åsikter/deklarera inkomster*)
- 5) svårigheter vid ordsammansättningar (*affärsbank/bank i affären*)

Det är emellertid viktigt att påpeka att en omständighet som i vissa fall stjälp ordförståelse i andra fall kan stödja och underlätta förståelsen av ord. Ljudlikhet är ett sådant exempel. Ordet *stabil* förknippas t.ex. helt riktigt med *stadig* och likaså *intern* med *inom* (Frick & Malmström 1983:160).

En tidig studie är också Josephson (1982) som visar att samhällsord, dvs. ord som förekommer i politiska, ekonomiska och samhällsrelaterade sammanhang, har visat sig vara svåra. Vidare har Ljungman (2012) undersökt vad som är typiskt för de lexikala enheter i Högscoleprovets orddel som flest provdeltagare svarar fel på. Fornnordiskt/germanskt ursprung, ålderdomlig prägel och låg frekvens är tre utmärkande drag hos de typiskt svåra orden.

De tre nämnda studierna fokuserar på ordförståelsen hos modersmålstalare av svenska. Inom andraspråksforskningen har man också försökt bringa klarhet i vad som utmärker just svåra (typer av) ord. De ord som orsakar feltolkningar

---

6 Frick & Malmström (1983) undersöker förståelsen av vad de beskriver som ”ganska lätta” och ”medelsvåra” ord som har med det sociala, politiska och ekonomiska livet att göra. Författarna är alltså i första hand intresserade av att undersöka informanternas passiva ordförråd och inte det aktiva. Informanterna utgörs av personer som ”tillsammans kan antas representera ett slags genomsnitt av svenska folket” (s. 8–9).



hos förstaspråkstalare är nämligen inte automatiskt av samma slag som de ord som ställer till problem för andraspråkstalare. I en norsk undersökning av ordförråd i läromedel (Golden & Hvenekilde 1983) framkommer t.ex. att ord som förknippas med olika skolämnena utan att vara egentliga facktermer vållar störst problem för andraspråkstalare, och inte de rena facktermerna som man kanske skulle kunna tro. Två exempel är *härskare* i historia och *gnida* i fysik. En förklaring till detta är att lärare per automatik har ett förstaspråksperspektiv och därför inte uppfattar de här orden som problematiska, dvs. okända eller svåra. Följden blir att icke-fackord, i motsats till fackord, inte förklaras (Golden 2006:117). Liknande slutsatser dras i det svenska projektet *Ord i läroböcker*, *OrdiL* (Lindberg & Johansson Kokkinakkis 2007).

Enström (2013b:183–192) diskuterar ords svårighetsgrad utifrån ett andraspråksperspektiv och sätter fokus på följande fyra egenskaper: frekvens, uttryckssida, innehållssida och kontext (efter Josephson 1982, se även Enström 1996:54–60). Frekventa ord som man ofta möter i sin vardag kan å ena sidan vara lätta att lära sig. Å andra sidan finns det ord som är mycket frekventa men som ändå kan orsaka stora svårigheter i både receptiva och produktiva sammanhang. En anledning kan vara att orden i sig själva är mångtydiga, en annan att orden kan ingå i många ordkombinationer där betydelsen kan ligga långt från det enskilda ordets betydelse. Högfrekventa verb är t.ex. i allmänhet en kategori ord som många gånger är svår att lära sig, t.ex. *göra*, *stå*, *ta*, *ha*, *hålla* och *dra* (Enström 2013b:184, jfr Viberg 2004). Verbet utgör kärnan i satsen och ”avgör såväl hur satsen ska konstrueras, vilka satsdelar en sats får innehålla, som vilket betydelsemässigt innehåll de omgivande satsdelarna kan ha, dvs. vilka betydelser som kan kombineras med varandra” (Enström 1996:59). Den betydelsemässiga avgränsningen mellan olika verb kan också vara problematisk eftersom de flesta verb har en grundbetydelse som kan modifieras genom tillägg av partiklar och prefix (se längre fram i detta avsnitt).

Svårigheter på uttryckssidan behandlas av Laufer (1997b:140–155), t.ex. uttal, ortografi, längd, morfologi samt böjnings- och avledningsrelaterad komplexitet. Exempelvis lärs ord vars fonem känns igen och som är lätta att uttala in snabbare än ord som har främmande uttal. Ord kan också vara svåra om de har ortografiska likheter med ett annat och kanske mer bekant ord, speciellt om orden avviker från varandra suffixmässigt (t.ex. eng. *comprehensive–comprehensible*). Ord kan också vara besvärliga när konsonanterna är samma men vokalerna är olika, eller tvärtom. Några svenska exempel är det redan nämnda ordparet *egal–legal* samt *förrätta–förorätta* och *rassla–prassla*. Vidare kan långa ord vara svårare att lära sig än korta, men bevisen för detta är allt annat än entydiga (jfr t.ex. de engelska ordparen *available–unavailable* och *bun–interdisciplinary*, se vidare Laufer 1997a). Att kunna identifiera ords ingående morfem kan förstås underlätta inläringen av dem. Men den här strategin kan också

vara vilseledande. Innebörden av morfemet *over* i ordet *overthrow* ('störta, kasta omkull') är t.ex. en helt annan än dess betydelse i ordet *overcook* ('överkoka'). Slutligen kan avledningar, speciellt om de innehåller svårtolkbara prefix eller suffix, vara besvärliga att förstå. Verbet *gå* ingår t.ex. i många prefixavledningar där betydelsen är alltifrån självklar, t.ex. *avgå, frångå, ingå* och *vidgå*.

Laufer (1997b) tar också upp innehållsmässiga egenskaper hos ord som kan bidra med inlärningssvårigheter, nämligen abstrakthet, användningsrestriktioner, idiomacitet och polysemi. Abstrakta ord är i allmänhet svårare att lära sig än konkreta ord (se vidare Enström 2013b:185). Universella abstrakta ord kan emellertid vara lättare att förstå och lära sig än mer språkspecifika konkreta ord. Med användningsrestriktioner avses t.ex. stilnivå och värdeladdning. I vilka sammanhang det är lämpligt att använda ett visst ord och när det inte alls passar kan vara en svår nöt att knäcka för en andraspråklärare. Till synes synonyma ord, där många men inte alla betydelsekomponenter sammanfaller, kan t.ex. vara utbytbara i ett sammanhang men inte i ett annat.

En viktig del av ordinlärningen består också av att förstå och lära sig vilka ord som kan kombineras med varandra och som genererar idiomatisk svenska. Inlärare kan ha speciellt svårt att identifiera flerordskombinationer i en text (se Prentice & Sköldbberg 2013), inte minst partikelverb. Möjligheterna att kombinera verb och partikel är stora och kan antingen generera en från det enkla verbet avvikande eller specificerande betydelse. Ett exempel på det förra är *komma* i förhållande till *omkomma*. En och samma partikel kan också ha flera olika betydelser (jfr *over* ovan). Vissa av dem kan dessutom användas både som lös och fast sammansättning, ibland med helt eller delvis skilda betydelser (jfr *avbryta–bryta av, utkomma–komma ut*).

Den potentiella svårigheten med polysemi blir uppenbar givet det faktum att uppemot 60 % av alla svenska ord har mer än en betydelse (Enström 2013b:187, Nagy 1997:64 anger en andel på 40 % för engelskans del). Detta kan ge upphov till skenbar genomskinlighet, dvs. man tror sig förstå ett ord men i själva verket har man endast en partiell kunskap om ordet och dess betydelse, användning och kombinationsmöjligheter med andra ord (se vidare Laufer 1997a:25–27). I många fall är det dessutom fråga om metaforiska användningar, vilka kan skilja sig mycket åt mellan språk och därför orsaka större inlärningssvårigheter för andraspråkstalare än för förstaspråkstalare. Slutligen verkar ord med många betydelsekomponenter ha en högre svårighetsgrad än ord med färre sådana komponenter, eftersom betydelsen då är mycket specifik (Enström 2013b:188, jfr t.ex. *darra, skälva, skaka*).

Vidare kan sammansättningar bereda svårigheter dels på uttryckssidan då det kan vara svårt att avgöra var gränsen går mellan sammansättningens ingående led (jfr *bild|rulle* och *bil|drulle*, se von Martens & Blensenius 2016), dels innehållsmässigt eftersom sammansatta ord ”till sin karaktär är flertydiga vad

gäller den exakta relationen mellan förled och efterled” (Enström 2013b:187). Sammansättningar behandlas vidare i avsnitt 3.1.2.1 Lemmasektion och i avslutning till Lexin i avsnitt 3.2.3.

Ett okänt ords betydelse är, enkelt uttryckt, lättare att uttyda ju mer känd kontext det omgärdas av. Nagy (1997:76–82) urskiljer tre typer av kunskap som bidrar till kontextbaserade tolkningar: 1) lingvistisk kunskap vilken inbegriper syntaktisk kunskap, kunskap om s.k. ”word schemas” (dvs. implicit kunskap om möjliga betydelser hos ett ord) samt kunskap om ordförrådet, 2) omvärldskunskap och 3) strategisk kunskap. Strategisk kunskap handlar om att ha en medveten kontroll över kognitiva resurser i samband med processen att uttolka betydelsen av okända ord i en text (Nagy 1997:82).<sup>7</sup>

För att till fullo kunna utnyttja de ledtrådar som en omgivande text innehåller krävs alltså mycket goda språkliga kunskaper. Modersmålstalare och mycket avancerade inlärare har följaktligen (mycket) bättre förutsättningar att utöka sitt ordförråd genom kontextinläring än inlärare på nybörjarnivå. Innan en inlärare börjar gissa sig till ett ords betydelse måste ordet i fråga dessutom ha identifierats som okänt. Det kan låta som en självklarhet, men som redan konstaterats kan man tro sig veta vad ett ord betyder utan att faktiskt göra det. Ledtrådarna i sig måste dessutom vara tillräckligt många och användbara, men så är förstås inte alltid fallet. De ledtrådar som återfinns kan i sig själva vara okända och således helt oanvändbara. Ledtrådarna kan också vara missledande, t.ex. om inläraren endast har partiell kunskap om orden bakom dem. Inlärare kan också vara förblindade av det som de förväntar sig att den aktuella texten ska handla om (se vidare Laufer 1997a:30).

Sammanfattningsvis kan sägas att det i vissa fall kan förutses vilka ord som är besvärliga att förstå och lära sig för inlärare, i viss utsträckning utifrån egenskaper hos inlärarnas förstaspråk, och i andra fall är den potentiella svårigheten hos ord mycket individuell och svårförutsägbar. Vi återkommer till frågan om svåra ord i avsnitt 7.4 som redogör för informanternas synpunkter på LSL, bland annat om (avsaknaden av) svåra ord.

Enström (1996:16) tar upp tre huvudsakliga inlärningsmetoder för uppbyggnad av ett ordförråd. Vid sidan av ordinläring i kontext (implicit ordinläring) avhandlas ordinläring med hjälp av ordböcker och ordinläring med hjälp av undervisning (jfr Enström 2013a:4–8 och 2016:25–34). Den största andelen ord, menar Enström (1996:17), lärs troligen in genom den första metoden (se ovan). Den andra metoden, ordinläring med hjälp av ordböcker, är

---

7 Nagy (1997:74) framhåller att den strategiska kunskapen är ett hjälpmedel i samband med ordförståelse och inte ordinläring. För att lära sig ett ord, ”to produce ownership”, krävs enligt Nagy att man stöter på ordet vid sju eller fler tillfällen.

aktuell vid såväl ordinlärning i kontext som vid ordinlärning i undervisning. En ordbok innehåller kunskap om ords yttre och inre struktur och hur ord är relaterade till varandra (jfr uppställningen från Stroud 1979). Ordboken utgör dels en källa till information, dels ett hjälpmedel vid inlärning (jfr Tarps kognitiva och kommunikativa funktioner, se avsnitt 2.3.3), speciellt inlärningsordböcker vars primära syfte är att underlätta ordförståelse och ordinlärning.

Nuförtiden har de flesta tillgång till elektroniska ordböcker via datorer och via sina mobiltelefoner. Det torde på ett mycket positivt sätt ha främjat just ordinlärning i kontext (man går på stan och stöter på ett nytt ord och kan enkelt slå upp ordet i en mobilapplikation). Om ordböcker ur ett andraspråksperspektiv handlar följande avsnitt.

### 3.1.2 Ordböcker ur ett andraspråksperspektiv

Den samlade lexikala kompetensen hos en språkbrukare utgör hens mentala lexikon. En allmän definitionsordbok skulle man kunna säga ger en kraftigt förenklad bild av en idealiserad språkbrukares mentala lexikon (se vidare Källström 2012:169).

Nation (2013:428) gör en uppställning av relationen mellan de informationskategorier som brukar ingå i en ordbok och de delar som innefattas i kunskapen om ett ord (se tabell 3:1, jfr Stroud 1979:174–175). Form, betydelse och användning delas upp i receptiv (R) och produktiv (P) kunskap.

De informationskategorier som kopplas till receptiv ordkunskap är etymologi, avledningar, exempel, betydelse och illustrationer. Till produktiv ordkunskap räknas bland annat informationskategorierna uttal, stavning, böjning, synonymymer, antonymer, kollokationer och uppgifter om stil och bruklighet. Ett urval av dessa kategorier beskrivs närmare i avsnitten 3.1.2.1–3.1.2.8 i relation till inlärningsordböcker och under avsnitten 3.2.2 och 3.2.3 i samband med beskrivningen av LSL.

När det gäller såväl explicit språkinlärning som implicit språktillägnande av ett andraspråk utgör ordböcker, i synnerhet tvåspråkiga sådana, viktiga hjälpmedel för att tillfredsställa behov i situationer av reception och produktion. Ordböcker betraktas ofta som auktoriteter vad gäller ords förekomst, betydelse och användning och människor har i allmänhet en mycket hög tilltro till att innehållet i en ordbok är absolut korrekt och sant (Enström 1996:21, se även Müller-Spitzer 2013:12). En ordbok ska ju beskriva språket, antingen i syfte att redovisa språkbruket (deskriptiv ordbok) eller i syfte att påverka språkbruket (normativ ordbok). Enligt Svensén avser deskriptiv ”lexikografens sätt att förhålla sig till det språkliga utgångsmaterialet och innebär att det ska råda över-

ensstämmelse mellan detta material och ordbokens beskrivning av det ”medan normativ har att göra med ”lexikografens förhållande till ordboksanvändaren och dennes behov av vägledning, främst vid textproduktion” (Svensén 2004:29). Få ordböcker är dock strikt deskriptiva eller normativa. Malmgren (2000:199) beskriver t.ex. förhållandet mellan deskriptiv och normativ som att en i huvudsak deskriptiv ordbok bör baseras på det faktiska språkbruket under hänsynstagande till språkssystemet, medan en i huvudsak normativ ordbok bör baseras på språkssystemet under hänsynstagande till bruket.

TABELL 3:I. *Ordboksinformation och vad det innebär att kunna ett ord. (Efter Nation 2013:428).*

Form	spoken	R	
		P	pronunciation, alternative pronunciations
	written	R	
		P	spelling, hyphenation (syllabification)
	word parts	R	etymology
		P	inflections, derived forms
Meaning	form and meaning	R	derived forms, etymology, examples
		P	
	concept and referents	R	meanings, illustrations
		P	examples
	associations	R	examples
	P	synonyms, opposites, superordinates	
Use	grammatical functions	R	
		P	grammatical patterns, examples
	collocations	R	
		P	collocations, examples
	constraints on use (register, frequency...)	R	
		P	frequency, register, style, etc.

Andraspråksinlärare och modersmålstalare skiljer sig inte så mycket åt i fråga om vilka ord de dagligen konfronteras med. I ordböcker har emellertid inlärare, i jämförelse med modersmålstalare, behov av mer information och fler förklaringar om det aktuella språket. Det gäller framför allt kulturspecifika ord. Vid inlärning på nybörjarnivå används gärna tvåspråkiga ordböcker men ju mer avancerade kunskaper som uppnås i målspråket desto större användning har inläraren av enspråkiga ordböcker (Svensén 2004:23–25). En enspråkig ordbok kan dock även på relativt tidiga stadier utgöra ett betydelsefullt komplement till den tvåspråkiga ordboken, förutsatt att definitionsspråket är möjligt att tillgodogöra sig.

I det följande går jag närmare in på de huvudsakliga informationskategorier

som de flesta allmänordböcker – oavsett funktion – brukar innehålla, samt illustrationer. Fokus ligger på enspråkiga inlärningsordböcker. Jag inriktar mig i det följande också huvudsakligen på internetordböcker, som ju är den ordbokstyp som undersöks i avhandlingen. I fråga om det lingvistiska innehållet är skillnaderna mellan en tryckt ordbok och en internetordbok emellertid inte så stora. Skillnaderna uppstår i synnerhet i frågor rörande innehållets åtkomstmöjligheter. Till detta återkommer jag framför allt i avsnittet som handlar om Lexin, 3.2.

### 3.1.2.1 Lemmaselektion

Vad gäller inlärningsordböcker presenterar Tarp (2006) vad han kallar en teori för lemmaselektion, dvs. urval av uppslagsord. Författaren uppehåller sig vid tre frågor:

- 1) Vilken empirisk bas ska ligga till grund för lemmaselektionen?
- 2) Vilka kriterier och principer ska vägleda lemmaselektionen?
- 3) Hur stort ska lemmabeståndet i inlärningsordböcker vara?  
(efter Tarp 2006:203)

När det gäller fråga 1 finns det enligt Tarp tre huvudsakliga källor från vilka tänkbara lemman väljs ut: (a) det egna ordförrådet ur vilket data utvinns genom introspektion, (b) andra ordböcker, läroböcker och dylikt samt (c) textsamlingar i form av korpusar. Med introspektion avses dels lexikografens iakttagelser av det egna språkbruket, dels data från informanter. Svensén (2004:50) gör i tillägg en åtskillnad mellan primärkällor, Tarps källa a och c, och sekundärkällor, Tarps källa b. Svensén tar också upp internet som en källa till lexikalisk information som torde sammanfalla med Tarps källa c) (2004:73–75.).

Lexikografen bör naturligtvis excerpera ett ordurval som är autentiskt i betydelsen belagt i (skriftliga) källor. I takt med den tekniska utvecklingen har korpusar fått en allt större betydelse i lemmaurvalsprocessen. Men även om moderna korpusar utgör ovärderliga hjälpmedel i det praktiska lexikografiska arbetet så ”spores samtidig en tendens til at reducere hele diskussionen om lemmaselektion til brugen og kvaliteteten af disse korpora, som i nogle tilfælde får næsten ophøjet status” (Tarp 2006:203). Tarps oro är i viss mån befogad. Hur stor en korpus än är så kommer det alltid att finnas ord, fraser och konstruktioner som inte återfinns i den (Svensén 2004:71–73). Det kan vara ord som är högfrekventa i allmänt talspråk men lågfrekventa i skriftspråkskorpusar. Det kan också vara ord som av någon annan anledning inte återfinns i en ”vanlig” korpus, t.ex. vissa fackspråkliga termer av allmänspråklig betydelse.

En (inlärnings)ordbok ska innehålla ett centralt och typiskt ordurval – ett basordförråd – och i fråga om principer för lemmaurvalet (dvs. Tarps fråga 2) är trots allt frekvens ett relevant kriterium. Frekvens ska dock inte vara det enda kriteriet för att ett ord ska uppnå lemmastatus. I en inlärningsordbok är t.ex. sådana ord som inryms under benämningen *sambällsord*, dvs. ord som behövs för att man som medborgare ska kunna orientera sig i det svenska samhället, viktiga att täcka in. Exempel på sådana är *dagis*, *kommunalval* och *moms* (Gellerstam 1999:6). Användarstudier kan också ge intressanta uppslag till principer för lemmaselektion. En traditionell studie är Summers (1988:113–114) som visar att modersmålstalare oftare slår upp svåra ord än andraspråksinlärare. Med svåra ord avser författaren t.ex. abstrakta samhällsord, nyord och ovanliga eller ålderdomliga ord. Andraspråksinlärare slår istället i högre grad upp högfrekventa ord, speciellt abstrakta sådana. Senare undersökningar av loggfiler har fokuserat på hur ofta frekventa ord slås upp (jfr de Schryver et al. 2006, Verlinde & Binon 2010 och Wolfer et al. 2014, se vidare avsnitt 4.3.3.2).

Vid sidan av (absolut och relativ) frekvens som urvalskriterium anger Tarp (2006:206) även relevans och konsekvens. Det finns t.ex. såväl systematiska som relevanta skäl till att låta samtliga årets månader ingå i lemmaurvalet, även om månadernas namn uppvisar mycket skiftande frekvenser i korpus. Ett annat exempel är antonyma ord. Om man låter ordet *dynamisk* ingå i lemmarurvalet finns det goda skäl att också låta ordet *statisk* göra det. Vidare finns det, som tidigare nämnts, ord som kan vara särskilt relevanta för den användargrupp som invandrare utgör. Det kan vara ord för lagar, institutioner och kulturella företeelser typiska för – i Lexins fall – svenska förhållanden (t.ex. *LO*, *allemansrätt*). Vidare kan (löst sammansatta) partikelverb orsaka stora svårigheter för inlärare, dels då det föreligger en ovanlig relation mellan uttryck och innehåll (se vidare Enström 2013b:190 och avsnitt 3.1.1), dels eftersom partikeln kan hamna långt efter verbet i en sats. Partikelverket blir då svårt att identifiera (Malmgren 2012:463). Det är dessutom viktigt att just sambandet mellan lösa och sammansatta partikelverb framgår klart i en inlärningsordbok (se Malmgren 2012:462–463).

I övrigt bör man vid lemmaurval beakta möjliga typer av lexikala former (t.ex. stavnings- och böjningsvarianter) och olika lexikala strukturer i form av bland annat förkortningar, affix och flerordsenheter. Vidare ingår inte enbart enskilda ord i urvalsprocessen utan lika angeläget är att man täcker in frekventa och relevanta konstruktioner, kollokationer och idiom (för definitioner av *kollokation* och *idiom* se Malmgren (2003) och Sköldberg (2004)). Sådana flerordsuttryck uppnår i regel inte lemmastatus men uppträder ofta som sublemman.

En typ av ord som i regel *inte* brukar behandlas i språkböcker är egennamn. De har ju emellertid också formella egenskaper (t.ex. stavning, böjning, uttal) som kan behöva beskrivas. Egennamn förekommer oftare i två-

språkiga ordböcker än i enspråkiga. Det kan också finnas semantiska motiv till att låta egennamn uppgå i lemmaurvalet och då är de mer aktuella för en enspråkig ordbok. Ett egennamn kan t.ex. börja användas appellativt och få en metaforisk betydelse (*en riktig Hitler*) eller bilda en avledning och då tillföras vissa bestämda konnotationer (*en lindgrensk saga*). En strikt åtskillnad mellan språkböcker och sakordböcker är därför varken eftersträvningsvärd eller speciellt användarvänlig (jfr avsnittet om funktionsläran, 2.3.3).

Slutligen varnar Tarp för vad han kallar för ett irrelevanskriterium vilket kan ”diktere frasortering af for eksempel egennavne, arkaiske og sjældne ord, vulgære ord, fagord eller ord fra et bestemt område” (2006:207). Han hänvisar till förordet i Michael Wests kända engelska enspråkig inlärningsordbok från 1935 där det konstateras att av utrymmesskäl utesluts ovanliga och tekniska ord som det är osannolikt att en invandrare kommer att stöta på (West 1935). Men det är ju omöjligt att förutspå vilka ord inläraren *inte* kommer att komma i kontakt med.

Lemmabeståndets storlek (dvs. Tarps fråga 3) bör enligt Tarp vara ett resultat av överväganden och beslut och inte omvänt, dvs. att man utifrån ett i förväg bestämt omfång formulerar kriterier för urval av lemman. Ordböcker som finansieras av kommersiella företag måste ofta rätta sig efter den senare premissen till följd av avgränsade ekonomiska och tidsmässiga villkor. Det finns dock inte några teoretiska argument för att begränsa orduvalet i en receptionsordbok menar Tarp, utan motiven till detta är endast pragmatiska. Tarp ställer upp åtta punkter som ska utgöra en systematisk metod för bestämning av selektionsprinciper och i sin tur lemmabestånd, t.ex. bestämning av de avsedda typerna av användarsituationer, bestämning av den avsedda användargruppens karakteristika och bestämning av användargruppens primära (funktionsrelaterade) behov (Tarp 2006:206). Tarp menar vidare att inlärningsordböcker med få lemman endast är hjälpsamma i undervisningssammanhang eller andra kontrollerade situationer, i synnerhet om inlärarna befinner sig på nybörjarnivå. När det gäller elektroniska ordböcker bortfaller i hög grad argumentet att av utrymmesskäl utesluta potentiella lemman, men ett kontrollerat och genomtänkt urval bör under alla omständigheter ändå göras.

En stor del av uppslagsorden i en ordbok utgörs av sammansättningar. Storleken på en ordbok beror därmed i hög grad på i vilken utsträckning man väljer att ansätta sammansättningar som lemman. Detta beror i sin tur t.ex. på om ordboken i fråga främst ska användas i receptions- eller produktions-situationer. För en receptionsordbok gäller ofta att genomskinliga sammansättningar inte ansätts som lemman, ofta grundat på antagandet att deras betydelse är lika med summan av delarna (kompositionalitetsprincipen). Det här har dock visat sig vara en sanning med modifikation. Det förutsätter t.ex. att betydelsen hos såväl sammansättningens förled som efterled är entydig, vilket ofta inte är fallet: ”Den utbredda polysemin hos många ordleder och



det faktum att det alltid finns flera möjliga relationer mellan lederna [...] gör att sammansättningar alltid blir principiellt flertydiga” (Svanlund 2009:34). Inte minst för en inlärare kan sammansättningar därför bereda svårigheter och vara ett skäl till att i en inlärningsordbok ansätta till synes genomskinliga sammansättningar som lemman (jfr avsnitt 3.1.1). För att ta ett exempel: förledet i sammansättningen *valnederlag* är flertydig. För en (vuxen) modersmålstalare känns emellertid sammansättningen igen från den politiska sfären och hen vet att det sannolikt rör sig om *val* i betydelsen ’politisk omröstning’ och inte ’däggdjur’. Denna erfarenhetsbaserade kunskap som modersmålstalare bär på kan t.ex. förmedlas i en inlärningsordbok genom en formell kommentar som den i SAOL14: ”Nästan alla sms. med *val-* hör till ’*val* 2” (2015:1485). Om inget annat anges så har förleden i en sammansättning med *val* just denna betydelse, dvs. ’politisk omröstning’. I en internetordbok skulle man kunna gå ett steg längre. En inlärare som stöter på en okänd sammansättning som visar sig saknas som lemma i ordboken skulle inte behöva stå rädlös. I ett första steg kan sammansättningen uppdelas maskinellt i sina två huvuddelar, något som nästan alltid går att göra entydigt (Malmgren 1995:302 baserat på Karlsson 1992). Utifrån statistiska underlag som beräknar sannolikheten att ett visst för- eller efterled har en viss betydelse skulle information kunna förmedlas till inläraren om vilken betydelse det är som rimligen åsyftas hos (leden i) sammansättningen, även om ordet *valnederlag* inte är upptaget i ordboken. En sådan ”dynamisk” ordbok skulle alltså även kunna ge information om ord som *inte* finns i ordboken.

Det talas gärna om att urval ska vara representativa. Ett representativt urval i icke-statistisk mening skulle kunna innebära att man täcker ”de typiska och centrala aspekterna av språket och ge[r] så många förekomster av ord och ordförbindelser som man anser sig ha tillräckligt stöd för i autentiskt material” (Svensén 2004:80). Ett representativt urval i denna mening betyder dock inte att just det urvalet är ändamålsenligt i förhållande till användarens behov. Det finns här en motsättning mellan ordboken som en dokumentation av språket och ordboken som ett hjälpmedel för att lösa språkliga problem (Béjoint 2000:184–187).

### 3.1.2.2 Uttal

Uttalsuppgifter tillfredsställer i synnerhet behov av produktion och förekommer därför i ordböcker som helt eller delvis är till för det ändamålet. Det inkluderar de allra flesta allmänordböcker. Vad gäller svenska enspråkiga L2-ordböcker behöver uttalet beskrivas på i princip alla ord medan det i enspråkiga L1-ordböcker endast behöver ges på ord som avviker från de svenska uttalsreg-

lerna (Svensén 2004:152). Uttalsuppgifterna ska, i likhet med allmänordböcker, ge upplysning om ett uttal som anses ”korrekt” (a.a. s. 143). Uttalet är i regel uppsvenskt (rikssvenskt) och ska vara representativt för ett vanligt och vårdat uttal av svenska ord. Det är ordens uttal i isolation som anges. Geografisk, stilistisk eller prosodisk uttalsvariation anges i regel inte. I inlärningsordböcker är det också viktigt att ge uppgift om uttal på vissa böjningsformer som avviker från grundordets uttal, t.ex. oregelbunden plural av substantiv.

Angivelser av uttal inkluderar information om vilket ljud det är fråga om, ljudets längd, tryckets placering och typ av tonaccent (Svensén 2004:147–148). Denna information förmedlas sedan visuellt med hjälp av diakritiska tecken eller genom en separat transkription (se jämförelsen mellan två ordboksartiklar i avsnitt 3.2.2). Antingen skrivs transkription med hjälp av ett givet system av symboler, t.ex. IPA-alfabetet, eller så används en ortografinära transkription. Det finns också exempel på kombinationer av dessa. I elektroniska ordböcker kan uttal också återges auditivt genom inspelat eller syntetiskt tal, något som numera är nästintill standard. Den visuella informationen är emellertid inte överflödigt utan de två informationssätten kan komplettera varandra (se vidare Lew under utg.).

### 3.1.2.3 Morfologi: böjning och ordbildning

Morfologi kan delas in i de två huvudkategorierna böjning och ordbildning (Svensén 2004:149, se även Golden 2011:48–51 för en översiktlig beskrivning av norsk morfologi, som har stora likheter med svensk).

Uppgifter om böjning tillfredsställer i första hand behov vid produktion (Tarp 2006:243). I en inlärningsordbok för produktion är det därför viktigt att ge utförliga böjningsuppgifter och allra helst explicit utskrivna former. Det gäller i synnerhet ett böjningssystem som det svenska som är mer intrikat än vad man kan tro (se vidare Berg & Cederholm 2001, Berg et al. 2008b). När det gäller reception kan böjningsuppgifter hjälpa användaren att nå fram till lemmat och därigenom ordboksartikelns innehåll (Svensén 2004:159). De allra flesta ordböcker inkluderar i regel böjningsinformation om såväl regelbundna böjningsmönster som avvikelser från dessa, dvs. oregelbundna former. I tryckta ordböcker innebär det ofta att oregelbundna böjningsformer återfinns som hänvisningslemman. Regelbundna böjningsparadigm kan förmedlas i anslutning till lemmat eller i olika typer av uppställningar i ordbokens utanförtext (Svensén 2004:161–163, Atkins & Rundell 2008:207). Det mest användarvänliga är rimligen att ge explicit böjningsinformation i anslutning till lemmat.

Med ordbildning avses bildning av sammansättningar och avledning. Medan avledning är en ordbildningsprincip som förekommer i de flesta språk

är möjligheterna att bilda sammansättningar jämförelsevis mycket stora i svenskan. Det gäller i synnerhet substantiv men även adjektiv och verb. Ett språks mönster för ordbildning underlättar andraspråksinläring – det är lättare att känna igen och minnas ord som består av återkommande delar (Källström 2012:169). Som vi redan vet kan emellertid betydelsen hos sammansatta ord, eller avledningar, inte alltid förstås med utgångspunkt i de ingående enheternas betydelse. Det är alltså inte så att mönstren för ordbildning utan undantag underlättar förståelsen och inläringen av ordförrådet (Källström 2012:176). I synnerhet ogenomskinliga sammansättningar och avledningar behöver således förklaras i inlärningsordböcker, men också ett urval av genomskinliga sådana (se vidare avsnitt 3.1.2.1 Lemmaselektion).

I en elektronisk ordbok kan alla böjningsformer och i artikeln ingående avledningar och sammansättningar vara sökbara och direkt leda till berörda ordboksartiklar, vad gäller böjning även sådana former som inte återges i artikeln (se avsnitt 3.2.4.1 om rättstavningsfunktionen i Lexin). Det elektroniska formatet gör det också möjligt att t.ex. klicka fram böjningsparadigm via ett popup-fönster.

#### 3.1.2.4 Betydelsebeskrivning

Det övergripande syftet med en betydelsebeskrivning är att möta ordboksanvändarens kommunikativa behov. Behoven kan se mycket olika ut men delas vanligen upp i receptiva behov och produktiva behov. Att på ett tillfredsställande sätt kombinera de två funktionerna i en och samma ordbok är inte lätt. Definitioner som är väl anpassade för receptiva behov är ”short, easy-to-follow, and therefore rather general” (Atkins & Rundell 2008:410). Produktion är en mer utmanande uppgift. För att kunna använda ett ord i produktion måste man kunna en hel del om ordet, alternativt ha en mycket klar uppfattning om den betydelse man vill uttrycka men saknar ord för (jfr Strouds uppställning i avsnitt 3.1.1). Definitioner som ska tillfredsställa produktiva behov måste därför vara mycket mer specifika och uttömmande. En ordbok får således prioritera den ena funktionen på bekostnad av den andra. I elektroniska ordböcker kan man lösa problemet genom att ha olika definitionsstilar för olika funktioner som man enkelt kan växla mellan.

Ett ords betydelse och användning kan beskrivas med hjälp av definition, synonym, exemplifiering, illustration eller översättning. I de flesta fall innehåller betydelsebeskrivningen en kombination av flera av dessa förklarings typer. I inlärningsordböcker för nybörjare är det ändamålsenligt att i möjligaste mån komplettera förklaringar med synonymer, exempel och illustrationer. Synonymer kan dock vara särskilt svåra för inlärare att bedöma eftersom full

synonymitet sällan förekommer inomspråkligt och inte heller mellan två språk (Enström 1996:24). Man måste också ha i åtanke att inlärarna oftast saknar kunskap om sociokulturella normer och kulturbundna företeelser. En fråga man bör ställa sig är dessutom om betydelsebeskrivningen i en inlärningsordbok för nybörjare ska begränsas till vissa grundbetydelser eller om så många betydelser som möjligt ska ingå (jfr Hulstijn & Atkins 1998:17). I det sammanhanget gäller det att finna en balans mellan att vara så exakt och noggrann i sin beskrivning som möjligt men samtidigt behålla en läsvänlig utformning av artikeln.

Det säger sig självt att definitioner ska vara begripliga. Det innebär att det språk som används för att förklara uppslagsord måste anpassas efter ordbokens syfte och med hänsyn till målgruppens (språkliga) kompetens. Polysema ord ska t.ex. inte användas i en otypisk betydelse. Definitionsspråket ska helst inte utmärka sig utan likna vanlig prosa så mycket som möjligt givet de förutsättningar som råder (Atkins & Rundell 2008:412).

En definition kan utformas på olika sätt. Ett traditionellt sätt att definiera på i allmänordböcker är genom att ange närmaste överbegrepp, *genus proximum*, och en eller flera bestämningar, *differentia specifica*. Definitionen består då av en fras vars huvudord tillhör samma ordklass som lemmat. Ordet *tax* i LSL definieras t.ex. med frasen 'en hund med lång kropp och korta ben'. Ett vanligt krav är också att definitionen är substituerbar. Det innebär att beskrivningen har en "sådan syntaktisk utformning att den kan sättas in på lemmats plats i löpande text" (Svensén 2004:283). Ett alternativt sätt att definiera introducerades i *Collins Cobuild English Dictionary* (1987). En så kallad Cobuild-definition är utformad som en hel mening. Ordet som förklaras ska återfinnas i definitionen. Ordet *cousin* ges t.ex. definitionen 'your cousin is the child of your uncle or aunt' och verbet *defeat* definieras 'if you defeat someone, you win a victory over them'. Syftet med en Cobuild-definition är att den ska vara enklare och begripligare än en traditionell definition, vilket ju är ändamålsenligt för en inlärare. Definitionen ska likna vanligt, naturligt språk som påminner om hur en lärare skulle svara om hen fick en fråga om betydelsen hos ett ord.

Oavsett vilket definitionstyp som används bör man inom ett ordboksprojekt ha kontroll över de ord som används i definitionerna. Man kan då använda sig av en i förväg bestämd *definitionsvokabulär*, dvs. en systematiskt utvald mängd av ord för att innehållsligt beskriva en större mängd av ord (Svensén 2004:303). Ett känt exempel på en ordbok som använder sig av en sådan *begränsad* definitionsvokabulär är den engelska enspråkiga ordboken *McMillan Dictionary* (2016, med första upplaga 2002). Målsättningen är där att varje definition ska vara så tydlig och enkel att förstå som möjligt och för att uppnå detta används en definitionsvokabulär på 2 500 ord. De utgörs av de vanligaste

och mest basala orden i engelska språket. Det är bara den mest grundläggande innebörden av ordet som får användas. Ordet *cold* används således alltid i betydelsen 'som har låg temperatur' och aldrig i betydelsen 'ovänlig'. I sällsynta fall används ord som inte finns med i definitionsvokabulären och det kan då handla om egennamn. För svenskans del är det svårt att komma ner i en definitionsvokabulär av McMillans storlek (som dock verkar vara betydligt större än de 2 500 orden, se Malmgren 2012:464). Anledningen till detta är framför allt att där svenskan bildar sammansättningar använder engelskan nominalfraser, vilket inte på ett lika naturligt och smidigt låter sig göras på svenska (jfr t.ex. sv. *sopbil* och eng. *garbage truck*). Malmgren (a.a. s. 463–464) menar därför att det för svenskans del är mer ändamålsenligt att sikta på en *kontrollerad* definitionsvokabulär än en begränsad sådan. En kontrollerad definitionsvokabulär innebär i LSL:s fall att orden som används i definitionerna finns som uppslagsord, eller åtminstone som morfologiska exempel (se vidare om Lexins uppdatering i kommande avsnitt).

Användarstudier har visat att inlärare kan ha mycket svårt att förstå och ta till sig definitioner i ordböcker (Wikström 1991:42–48, Nesi & Meara 1994, se även Nesi 2000b som undersöker läsbarheten hos olika definitionsstilar). Frågan är om det har med bristande språkkunskaper att göra eller med det sätt varpå definitioner skrivs. Man kan tänka sig att båda tolkningarna har relevans.

### 3.1.2.5 Språkprov

Språkprov är enligt Malmgren (1994:108) ”fraser (eller sammansättningar och avledningar) som återges inom ramen för en ordboksartikel och som innehåller uppslagsordet och har till uppgift att illustrera det i någon av dess betydelsenyanser eller användningar”. Liknande definition återfinns hos Svensén (2004:343) som dock inte inkluderar sammansättningar och avledningar (morfologiska språkprov) i sin begreppsförklaring (jfr Atkins & Rundell 2008:453ff.). Språkprovet ska fungera som ett stöd till definitionen och innehålla upplysningar om konstruktionsmöjligheter, typiska kollokationer och olika slag av associationer och stämningar som lemmat ger upphov till.<sup>8</sup> I bästa fall innehåller ett språkprov också relevanta encyklopediska uppgifter

---

8 Man kan ha olika åsikter om huruvida språkprov är en särskild informationstyp. Svensén (2004:345) menar att språkprov realiserar eller åskådliggör olika informationstyper och ser alltså inte språkprov som en särskild informationskategori. För den här avhandlingens vidkommande är frågan om språkprovets status som informationstyp inte av avgörande betydelse, men av praktiska skäl rubriceras det som en separat informationskategori.

som kan bidra till bättre förståelse av ordets betydelse och bruk (se vidare Svensén 2004:348–351). I språkprovets uppgift ingår vidare att förutom det typiska bruket även exemplifiera typiska och frekventa avvikelser från detta bruk (jfr avsnitt 3.1.2.7 om bruk och stil). Språkprovet har också en viktig uppgift i att fungera som en motivering till betydelseindelningen (se Malmgren 1994:110–111).

Atkins & Rundell (2008:458–461) ställer upp tre kriterier som mått på goda språkprov. De ska vara 1) typiska och naturliga, 2) informativa och 3) begripliga. Naturlighet beskrivs som ett mer intuitivt mått på ett bra språkprov men viktigt är förstås att den användning som illustreras är frekvent. Med informativa språkprov menas att de ska komplettera definitionen och hjälpa användaren att bättre förstå det aktuella ordets betydelse och användning. De får absolut inte motsäga det som framgår av definitionen. Vidare ska språkprov vara begripliga, dvs. inte bereda läsaren några svårigheter vad gäller deras betydelse som helhet eller de ingående ordens betydelse. Man bör således undvika svåra eller tvetydiga ord och uttryck.

Förutom att vara naturliga/typiska, informativa och begripliga, ska språkprov i idealfallet representera autentiskt bruk och det är numera praxis att använda sig av elektroniska korpusar vid ordboksframställning (Atkins & Rundell 2008:457, se även Kilgariff et al. 2008 om automatisk generering av goda språkprov). Belägg som uppfyller samtliga kriterier kan dock vara svår- funna i korpusar och faktorer såsom ordbokens funktion och de resurser som finns tillgängliga leder till olika lösningar i olika ordböcker. Även den tid som ett ordboksprojekt har till sitt förfogande utgör en påverkansfaktor. I synnerhet i tryckta ordböcker ska språkproven dessutom uppta så lite plats som möjligt, vilket är ett av skälen till att autentiska exempel redigeras. I och med de elektroniska ordböckernas intåg har förutsättningarna förändrats vad gäller t.ex. språkprovs längd och antal. Där har man också möjlighet att länka ordboken till olika korpusar och användaren kan skapa sig en egen uppfattning av bruket. Trots det obegränsade utrymmet i elektroniska ordböcker krävs ändå en genomtänkt framställning av det man vill beskriva eller, som här, illustrera med goda språkprov.

Förespråkarna för autentiska språkprov anser att ”konstruerade/bearbetade språkprov [är] mindre typiska och framför allt mindre naturliga” och att även små ändringar i syfte att göra dem mer lättillgängliga eller regelbundna kan beröva dem deras karaktär, och i värsta fall vilseleda användarna (Svensén 2004:346). Om man strikt ska följa kravet på autenticitet så måste emellertid många bra belägg ratas på grund av t.ex. längd eller distraherande element. Förespråkarna för denna hållning menar dock att korpusarna numera är så stora att man kan hitta goda språkprov till de flesta ord, även till sådana som är mindre frekventa. Det hela är dock inte så enkelt, vilket Kilgariff et al.

(2008:426) konstaterar: ”Even in the very large corpora available today, it is hard to find whole sentences which perfectly meet all the relevant criteria, and some degree of editorial intervention is usually needed”.

Förespråkarna för redigerade språkprov anser att språkproven ska illustrera lemmanas kärnbetydelse(r) och att ordboksanvändaren utifrån denna information ska kunna sätta in ordet i passande kontexter med passande grammatisk form (Schemann 1989:1026). Språkproven ska utgöra ett mönster, en prototyp, som i sin tur kan bidra till förståelsen av ordets användning i autentiska sammanhang. Nikula (1995:313) uttrycker det som att språkproven tydliggör en regelformulering och att man slår upp ett ord för att få reda på de idiosynkratiska regler som gäller användningen av detta ord. Korpusen utgör också först och främst en källa av bevis och inte en källa av färdiga språkprov (Atkins & Rundell 2008:458).

Förespråkarna för de två olika språkprovstyperna, dvs. autentiska och redigerade, kan sägas representera två olika skolor. Oavsett till vilken skola man ansluter sig är en avgörande faktor som sagt vilken ordbokstyp man har att göra med. För inlärningsordböcker gäller i hög grad att de ord som omger uppslagsordet inte får försvåra tolkningen av språkprovet. I en inlärningsordbok för receptionsändamål bör språkproven också utgöras av hela meningar och inte av infinitivfraser. Språkprovet ska stödja och konkretisera definitionen och ”[m]ånga gånger kan ett bra exempel bidra mer till förståelsen av ett ord än en kanske abstrakt definition” (Malmgren 2012:460). Att upprepa ord är också ett sätt att underlätta och påskynda inläringen av nya ord och det kan också utnyttjas vid utformningen av språkprov.

### 3.1.2.6 Konstruktioner

Med konstruktionsuppgifter avses här framför allt uppgifter om (syntaktisk) valens, dvs. (precisering av) vilka bestämningar som ett ord kräver för att uttrycken där det ingår ska bli korrekta och meningsfulla (Svensén 2004:186, jfr Malmgren & Toporowska Gronostaj 2009:181). Utöver obligatoriska bestämningar kan ord också ha optionella bestämningar. Valens inkluderar också information om vilken grammatisk form orden har i de olika uttrycken.

Konstruktionsuppgifter kan ges för substantiv, adjektiv och verb. Enligt Svensén (2004:188) är verb den ordklass som blivit grundligast beskriven i ordböcker, framför allt på grund av sin centrala ställning i satsen (se avsnittet om svåra ord ovan). I enspråkiga inlärningsordböcker är informationen framför allt inriktad på produktion på främmandespråket.

Att utforma ett system för att förmedla syntaktisk information som kombinerar ”descriptive power with accessibility” (Atkins & Rundell 2008:401)

har emellertid visat sig vara besvärligt, i synnerhet ifråga om verbens valensbeskrivning. Särskilt brittiska inlärningsordböcker har kännetecknats av att använda sig av mer eller mindre intrikata kodsystem. I ordboksartiklarna finner man en kombination av bokstäver (vanligen versaler) och/eller siffror. En kodnyckel återfinns sedan vanligen i ordbokens utanförtext. Koderna har kritiserats för att vara alltför komplicerade (Malmgren & Toporowska Gronostaj 2009:185). Användarna verkar ha svårt att ta till sig informationen eller så är de ovilliga att lägga ned den tid som krävs för att förstå dem. På senare tid har man inom (den brittiska) inlärningslexikografin därför sökt sig bort från formella syntaktiska kodsystem mot förenklade och mer användarvänliga sätt att beskriva ords konstruktionsmöjligheter (Atkins & Rundell 2008:401). I elektroniska ordböcker är dessutom tillgången till utrymme mindre begränsad än i tryckta ordböcker, och koderna behöver inte vara lika komprimerade (se vidare Malmgren & Toporowska Gronostaj 2009:194). En form av förenkling är att använda sig av metaspråklig klartext. Det anses mer lättillgängligt än ett kodsysteem men beskrivningen riskerar samtidigt att bli mindre precist (Svensén 2004:192). Ytterligare ett sätt är att använda sig av *döda* eller *levande språkprov*. Ett dött språkprov innehåller endast satsens nödvändiga komponenter. Lexinordböckerna innehåller t.ex. döda språkprov. Under lemmat *sammenfatta* ges konstruktionsuppgiften 'ngn sammenfattar (ngt/SATS)' (se vidare avsnitt 3.2.3). Levande språkprov utnyttjar istället personliga pronomen och typiska nominalfraser. När det gäller verb används ofta finita former istället för infinitiven. Konstruktionsuppgifter kan också vara inbyggda i definitionen såsom i Cobuild-ordböckerna (se avsnittet om betydelsebeskrivning ovan).

### 3.1.2.7 Bruk och stil

En inlärningsordbok behöver dessutom ge upplysning om ords eventuella användningsrestriktioner ifråga om bruk och stil. Denna typ av information inbegrips i termen *diasystematisk markering* (Svensén 2001:378) och innebär att "en viss lexikalisk enhet i något avseende avviker från huvudmassan av de enheter som beskrivs i en ordbok". I beskrivningen av ordet bör då ingå uppgifter om dess användningsrestriktioner. I SAOL14 (2015:XXIX) har man t.ex. följande fem olika kriterier: fackspråkighet, ålderdomlighet, vardaglighet, regionalitet och (negativ) värdering. I den aktuella ordboksartikeln ges sedan verbala markörer i förkortad form. Några (utskrivna) exempel på sådana är: flygteknik, språkvetenskap, något ålderdomligt, starkt vardagligt, högtidligt, provinsial, nedsättande.



### 3.1.2.8 Illustrationer

Illustrationer befattar sig med världen och inte med det språkliga tecknet och är till sin karaktär encyklopediska (Svensén 2004:363). En bild behöver inte processas på samma sätt som en verbal beskrivning utan har en mer omedelbar verkan. Inlärningseffekten förstärks därmed av en bild som ”på ett helt annat sätt än en språklig beskrivning kan vädja till läsarens tidigare erfarenheter av världen och framkalla aha-upplevelser” (a.a. s. 364). I språkböcker utgör illustrationer ett visuellt stöd till den verbala ordboksartikeln. De förekommer framför allt i inlärningssammanhang. De vanligaste bilderna illustrerar konkreta substantiv/företeelser men även adjektiv, verb och funktionsord exemplifieras ibland med bilder. Illustrationer används framför allt till att beskriva utseende särskilt hos substantivens referenter, och mer sällan deras funktion.

Illustrationen ska samspela med uppslagsordet och dess definition och språkprov. En bild kan t.ex. vara mer beskrivande och förmedla fler egenskaper hos lemmat än vad den verbala definitionen har möjlighet att göra. En bild kan också spara utrymme då en tillräckligt utförlig definition skulle ta alltför mycket plats. Bilder är dessutom effektiva ”när det gäller att beskriva vissa rumsliga, tidsmässiga etc. sammanhang och olika begreppsrelationer” (Svensén 2004:364). Det kan också vara särskilt fördelaktigt vid språkinlärning att genom bilder visa grupper av företeelser inom ett och samma ämnesområde (se figur 3:1 nedan).

Urvalet av lemmat som ges en illustration bör vara noga genomtänkt och motiverat, och inte göras på måfå (jfr Stein 1991:105). Man bör ställa sig frågan ”varför en viss betydelse kan behöva stöd av en illustration, vilken typ av illustration som är den optimala och framför allt om en illustration alls är möjlig att åstadkomma” (Nikula 2008:305). En illustration krävs när ”betydelsen av ett ord inte kan anges tillfredsställande verbalt så att ordboksanvändaren kan identifiera referenten i den utomspråkliga verkligheten” (a.a. s. 307). En bild utgör emellertid inte alltid en odiskuterbar återgivning av ett ord. Hur ser t.ex. det prototypiska bordet ut? Att finna fram till en allmängiltig bild av ett föremål är något av en utmaning i sig. Dessutom är naturligtvis ordbokens målgrupp en väsentlig faktor att beakta. För inlärare spelar bilder en avgörande roll, framför allt för nybörjare och då en tvåspråkig ordbok saknas.<sup>9</sup> För elektroniska ordböcker är situationen kanske mer fördelaktig vad gäller ekonomiska och utrymmesmässiga faktorer. Valet av vilka lemmat som ska illustreras och hur illustrationerna ska se ut måste ändå fortfarande göras. Vad gäller

---

9 Två- eller flerspråkiga bildordböcker är också utomordentliga hjälpmedel i produktionssituationer, t.ex. när man undrar hur ett verktyg benämns på ett annat språk.

det tekniska utförandet kan man diskutera om fotografier eller teckningar är mest fördelaktiga. Enligt Nikula (2012:111) är man numera relativt ense om att ”teckningar är att föredra framför foton eftersom en väljord teckning kan utgöra ett destillat av de egenskaper som anses vara typiska för det avbildade, dvs. det prototypiska föremålet”. Diskussionen är relevant även vad gäller elektroniska ordböcker, om än inte i lika hög grad.

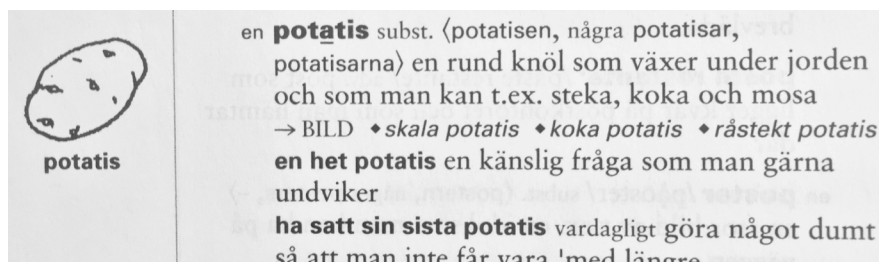


FIGUR 3:1. Scenisk illustration över lemmat vardagsrum. Lexin/Bildteman.

Det finns förstås olika typer av illustrationer (Svensén 2004:369–377). En *unikal illustration* utgörs av ett enda föremål. Om föremålet inte följs av annat än en rubrik utgår man från att bilden är prototypisk (a.a. s. 370). En *strukturillustration* visar föremålet i sitt normala sammanhang. Om det inte är uppenbart vilken del av bilden som beskriver lemmat måste det aktuella föremålet markeras. Det finns också motsvarande *funktionsillustration* där föremålets normala funktion förs fram. Handlingar och skeenden kan illustreras genom en *sekvensillustration*. Då visas föremålet eller personen i flera positioner eller utvecklingsstadium (a.a. s. 373). I en inlärningsordbok kan så kallade *nomenklatoriska* och *sceniska illustrationer* vara särskilt ändamålsenliga. De förra innehåller centrala föremål och begrepp inom ett specifikt ämnesområde, t.ex. inom olika sporter, och de senare visar en hel miljö med typiska föremål och aktiviteter, t.ex. ett vardagsrum (se figur 3:1 från Lexins bildtema).

Det kanske är en självklarhet att man bör använda sig av tidsenliga bilder. I en inlärningsordbok innebär det allt som oftast att bilderna ska vara samtida.

Den uppenbara och mest pedagogiska placeringen av en illustration är i mycket nära anslutning till det lemma som illustreras. Figur 3:2 visar hur illustrationerna används i *Natur och Kulturs Svenska Ordbok* (2001), NoKSO (se vidare Martola 2002:364–365). Där är bilderna placerade på samma uppslag som det lemma som illustreras. I artikeln får användaren dessutom information om att det till aktuellt lemma finns en illustration genom markeringen → BILD.



FIGUR 3:2. Uppslagsordet potatis och dess illustration (NoKSO 2001).

På så sätt kan ett samspel uppstå mellan den verbala och den illustrativa beskrivningen av lemmat. I en elektronisk ordbok kan man, som i LSL, ha en länk från ordboksartikeln till en bild, vilket framgår av exemplet med vardagsrummet ovan (se figur 3:1). Det elektroniska formatet öppnar också upp för möjligheten att ha rörliga bilder. I just LSL illustreras t.ex. som redan nämnts ca 700 verb med hjälp av korta videofilmer.

I tvåspråkiga ordböcker är det mera ovanligt med bilder men de kan vara särskilt effektiva vid beskrivningen av kulturspecifika ord (Svensén 2004:377).

## 3.2 Lexin – inlärningsordböcker för invandrare

### 3.2.1 Projektet Lexikon för invandrare

Invandringen till Sverige från icke-skandinaviska och icke-europeiska länder ökade kraftigt under 1970- och 80-talen och stora grupper av människor kom i behov av svenskundervisning. Tillgången till ordböcker mellan svenska och de nyanlända invandrarnas språk var vid denna tid emellertid mycket begränsad, för att inte säga obefintlig. Samtidigt var intresset från förlagens sida svalt. Förmodligen såg man inga större kommersiella vinster i att producera ordböcker för dessa språk. Staten tog i det läget på sig ansvaret att råda bot på

denna ordboksbrist. År 1977 initierade Skolöverstyrelsen därför ett forsknings- och utvecklingsprojekt med syfte att framställa lexikon anpassade till språkundervisningen av invandrare.<sup>10</sup> Utgångspunkten var riksdagens beslut att ”samhället ska främja aktiv tvåspråkighet och underlätta språkinläringen för invandrare” (Beijer & Fre Woldu 1997:3). Lexin-projektet var således ett svar på det ökade behovet av inlärningsordböcker och tvåspråkiga ordböcker hos dessa grupper av människor. Tanken var att lexikonen skulle kunna användas av invandrare med ”begränsad skolbakgrund och utan tidigare erfarenhet av lexikonanvändning” (ibid.). Den primära målgruppen definierades som ”nyttillkomna invandrare och flyktingar som ska lära sig svenska” (a.a. s. 4). Bra tvåspråkiga lexikon är essentiella hjälpmedel när man ska lära sig ett nytt språk och det unika med Lexin-ordböckerna är enligt Beijer & Fre Woldu att de har utarbetats efter den ovan definierade målgruppens förutsättningar.

Projektet påbörjades av Statens institut för läromedel och Statens invandrarverk. Institutionen för språkvetenskaplig databehandling (Språkdata) vid Göteborgs universitet fick, under ledning av Martin Gellerstam, i uppdrag att utarbeta det svenska ordboksunderlaget. Det uppdraget sades dock upp 1995. Sedermera övertog Statens institut för läromedel huvudansvaret för att i samråd med Skolöverstyrelsen ta fram lexikon för invandrare. Från 1991 låg ansvaret hos Skolverket och från 2003 hos Myndigheten för skolutveckling. Sedan 2007 ligger ansvaret hos Språkrådet, den språkvårdande avdelningen inom Institutet för språk och folkminnen ([www.sprakochfolkminnen.se](http://www.sprakochfolkminnen.se)).

Ett viktigt mål med projektet var att ta fram en noga övertänkt och målgruppsanpassad svensk uppsättning av uppslagsord försedd med olika typer av lingvistisk information. Detta material skulle sedan ligga till grund för de tvåspråkiga ordböckerna. Tanken var också att lemmauppsättningen i de tvåspråkiga versionerna skulle kunna ”vändas” så att det främmande språket också blev källspråk i de tvåspråkiga lexikonen (Hult et al. 2010, jfr Martin & Tamm 1996 och Martin 2007 om OMBI, ett projekt liknande Lexin). Det här senare syftet har emellertid inte realiserats utan har fått stå tillbaka för det större behovet av att framställa nya lexikon med svenska som källspråk (Gellerstam 1999:3–4). Tack vare digitaliseringen av ordböckerna kan de tvåspråkiga ordböckerna dock i viss mån användas som L1–L2-ordböcker, dvs. den arabisk-svenska ordboken kan användas (av svenskar) för reception av arabiska och (av arabisktalande) för produktion av svenska (jfr avsnitt 2.3.2).

Det svenska materialet resulterade för det första i en svensk enspråkig ordbok, *Svenska ord med uttal och förklaringar* (1984). Innehållet i denna pre-

---

10 Se Tarp (1999:107–108) för en diskussion av termen *invandrarordbok* i förhållande till *inlärningsordbok*.

senteras närmare i avsnitt 3.2.2, under rubriken ”Lexins svenska lexikon 3”. Utifrån det svenska materialet har hitintills ett tjugotal tvåspråkiga lexikon utarbetats. Det har också framställts en övningsbok till Lexin.

Den fritt tillgängliga internetversionen av Lexin har funnits sedan 1995. I tabell 3:2 ges en förteckning över aktuella Lexin-lexikon och deras åtkomstmöjligheter. Uppgifterna är hämtade från Lexins och Språkrådets webbsidor i juli 2016.

TABELL 3:2. *Tillgängliga och planerade Lexin-lexikon (2016-07-22).*

<b>Lexin-lexikon, från svenska till:</b>	<b>Antal lemman</b>	<b>Tryckt version</b>	<b>Internet- version</b>
Albanska	28 500	X	X
Arabiska	28 500	X	X
Azerbajdzjanska	28 500		X
Bosniska	28 500	X	X
Finska	28 500	X	X
Grekiska	28 500	X	X
Kroatiska	28 500	X	X
Nordkurdiska	28 500	X	X
Sydkurdiska	28 500	X	X
Persiska	28 500	X	X
Romska (arli)	28 500	X	X (under arbete)
Romska (kelderash)	–	X (under arbete)	
Romska (lovari)	–	X (under arbete)	
Ryska	28 500		X
Serbiska (latinskt)	28 500	X	X
Serbiska (kyrilliskt)	28 500		X
Somaliska	28 500	X	X
Spanska	28 500	X	X
Thailändska	–	X (under arbete)	
Turkiska	28 500	X	X
Makedonska	17 000	X	
Rumänska	17 000	X	
Tigrinska	7 000	X	X
Amhariska	5–7 000		X
Pashto	5–7 000		X
Turabdinska	4 500	X	
Vietnamesiska	5 000	X	
LSL3 (1995)	28 500	X	
LSL4 (2011)	31 500		X

De första tvåspråkiga lexikonerna som publicerades var ett svensk-turkiskt, ett svensk-serbokroatiskt och ett svensk-kroatiskt. De färdigställdes 1984–85. Dessa var i första hand passiva L1–L2-ordböcker, inte aktiva L1–L2-ordböcker (men kunde ändå användas som sådana). De innehåller t.ex. inga uttals- eller böjningsuppgifter för L2-orden.

Med undantag för det svensk-engelska lexikonet (som sedermera togs bort, se nedan) har ordböckerna riktats till större invandrargrupper i Sverige. Den svensk-engelska ordboken var avsedd som ett komplement för de grupper av invandrare som inte hade någon ordbok men som kunde tillräckligt mycket engelska (Gellerstam 1999:4). Vid urvalet av lexikonspråk har man tagit hänsyn till flera faktorer:

De faktorer som ligger till grund för bedömningen är bl.a. antal personer i Sverige som har språket i fråga som modersmål/umgängesspråk, populationens ålderssammansättning och utbildningsbakgrund, läsfärdighet på modersmålet hos de vuxna, antalet elever som deltar i hemspråksundervisningen samt eventuell tillgång till lexikon till och från andra språk (exempelvis engelska) om dessa språk varit undervisningsspråk i ursprungslandet. (Beijer & Fre Woldu 1997:11.)

Det kommer naturligen alltid att finnas många fler minoritetsspråk än tvåspråkiga Lexin-lexikon. Parkvall (2016) konstaterar t.ex. att det talas runt 200 språk i Sverige idag. För dessa användare är den enspråkiga ordboken, alltså LSL, ett än mer viktigt hjälpmedel.

Lexin-projektet utvecklades senare i separata nordiska Lexin-projekt. I början av 1990-talet ställdes nämligen den svenska databasen till förfogande för de övriga nordiska länderna så att de kunde framställa sina egna Lexin-lexikon. Med tanke på det nära släktskapet mellan svenska, norska, danska och i viss utsträckning även isländska, antog man att en betydande del av Lexin-materialet kunde återanvändas. Man föreställde sig att om man kunde hitta ekvivalenter till det svenska materialet i de främmande språken så borde det också vara möjligt att byta källspråken och återanvända översättningar som redan gjorts till de aktuella målspråken. Det skulle spara både tid och pengar, vilket är av största intresse när det gäller att framställa lexikon, i synnerhet lexikon på invandrarspråk som har relativt få talare om man ser till dem som bor i de skandinaviska länderna. De nordiska Lexin-projekten lyckades rätt så bra, framför allt i Norge (se vidare Holmboe 1999, Bjørneset & Svavarsdóttir 2005:52–53, 63, Hovdenak 2008, Hult et al. 2010 samt temanumret av *LexicoNordica* 6).

Vad gäller den tekniska uppläggningsen av projektet var datatekniken en viktig grundval för arbetet, t.ex. vid insamling av ordförrådet i en databas, inkodning av det svenska ordförrådet, typografering och sättning. Databehandling kom också att utnyttjas i en rad olika arbetsmoment såsom insamling av språkligt referensmaterial, urval av det svenska ordförrådet, urval av delordförråd för olika

stora ordböcker, konsekvenskontroller och utskrifter (Gellerstam 1995:13). Det kan tyckas vara en självklarhet att utnyttja tekniken på detta sätt, men Lexin-projektet påbörjades i slutet av 1970-talet och många av de tekniska möjligheterna vi nu tar för givna var inte tillgängliga då.

### 3.2.2 *Lexins svenska lexikon 3, LSL3*

Den första upplagan av den svenska lexikondelen, *Svenska ord – med uttal och förklaringar* (1984), utarbetades som sagt vid Institutionen för språkvetenskaplig databehandling vid Göteborgs universitet under ledning av Martin Gellerstam. Ordboken har sedan dess utkommit i ytterligare tre upplagor (1992, 1995, 2011). Den senaste uppdateringen genomfördes endast för internetversionen och i samband med detta arbete döptes lexikonet också om från *Svenska ord* till *Lexins svenska lexikon*, LSL (se avsnitt 3.2.3). Som redan framgått används här benämningarna LSL3 och LSL4 för att skilja de två senaste versionerna åt.

LSL är varken ensidigt inriktad på produktion eller reception utan innehåller komponenter som ska stödja båda funktionerna. Behoven av produktion tillfredsställs t.ex. genom språkexempel och syntaxinformation och behoven av reception framför allt genom definitionerna. (De tvåspråkiga ordböckerna är mer utpräglade receptionsordböcker.) Målgruppen för ordboken är som redan påpekats ovana lexikonanvändare och därför är ordboksartikeln ”spatiös, böjningsangivelserna utskrivna, exempelmeningar tydligt markerade och uttal angivet på varje ord” (Gellerstam 1999:5). Det finns också en föreställning om att målgruppsanvändaren inte har någon erfarenhet av studier i främmande språk och ofta också begränsad läsfärdighet även på modersmålet. Men medan de tvåspråkiga Lexin-ordböckerna är riktade till *absoluta nybörjare* i svenska är målgruppen för den enspråkiga svenska ordboken som sagt personer som har ett annat modersmål än svenska och som är *relativt avancerade nybörjare*.

Uppslagsorden i LSL utgörs av följande ordtyper (efter Gellerstam 1999:6):

- a) ett frekvensbaserat skriftspråkligt basordförråd (huvudkälla Allén 1977);
- b) ett ”praktiskt” ordförråd hämtat ur ordindex till ett antal läromedel för invandrarundervisning;
- c) samhällsord, dvs. ord som behövs för att man som medborgare skall kunna orientera sig i det svenska samhället (*moms, daghem, kommunalval* etc.);
- d) diverse vardagliga och talspråkliga ord, ”svåra” ord enligt ordförståelseundersökningar (enligt Frick & Malmström 1983).

Listan täcker in de tre huvudsakliga källor för lemmaurval som tas upp i avsnitt 3.1.2.1. Det första lemmaurvalet bestod av ca 17 000 ord, varav ca 1 000 var samhällsord, och antalet ord har sedan utökats för varje upplaga. Tredje upplagan uppges ha ca 28 500 uppslagsord.<sup>11</sup> Cirka 1 800 ord illustreras och ingår som en bildbilaga i ordboken.

I tabell 3:3 beskrivs de informationskategorier som ingår i LSL.

TABELL 3:3. Informationskategorier i LSL (efter Gellerstam 1999:7).

Formkategorier	Stavning Ordböjning Uttal Grammatisk kommentar
Innehållskategorier	Ordförklaring Sakupplysning Synonymi/Antonymi Bilder
Kontextbeskrivning	Stil Användningsområde Syntaktisk beskrivning
Språkexempel	Meningar och meningsfragment Idiom Sammansättningar

Enligt Malmgren (1999:81) är det inte mycket som skiljer LSL från en konventionell svensk definitionsordbok: ”De inslag som ger ordboken karaktär av ’invandrarordbok’ är, bortsett från förekomsten av ett litet antal speciella samhällsord, främst två: bildtemana och principen att ge presensformen av verben som uppslagsform istället för infinitiven”. Vissa informationskategorier i LSL har emellertid drag som skiljer den från andra (inlärnings)ordböcker. En komprimerad form av artikeln för verbet *vänta* i LSL3 (1995) respektive *Natur och Kulturs Svenska Ordbok* (2001), NoKSO, återges på nästa sida. NoKSO får här representera en traditionell inlärningsordbok (jfr NoKSO 2001:IX–X).

---

11 Antalet lemma är enligt omslagstexten till andra och tredje upplagan 25 000. Siffran 28 500 är resultat av en generös tolkning av vad som räknas som lemma, men är den siffra som brukar anges i Lexin-sammanhang.



LSL3 (1995)

**väntar** [²ven:tar] *väntade, väntat, vänta(!)* verb

**1** låta tiden gå tills något inträffar:

<A väntar (på) B/x>

- vänta på bussen

- vänta på besked

- vänta barn (=vara gravid)

- väntan

**2** tro att något ska hända:

<A väntar x/att+S>

- jag väntade mig inte att du skulle

vara intresserad

- kostnaderna väntas öka

NoKSO (2001)

**vänta**<sup>v</sup> verb <väntar, väntade, har väntat; vara väntad, vänta!>

**1** inte göra något förrän något händer eller någon kommer

• stå i kö och vänta • vänta tills huvudvärken går över

**2** tro eller vara säker på att något ska hända

• Vädret väntas bli bättre imorgon. • De kom tidigare än väntat. • Det var väntat att de skulle vinna.

**vänta sig något** hoppas på något • Vi väntar oss en trevlig kväll.

**vänta barn** vara gravid • Hon väntar barn och ska föda i mars.

**vänta med att göra något** inte göra något nu utan senare • Jag väntar med att ringa till imorgon.

**vänta på någon/något** inte göra något förrän något/någon kommer • vänta på bussen

**vänta 'in någon/något** vänta tills någon/något kommer • Bussen väntade 'in tåget.

**vänta 'ut någon/något** vänta ända tills någon kommer eller något tar slut • vänta 'ut regnet

FIGUR 3:3. *Artikeln väntar i LSL3 och vänta i NoKSO, i komprimerad form.*

Det kanske mest iögonfallande draget är, vid sidan av artiklarnas olika längd, att i LSL3 står verbens uppslagsord som sagt i presensform. Det viktigaste skälet till valet av presensformen som uppslagsform är enligt Gellerstam ”uttryckliga önskemål från invandrarlärare som länge använt presens som ’kom ihåg-form’” (Gellerstam 1999:8, se även Gull & Klintonberg 2004). Ett viktigt argument som lyfts fram är att det utifrån presensformen, en finit verbform, är lättare att producera övriga former av verbet, än om man utgår från verbets grundform.<sup>12</sup>

12 Det finns också en preferens hos infödda talare som interagerar med L2-inlärare att använda

## Argumentet får stöd av Zgusta (1971):

The guiding principle for the choice of the canonical form is that it should be as good a starting point for the construction of the other forms of the paradigm as possible. In the majority of languages, the canonical forms are by now fixed by tradition and the lexicographer will do well not to abandon the tradition unless the reasons for a change are really imperative and the lexicographer is preparing a dictionary which will really command an authoritative influence. (Zgusta 1971:120.)

Det bör dock tilläggas att argumentet att ha presensformen som uppslagsform snarare är produktionsmotiverat än receptionsmotiverat (Malmgren 1999:82, se även Garlén 1985). Eftersom åtminstone de tvåspråkiga Lexin-ordböckerna i första hand är receptionsordböcker är beslutet att välja en från traditionen avvikande uppslagsform i viss mån anmärkningsvärt.

Ordböckerna har vidare valt olika sätt att beskriva uttalsuppgifter. I LSL anges uttal medelst ortografinära transkription. I NoKSO markeras uttalet (sparsamt) i lemmat. Grav accent anges i LSL med siffran två före transkriptionen och i NoKSO med ett upphöjt litet v i slutet av lemmat. Därtill markeras i NoKSO betoning i sublemman som utgörs av partikelverb. Betoningen markeras dessutom i de efterföljande språkexemplen (se t.ex. partikelverbet *vänta in* i NoKSO-artikeln).

Ordklassen skrivs som synes ut i båda ordböckerna men har olika placering – i LSL före böjningsformerna och i NoKSO efter.

När det gäller ordens böjning har LSL3 utöver presensformen, dvs. lemmats form, även preteritum, supinum och därefter infinitiv och imperativ. De två sista formerna upprepas inte utan imperativformen indikeras av ett utrops-tecken inom parentes såsom i exempelartikeln. Om imperativformen avviker från infinitiven skrivs båda formerna ut. För verbet *söker* anges t.ex. formerna som följer: *sökte, sökt, sök!, söka*. I NoKSO skrivs böjningsformerna också ut explicit och är därtill fler. Här anges även perfekt participformen, i det här fallet *väntad*. Framför infinitiv-, supinum- och participformen står därtill småorden *har* och *är* som indikerar vilken böjningsform det är fråga om.

Båda ordböckerna har höga ambitioner att formulera ordförklaringar på en lättillgänglig svenska vilket t.ex. innebär att man försöker undvika långa sammansättningar eller tunga substantivfraser (NoKSO 2001:XIII, se vidare Loenheim 2012). Av artiklarna i figur 3:3 framgår att definitionen i NoKSO är något mångordigare än LSL3. Vidare presenteras och beskrivs lemmats två betydelsemoment på liknande sätt i de två ordböckerna. Utöver ordförklaringar

---

presensformen framför preteritum (se vidare Abrahamsson 2009:175–176).

ges i LSL i vissa fall även sakupplysningar, dvs. encyklopediska uppgifter (jfr diskussionen om språk och sak i kapitel 2). Under lemmat *läkarintyg* finner vi t.ex. efter definitionen en upplysning om att Försäkringskassan kräver intyg om man är sjuk längre än ett visst antal dagar för att man ska få fortsatt sjukpenning.<sup>13</sup> I NoKSO:s artiklar kan den här typen av samhällsinformation också framgå men den markeras inte explicit som i LSL.

Efter definitionerna följer konstruktionsuppgifter i LSL3. De beskrivs med hjälp av döda språkprov (se avsnitt 3.1.2.6). För verben ges uppgifter om så kallade medspelare i satsen (t.ex. objekt, se artikeln *väntar* ovan). För verbet *väntar* i betydelsemoment 1 står det 'A väntar (på) B/x', där A och B står för mänskliga referenter och x och y för ting och abstrakta referenter. För en utförligare beskrivning av ordboksstrukturen i stort, se Gellerstam (1999) och specifikt syntaxmarkeringen, se Gellerstam (1988). Några sådana explicita valensuppgifter återfinns inte i NoKSO.

Därefter kommer språkexemplen. I LSL kan de bestå av satser, satsfragment, idiom, sammansättningar och avledningar (se tabell 3:3). För idiom ges en förklaring inom parentes (se *vänta barn*). Sammansättningar – där uppslagsordet utgör förled eller efterled – nyttjas till att exemplifiera uppslagsordets olika huvudbetydelser. I artikeln för lemmat *affär* ingår t.ex. efterledssammansättningarna *miljonaffär*, *tobaksaffär* och *kärleksaffär* (Gellerstam 1999:8). Artikeln *väntar/vänta* har inga sammansättningar. Däremot innehåller artikeln i LSL3 en avledning i form av ett verbalsubstantiv, *väntan*. I NoKSO skiljs språkexempel i form av satser eller satsfragment och idiomliknande uttryck åt. De senare utgör sublemman vilka markeras med fetstil. De följs av både en förklaring och ett exempel. En stor del av sublemmana är partikelverb och som vi vet framgår i dessa fall också frasens betoning i sublemmat och det efterföljande språkexemplet (jfr avsnitt 3.1.1 om inlärares svårigheter att identifiera flerordskombinationer, i synnerhet partikelverb).

Båda ordböckerna innehåller illustrationer. I LSL återfinns illustrationerna i en bildbilaga med 1 800 bilder ordnade efter teman, t.ex. kroppsdelar, djur och växter, verktyg. I internetversionen finns länkar från artiklarna till bildbilagan. I NoKSO finns 550 teckningar utspridda i hela ordboken (se figur 3:2).

Som vi ska se är LSL4 på många sätt mer lik NoKSO än sin föregångare LSL3.

---

13 I en statisk ordbok som LSL kan det emellertid vara vanskligt att vara alltför precis när det gäller den här typen av sakinformation eftersom den snabbt kan bli inaktuell.

### 3.2.3 Lexins svenska lexikon 4, LSL4

År 2008 fick Lexikaliska institutet vid Institutionen för svenska språket vid Göteborgs universitet, under ledning av professor Sven-Göran Malmgren, i uppdrag att uppdatera LSL3. Arbetet utfördes under åren 2008–2011. Inför uppdateringsarbetet beslutade redaktionen att stärka LSL:s karaktär av inlärningsordbok, t.ex. genom att utforma ordboksartikelns struktur och innehåll på ett än mer explicit sätt. Dessutom skulle man låta receptionsaspekten ha företräde vid de tillfällen då den kom i konflikt med produktionsaspekten, kanske främst av den enkla anledningen att ”man måste förstå ett ord innan man kan använda det aktivt” (Malmgren 2012:456). Ungefär 3 000 nya uppslagsord tillkom vid uppdateringen (med ljudfiler). Vid lemmaurvalsprocessen utgick man bland annat från frekvenslistor innehållande uppslagningar i (internetversionen av) LSL3 som inte gett träff i ordboken. Uppdateringen resulterade i en internetversion och genererade ingen ny tryckt ordbok.

I det följande illustreras uppdateringsarbetet genom en jämförelse av artikeln för ordet *väntar* i LSL3 (se figur 3:4) och motsvarande beskrivning av ordet i LSL4 (se figur 3:5) (efter Malmgren 2012). Artiklarna återges i komprimerad form och överensstämmer inte helt med sitt utseende i webbgrensningen.

LSL3 (1995):  
**väntar** [²ven:tar] *väntade, väntat, vänta(!)* verb  
**1** låta tiden gå tills något inträffar:  
<A väntar (på) B/x>  
- vänta på bussen  
- vänta på besked  
- vänta barn (=vara gravid)  
- väntan  
**2** tro att något ska hända:  
<A väntar x/att+S>  
- jag väntade mig inte att du skulle vara intresserad  
- kostnaderna väntas öka

FIGUR 3:4. Artikeln *väntar* i LSL3, i komprimerad form.

Som framgår av artikeln *väntar* i LSL4 kvarstår presensformen som uppslagsform för verben (av samma skäl som de från början infördes).

Uttalsangivelserna har utökats på så sätt att alla böjningsformer med oförutsägbart uttal har försetts med uttalsuppgifter. Uttalet av samtliga böjningsformer ska framgå genom tilläggsinformation och enkla regler.

Böjningsformerna för verben har som synes blivit fler i LSL4 jämfört med LSL3, och de är också mer explicit beskrivna. Förut angavs infinitivformen till-

sammans med (den eventuella) imperativformen, dvs. *vänta (!)*, som också stod sist i raden av böjningsformer. I LSL4 placeras infinitivformen först. Därefter följer preteritum, supinum, perfektparticip och imperativ, såsom det ser ut i artikeln *väntar*. Småorden *att*, *har* och *är* har lagts till och indikerar vilken böjningsform det är fråga om (en idé inspirerad av *Natur och Kulturs Svenska Ordbok* 2001, se artikeln *vänta* i föregående avsnitt). I de fall verben saknar perfektparticipform och/eller imperativform anges det explicit med en platt-hållarmarkör (–), se *väntar sig*. För substantiven gäller att bestämd form plural numera alltid visas, i motsats till vad som var fallet i LSL3.

LSL4 (2011):

**väntar** [<sup>2</sup>ven:tar] verb, att *vänta*, *väntade*, har *väntat*, är *väntad*, *vänta!*

1 låter tiden gå tills något inträffar, avvaktar:

- *de fick vänta på bussen i tio minuter*

- *han väntade på besked från sin läkare*

<ngn väntar (på) ngn/ngt>

- **väntar barn** är gravid

2 tror att något ska hända:

- *han väntade inte att han skulle komma in på högskolan*

- *kostnaderna väntas öka nästa år*

<ngn väntar ngt/att+S>

**väntan** [<sup>2</sup>ven:tan:] subst. *väntan*, –, –

handlingen att vänta (på någon eller något):

- *i väntan på bussen läste hon en bok*

- *det blev en lång väntan i kön innan biljettkassan öppnade*

**väntar sig** [<sup>2</sup>ven:tarsej] verb att *vänta sig*, *väntade sig*, har *väntat sig*, –, *vänta dig!*

tror att något kommer att hända

- *de hade väntat sig seger men blev besvikna*

<ngn väntar sig ngt/SATS>

FIGUR 3:5. *Artikeln väntar, väntan och väntar sig i LSL4, i komprimerad form.*

Vad gäller definitionerna har flera stora förändringar genomförts. För det första valde man att i verbdefinitionerna använda presens för det första verbet (i definitionen), dvs. 'låter/tror etc.' och inte 'låta/tro etc.' i fallet *väntar*. För det andra har flertydiga ord i definitionerna markerats med en siffra vilken anger vilken betydelse av orden som åsyftas. För det tredje har man infört underavdelningar av betydelsemoment och har som krav att "det alltid ska framgå vilken betydelse ett exempel illustrerar" (Malmgren 2012:549). För det fjärde har ambitionen i arbetet med uppdateringen varit att i möjligaste mån jämte en definition även ha synonymer till uppslagsorden, se *väntar* betydelse 1 i LSL4 som får synonymen *avvaktar*. Det är en anpassning till LSL:s roll som i första

hand en receptionsordbok och utgör också ett bra komplement till de tvåspråkiga lexikonerna. Sist men inte minst har en stor förändring genomförts vad gäller definitionsvokabulären (jfr Pálfi & Tarp 2009). I LSL4 ska alla orden i definitionerna (inklusive sammansättningar och avledningar) finnas som uppslagsord eller åtminstone som morfologiska exempel, och hittas med en enkel klickning från (nästan) vilken böjningsform som helst. Över huvud taget kan man slå på (nästan) vilken böjningsform som helst och komma till grundformen och artikeln. Språket i definitionerna ska också vara enklare och alltför svåra ord undviks i språkproven. Vad som anses vara svåra ord kan dock, som vi redan har sett, i sig vara en besvärlig fråga (se avsnitt 3.1.2).

Loenheim (2012:309) menar att LSL kan nå längre i sin målgruppsanpassning genom att dels ha ett friare förhållningssätt till substituerbarhetsprincipen, dels göra definitionerna än mer konkreta, direkta och lättillgängliga, framför allt med hjälp av längre omskrivningar (vilket ju internetformatets större utrymme inbjuder till).

Morfologiska språkprov, sammansättningar, visar hur grundordet förändras som förled i sammansättningar och de bidrar också till att åskådliggöra betydelsebeskrivningen (Malmgren 2012:459–460). De spelar således en viktig roll i en ordbok som LSL. I LSL3 förekom icke-transparenta sammansättningar som morfologiska exempel, utan att vidare förklaras (t.ex. *gatukök*). I LSL4 återfinns de här språkproven som egna uppslagsord med fullständig artikelinformation.

De syntaktiska språkproven ska stödja och konkretisera definitionerna. Antalet syntaktiska språkprov är större i LSL4, i genomsnitt ett exempel per uppslagsord, och de har också i de flesta fall byggts ut till fullständiga meningar vilket syns tydligt i artikeln *väntar*. Ifråga om autenticitet har kravet på en pedagogisk framställning fått företräde (se avsnitt 3.1.2.5).

Angående konstruktionsuppgifter har notationen för verben och adjektiven förändrats. Som framgår av artikeln *väntar* (betydelse 1) har ’<A väntar (på) B/x>’ ersatts av <ngn väntar (på) ngn/ngt> (jfr Svensén 2004:192). Konstruktionsuppgifterna har dessutom omplacerats och återfinns nu istället *efter* språkexemplen. Den nya notationen är tydligare men beskrivningen har också blivit mindre precis. Det framgår inte längre lika tydligt att de två mänskliga referenterna, i LSL3 A och B, är två olika personer (se avsnitt 3.1.2.6 Konstruktioner). Därutöver kan man undra varför ordboksredaktionen har valt att ha förkortningar i en webbordbok, t.ex. *ngn* (dvs. någon) och *ngt* (dvs. något). Det utrymme som finns till förfogande i en internetordbok, i synnerhet på en datorskärm, kan i det här fallet utnyttjas till att ge ordboksartikeln ett läsvänligare utseende. Man kan tänka sig att förkortningarna helt enkelt har fått stå kvar av gammal vana.

Idiomen har uppgraderats till ett slags sublemman vilket innebär att de står i fetstil och placeras efter valensinformationen, under det betydelsemoment de

tillhör. Idiomen tas nu också upp under alla de ”tunga” orden. En identisk beskrivning av idiomet *väntar barn* återfinns alltså även i artikeln för *barn*. För idiom som innehåller verb, som i exemplet *väntar barn*, gäller att verben står i presens i både sublemmaformen och i definitionen.

I LSL3 är vidare alla äkta sammansättningar försedda med lodstreck (cykel|bana) vilket ”ger användaren värdefulla upplysningar om deras struktur” (Malmgren 2012:457). I LSL4 införs av den anledningen lodstreck även vid sammansättningar med adverb eller preposition som förled (för|aning).

Därtill har avledningen *väntan* i LSL3 avlägsnats från ordboksartikeln och istället fått status som lemma. Detsamma gäller för det reflexiva verbet *väntar sig*. Orsaken till utbrytningen är att de reflexiva verben många gånger avviker starkt från det enkla verbet såväl semantiskt som syntaktiskt. Lemmastatusen möjliggör en mer korrekt beskrivning av dem.

I LSL4 har också partikelverben fått en utförligare beskrivning vad gäller förhållandet mellan lös och fast sammansättning (jfr Källström 2012:177 som redogör för anledningen till att partikelverb kan vara besvärliga utifrån ett inlärningsperspektiv).

I Malmgren (2012) ges en fullständigare beskrivning av uppdateringsarbetet. En omsorgsfull granskning av *Lexins svenska lexikon* är Loenheim (2012).

### 3.2.4 Mer om Lexins internetversion

#### 3.2.4.1 Webbgränssnittet

Den fritt tillgängliga Internetversionen av Lexin-ordböckerna lanserades som sagt 1995 och kompletterades under de efterföljande åren med allt fler tvåspråkiga lexikon. Det svensk-svenska lexikonet gjordes tillgängligt först 2005. Då såg webbsidan ut såsom figur 3:6 visar.

I enlighet med vad som redan sagts om artiklarnas utseende ges först information om uttal, som man även kan lyssna till, böjningsformer och ordklassstillhörighet. Sedan följer en betydelsebeskrivning och som synes har ordet *nyckel* både en bokstavlig och en bildlig betydelse. Det är även möjligt att se en bild på uppslagsordet via länken ”bildtema 10”, där man visuellt kan ta del av i vilka sammanhang en nyckel används. Vidare ges syntaktiska och morfologiska exempel samt sammansättningar och avledningar. I nedre delen av spalten till vänster om uppslagsordet finns länkar till de övriga lexikonerna i Lexin-serien. Så här ser alltså ordboken ut vid tidpunkten för insamlingen av delmaterialet från 2007/2008.



FIGUR 3:6. Webbgränssnitt före uppdateringen: uppslagsord nyckel.

Sökfunktionen i lexikonet har utvecklats av Viggo Kann och sedermera också Joachim Hollman vid Institutionen för numerisk analys och datalogi (NADA), KTH. Webbplatsen har utvecklats av Joel och Jonas Sjöstrand. Skolan för datavetenskap och kommunikation vid KTH står för närvarande för driften av serverdatorn.

Webbsidans utseende har förändrats två gånger. Första gången i anslutning till uppdateringsprojektet. Då såg webbsidans startsida ut som i figur 3:7.



FIGUR 3:7. Webbgränssnitt efter uppdateringen 2011: startsida.



Anledningen till att webbsidan förändrades ytterligare var ändrade förhållanden i Språkrådets organisation. Färgsättningen har förändrats och KTH:s logotyp har tagits bort samt länken med en bild på Lexins övningsbok. I nuvarande skick ser webbsidan ut som såsom förevisas i figur 3:8, med en uppslagning på ordet *nyckel*.

The screenshot shows the Lexin website interface. At the top, there is a header with the text "Lexin – ett samarbete mellan Institutet för språk och folkminnen och Kungliga tekniska högskolan" and a link "[Har du problem med vissa tecken på Iphor Ipad?]". Below the header is a navigation bar with buttons for "Slå upp", "Bilder och filmer", "Nätbokhandel", "Om Lexin", "Kontakt", and "Andra nätlektioner".

The search interface on the left includes radio buttons for "Till", "Från", and "Båda", a "Välj språk" dropdown menu, and a search input field containing "nyc". A dropdown menu below the search field lists suggestions: "nyckel", "nycklar", "nyck", "nyckelpiga", and "nyckelben". To the right of the search field, there are buttons for "pl" and "?", and a "nyckel" button. Below the search field is an "Inställningar" button.

The main content area on the right displays the definition for "nyckel":

- nyckel** [ˈnykɛl] LYSSNA subst. (nyckeln, nycklar, nycklarna)
- redskap att öppna eller stänga lås med
- Sammansättningar:
  - nyckel|knippa
  - bil|nyckel
  - dörr|nyckel
- Exempel:
  - han vred om nyckeln och öppnade dörren
- <även bildligt> verksamt medel (1)
- Exempel:
  - hårt arbete är nyckeln till framgång
- <i sammansättningar även> viktig, central
- Sammansättningar:
  - nyckel|person
  - nyckel|roll

FIGUR 3:8. *Det senaste gränssnittet: uppslagsordet nyckel. Hämtad 2016-07-22.*

Så här ser alltså ordboken ut vid tidpunkten för insamlingen av delmaterialet från 2011. I samband med uppdateringen av den svenska ordbasen infördes även nya sökfunktioner på webbsidan. För det första blev många fler ordformer sökbara, inte bara de böjningsformer som anges i artiklarna utan även sådana former som inte anges explicit. Dessutom blev de sammansättningar och avledningar som återfinns i ordboksartiklarna sökbara. Detta har i hög grad förbättrat träffstatistiken (Lexinstatistik 2016, se också kapitel 6). Före uppdateringen möttes man t.ex. av uppgiften att ordet *bilnyckel* inte fanns i lexikonet, trots att det faktiskt gjorde det. Det stod (och står) som ett morfologiskt exempel under uppslagsordet *nyckel*. För det andra infördes en rullista med möjliga lemman som börjar på de bokstäver användaren har skrivit in i sökrutan. För det tredje är alla ord i artiklarna klickbara. Om en användare läser definitionen för ordet *nyckel* och undrar över ordet *redskap* kan hen enkelt slå upp det ordet genom att dubbelklicka på det.

Uppslagningar på homografa former ger träff på samtliga berörda lemman. En sökning på strängen *cyklar* genererar alltså följande sökresultat (se figur 3:9):

- ▣ **cyklar** [ˈsyk:lɑr] LYSSNA verb  
(att cykla, cyklade, har cyklat, är cyklad, cyklat)
- ▣ åker cykel  
VISA FILM
- Exempel:
  - hon cyklar till och från jobbet varje dag
- Konstruktioner:
  - ngn cyklar
- Uttryck:
  - **är ute och cyklar** misstar sig helt
- ▣ **cykel (1)** [syk:el] LYSSNA subst.  
(cykeln, cyklar, cyklarna)
- ▣ ett fordon med två hjul och pedaler som man trampar fram  
BILD SVENSKA
- Sammansättningar:
  - cykel|däck
  - cykel|hjäl
  - barn|cykel
  - racer|cykel
- Exempel:
  - en cykel med fem växlar

FIGUR 3:9. Resultat av uppslagning på söksträngen cyklar i LSL4. Hämtat 2016-11-18.

I LSL4 är som redan framgått fler böjningsformer sökbara, t.ex. substantiv i bestämd form plural. En uppslagning på ordformen *fiskarna*, som inte genererar någon direkt träff i LSL3, resulterar således i LSL4 i direkt träff på två lemman, nämligen *fisk* och *fiskare*. Det utökade antalet sökbara böjningsformer leder följaktligen till att även antalet homografa ordformer som ger träff i LSL är fler.

### 3.2.4.2 Rättstavningsfunktionen i Lexin

Internetversionen av LSL har en rättstavningsfunktion. Om en söksträng inte ger en direkt träff i LSL reagerar rättstavningsfunktionen på ettdera av följande tre sätt:

- (1) Man möts av beskedet ”Ordet x finns inte i lexikonet.”
- (2) Man möts av beskedet ”Ordet x finns inte i lexikonet, men följande liknande ord finns: y, z, å etc.”
- (3) Sökordet korrigeras automatiskt till ett existerande uppslagsord, som slås upp.

Ordförslagen som presenteras i Lexins rättstavningsfunktion är baserade på editeringsavstånd. Editeringsavståndet mellan två ord är det minimala antalet bokstavsoperationer som krävs för att omvandla det ena ordet till det andra. Det finns tre tillåtna bokstavsoperationer: a) ta bort en av bokstäverna i ordet, b) lägga till en bokstav någonstans i ordet och c) byta ut en bokstav i ordet mot en annan bokstav. Ordet *alroitm* omvandlas t.ex. till *algoritm* genom att bokstaven r byts ut mot g (regel 3) och bokstaven r skjuts in efter bokstaven o (regel 2). Kedjan *alroitm* → *algoitm* → *algoritm* visar att editeringsavståndet mellan *alroitm* och *algoritm* är högst 2. Eftersom det inte går att förändra *alroitm* till *algoritm* i en enda bokstavsoperation så är editeringsavståndet mellan orden precis 2 (Kann 2016).

Nästan alla seriösa elektroniska ordböcker av idag har en rättstavningsfunktion. Den kan vara mer eller mindre utvecklad. Vad gäller rättstavningsfunktionen i Lexin finns en del övrigt att önska. Den borde framför allt ta hänsyn till ljudlighet. Om man t.ex. vill slå upp ordet *journalist* men slinter på tangentbordet och istället skriver *\*kournalist* automatkorrigeras man till uppslagsordet *journalist*. Om man däremot är osäker på stavningen och gör en sökning på *\*schornalist* möts man av informationen att ordet inte finns i lexikonet. Det kan tyckas anmärkningsvärt att en inlärningsordbok inte tar hänsyn till ljudlighet. Andraspråksinlärare torde vara särskilt behjälpta av en rättstavningsfunktion som tar hänsyn till uttalsmässiga likheter mellan olika ord(former). En mycket fin rättstavningshjälp har SAOL Plus, cd-versionen av SAOL13 (se vidare Berg et al. 2008a). Den tar hänsyn till såväl editeringsavstånd som ljudlighet.

En annan sak man kan anmärka på är den ordning i vilken ordförslagen presenteras. Man skulle t.ex. önska att de mest relevanta uppslagsorden kommer bland de första förslagen. Ett exempel är söksträngen *twitter* som ger tolv ordförslag varav de två sista är *twittra* och *twittrar*. När ordförslagen är få är det här inget problem, men när förslagen blir 20–30 stycken (eller ännu fler) blir ordningsföljden avgörande med tanke på ordbokens användarvänlighet. I nuläget verkar nämligen de förslag som har editerats minst komma först i förhållande till sökordet, vilket alltså inte behöver betyda att de är relevantast. Det verkar heller inte finnas någon rimlig gräns för hur många förslag som ges (se vidare kapitel 6).

### 3.2.4.3 Användningen av lexikonet

Internetversionen av Lexin har som redan konstaterats funnits gratis tillgänglig sedan 1995. Från början låg det svensk-engelska och det svensk-finska lexikonet uppe. Webb sidan har sedan successivt utökats med de andra Lexin-

lexikonen. LSL tillgängliggjordes 2005. Statistik på antal uppslagningar har samlats in sedan 1997. Just det året gjordes totalt 3,3 miljoner uppslagningar i ordböckerna. Under de första fem åren tiodubblades antalet uppslagningar och 2001 gjordes 31,1 miljoner uppslagningar. 2006 var den siffran uppe i 221 miljoner. Den högsta noteringen hittills (t.o.m. 2015) är 342 miljoner från år 2008. Så länge det svensk-engelska lexikonet fanns tillgängligt utfördes majoriteten av uppslagningarna i detta lexikon. I samband med att det togs bort och flera andra förändringar av webbsidan genomfördes minskade antalet uppslagningar till 135 miljoner år 2010. Målgruppen för det svenska-engelska lexikonet var invandrare som hade engelska som modersmål eller de som haft engelska som undervisningsspråk i sitt hemland. Även om det länge hade funnits en god tillgång på lexikon mellan svenska och engelska så menade man att ”även dessa invandrare och flyktingar behöver den information som finns just i Lexin” (Beijer & Fre Woldu 1997:17). Man menade vidare att många invandrare dessutom hade ett modersmål där antalet som talar detta språk i Sverige var så litet, att det inte var aktuellt med lexikonutveckling. Som sekundär målgrupp såg man vidare svenskspråkiga ungdomar som studerar utomlands och även engelskspråkiga universitetsstuderande som läser svenska upp till högskolenivå, vilka är i behov av den typ av information som tillhandahålls i Lexin-lexikonen. Förklaringen till att man sedan trots allt valde att avlägsna det svensk-engelska lexikonet kan man läsa på Lexins webbsida:

Det engelska lexikonet är i behov av revidering. Men som statlig myndighet har vi begränsade ekonomiska resurser. Vi väljer därför att prioritera minoritetsspråk där det saknas lexikon. I dag finns dessutom flera omfattande engelska lexikon på nätet, vilket gör att behovet inte är lika stort som när engelska Lexin lanserades för femton år sedan. (Lexin Frågor och svar 2016.)

Användningen av Lexin-ordböckerna har numera åter ökat till nivåer som då det svensk-engelska lexikonet ingick i Lexin och 2015 gjordes totalt 330 miljoner uppslagningar (Lexinstatistik 2016). Enligt Domeij & Karlsson (2013:531) är Lexin en av Sveriges mest besökta webbtjänster över huvud taget.

Tabell 3:4 redovisar antal uppslagningar i LSL per månad för åren 2008, 2011, 2014 och 2016, dvs. före och efter uppdateringen samt mer aktuella uppgifter. Uppslagningarna i LSL brukar utgöra i genomsnitt cirka 4–5 % av det totala antalet uppslagningar i Lexin-lexikonen.

Under första halvåret 2011 (se tabell 3:4) råkade webbsidan ut för en systemkrasch vilket är anledningen till de uteblivna uppgifterna i tabellen. Undantaget detta halvår har antalet uppslagningar ökat kraftigt från 2008 till 2014, t.ex. är antalet uppslagningar i april 2014 74 % fler än i april 2008.

TABELL 3:4. *Antal uppslagningar i Lexins svenska lexikon.*

Månad	2008	2011	2014	2016
Januari	617 903	Ingen uppgift	1 261 610	1 437 888
Februari	598 760	222 625	1 223 816	1 626 952
Mars	597 964	243 435	1 368 451	1 538 203
April	675 868	Ingen uppgift	1 176 744	1 609 638
Maj	572 829	Ingen uppgift	1 120 522	1 563 138
Juni	293 354	Ingen uppgift	727 211	944 828
Juli	206 550	278 564	651 761	707 676
Augusti	330 506	491 487	926 972	1 103 075
September	637 968	941 500	1 389 781	1 602 775
Oktober	637 883	890 650	Ingen uppgift	1 630 544
November	655 440	840 238	1 483 272	
December	480 050	618 143	Ingen uppgift	
<b>Totalt</b>	<b>6 305 575</b>	<b>4 526 642</b>	<b>11 330 140</b>	

I jämförelse med andra nordiska enspråkiga internetordböcker är LSL en frekvent använd ordbok. Lorentzen & Theilgaard (2012) rapporterar 194 660 uppslagningar i *Den Danske ordbog*, insamlade under en månads tid. I Nygaard & Fjelds (2008) analys av uppslagningar i *Bokmålsordboka* är det genomsnittliga antalet uppslagningar per månad drygt 910 000. Användningen i form av lexikonuppslagningar har sedan studiernas genomförande med all säkerhet ökat i båda ordböckerna.



## 4. Användarforskning: metoder och resultat

I det här kapitlet diskuteras inledningsvis de metoder som användarforskningen har till sitt förfogande. Fokus ligger sedan på de metoder som nyttjas i denna avhandling, nämligen *frågeformulär* och *observation*. Därefter redogör jag för resultat från användarstudier som använt sig av de aktuella metoderna, dels traditionella undersökningar, dels undersökningar av ordböcker i elektroniskt format med utgångspunkt i internetordböcker.

### 4.1 Möjliga tillvägagångssätt

De som bedriver användarforskning inom lexikografi, dvs. undersöker ordboksanvändare och ordboksanvändning, har inspirerats av den metodik som används inom samhällsvetenskaplig forskning (Lew 2002, Tarp 2008a). Inom samhällsvetenskapen är *surveyundersökning* en etablerad term. Tre vanliga metoder vid surveyundersökningar är *intervju*, *frågeformulär* och *observation* (Denscombe 2016:27–29). En intervju lämpar sig väl om man är intresserad av mer djupgående aspekter av ett fenomen. Det är dock tidsödande att genomföra intervjuer vilket innebär att dessa studier ofta har ett relativt litet antal informanter. Ett stort antal informanter kan man däremot nå med ett frågeformulär eller (systematisk) observation. Dessa två metoder behandlas utförligare i nästkommande avsnitt.

Andra möjliga forskningsstrategier är *protokoll*, *experiment* och *test* (Tarp 2008a:18–21). Med ett protokoll kan man få tillgång till både yttre och inre

aspekter av ett användningsbeteende, t.ex. vilken ordbok en användare väljer i en viss situation och varför hen valde att använda just den ordboken för att lösa sitt språkliga problem. Nackdelen är att den här typen av metod tar mycket tid i anspråk vilket begränsar antalet informanter. Därigenom kan resultaten sällan generaliseras. Vid ett experiment är man intresserad av att pröva en hypotes och/eller undersöka effekten av en viss faktor, t.ex. i hur hög grad konsultation av en ordbok kan hjälpa en användare att tillfredsställa sina lexikografiskt relevanta behov. Test ingår som en integrerad del av experiment (se vidare Tarp 2008a:21).

Numera genomförs forskning många gånger med hjälp av internet. När det gäller surveyundersökningar har internet flera fördelar. Det är tids- och kostnadsbesparande, datahanteringen är smidigare, det medger en bred geografisk täckning och det är dessutom miljövänligt ur ett resursperspektiv (Denscombe 2016:36).

## 4.2 Valda tillvägagångssätt

### 4.2.1 Surveyundersökning genom frågeformulär

Bruk av frågeformulär har länge varit populärt och frekvent använt i lexikografiska användarstudier. De första studierna om användare och deras användning av ordböcker genomfördes med hjälp av just enkäter (Quirk 1973, Tomaszczyk 1979 och Béjoint 1981 för att nämna några).

Fördelen med en enkät, eller ett frågeformulär, är att man kan undersöka många informanter och få kvantitativa data som kan beräknas statistiskt. Datainsamlingen kan dessutom göras på relativt kort tid. Formulären kan distribueras som postenkäter hem i brevlådan eller i elektroniskt format, vanligen som webbaserat frågeformulär. I jämförelse med postenkäter kan frågeformulär i elektroniskt format distribueras snabbt och till en rimlig kostnad. Informanten klickar för svarsalternativen och kan sedan med en enkel knapptryckning skicka in enkäten. Svaren kan därefter automatiskt och med stor precision läsas in i ett statistik- eller kalkyleringsprogram, i synnerhet om man har använt sig av slutna svarsalternativ. Ett webbaserat formulär har dessutom stora fördelar när det gäller design, färg och form som kan varieras i det oändliga (Denscombe 2016:37).

En enkät passar bra till faktafrågor om t.ex. ålder, kön, sysselsättning och frågor rörande människors åsikter. Enkäter har även använts för beteenderelaterade spörsmål, men med begränsad framgång. Hatherall (1984) brukar ci-



teras för att understryka svårigheten i att avgöra sanningshalten i svar på frågor av det slaget (jfr Denscombe 2016:31).

Are subjects saying here what they do, or what they think they do, or what they think they ought to do, or indeed a mixture of all three? Do they all define the categories in the same way – and in the same way as the researcher? When all is said and done, do we not, on this basis, arrive at a consensus on how subjects are likely to behave when faced with a particular questionnaire, rather than authentic data on what they use the dictionary for? (Hatherall 1984:184.)

Den här omständigheten är på intet sätt unik för studier i ordboksanvändning utan gäller för alla studier där informanter är inblandade. Som framgår av citatet tillkommer det besvärande faktum att enskilda individer kan lägga olika betydelse i en och samma formulering. Ett belysande exempel är den relativt vaga formuleringen *högt förtroende* i en fråga som ”Har du högt förtroende för regeringen?”. Sanningshalten i svaren kan alltså vara mycket hög men informanterna kan ha svarat utifrån helt eller delvis skilda tolkningar av betydelsen hos den aktuella frasen.

Enkäter passar alltså mindre bra när människor ska uppge hur de brukar göra eller bete sig i specifika situationer och vid andra minnesrelaterade spörsmål, såvida det inte är ämnet för undersökningen. Det är dessutom svårt att kontrollera med vilken precision och uppriktighet informanterna uppger sina svar. En annan svaghet med ett frågeformulär, i kontrast till en (djup)intervju, är att man ”tenderar att gå miste om djupet till förmån för bredden när det gäller de data som produceras” (Denscombe 2016:59). Forskaren har vidare begränsade möjligheter att förklara eller omformulera en fråga i de fall informanterna inte förstår vilken information som efterfrågas. Det finns heller ingen möjlighet att ställa följdfrågor.

Lew (2002) diskuterar utnyttjandet av frågeformulär vid studier av ordboksanvändning. Han polemiserar bland annat mot Hatherall (1984) som kritiserat metoden för att brista i reliabilitet. Reliabiliteten rör mätningens kvalitet och har att göra med upprepbarhet, pålitlighet (t.ex. att forskaren besitter tillräcklig och rätt kompetens) och tillförlitlighet (t.ex. svarsfrekvens). Lew menar att den metod som Hatherall istället förespråkar, direkt observation, lider av samma brister. Lew (2002:268) förespråkar istället *methodological triangulation/methodological plurality*, dvs. tillämpning av olika metoder samtidigt. På svenska används termen *metodtriangulering* eller *metodkombination* (se vidare avsnitt 4.2.3).

Många enkätstudier har kritiserats starkt för tillkortakommanden i själva genomförandet av undersökningen, t.ex. vad gäller urvalsförfarande och frågeformuleringar, och vidare för att resultat misstolkas och inte kan generaliseras (Hatherall 1984:184, Tarp 2008a:16–17). Ett exempel på en äldre studie som

uppvisar brister i samband med genomförandeprocessen är Barnhart (1962). Barnhart distribuerade 108 enkäter till 99 universitet i USA. Studiens mycket knapphändiga beskrivning lyder:

The teachers were asked to rate six types of information commonly given in college dictionaries according to their importance to the college freshmen. Their replies indicate that the college freshman uses his dictionary most frequently for meaning and almost as frequently for spelling. Pronunciation is third with synonym studies and lists, usage notes, and etymologies far behind. (Barnhart 1962:162–163)

Studien visar uppenbara brister i reliabilitet. Lärare får svara för studenternas användning. I bästa fall har vi att göra med andrahandsinformation, dvs. att lärare har frågat sina studenter vilken information de oftast eftersöker i en ordbok. I värsta fall har lärare fyllt i enkäten endast utifrån sina personliga uppfattningar om studenternas ordboksanvändning. En undersökning där istället validiteten, dvs. relevansen av den insamlade datan, för det givna problemet kan ifrågasättas är Quirk (1973). Han vill undersöka den engelska enspråkiga ordbokens status och sociala inflytande i Storbritannien. Syftet konkretiseras i trettio frågor som handlar om när informanterna senast använde en ordbok, genomsnittlig användningsfrekvens, anledning till den senaste användningen och av vilken anledning ordboken brukar användas. Det är inget fel på frågorna i sig men man kan undra om Quirk utifrån svaren verkligen kan uttala sig om den engelska enspråkiga ordbokens status och sociala inflytande i Storbritannien.

Lew (2002:270) ställer samman ett antal faktorer att beakta när man planerar att genomföra en enkätundersökning:

Do:

- write your questionnaires in the subject's native language
- pay attention to clean, unambiguous graphical layout
- consider in each case whether multiple-choice or open-ended or mixed question format is most appropriate
- decide before the design is complete how the results will be coded and processed
- screen your questions and multiple-choice answers for possible bias
- ask a colleague or two to read through a draft of your questionnaire
- pilot your questionnaire
- allow appropriate time for your questionnaire to be completed

Don't:

- use technical language that subjects might not understand
- use complex syntax
- use negatives in questions
- let page breaks split questions
- put nonessential questions in the questionnaire just because others had them
- give away your own position or preference in any way

Lew konstaterar slutligen att "[u]ltimately, questionnaires are only as good – or bad – as you make them" (2002:270). Nuförtiden är studier av ordboksanvändare och ordboksanvändning på det hela taget mer genomtänkta och användarforskare har, i den mån det går, tagit hänsyn till sådana faktorer som Lew listar.

Att använda ett frågeformulär via internet istället för i pappersformat verkar ha positiv effekt på den information som tillhandahålls av informanterna. Inför planeringen av en ny utgåva av *Macmillan English Dictionary* (MED) genomförde förlaget en omfattande enkätundersökning genom ett webbaserat frågeformulär som gav mer än 1 300 svar. Förlagets erfarenhet var att människor verkar mer villiga att bidra till forskning som bedrivs på internet (Atkins & Rundell 2008:30) än genom det traditionella pappersformatet. Det förutsätter dock att webbsidan har tillräckligt många besökare som därtill är villiga att delta i en enkätundersökning.

#### 4.2.2 Surveyundersökning genom observation

Det finns två huvudtyper av observation: *deltagande* och *systematisk*. Den förra genererar kvalitativa data och den senare kvantitativa (Denscombe 2016:293). Metoden bygger på forskarens direkta observation och inte på vad någon uppger i ett frågeformulär eller i en intervju. Till en observationsstudie samlas data in ute på fältet i verkliga situationer och inte i ett laboratorium. Det som sker hade skett även om ingen hade observerat det. Observationen utförs alltså i en naturlig miljö vilket är själva poängen med denna metod. Det är ett autentiskt beteende som iakttas. Vi har alltså att göra med förstahandsinformation.

Även denna metod har naturligen sina svagheter. Vad gäller observationsstudier är forskarens perception en förutsättning men också en källa till problem. Personliga faktorer kan inverka negativt på forskarens uppfattningsförmåga och i värsta fall leda till att insamlade data blir opålitliga (Denscombe 2016:315). Det som rent objektivt kan observeras riskerar alltså att skymmas av forskarens subjektiva önskan om eller förväntan över vad som ska ske. Ett

vanligt tillkortakommande med observationsstudier är vidare att det bara är just beteendet som kan studeras och inte avsikten med beteendet. I samband med observationsstudier brukar man också tala om *observer's paradox* (Labov 1972:209), dvs. det faktum att metoden förutsätter en närvarande forskare som observerar det som sker men att människor kan ha svårt att agera naturligt om de vet att de blir iakttagna. Det här problemet bortfaller emellertid till viss del i samband med forskning på internet, t.ex. vid registrering och insamling av uppslagningar i en ordbok (jfr Hulstijn & Atkins 1998:14).

När det gäller observationsstudier utgör internet inte bara ett nytt medium att beforska utan det erbjuder också nya möjligheter till datainsamling. För den lexikografiska användarforskningens vidkommande handlar det framför allt om registrering av loggfiler. I Almind (2008:34) redogörs för vad en loggfil är:

En teknisk præcis beskrivelse af en logfil er, at den er en optegnelse over tid af hændelser og transaktioner enten (a) internt i en enhed, kaldet en systemlog, eller (b) mellem mindst to enheder i et netværk, kaldet en transaktionslog.

Enligt Almind (ibid.) är det framför allt *transaktionsloggen* som är av lexikografiskt intresse. Transaktionsloggen innehåller data om vilken webbläsare som används, vilka lexikonuppslagningar som görs och vilka länkar användarna aktiverar och när. För att identifiera enskilda användare och deras ordboksanvändning bör loggfilen registrera följande uppgifter (a.a. s. 37): 1) brukaridentitet, 2) söksträng, 3) sökmetodik och 4) sökväg.

När det gäller (1) *brukaridentitet* kan enskilda brukare identifieras med hjälp av ip-adress<sup>14</sup>, cookie<sup>15</sup> eller inloggning. Att hävda att en ip-adress är detsamma som en användare kan dock vara mycket vanskligt. En ip-adress kan vara fast och då knyts till en specifik dator. Oftast är dock adresserna inte fasta utan kan variera från ett användningstillfälle till ett annat. Datorer kan också dela adress men utifrån uppfattas som en enda enhet. En specifik dator används dessutom ofta av mer än en person. Oavsett om ip-adressen är fast eller inte är det alltså omöjligt att veta hur många unika användare som döljer sig bakom den. Att använda sig av cookies eller inloggning är ändå mer tillförlitliga strategier om man vill identifiera enskilda användare. Det är emellertid inte alltid praktiskt genomförbart, eller ens tillåtet, att använda sig av dem.

---

14 IP står för Internet Protocol. *Ip-nummer*, *ip-adress*, är ett identifikationsnummer som är unikt för varje dator som är ansluten till internet (*Nationalencyklopedin*, IP-nummer).

15 En *cookie* är (inom datortekniken) en textfil med information som läggs in och sparas på den egna datorns hårddisk när man besöker vissa webbplatser på internet (*Nationalencyklopedin*, cookie).

Almind (2008:40) definierar (2) *söksträng* som ”en række tegn, der kan bestå af tal, bogstaver, mellemrum, kommaer etc.”. I en internetordbok skrivs söksträngen in i ett sökfält på webbsidan och sökningen utförs vanligen genom en knapptryckning. Resultatet är en lexikonuppslagning.

Med (3) *sökmetodik* avses aktivering av övriga sökfunktioner, t.ex. klickningar på länkar och andra knappar som webbsidan innehåller.

Slutligen har (4) *sökväg* att göra med datum och klockslag för samtliga typer av transaktioner. Enligt Almind (2008:39) finns det, med vissa förbehåll, goda förutsättningar att identifiera enskilda användare, deras sökbeteende och sökbehov med hjälp av den information som kan fås genom transaktionsloggen (se vidare avsnitt 4.3.3.2).

Loggfiler kan generera ett stort kvantitativt underlag för undersökning av ordboksanvändning. Styrkan med loggfiler är också att det material som genereras i hög grad är autentiskt, vilket har varit ytterst problematiskt att uppnå vid traditionella undersökningar av användningen av tryckta ordböcker. En annan fördel är att forskarens påverkan är liten genom att registrering av användarnas aktiviteter sker utan forskarens (fysiska) närvaro (jfr dock avsnittet om etiska aspekter, 5.1.4).

Loggfilsanalyser fick ett enormt genomslag i början av 2000-talet, inte minst tack vare de Schryver & Joffe (2004). Internetordböcker gav lexikografer nya möjligheter att undersöka det autentiska bruket av ordböcker. Nu skulle äntligen sanningen om hur ordböcker används komma fram: Vilka ord slås upp? Hur ofta? Vilka ord slås aldrig upp? Inom praktisk lexikografi var loggfiler framför allt ett fantastiskt verktyg för att upptäcka lemmaluckor (jfr Nygaard & Fjeld 2008). För den teoretiska lexikografins vidkommande bidrog loggfilsanalyserna till vidareutvecklingen av användarforskning som vetenskapligt undersökningsområde.

Snart visade det sig emellertid att metoden hade begränsningar och nyttan av loggfilsanalys som ensamt verktyg problematiserades. Loggfiler avslöjar ingenting om anledningen till ordboksanvändningen eller vilken specifik typ av information som eftersöks gällande det uppslagna ordet. De kan heller inte avslöja någonting om vad användaren gör med det sökresultat som en uppslagning genererar. I takt med att det utvecklades allt mer avancerade gränssnitt, där användaren t.ex. måste ange i vilket syfte konsultationen sker, kunde man emellertid komma en bra bit längre i analysen (se t.ex. Bergenholtz & Johnsen 2007 som anlägger ett funktionsorienterat perspektiv på loggfilsundersökningar). Användarna var dock i mångt och mycket fortfarande anonyma.

Observationsstudier kan också genomföras med hjälp av ögonrörelsemätning, så kallad *eye-tracking*. Det är en traditionell metod som bland annat har använts inom psykologisk forskning. Eye-tracking kan med fördel användas vid internetbaserade observationsstudier för att undersöka ögonens rö-

relsemönster på bildskärmen, t.ex. i samband med användarstudier om ordboksanvändning. Metoden är tidskrävande, informanturvalet är ofta litet och användarsituationen är mer eller mindre konstruerad, men intressanta kvalitativa data kan utvinnas (se t.ex. Tono 2011, Køhler Simonsen 2011).

### 4.2.3 Metodkombination

Att kombinera olika metoder inom en och samma undersökning är ingen ny företeelse. På senare tid har emellertid detta tillvägagångssätt kommit att betraktas som en forskningsstrategi i sin egen rätt med ett vedertaget namn och vetenskaplig trovärdighet, och inte bara som en kombination av traditionella strategier (Denscombe 2016:229). *Metodkombination* innefattar såväl kombination av forskningsdesign, forskningsstrategier och analysmetoder (a.a. s. 221). Denscombe tar sin utgångspunkt i samhällsvetenskapen. De flesta element i hans forskningshandbok är dock direkt applicerbara på humaniora.

Att använda sig av flera olika material anses fördelaktigt när det gäller kvaliteten på de data som produceras. Vanligen kombineras kvalitativa och kvantitativa metoder, data eller forskning. Genom att kombinera olika metoder för att undersöka samma sak kan man öka datans träffsäkerhet. Om de olika metoderna ger data som överensstämmer med varandra kan forskaren med större övertygelse anta att resultaten är riktiga. Bruket av flera metoder kan dessutom ge en mer fullständig och komplett bild av det som undersöks. Slutligen kan kontrasterande metoder användas för att föra analysen framåt. En metod kan användas för att skaffa fram information som kan undersökas med en annan, kanske ny, metod.

När man kombinerar olika metoder inom en och samma undersökning är det viktigt att förklara ordningsföljden mellan de olika metoderna. Metoderna kan användas sekventiellt (först den ena, sedan den andra) eller simultant (parallellt genomförande). Forskaren bör också ha klart för sig vilken vikt som läggs vid respektive metod. Den ena kanske betraktas som huvudsaklig och den andra som underordnad eller så kan de vara likvärdiga. Detta måste framgå.

Inom metodkombination är en metod inte bra eller dålig i sig utan det handlar om hur användbar den är för den specifika fråga som undersöks. Det finns t.ex. frågor om ordboksanvändning där en enkät, och inte direkt observation, är den mest lämpliga metoden, t.ex. huruvida man använder ordböcker i hemmet, skolan etc. eller vilken inställning man har till (att använda) ordböcker. Omvänt finns det information om sökbeteende där direkt observation ger mer sanningsenliga svar på frågor om hur ordboksanvändning sker, än vad en fråga skulle kunna ge.

Anledningen till att metodkombination används är alltså att man önskar koppla samman metoderna och de data de genererar på ett fördelaktigt sätt. Men det finns nackdelar med de flesta strategier och så även med metodkombination. Att använda sig av fler än en metod kostar mer och tar mer tid i anspråk. Forskaren måste dessutom läsa in sig på fler metoder. Det är heller inte säkert att strategin ger önskat resultat, dvs. att fynden från de olika metoderna bekräftar varandra. Dessutom finns det en risk att åtskillnaden mellan kvantitativa och kvalitativa metoder, data etc. förenklas.

En viktig term inom metodkombination är *triangulering* (jfr Lew 2002). Den har sitt ursprung inom bland annat lantmäteriet där det handlar om att exakt lokalisera en punkt genom att betrakta den från två andra kända positioner. Triangulering har en särskild betydelse inom samhällsvetenskapen och innebär ”en grund för antagandet att det går att förstå ett forskningsämne bättre om det betraktas ur mer än ett perspektiv” (Denscombe 2016:221). Det finns olika sätt som triangulering kan användas på i anslutning till metodkombination. *Metodologisk triangulering mellan metoder* är det vanligaste tillvägagångssättet (jfr Lew 2002). Tanken är då att de ingående metoderna ska vara så olika som möjligt, t.ex. intervju och observation. Ett annat sätt är *datatriangulering* där man använder kontrasterande informationskällor, t.ex. data från olika tidpunkter eller olika informanter. I delstudie 3 kommer jag att använda mig av en variant på datatriangulering (se vidare avsnitt 5.3).

#### 4.2.4 Urval – sannolikhet och representativitet

Ett generellt problem vid användarstudier är svårigheten att uppnå ett *sannolikhetsurval*, dvs. att urvalet är slumpmässigt ur en viss population (Tarp 2008a:23). Endast i en undersökning baserad på ett slumpmässigt urval kan resultaten generaliseras. Orsakerna till att ett sådant urval är svårt att uppnå är att det inte finns tillräcklig information om populationen (vilka användarna är eller hur många de är) för att man ska kunna göra ett sannolikhetsurval. Det kan också vara svårt att komma i kontakt med urvalet genom den konventionella tekniken för sannolikhetsurval (jfr Atkins & Varantola 1998a:21–22). Problematiken kring urvalsförfaranden är emellertid inte unik för humaniora utan drabbar också t.ex. samhällsvetenskap.

Termen *population* har en särskild betydelse i detta sammanhang. Den syftar på ”alla de människor, händelser eller saker som ingår i den kategori som utforskas”, och inte på ett helt lands befolkning (Denscombe 2016:46). Det man ska fråga sig som forskare är alltså vilka människor etc. som kommer att vara relevanta för just den aktuella studien och ge värdefulla data för det specifika ämne som undersöks.

Ett *icke-sannolikhetsurval* innebär att ”forskaren har en viss bestämmanderätt eller valfrihet i något skede av urvalsprocessen” (Denscombe 2016:64). Ett sådant urval kan användas när forskaren finner det svårt eller icke önskvärt att göra ett sannolikhetsurval. En sådan situation kan t.ex. uppstå när internet är objektet för surveyundersökningen. Det finns ett tiotal olika tekniker för att ta fram icke-sannolikhetsurval. Denscombe nämner bland andra vad han kallar *subjektivt urval* (jfr Holme & Solvang 1997:183–184). Det är en typ av urval som är lämplig i anknytning till studier med hjälp av internet och används vid situationer då forskaren har en viss förkunskap om de människor eller företeelser som ska undersökas och medvetet väljer ut de undersökningsobjekt som troligast ger mest värdefulla data (Denscombe 2016:74–75). Ett av de tillfällen då webbaserade surveyundersökningar fungerar bäst är enligt Denscombe när det går att använda självselektion. Det ger nämligen tillgång till människor eller företeelser som är särskilt relevanta för undersökningens syfte. Denscombe beskriver det vidare som att de olika teknikerna leder till urval som är mer eller mindre slumpmässiga. Forskaren måste hitta fram till det mest ändamålsenliga urvalsstrategin utifrån de praktiska förutsättningar som råder (se vidare a.a. s. 85–87).

I majoriteten av de studier som presenteras i kommande avsnitt, om inte alla, har vi att göra med icke-sannolikhetsurval (jfr Tarp 2008a som ger en översikt över tillgängliga metoder inom lexikografisk användarforskning). Så är också fallet i min undersökning (se vidare kapitel 5, Metod och material). Även studier baserade på ett icke representativt urval kan emellertid bidra med relevant information om tänkbara tendenser som kan gälla för större grupper och utgöra ett underlag för kritik och vidare utveckling av tidigare studier och teorier (Sveningsson et al. 2003:66).

### 4.3 Resultat av tidigare användarstudier

Atkins & Rundell (2008:30) nämner två typer av användarundersökningar inom ordboksvärlden – marknadsundersökningar och vetenskapliga undersökningar. De förra genomförs i regel av ordboksförlag och görs sällan allmänt tillgängliga. Det betyder på intet sätt att dessa studier är mindre vetenskapliga men eftersom de är hemliga får forskarsamhället sällan kunskap om dem. Det är alltså i regel undersökningar av det senare slaget som utgör underlag för översikter som denna (jfr dock Holmer & Sköldberg 2016). Användarstudier fokuserar i allmänhet på ordboksanvändare i någon typ av studiemiljö – inte minst för att ordboksanvändning i mycket stor utsträckning sker i utbildningssammanhang – och framför allt på studenter med engelska som främmandespråk.



På senare tid har emellertid andra grupper av användare undersökts, t.ex. översättare och bibliotekarier, och även andra språk. Inte minst gör internetformatet det enklare att nå ut till en mer heterogen grupp användare.

Som redan konstaterats får användaren under 1980-talet en allt mer framträdande roll inom lexikografin som helhet (se kapitel 2). Ditintills hade bara enstaka användarstudier genomförts, t.ex. Grinstead (1915) och Barnhart (1962). Nu börjar det dyka upp allt fler undersökningar av användare och användning av ordböcker. Dessutom efterfrågas ordböcker som är speciellt anpassade till och inriktade på inlärningssituationer. Befintliga en- och tvåspråkiga ordböcker hade inte tidigare framställts utifrån just inlärares behov.

Firstly, it is time to replace impressionistic opinions and speculations by factual evidence based on empirical research. Secondly, such studies have already confirmed the hypothesis that the foreign language learner needs a new interlingual dictionary to meet his particular needs. Thirdly, the user perspective can help determine real needs of real people in contexts of communicative conflict and deficit. Fourthly, the status of the bilingual learner's dictionary must be assessed in relation to the monolingual dictionary and in relation to other learning aids. (Hartmann 1983:199.)

Enligt Svensén (2004:24) är det få ordbokstyper som har utforskats så mycket som enspråkiga ordböcker för L2-inlärning och ”det finns knappast någon aspekt av dessa ordböckers innehåll, uppbyggnad, funktion och användning som inte har behandlats av ett flertal författare”. De flesta användarstudier av inlärningsordböcker handlar om rätt så avancerade inlärare och studenter. Studier med mindre kvalificerade informanter (non-professionals) råder det däremot en brist på (jfr översikten hos Welker 2006).

Fram till slutet av 1990-talet är den empiriska användarforskningen enligt Hulstijn & Aktins (1998) ganska slumpartad. De efterlyser en ”systematic study of the way in which various variables interact when dictionary users consult a dictionary which contains complex information” (1998:16). Genom en mer teoribaserad ansats utformar författarna ett förslag till att göra forskning om ordboksanvändning till ett sammanhängande och systematiskt forskningsfält, i syfte att få mer insikt om den kognitiva processen i samband med att en ordbok används.

Enligt Svensén (2004:534–537) inbegriper de flesta studier någon eller några av följande frågeställningar:

- Vilken ordbok (t.ex. enspråkig eller tvåspråkig) används i vilken situation (t.ex. reception, produktion, översättning)? Vilken ordbok är effektivast?

- Vilka behov av information har användarna? I vilka situationer uppstår t.ex. behovet av att använda en ordbok? Vilken typ av information eftersöks? Vilka andra hjälpmedel används för att lösa språkliga problem?
- Vad har användarna för konsultationsfärdigheter? Hur letar sig användaren fram till det användarna söker efter i en ordboksartikel?
- Vilka attityder och vanor har människor vad gäller användningen av ordböcker?

De ökade intresset för användarstudier återspeglas också i den tongivande Euralex-konferensen som äger rum vartannat år. Från dess start 1983 och fram till och med 2016 har antalet användarstudier stadigt ökat. I förordet till konferensrapporten från den första Euralex-konferensen 1983 står att läsa: "Some of the themes were partly predictable, such as the emphasis in all sections on the needs of the dictionary user" (Hartmann 1984b: Preface). I denna konferensrapport återfinns dock bara ett bidrag som redogör för en användarstudie, nämligen Glyn Hatherall's "Studying dictionary use: some findings and proposals" (1984). Under 1980-talet var det totala antalet användarbidrag i konferensvolymerna bara en handfull, under 1990-talet ökade de till tio bidrag, under 2000-talet uppgick de till 19 bidrag och under 2010-talet handlar det hittills (oktober 2016) om 27 studier i detta ämne. Från och med 1998 har varje konferens tillägnat ordboksanvändning en separat sektion. Och vid Euralex-konferensen i Bolzano 2014 var mottot för hela konferensen "The user in focus".

Det första mycket uppmärksammade bidraget som handlar om användning av ordböcker i elektroniskt format är de Schryver & Joffe (2004). År 2009 hölls den första konferensen som enbart inriktade sig på elektronisk lexikografi, eLex (Granger & Paquot 2010).

I följande tre avsnitt redovisas ett urval av traditionella enkätstudier (4.3.1), traditionella observationsstudier (4.3.2) och studier av ordböcker i elektroniskt format (4.3.3). Mer omfattande översikter över forskning om ordboksanvändning ges t.ex. av Dolezal & McCreary (1999), Nesi (2000b:3–56), Lew (2004:17–54) och Welker (2006).

### 4.3.1 Några traditionella enkätstudier

Welker (2006) är ett relativt färskt översiktsverk (dock författad på portugisiska) över befintliga användarstudier (jfr Dolezal & McCreary 1999, se också Tarp 2010 som recenserar Welkers bok, på engelska). Sammanlagt 220 studier summeras och boken inkluderar också en bibliografi med 450 referenser. Inte mindre än 94 av dessa baseras på frågeformulär. I sin sammanställning skiljer

Welker på studier om modersmålsanvändares användning (19 studier) och andraspråksanvändares användning av ordböcker (75 studier). Bland de förra finner vi t.ex. Quirk (1973), Kipfer (1987) och Benbow (1990). Att just dessa studier brukar återkomma i referenslistor har delvis att göra med att antalet informanter i jämförelse med andra studier är stort (Quirk=220, Kipfer=292, Benbow=497). Studierna av andraspråksanvändares bruk av ordböcker utgör som synes en mycket mer omfattande samling. Här återfinns bland andra Tomaszczyk (1979), Béjoint (1981), Hartmann (1983), Atkins & Varantola (1998a) och Lew (2004). I följande genomgång inriktar jag mig huvudsakligen på dessa. De har fått stort genomslag i senare forskning och är relevanta med tanke på min undersökning.

Den första större studien om icke-modersmålstalares ordböcksanvändning är Tomaszczyk (1979). Hans enkät innehåller 57 frågor om personlig bakgrund, tidigare erfarenhet av språkinlärning, nuvarande språkanvändning, användning av ordböcker och utvärdering av innehållet i ordböcker (a.a. s. 104). Av de 1 000 distribuerade enkäterna återsändes 449 stycken ifyllda (ett icke-sannolikhetsurval med stort bortfall). Informanterna utgörs av dels 284 studenter på universitetsnivå (varav de flesta är polska främmandespråksinlärare), dels lärare och översättare. Enspråkiga ordböcker får bäst omdömen men tvåspråkiga används mest av alla, i synnerhet L2–L1-ordböcker, oavsett vilken inlärningsnivå användarna befinner sig på. Ju bättre kunskap informanterna har i det främmande språket, desto oftare använder de enspråkiga ordböcker (jfr Svensén 2004:23–25). Informationskategorier som eftersöks är i fallande ordning betydelse (av innehållsord), avstavning, stavning, stil och bruklighet samt grammatik i förbindelse med reception och produktion. Majoriteten av informanterna anser vidare att bilder och teckningar skulle underlätta förståelsen av vissa ord. Vid inköp av ordböcker går informanterna i första hand efter antal uppslagsord, i andra hand efter någons rekommendation och i tredje hand på priset. Även andra variabler är inblandade, t.ex. papperskvalitet, författare och layout.

Två tredjedelar av informanterna i Tomaszczyks studie väljer att kommentera sina svar. Ungefär 80 % av kommentarerna rör tvåspråkiga ordböcker. Det vanligaste önskemålet är mer information om allt. Mer specifikt önskas t.ex. fler språkexempel och åtminstone lite information om ords kollokationsbenägenhet. De mer avancerade informanterna verkar veta vad de kan förvänta sig av sina ordböcker och verkar få ut mest av dem, medan informanter på nybörjar- eller avancerad nybörjarnivå inte känner till ordböckerna tillräckligt. De har ofta orimliga och motstridiga krav på dem. Inte bara lexikografer kan ha nytta av studiens resultat enligt Tomaszczyk, utan också språklärare. Och det studenter visar sig behöva är inte uppmuntran till att använda ordböcker utan mer instruktioner i hur man får ut det mesta möjliga av de ordböcker som finns tillgängliga (Tomaszczyk 1979:117).

Åtskilliga av Tomaszczyks resultat bekräftas i senare undersökningar, t.ex. av Béjoint (1981). Béjoint redogör för en enkätstudie med 122 deltagare. Informanterna är franska studenter som läser engelska på universitetsnivå. Studien koncentrerar sig på användningen av enspråkiga engelska allmänordböcker och syftet med undersökningen är att ta reda på användarnas lexikografiska behov och konsultationsfärdigheter. Betydelse är den informationskategori som informanterna oftast är intresserade av och därefter syntaktisk information och synonymer. Informanterna använder i synnerhet ordböckerna i situationer av skriftspråklig reception. I likhet med vad många användarstudier visar har ordböckernas användarvägledning mycket sällan konsulterats. Kodsystém som eventuellt finns i ordböckerna utnyttjas också i liten utsträckning.

Flera frågor i Béjoints studie förutsätter att informanterna minns tillfällen då de inte hittar vad de söker. Många informanter kommer över huvud taget inte ihåg sådana situationer (jfr avsnitt 4.2.1). Kommentarererna kan dock handla om otillfredsställande definitioner och ord som saknas. Slangord, amerikanismer, tekniska ord och sammansättningar är de typer av ord som nämns oftast. Informanterna uppger att de ofta slår upp idiom och ibland encyklopediska eller kulturspecifika ord, förkortningar och slangord. Däremot slår de aldrig upp formord, tabuord, egennamn eller (vad de uppfattar som) vanliga ord. Slutligen använder informanterna bildillustrationer förvånande sällan.

Hartmann (1983) sätter fokus på behovet av specifika tvåspråkiga inlärningsordböcker för andraspråksinlärare. Inlärare av främmande språk förtjänar en egen ordbokstyp som måste vara skild från historiska och terminologiska ordböcker och översättningsordböcker. Hartmann finner det frustrerande hur lite objektiv kunskap vi har om varför, hur och hur ofta ordböcker används i inlärningsprocessen av ett annat språk. De 185 informanterna i Hartmanns studie utgörs av engelska inlärare av tyska men också av lärare. Majoriteten av studenterna har läst tyska i tre eller fyra år. Av studien framkommer att tvåspråkiga ordböcker används i större utsträckning än enspråkiga. Men 40 % (inklusive lärarna) rapporterar tillfällen där en enspråkig ordbok är mer användbar än en tvåspråkig. Mest eftersökta informationskategori är återigen betydelse men också information om grammatik, användning i kontext, stavning och synonymer efterfrågas. I likhet med i Béjoint (1981) eftersöks kulturspecifika och encyklopediska angivelser, vanliga ord samt slangord. Dessutom ägnar också Hartmanns informanter mycket begränsad tid till att lära sig koder, symboler eller förkortningar i ordböcker. Orsaker till missbelåtenhet är framför allt ord och betydelser som saknas och röriga eller alltför långa ordboksartiklar. Det vanligaste förbättringsförslaget är att det ska infogas fler exempel. Avslutningsvis hänvisar Hartmann till Wiegand (1977) som påpekar vikten av att i just inlärningsordböcker i större grad ta hänsyn till inlärarens

behov och fokusera på de verkliga konflikter (jfr Wiegand i avsnitt 2.3.1) som driver människor till att använda ordböcker istället för att endast katalogisera problem som rör lexikografers arbete (Hartmann 1983:196). Det är detta som Tarp (2006) senare diskuterar i samband med extra-lexikografiska situationer, en central term i Århuskolans funktionslära (se avsnitt 2.3.3).

Atkins & Varantola (1998a) är en av de allra första riktigt stora användarstudier sett till antalet informanter. Initialt distribuerades 1 600 enkäter. I studien ingår sedan ett antal av 723 informanter. Även om bortfallet är stort är deltagarna i studien många fler än i tidigare studier. Författarna undersöker ordboksanvändning hos inlärare av engelska som främmande språk. Informanterna fördelar sig på fyra källspråk: franska, tyska, italienska och spanska. De övergripande forskningsfrågorna handlar om hur effektiva ordböckerna är med att hjälpa inlärare i situationer av reception, översättning och produktion, huruvida tvåspråkiga och enspråkiga ordböcker är lika effektiva hjälpmedel, vilka attityder studenterna har till de två ordbokstyperna och huruvida studenterna fått instruktion i hur man använder ordböcker. Undersökningen genomförs med hjälp av tre olika frågeformulär: ett för användarprofilering, ett för att bestämma studenternas kunskapsnivå i engelska och ett för att genomföra testet i ordboksanvändning. Användarprofileringen inbegriper frågor om bland annat modersmål, antal år av engelskstudier, anledningen till studierna, vilka läroböcker som används, vilka ordböcker studenterna äger och vanor i ordboksanvändning (jfr Atkins & Varantola 1998b). Med hjälp av användarprofileringen indelas de 723 informanterna i fyra kategorier utifrån deras uppmätta L2-kunskaper. En ytterligare indelning görs också efter informanternas modersmål. De flesta informanterna har studerat engelska i 5–9 år. De spanska informanterna är i majoritet bland dem som har studerat engelska i 10 år eller mer. Av användarprofileringen framkommer också att 60 % av informanterna aldrig har fått några instruktioner om hur man använder ordböcker. Bland de 14 % som faktiskt har fått noggranna och systematiska instruktioner är de flesta italienska och tyska informanter. De resultat som framkommer av användarprofileringen ställs sedan mot resultaten från ordboksanvändningstestet (se vidare Atkins & Varantola 1998a:27–52).

En senare större användarstudie är Lew (2004). I hans undersökning ingår 712 polska inlärare av engelska. Två frågeformulär används i kombination med ett test. Ett av formulären låter Lew informanternas lärare fylla i, delvis i syfte att ta reda på deras bedömning av studenternas kunskapsnivå i engelska och vilka läroböcker som används. Formuläret innehåller också några frågor om informanternas ordboksanvändning. Det andra formuläret fylls i av informanterna och syftet är även där att ta reda på informanternas utbildningsbakgrund och kunskapsnivå i engelska (jfr avsnitt 5.5 om frågeformuläret i min undersökning). Syftet konkretiseras i frågor som hur många år de studerat engelska,

antal lektionstimmar per vecka och hur de skattar sina engelskkunskaper i förhållande till sina klasskamrater. Majoriteten av informanterna är studenter i högstadie- och gymnasieåldern men även representanter från yrkesskolor och universitet återfinns bland deltagarna. Informanterna med mindre avancerade engelskkunskaper är fler än de som har bedömts ha mycket avancerade engelskkunskaper. Paradoxalt nog är de informanter som har studerat engelska i 4–6 år (eller mer) fler än de informanter som har studerat engelska i 1–3 år. På frågan hur informanterna uppskattar sin engelskanivå i förhållande till sina klasskamrater svarar inte helt oväntat många att de ligger på en genomsnittlig nivå (se vidare Lew 2004:71). I likhet med Atkins & Varantola (1998a) samkör Lew sedan svaren vilket resulterar i att informanterna indelas i fem olika grupper utifrån kunskapsnivå i engelska. Grupperna innehåller ungefär lika många informanter.

I de senare studierna används enkäten som ett underlag till användarprofilering (jfr dock Hatherall 1984) och metoden utnyttjas i kombination med andra tillvägagångssätt, företrädesvis observation och test (såsom i min undersökning). För att vara studier av tryckta ordböcker gäller de imponerande många informanter. Vad man däremot kan ha synpunkter på är de ibland relativt stora informantbortfallen som inte vidare diskuteras.

### 4.3.2 Några traditionella observationsstudier

Viktigare observationsstudier av icke-modersmålstalets ordboksanvändning är t.ex. Hatherall (1984), Atkins & Varantola (1998b), Varantola (1998), Wingate (2002) och Nesi & Haill (2002).

Hatherall (1984) redogör kort för en användarstudie där han kombinerar en enkätundersökning med direkt observation. En grupp främmandespråksstudenter fick i uppgift att översätta en text från främmandespråket till tyska. Undersökningssituationen var i hög grad kontrollerad. Informanterna fick en timma på sig och skulle under tiden ange vilken ordbok och upplaga de använde, vilka ord som slogs upp och vad de letade efter. Efter uppgiften fylldes ett frågeformulär i som innehöll frågor om hur de hade genomfört uppgiften, hur bra de tyckte att det gick, om tidsaspekten inverkade på tillvägagångssättet m.m. Hatherall övergår sedan till att ha synpunkter på vilka aspekter som man bör ha i åtanke vid genomförandet av framtida studier. En viktig aspekt är att informantgrupperna bör vara så homogena som möjligt, t.ex. vad gäller den kunskapsnivå de uppnått i det främmande språket. Av Hatheralls undersökning framgår att de avancerade studenterna oftare använder ordböcker än de mindre avancerade studenterna. Informanterna, främst de mindre avancerade studenterna, visar också på en allmän tendens att översätta texten ord

för ord. Hatherall ifrågasätter slutligen vilken egentlig kunskap lexikograferna har om användarnas behov: "The compiler makes decisions based on a view of himself as user" (1984:183). Ordboksredaktören med sin språkliga och lexikografiska kompetens är dock i de flesta fall inte särskilt representativ för de faktiska användarna, avslutar Hatherall.

Atkins & Varantola (1998b) undersöker avancerade inlärares konsultationsstrategier och vill ta reda på vad som faktiskt sker när man använder en ordbok, t.ex. vilken typ av information som eftersöks, om man hittar den information man söker efter, vad man gör när ordboken inte ger önskad information och när man väljer att använda en enspråkig L2-ordbok. Ambitionen är att observera konsultationsprocessen i en så naturlig användningssituation som möjligt. Informanterna får utföra en översättningsuppgift (till eller från modersmålet) och slå upp de ord de finner nödvändiga att slå upp bland ett urval av 100 (!) ordböcker. Observationen sker genom att en person sitter bredvid varje informant och noterar varje steg av konsultationsprocessen, och motiven bakom de olika besluten i denna process, med hjälp av ett formulär ('recording sheet'). Formuläret innehåller 10 frågor, bland andra: Vilket ord slås upp? Varför slås det upp? Hittas den önskade informationen? Om ja, var (i ordboksartikeln) hittas denna information? Vad gör informanten sedan? En viktig distinktion görs mellan uppslagning och sökning. En uppslagning innebär endast att just ett ord slås upp i en ordbok. En sökning innefattar alla uppslagningar som görs för att lösa ett specifikt språkligt problem. Ett formulär per uppslagning fylls i.

Totalt genomför de 103 informanterna 1 000 uppslagningar fördelat på 574 sökningar. Mer än hälften av sökningarna (57 %) innefattar endast en uppslagning, och en fjärdedel (25 %) innefattar två uppslagningar. Det innebär att omkring 18 % av informanternas söksessioner inbegriper fler än två uppslagningar. Informanternas egen skattning av andelen lyckade sökningar är högre för reception, dvs. översättning L2-L1, och högre för tvåspråkiga ordböcker. Författarna finner bl.a. att inlärare använder en tvåspråkig ordbok tre gånger så ofta som en enspråkig ordbok. I 40 % av de 1 000 uppslagningarna anses uppslagningen misslyckad. Den höga siffran kan inte bara förklaras med brister hos ordböckerna, menar författarna, utan beror också på olämpliga sökstrategier och orealistiska förväntningar hos användarna (jfr Hult 2008b som i sin loggfilsstudie kommer fram till liknande resultat, se avsnitt 4.3.3.2). På det hela taget är emellertid 59 % nöjda med den information de erhåller efter att en sökning har slutförts. De enspråkiga engelska inlärningsordböckerna används oftare för att finna passande kollokationer eller för att kontrollera en översättning, men detta först efter att användarna har använt en tvåspråkig ordbok för att finna en ekvivalent. Det är således den produktionsrelaterade användningssituationen som är i fokus. Liknande resultat framkommer för övrigt hos Varantola (1998). En omständighet som gör intryck på författarna är just att avancerade och skickliga

L2-inlärare i så stor utsträckning använder ordböcker för att försäkra sig om att man gör en korrekt översättning, i synnerhet när det gäller kollokationer.

Wingate (2002) undersöker ordboksanvändning i samband med läsförståelse hos inlärare på fortsättningsnivå ('intermediate students') i tyska i Hong Kong. En jämförelse mellan en traditionell tvåspråkig ordbok och en enspråkig tysk ordbok för främmandespråksinlärare visar att de misslyckade uppslagningarna har olika orsaker beroende på vilken ordbok som användes. Beträffande den tvåspråkiga ordboken har majoriteten av misslyckade uppslagningar sin förklaring i bristfälliga sökstrategier, t.ex. att informanten bara läser första delen av ordboksartikeln eller att hen missförstår eller ignorerar symboler och förkortningar. När det gäller den enspråkiga ordboken beror misslyckade uppslagningar oftast istället på svårigheter med definitionsspråket eller definitionsstrukturen (jfr Bogaards 2003:200). Wingate genomför också en kompletterande undersökning där hon formulerar nya definitioner i Cobuild-format (se avsnitt 3.1.2.4). De jämförs sedan med existerande definitioner i en "tänka-högt"-undersökning och ett kvantitativt experiment. Hon kommer fram till att Cobuild-definitioner är mer effektiva, men bara för inlärare med hög språklig förmåga (jfr Sinclair 2004). För inlärare med begränsad språklig förmåga är dessa definitioner snarare mindre effektiva (Wingate 2002:151–152). De leder inte heller till bättre förståelse när det gäller ovanliga betydelser av polysema ord (se vidare Wingate 2002:98–207). I övrigt leder definitioner som följs av exempel till bättre förståelse av ett ord.

Som redan konstaterats är det besvärligt att observera naturlig ordboksanvändning, inte minst för att bruk av ordböcker ofta är en rätt så privat angelägenhet. Det är normalt inte något man gör i grupp. Nesi & Haill gör i sin studie (2002) ett försök att skapa en undersökningssituation som är mer autentisk. Autentisk innebär här att informanterna tillåts använda vilken ordbok de vill, att de får slå upp de ord de önskar i en egen vald text och i sin egen takt. Datainsamlingen görs under tre år med 89 internationella studenter som läser engelska på universitetsnivå. Syftet med studien är inte att generera statistiskt jämförbara data, utan snarare att få en uppfattning om internationella studenters verkliga användning av ordböcker i situationer av reception. Författarna önskar också ge en översikt över vilka problem informanterna erfar, vilka typer av texter de läser, vilka typer av ord de slår upp och deras attityder till själva konsultationsprocessen. Då bara några få informanter använder sig av tvåspråkiga ordböcker över huvud taget undersöker författarna framför allt användningen av enspråkiga ordböcker. På det hela taget utgörs de uppslagna orden av en typ av ord som studenterna förväntas behärska på den nivå engelskstudier bedrivs. Lärarna tror sig alltså inte behöva förklara de här orden under lektionerna. Det absolut vanligaste problemet i användningssituationen är att informanten väljer fel lemma eller sublemma, och då vanligen på grund



av svårigheter att identifiera ordets ordklassstillhörighet (jfr avsnitt 3.1.1 om svåra ord). De flesta problemen kan alltså föras till informanternas bristfälliga sökstrategier: "[They] did not examine the context of unknown words sufficiently well, and jumped to conclusions regarding word meaning" (Nesi & Hail 2002:295). Men också ordböckernas innehåll och struktur kan förklara varför informanterna stöter på problem, t.ex. genom missledande eller svårttydda definitioner. På det hela taget uppvisar informanterna en tendens att inte vilja medge att de möter några som helst problem i användandet av ordböcker.

Sammanfattningsvis kan man om de traditionella observationsstudierna säga att det är svårt att komma undan en konstruerad användningssituation. Dessa skapade situationer är i synnerhet problematiska för de forskare som vill undersöka just naturligt ordboks beteende.

Vi har hittills fokuserat på användarstudier där tryckta ordböcker har stått i fokus. När elektroniska ordböcker börjar utvecklas riktas förstås uppmärksamheten mot användningen av dessa (se Nesi 1999). Till en början genomförs flera studier som jämför användningen och användbarheten hos tryckta ordböcker i förhållande till elektroniska ordböcker (t.ex. Nesi 2000a, Weschler & Pitts 2000, Koyama & Takeuchi 2004, Kobayashi 2008). Även om det elektroniska formatet har vissa uppenbara fördelar, bland annat genom att själva konsultationsprocessen kan genomföras på mycket kortare tid, visar sig studierna leda till delvis motstridiga resultat (Dziemianko 2012). I Dziemianko (2010) är t.ex. den elektroniska ordboken signifikant mer användbar i både situationer av reception och produktion: "More precisely, the proportion of correctly explained target items was significantly higher in the electronic-dictionary condition than in the paper-dictionary condition" (Dziemianko 2012:323). Dziemiankos resultat står emellertid i kontrast till de slutsatser som redovisas i en liknande studie av Nesi (2000a).

I följande avsnitt ligger fokus enbart på användning av elektroniska ordböcker.

### 4.3.3 Studier av ordböcker i elektroniskt format

Övergången från tryckta till digitala ordböcker, dvs. där slutprodukten är av ett helt annat format, förändrar i grunden de ursprungliga betingelserna för ordboksanvändandet (jfr Lew 2013, Lew & de Schryver 2014). För att få tillträde till internet över huvud taget behöver man naturligtvis ha tillgång till en elektronisk enhet (dator, surfplatta, smarttelefon etc.) som har (bra) uppkoppling. För att söka i en internetordbok behöver man inte kunna alfabetet. I ett sökfält skriver man in det ord man vill slå upp och trycker på sökknappen, alternativt Enter-tangenten. Man behöver heller inte veta exakt hur

ett ord stavas eller hur dess uppslagsform ser ut. De flesta internetordböcker av idag har som tidigare nämnts en rättstavningsfunktion. Det är dessutom ibland möjligt att slå upp ett ord genom att helt enkelt läsa upp det, dvs. genom så kallad röststyrning (se vidare Køhler Simonsen 2014:428, Holmer & Sköldberg 2014:248). Nuförtiden är också i regel ett uppslagsords samtliga böjningsformer sökbara. När det gäller oregelbundna böjningsformer är de sökbara även i de flesta tryckta ordböcker, men i en elektronisk ordbok kan användaren automatiskt, eller genom en enkel knapptryckning, omgående bli dirigerad till rätt ordboksartikel. En annan skillnad är att en uppslagning i en internetordbok många gånger endast resulterat i att det uppslagsord som slås upp visas. I en tryckt ordbok får man som användare tillgång till omgivande ordboksartiklar, dvs. hela uppslaget. Inte sällan utgörs de omgivande artiklarna av besläktade ord till sökordet, vilket kan komma användaren till gagn.

Ordboksanvändningen på internet har på ett dramatiskt sätt även vidgat utbudet av källor till lexikografisk information. En ordbok är inte alls det självklara referensverk till vilket man först vänder sig när man vill veta något om ett ord. Den naturliga startpunkten är för många användare idag en sökmotor, vanligen Google. Googlandet *kan* resultera i en ordbokskonsultation men det kan lika gärna leda till att man försöker lösa det språkliga problemet med hjälp av en helt annan (mer eller mindre seriös) typ av informationskälla (på internet).

Den digitala revolutionen har skapat förutsättningar för användaren att bli mer deltagande i det praktiska lexikografiska arbetet. Den ökade omsorgen om användaren och hens möjlighet att aktivt bidra i utvecklingen av bättre lexikografiska produkter är alltså inte (bara) ett resultat av lexikografers goda vilja. Användaranpassning, *customization*, är möjlig från två huvudsakliga håll. Användaren väljer och ändrar själv bland potentiella lexikografiska data och det sätt varpå de presenteras. Användaren kan också ges möjlighet att lämna egna bidrag (jfr *crowdsourcing*, t.ex. idén bakom Wikipedia). I annat fall ”gör datorn jobbet” genom att följa och logga specifika användares sökningar och sökbeteenden (se vidare de Schryver & Joffe 2004 och Kwary 2012). Kombinationer av detta är också möjliga.

Ambitionen med denna avhandling är att göra en kartläggning av användarna och deras användning av internetordboken *Lexins svenska lexikon*. Det förutsätter ett omfångsrikt material och de valda tillvägagångssätten – enkätundersökning genom ett webbaserat frågeformulär och observation genom loggfilsanalys – lämpar sig således bra för studiens ändamål. I det följande redovisas resultat från studier som har använt sig av dessa metoder. Fokus ligger som sagt på användningen av internetordböcker (se Lew 2015 för en översikt över forskning i användning av internetordböcker). På senare tid har också flera användarundersökningar av ordböcker i appformat genomförts (se t.ex. Køhler Simonsen 2014, 2015, Holmer et al. 2015).

#### 4.3.3.1 Webbenkäter

Användarstudier på internet som utnyttjar frågeformulär är veterligen relativt få till antalet. Ett (vad det verkar) tidigt arbete är den pilotstudie som tillhandahåller jämförelsematerial i min avhandling (Hult 2008a). Pilotstudien följs 2011 upp av ytterligare en webbenkät som utgör ett centralt material i avhandlingen. Samma år genomförs också en webbenkätstudie på den danska ordboksportalen ordnet.dk (Lorentzen & Theilgaard 2012). Denna ordboksportal innehöll vid studiens genomförande tre ordböcker. Användarstudien fokuserar på en av dem, nämligen *Den Danske Ordbog* (DDO). Enkätstudien görs i samband med en loggfilsanalys av samma ordbok (se kommande avsnitt). Studien omfattar 1 082 enkätsvar som samlas in under tre veckor. Det är frivilligt att delta i undersökningen. Frågeformuläret innehåller frågor om användaren (ålder, kön, utbildning etc.), användningen (sökfrekvens, typ av användningssituation, typ av information som efterfrågas etc.) och avslutningsvis en möjlighet att lämna en kommentar. Könsfördelningen hos informanterna är helt jämn. Medelåldern är 37 år och spannet mellan äldst och yngst är brett. Ordboken attraherar således en bred användargrupp. Informanterna är välutbildade, över hälften av dem som är äldre än 25 år har en universitetsutbildning. Eftersom enkäten är frivillig drar Lorentzen & Theilgaard slutsatsen att den lockar till sig informanter som är vana användare vilka också är mer intresserade, välorienterade och engagerade i webbsidan. Över 90 % av informanterna visar sig också vara frekventa användare. Ordboken används på jobbet eller i skolan. Enligt författarna är det också förvånansvärt många som använder DDO i hemmet, speciellt förstagångsbesökarna. Icke förvånande är att informanterna oftast är ute efter information om betydelse och användning samt stavning. Etymologi är också en relativt populär informationstyp för både vana och nya användare medan information om ett ords existens, fasta uttryck och böjning verkar intressera vana användare i större utsträckning. Kommentarsfältet ger vidare värdefull information om vad informanterna saknar i ordboken. Mest önskar sig informanterna fler synonymer, idiom och fasta uttryck, språkexempel, etymologi samt fler fackord och främmande ord.

Enkätstudien följs senare också upp av en kvalitativ användarundersökning genom ett ”tänka-högt”-protokoll med tio informanter. Den bekräftar till viss del den information som framkommer av enkätstudien men ger i synnerhet värdefull information rörande webbsidans design och struktur (se vidare Lorentzen & Theilgaard 2012). Författarna ser också nyttan av att kombinera en enkätstudie som den här med uppgifter om användarnas faktiska sökningar genom loggfilsanalyser. Tillsammans kan materialen ge goda uppslag till förbättringar av ordboken (jfr de Schryver & Joffe 2004).

Relativt nyligen har också användningen av de norska Lexin-ordböckerna (se avsnitt 3.2.1 om Lexin på export) undersökts (Rambøll 2012:1–4, se också Bjørneset 2016). Studien utfördes på uppdrag av norska Utdanningsdirektoratet. Undersökningen riktar sig till tre användargrupper: skolledare, lärare och användare via webbsidan. Endast för den sistnämnda gruppen används ett webbformulär och då genom ett popup-fönster på Lexins webbsida. Formuläret tillgängliggjordes på sex språk (norska, engelska, arabiska, polska, urdu och somaliska). En extern översättningsbyrå anlätades till de icke-norska språken. Studien resulterar i drygt 1 100 enkätsvar insamlade under ca fyra veckor. Av antalet unika ip-adresser (se avsnitt 5.1.2) att döma besvaras enkäten vid ungefär ett av åtta tillfällen som popup-fönstret presenteras. Drygt hälften av enkäterna fylls i på norska, en fjärdedel på engelska och därefter i fallande ordning: polska 93, arabiska 70, somaliska 54 och urdu 2. Enligt studien har den genomsnittlige Lexin-användaren bott i Norge i 2–5 år.

Användningen av den norska internetversionen har ökat kraftigt för varje år sedan lanseringen 2002. Nästan 60 % av informanterna uppger vidare att de använder sig av den nynorsk-engelska ordboken och många saknar den borttagna bokmål-engelska ordboken. I övrigt används mest ordböckerna från norska till arabiska, polska, somaliska, tigrinska och ryska. Lexin används i samband med undervisningssituationer. Brukarna är i hög grad nöjda med webbsidans gränssnitt som de tycker är enkelt att använda. Tillfredsställelsen med Lexin i sin helhet är störst bland brukare som använt Lexin i tre år eller mindre. Det är på intet sätt något anmärkningsvärt resultat. Ju mer norska man lär sig desto större blir behovet av en mer avancerad ordbok. De flesta norska Lexin-ordböcker har inte mer än ca 11 000 uppslagsord. Önskemål inför framtiden inkluderar fler internetbaserade ordböcker, fler nya språk samt att den bokmål-engelska ordboken återkommer. Även fler sökmöjligheter efterfrågas, t.ex. att kunna söka från främmandespråket till norska. I likhet med vad som framgår av de flesta användarundersökningar önskas också ett utökat urval av ord och uttryck, däribland fackord och tekniska termer (jfr Hult 2008a där informanterna ger uttryck för liknande önskemål). Av undersökningen framkommer också att Lexin anses vara ett unikt språkligt verktyg och att det inte verkar finns några liknande hjälpmedel för språkinläring på den norska marknaden. Det finns dock ett behov av att göra Lexin mer känt bland lärare och skolledare, inte minst i syfte att fånga upp presumtiva användare i form av invandrare och språkliga minoriteter.

Müller-Spitzer (2013) redogör för en enkätstudie som har en lite annorlunda utformning än de studier som hittills presenterats. Syftet är att ta reda på i vilka situationer behovet av att använda en ordbok uppstår. Hon vill få en inblick i hur användare reflekterar kring sin användning av ordböcker, i synnerhet med avseende på ”the external conditions and situations in which a

dictionary consultation is embedded” (a.a. s. 2). Situationen likställs med det som Tarp (2008b:44) kallar extra-lexikografiska situationer. Det finns bara få användarstudier som koncentrerar sig på den yttre kontexten i vilken lexikografiskt relevanta behov uppstår eftersom, menar Müller-Spitzer, data av den här typen är svåra att få fram. I webbenkäten uppmanas informanterna att svara på en öppen fråga om de situationer i vilka de skulle använda en ordbok. Författaren önskar samla in data ”in an exploratory way” (2013:3) och ett öppet svarsalternativ erbjuder större anonymitet och kan på så sätt generera mer ärliga svar. Öppna svarsalternativ ger också större möjlighet för informanterna att motivera sina svar. Författaren vill heller inte låsa sig till traditionella lexikografiska termer som *produktion* och *reception* och dylikt. Hon vill istället fokusera på de mer okonventionella omständigheter i vilka ordboksanvändning uppstår. Det är något av ett nytt grepp inom användarforskningen. Enkäten innehåller därtill uppgifter om informanternas personliga, akademiska och professionella bakgrund.

Totalt ingår 684 informanter i Müller-Spitzers studie. Enkäten besvaras framför allt av språkvetare, översättare och språkvetarstudenter och i mindre utsträckning av lärare och inlärare. Kvinnor besvarar undersökningen i större utsträckning än män. Enligt författaren är det större chans att få mer detaljerade och bättre svar på öppna frågor i webbenkäter än i pappersenkäter, i synnerhet när själva kommentarsfältet är stort (Müller-Spitzer 2013:5). Det visar sig också att många informanter bidrar med utförliga kommentarer. I genomsnitt innehåller kommentarerna 37 ord och den längsta kommentaren innehåller 448 ord. Hälften av informanterna har skrivit mellan 15 och 47 ord (jfr kapitel 7, avsnitt 7.4). Kommentarer delas in efter om de innehåller ord som har med textproduktion, textreception eller översättning att göra samt en övrig kategori. De fyra kategorierna överlappar i vissa fall då en kommentar kan innehålla uppgifter som rör mer än en kategori (jfr Hult 2014). Mer än hälften av kommentarerna visar sig handla om textproduktion. Intressant är att hela 41 % av kommentarerna innehåller beskrivningar av användningssituationer som hänförs till kategorin Övrigt. Författaren finner att kommentarerna ger många intressanta och detaljerade uppgifter om användningssituationer. En av de kortare kommentarerna återges nedan. Frågan handlade alltså om i vilka situationer behovet av att använda en ordbok uppstår.

For reading articles online, for writing and translating online, for double-checking dubious Scrabble offerings played on a gameboard in another room, etc. (Müller-Spitzer 2013:5.)

Det här mer förutsättningslösa sättet att närma sig användarna och deras situationer av ordboksanvändning är intressant. Metoden kan uppenbarligen generera relevant information om potentiella ordboksanvändares mycket specifika

användningssituationer. Det verkar dock som om de flesta av kommentarerna likväl kan insorteras under de klassiska beteckningarna produktion, reception och översättning.

I motsats till Müller-Spitzer anknyter Holmer & Sköldberg (2016) till situationer av reception och produktion. De undersöker användningen av den fritt tillgängliga och kommersiella ordboksportalen synonymer.se. Undersökningen sker inom ramen för ett samarbete med företaget som driver ordboksportalen. Synonymer.se är en frekvent använt webbsida med två miljoner sidvisningar per månad (uppgifter från 2013). Under den aktuella insamlingsperioden, en vecka i oktober 2013, är antalet unika webbläsare ca 280 000 och i genomsnitt görs 3,4 uppslagningar per söksession (a.a. s.220). Två typer av material kombineras, webbsidesstatistik genom Google Analytics och material generat från en relativt utforskad metod som bygger på så kallade popup-frågor, dvs. korta frågor som plötsligt visar sig på skärmen. I det här fallet riktas undersökningen mot frekventa användare och deras användning av synonymer.se. Frågeformuläret innehåller fyra frågor. De handlar om typisk användarsituation, lexikografiska behov, nöjdhet med sökresultatet och modersmål. I syfte att få många informanter som dessutom svarar på samtliga frågor väljer författarna att begränsa antalet frågor till fyra. I media rapporteras vid den här tiden om svenskars enkättrötthet och den påföljande risken att alltför stora informantbortfall kan underminera resultaten från statistiska (opinions)undersökningar. Insamlingen resulterar i 1 060 inskickade enkäter. Det visar sig att 70 % av informanterna använder ordboken i samband med att de studerar. Totalt 60 % använder den för att variera sitt språk, dvs. så som många synonymordböcker är tänkta att användas. Anmärkningsvärt är att 30 % av informanterna uppger att de använder ordboken för reception. Användarna av synonymer.se är i regel nöjda med sökresultaten. Författarna reserverar sig dock för att svaren kan präglas av just det sökresultat informanterna har framför sig när de besvarar popup-frågan. En övervägande del av informanterna har svenska som modersmål (80 % av 770 svarande på denna fråga). Av de svenskinlärare som besvarar enkäten använder förhållandevis många ordboken i situationer av reception. Holmer & Sköldberg menar att dessa användare nog skulle vara mer hjälpta av en definitionsordbok. En tid efter undersökningstillfället kompletteras också webbsidan med en sådan ordbok.

#### 4.3.3.2 Loggfilsanalyser

I Almind (2008) ges en utförlig beskrivning av olika typer av loggfiler och hur de kan användas till att undersöka ordboksanvändning (se avsnitt 4.2.2).

Inom praktisk lexikografi har framför allt loggfilernas potential att avslöja lemmaluckor i ordboken lyfts fram (se t.ex. de Schryver & Joffe 2004, Nygaard & Fjeld 2008). Loggfiler kan också utnyttjas till att mäta effekten av revideringar, t.ex. genom att jämföra träffprocenten före och efter en uppdatering. En tidig loggfilstudie som gör just detta är Rogström (2000). Rogström analyserar loggfiler från internetversionen av *Svenska Akademiens ordbok* (SAOB), en omfattande historisk ordbok (dock ännu ej färdigställd). För närvarande (juli 2016) innehåller SAOB ca 466 000 uppslagsord. Syftet med undersökningen är att få ett underlag till en förbättring av SAOB:s gränssnitt. Materialet utgörs av användarloggar från två olika tillfällen under ett år, en dag i februari 1999 och en dag i oktober 1999. Däremellan genomförs vissa förbättringar av webbsidan. Majoriteten av sökningarna i det första materialet görs i SAOB:s enkla sökgränssnitt. Över 50 % av dessa sökningar ger inte någon träff i ordboken. Om man ser till de uppslagningar som sker i det avancerade sökgränssnittet är det hela 68 % som inte ger någon träff. De vanligaste feltyperna är att ordet man söker på inte finns i ordboken, att stavningen är felaktig eller att sökordet står i en böjningsform. Enligt Rogström är det framför allt okunskap om vilken typ av ord man kan förväntas hitta i ordboken och stavningen av orden som är förklaringen till de många misslyckade sökningarna. Som ett resultat av studien ändras gränssnittet. Ett rättstavningsprogram implementeras som ger förslag på rätt stavning av vanliga felstavningar. Av det andra materialet framgår att felprocenten minskat till ca 30 %, framför allt tack vare rättstavningshjälpen. Användarna får i det nya gränssnittet också tydligare information om möjliga förklaringar till att vissa söksträngar inte ger någon träff, t.ex. att ordboken (vid undersökningstillfället) bara innehåller ord mellan A-STÅ.

de Schryver & Joffe (2004) undersöker loggfiler i en tvåspråkig internetordbok mellan ett sydafrikanskt bantuspråk, Sesotho sa Leboa, och engelska. Ordboken innehåller 25 000 poster på Sesotho sa Leboa och 28 000 på engelska, dvs. den är nästan i samma storlek som LSL. En överraskande hög andel av sökningarna består av felstavade ord. Författarna redogör för hur loggfilerna regelbundet utnyttjas under projektets gång. Ett sätt är att frekventa icke-träffar, dvs. sökningar som inte återfinns i ordbokens lemmauppsättning och som alltså visar sig vara lemmaluckor i ordboken, läggs till när ordboken uppdateras. Andelen icke-träffar sjunker från 33 % till 25 % vid en uppdatering av lemmaurvalet. I studien redogör de Schryver & Joffe även mer detaljerat för hur de följer specifika användares sökningar och därmed kan undersöka deras individuella sökstrategier. Författarna lägger därtill upp ett formulär online

där användare kan ge feedback på ordboken. De påpekar till sist hur kontinuerlig analys av loggfiler tillsammans med användarnas feedback kan utnyttjas till att anpassa innehåll och utformning efter brukarnas behov (jfr Bergenholz & Johnsen 2007).

Vidare undersöker Bergenholtz & Johnsen (2005) loggfilerna från *Den Danske Netordbog*, en enspråkig internetordbok med (vid undersökningstillfället) 109 000 lemman. De konstaterar följande:

Analysis of log files reveal exactly which lemmas and which types of information have been requested, and, perhaps more significantly, which lemmas and which types of information have been requested but were not found in the dictionary. (Bergenholtz & Johnsen 2005:117.)

Med hjälp av loggfilerna kan författarna ta reda på vilka typer av problem användarna stöter på när de söker på ett ord och inte får någon träff i ordboken. Andelen icke-träffar är ca 20 %. Loggfilerna visar t.ex. att användarna söker på passiv- och imperativformer av verb utan att få träff. Det finns alltså ett behov av att kunna utföra den typen av sökning och ordboksredaktörer bör därför fundera över om inte dessa former ska göras sökbara. Studien visar också att majoriteten av de 100 vanligaste misslyckade sökningarna beror på felstavningar. Det kan finnas olika orsaker till att man stavar fel, t.ex. slarv, okunskap eller dyslexi. En stor del av felstavningarna bland de undersökta loggfilerna beror enligt författarna på en uttalsenlig stavning (a.a. s. 131). En annan orsak till att man inte får träff är att man inte vet om ett uttryck skrivs som ett eller två ord. Icke-träffarna består också av ord som inte finns, vilket exempelvis kan bero på att ordet har försetts med suffixet *-ning* (*beskrivning*) i stället för det i danskan korrekta *-else* (*beskrivelse*). Användarna verkar också vara osäkra på när de ska använda foge-s.

Ett användningsområde för loggfiler är som sagt upptäckten av lemmaluckor genom att man ser vilka sökningarna som inte får träff. Många av de ord som användarna inte finner i *Den Danske Netordbog* är datarelaterade, finansiella, juridiska och medicinska termer och begrepp. De vanligaste orden inom respektive fackområde, de som redan används frekvent i allmänspråket, bör enligt författarna ingå i ordboken (Bergenholtz & Johnsen 2005:137–138).

de Schryver et al. (2006) redogör för ett projekt där syftet är att ta fram en helt och hållet korpusbaserad swahili-engelsk internetordbok. Till skillnad från traditionella ordböcker lemmatiseras i den här ordboken inte bara de swahiliska ordens stamformer utan man ansätter också frekventa ortografiska former som lemman. Vid undersökningstillfället innehåller ordboken 6 500 poster på swahili och därtill ett engelskt sökindex på 11 500 ord. Författarna vill ta reda på om det utökade lemmaurvalet har positiv effekt på hur väl användare lyckas hitta fram till de ordboksartiklar som eftersöks. Genom att följa och registrera



användarnas aktiviteter på individnivå menar författarna att man kan få mer tillförlitliga data om enskilda användares sökbeteende i en elektronisk ordbok. De skriver:

For the first time, truly unobtrusive data will be collected and true look-up behaviour in an electronic environment will be recorded. (de Schryver et al. 2006:69.)

Ordboken beskrivs som ett pågående projekt som hela tiden förbättras baserat på de uppslagningar och det sökbeteende som syns i loggfilerna. Nya programvaror utvecklas successivt. Över en halv miljon uppslagningar ingår i loggfilstudien. Uppslagningarna jämförs bl.a. med de korpusframtagna frekvenserna. Författarna uppehåller sig i synnerhet vid frågan om användare slår upp vanliga ord. I så fall skulle en korpusbaserad taktik förordas vid lemmaurvalsförfarande. Vid första anblicken verkar det vara just så, dvs. att ord som har höga frekvenser i korpusen ofta slås upp i ordboken. Men det visar sig vara en inte helt entydig bild. När det gäller de 3 000 vanligaste swahiliorden råder det överensstämmelser mellan korpusrankningen och uppslagningsrankningen. Därefter minskar överensstämmelsen radikalt och antalet korpusbelägg ger ingen indikation på att ett ord av frekvensskäl bör ingå i ordboken. Författarna ställer i artikelns rubrik frågan om användarna verkligen slår upp vanliga ord. Svaret är alltså både ja och nej (jfr Verlinde & Binon 2010:1148).

Hult (2008b) redogör för en studie av ett dygns uppslagningar i LSL3. Målet med undersökningen, som är ett slags pilotstudie inför avhandlingen, är att få kunskap om användarnas faktiska sökbeteende och därigenom också få en uppfattning om vad de förväntar sig att LSL3 ska innehålla. Studien avgränsas till att undersöka de sökningar som inte ger någon träff i ordboken. Det visar sig nämligen att andelen icke-träffar är så hög som 41 %. Detta är intressant då ordboken har en specifik målgrupp och en utformning som i hög utsträckning anpassats efter denna målgrupp. Figur 4:1 visar gränssnittets utseende vid tidpunkten för studien.

Antalet uppslagningar den dagen loggfilerna registreras uppgår till nästan 17 000, varav alltså 41 % utgörs av icke-träffar (jfr de Schryver & Joffe 2004, Bergenholtz & Johnsen 2005).

För det första redogör Hult för de olika slag av sökningar som resulterar i icke-träff. Över hälften av icke-träffarna utgörs av rättstavade enkla ord, avledningar och sammansättningar. Användare söker på såväl grundform som pluralform och bestämda former. De utländska ord som återfinns i loggfilerna är till 90 % engelska.



FIGUR 4:1. Exempel på uppslagsordet nyckel. Hämtad 2007-08-28.

Intressanta är de så kallade encyklopediska och googleinspirerade sökningarna, t.ex. geografiska platser och namn på artister (*Tanzania, Aretha Franklin*) respektive flerordssökningar där användaren verkar vara ute efter information om något specifikt snarare än en definition av den sökta strängen (*VM i innebandy 2006, de fyra grundlagarna*), en företeelse som också lyfts fram av de Schryver et al. (2006:71). Andra söksträngar som ofta resulterar i icke-träff är tvåordslemman och mer eller mindre fasta konstruktioner som t.ex. idiom och kollokationer samt tillfälliga konstruktioner. Enligt Hult är felaktigt särskrivna uttryck färre än man hade kunnat förvänta. Författaren drar slutsatsen att den höga andelen icke-träffar dels beror på det relativt låga lemmaurvalet (omkring 28 500), dels på att urvalet är gjort med tanke på nyanländas lexikografiska behov. LSL3 verkar attrahera en mycket mer heterogen grupp än så.

Hult diskuterar också det höga antal ordförslag som många gånger presenteras för användaren vid en icke framgångsrik sökning. Hon efterlyser ett mer avancerat verktyg baserat på såväl editeringsavstånd som uttal, för att kunna uppnå ett mer relevant urval av förslag på ord som användarna kan ha menat (jfr Měchura 2008:1298, se även Lew & Mitton 2011 om rättstavningshjälpens funktionalitet i några stora engelska inlärningsordböcker på internet). Loggfilerna indikerar också att användarna tenderar att upprepa sina sökningar på olika sätt, i de fall deras första sökning leder till en icke-träff. De två vanligaste strategierna är att pröva en annan stavning (ortografisk ändring) och att ta bort eller lägga till en ändelse (morfologisk ändring) (jfr Měchura 2008:1299 och Rogström 2000:260).

Underlaget till studien består av uppslagningar från en dag och får anses som relativt litet. Man kan heller inte med säkerhet knyta de upprepade sök-

ningarna till en och samma person eftersom det inte var möjligt att identifiera varje specifik användare. Det är dock uppenbart att användarna har höga förväntningar på ordbokens innehåll.

Nygaard & Fjeld (2008) presenterar en loggfilsstudie av internetversionen av *Bokmålsordboka*, en norsk enspråkig allmänordbok som är gratis tillgänglig. Vid tidpunkten för undersökningen görs ca en miljon uppslagningar per månad. Loggfils materialet består av drygt 10 miljoner uppslagningar mellan januari och november 2007. Uppemot 60 % av uppslagningarna ger träff. Ytterligare en dryg miljon uppslagningar görs på böjda former där användarna automatiskt hänvisas till en grundform. De kvarvarande uppslagningarna, nästan en tredjedel, leder dock inte till någon träff. Författarna fokuserar på 500 slumpmässigt utvalda misslyckade sökningar och undersöker vad som kan tänkas ligga bakom dem. Orsakerna till icke-träff sammanfattas i fem typer: stavfel, lemmaluckor, fel språk eller målform, sär- eller sammanskrivningar och egennamn (a.a. s. 62). Klart vanligast är stavfel, framför allt bruk av enkel konsonant istället för dubbel, sammanskrivning av flerordsuttryck och felaktig vokal vid det öppna e-ljudet. Näst vanligast är rättstavade norska ord som alltså utgör lemmaluckor. Det kan vara regelrätta avledningar (som lätt skulle kunna hänvisa till grundordet), självförklarande sammansättningar och sammanskrivningar. Det kan också vara äkta lemmaluckor i form av nyord. De flesta icke-träffarna som utgörs av utländska ord är helt enkelt sökningar på nynorska ord. Vad gäller felaktig sär- eller sammanskrivning är t.ex. adverbialfraser som består av prepositionsuttryck vanliga, t.ex. *imorgen*. Icke-träffar som utgörs av proprier visar att många användare inte har klart för sig skillnaden mellan ordbok och lexikon.

Analysen av de misslyckade sökningarna leder fram till förbättringsförslag. Enligt Nygaard & Fjeld kan man för det första inkorporera en hänvisningsfunktion för att vägleda användaren rätt, t.ex. från avledningar till det icke avledda grundordet som ingår i lemmaurvalet. För det andra går det att implementera en rättstavningskontroll. Den bör anpassas efter vanliga norska stavfel. Slutligen kan analysen av icke-träffar ge goda uppslag till nya lemman.

Tidigare beskrevs loggfiler och dess potential med utgångspunkt i Almind (2008, se avsnitt 4.2.2). I samma artikel ger Almind också ett par exempel på hur själva analysen av en loggfil kan utföras. Loggfiler från en dansk idiomordbog, *Ordbogen over Faste Vendinger*, registrerades under ett års tid. I ett av Almind exempel visar sig ip-adressen leda till en bostadsrättsförening (Almind 2008:47). Det innebär att det är osäkert om det är tal om en enda person bakom användningen. Den påföljande analysen av söksträngar, sökmetodik och sökväg som knyts till den aktuella ip-adressen ger dock vid handen att det sannolikt handlar om just en enskild användare. Almind identifierar 11 sökvägar fördelat på 30 lexikonuppslagningar (Almind 2008:37–42). Den första

sökvägen är från 8 mars 2007 och den sista från 8 december samma år. Det visar sig t.ex. att just den här personen nästan uteslutande gör *en* uppslagning per konsultationstillfälle, att hen ofta söker i webbsidans standardläge (för receptionssituationer) och inte använder sig av den utökade sökfunktionen (för produktionssituationer).

I det andra loggfilsexemplet pekar ip-adressen mot en dansk reklambyrå. Företaget har många anställda men Almind (2008:50) tolkar det som att de 200 sökningar som analyseras hör till en och samma person. Hen är förmodligen professionell textförfattare. Hen ger intryck av att vara van vid att använda sökmaskiner och verkar också vara klar över sina förväntningar av ordboken. Personen ifråga aktiverar t.ex. aldrig användarhandledningen. Uppslagningarna sker snabbt, nästan hektiskt, och är tematiskt sammanhängande (exempelvis söksträngarna *grund, hus, bygge* och *landskap*). Hen arbetar koncentrerat och målmedvetet, ingen sökväg varar t.ex. längre än 30 minuter. Alla sökningar görs mellan ca 10 och 16 på dagen. Denna förmodade textförfattare verkar dock inte använda ordboken efter lexikografens förväntningar. En författare skriver texter och befinner sig alltså i en produktionssituation. Ändå används bara standardläget som är tänkt att användas i situationer av reception. Almind spekulerar över om det kan bero på att den utökade sökfunktionen tar längre tid och ger för många sökresultat. Det har en yrkesman på en reklambyrå med överhängande deadline inte tid att vänta på eller ta till sig. Almind medger att analysen av loggfilerna och därigenom användares sökbeteende förutsätter att forskaren är såväl flitig som fantasifull i sin analys. Det framgår emellertid klart att man med den här typen av tidsbaserad loggfilsanalys kan komma närmare enskilda brukares sökbeteende (och få en upplevelse av att, osynlig, se dem över axeln) än vid rent statistiska analyser av frekventa uppslagningar.

Tidigare presenterades en enkätstudie av Lorentzen & Theilgaard (2012) om användarna av *Den Danske ordbog*, DDO (se avsnitt 4.3.3.1). Den genomfördes i kombination med en undersökning om själva användningen av ordboken. Data samlas in vid två tillfällen, september 2010 och januari 2012, genom verktyget Google Analytics och från ordbokens loggfiler. Författarna är bl.a. intresserade av att få veta hur användarna når fram till ordbokens webbsida, t.ex. genom att undersöka vilka söksträngar som effektivast leder fram till ordboken. Det visar sig att metatermer (t.ex. *ordbog*) är effektivast men vanligast är att användarna söker på de ord eller strängar av ord som de faktiskt vill slå upp i ordboken. Dessa söksträngar jämförs sedan med söksträngarna från ordbokens sökfält för att ta reda på om användarna söker på ett annorlunda sätt i ordboken än i sökmotorerna. Det visar sig att söksträngarna skiljer sig åt. Söksträngarna i sökmotorerna består till 40 % av mer än ett ord. Motsvarande siffra för webbsidan är 4 %. Vidare är andelen misslyckade sökningar ca 20 % och utgörs också här till största delen av stavnings- och skrivfel och böjnings-

former. I övrigt återfinns bland icke-träffarna i fallande ordning främmande ord, egennamn, sammansättningar, nyord, flerordsuttryck, förkortningar och avledningar (jfr Nygaard & Fjeld 2008:63–70 och Hult 2008b ovan). Mot bakgrund av den första loggfilsanalysen genomförs tre förändringar av DDO:

- (1) Samtliga ordbokens sidor indexeras i sökmotorer, istället för endast ett urval. Det innebär att man via Google kan hänvisas direkt till ett uppslagsord och inte först till ordbokens ingångssida.
- (2) Fler böjningsformer görs sökbara, t.ex. presens participformer av verb och genitivformer av alla substantiv.
- (3) Sökresultatet i ordboken ges en mer central position för att användarna tydligare ska uppfatta förslagen som ges vid en icke-träff.

Huruvida dessa förändringar leder till önskat resultat undersöks vid det andra insamlingstillfället. Det visar sig att den utökade indexeringen slår igenom stort. De dagliga besöken ökar från omkring 3 000 till uppemot 20 000. Från att hälften av besökarna når ordboken genom en sökmotor kommer nu 84 % av användarna från en sökmotor. Även de adderade böjningsformerna har effekt – användarna får nu antingen träff på böjningsformerna eller hänvisas till andra förslag på uppslagsord.

de Schryver et al. (2006) ställer sig frågan om användare verkligen slår upp vanliga ord. Wolfer et al. (2014) följer upp detta spörsmål och polemiserar delvis mot de Schryver et al. De jämför loggfilerna från två olika typer av tyska internetordböcker, varav den ena är tyska *Wiktionary*, med en tysk referenskorpus (jfr Trap-Jensen et al. 2014 för en liknande dansk jämförelse). Wolfer et al. slår för det första fast att vanliga ord i korpusen ofta slås upp i ordböcker. Men de visar också hur ords tidsmässiga sociala relevans ("temporal social relevance") påverkar ordens skiftande frekvenser i uppslagsloggarna. Författarna identifierar tidpunkter då vissa ord slås upp ovanligt ofta (relativt frekvenser över tid) och fokuserar på ord som har kortsiktiga söktoppar. De kan vara ord som uppmärksammas och diskuteras i sociala medier, t.ex. i anknytning till politiska händelser, men inte enbart. Wolfer et al. menar att detta är en viktig faktor att ta hänsyn till och föreslår en metod för lemmaurval som tar hänsyn till korpusbeläggens frekvensmässiga växlingar och trendkänslighet.

#### 4.3.4 Sammanfattning

I detta kapitel konstateras inledningsvis att de som bedriver användarforskning inom lexikografi, dvs. undersöker ordboksanvändare och ordboksanvändning, har inspirerats av den metodik som används inom samhällsveten-

skaplig forskning (jfr Lew 2002, Tarp 2008a). De valda tillvägagångssätten i min undersökning är (webbaserat) frågeformulär och observation (genom registrering av loggfiler), två vanliga metoder vid surveyundersökningar. Frågeformulär lämpar sig väl för faktafrågor och frågor rörande människors åsikter. Beteenderelaterade spørsmål undersöks i sin tur med fördel genom direkt observation. Registrering av lexikonuppslagningar i en ordbok är på så vis ett utmärkt sätt att få tillgång till autentiska uppgifter om ordboksanvändning på internet. En tredje metod som används i denna undersökning är metodkombination genom datatriangulering. Det innebär i det här fallet att data som genereras från de förstnämnda metoderna kombineras och genererar ett tredje material. Syftet med datatrianguleringen är huvudsakligen att få en mer komplett och nyanserad bild av användningen av LSL.

Användarstudier har i allmänhet fokuserat på ordboksanvändare i någon typ av studiemiljö – inte minst för att ordboksanvändning i mycket stor utsträckning sker i utbildningssammanhang – och framför allt på studenter med engelska som främmandespråk. På senare tid har emellertid andra grupper av användare undersökts, t.ex. översättare och bibliotekarier, och även andra språk. Inte minst gör internetformatet det enklare att nå ut till en mer heterogen grupp användare.

Utmärkande för min undersökning är att den ordbok som undersöks riktar sig till inlärare på nybörjarnivå som lär sig svenska som ett andraspråk, och inte som ett främmandespråk.

## 5. Metod och material

I detta kapitel beskrivs först processen kring undersökningens datainsamling, dvs. val av metod, insamlingsförfarande och metodologiska överväganden. Därefter beskrivs de tre material som senare ligger till grund för de tre resultatkapitlen. Materialinsamlingen har gjorts med hjälp av Viggo Kann, professor i datalogi vid KTH, Stockholm, och webbansvarig för Lexins lexikon. Materialbearbetningen har sedan genomförts med hjälp av Christian Sjögreen, systemutvecklare vid Lexikaliska institutet, Institutionen för svenska språket vid Göteborgs universitet.

Att kombinera olika metoder inom ramen för en och samma undersökning har förordats av flera lexikografer, t.ex. Lew (2002:268) och Tarp (2008a:27) (se avsnitt 4.2.1). I enlighet med traditionella forskningsstrategier används i denna undersökning en kombination av två metoder för att undersöka dels *användningen* av LSL, dels *användarna* av LSL. Det görs medelst metoderna loggfilsanalys och webbaserat frågeformulär. Därtill använder jag mig av metodkombination genom datatriangulering. Denna strategi innebär att data från de nämnda metoderna länkas samman och genererar ett nytt material. Ordningföljden är parallell och de två metoderna är likvärdiga (se avsnitt 4.2.3 Metodkombination).

## 5.1 Metod i delstudie 1: Loggfilsanalys

### 5.1.1 Insamlingsförfarande

Delstudie 1 syftar till att undersöka den autentiska användningen av LSL före och efter uppdateringen 2008–2011 (se avsnitt 3.2.3). Metoden som används är loggfilsanalys. En mycket stor fördel med den här metoden är att det finns en likhet mellan undersökningssituationen och en verklig brukssituation, representerad av faktiska användare och deras aktiviteter på LSL:s webbsida.

Undersökningen baserar sig på loggfiler insamlade vid två tillfällen, 2008 och 2011. Som framgår av Almind (2008:34, se även avsnitt 4.3.3.2) finns det olika typer av loggfiler. I denna undersökning ingår loggfiler som utgörs av dels rena *lexikonuppslagningar*, dels *navigeringsval* på LSL:s webbsida. Jag väljer att kalla dem loggfiler av typ a respektive typ b.

Med *lexikonuppslagningar* menas de söksträngar som användare skriver in i webbgränssnittets sökfält och som slås upp genom att de trycker på knappen ”Slå upp” eller genom att trycka på skrivbordets Enter-tangent (se figur 5:1). Sammanlagt registrerades 359 411 lexikonuppslagningar vid de två insamlingstillfällena.

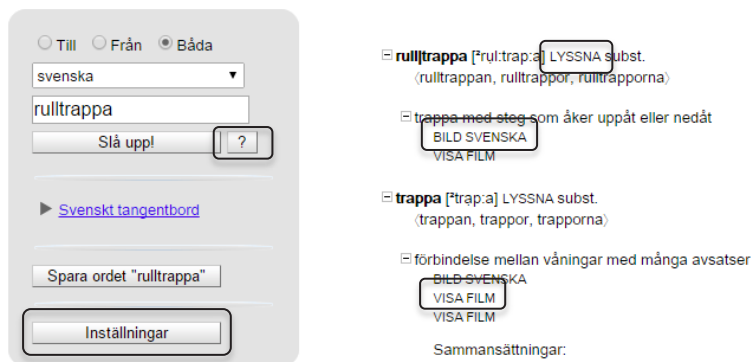


FIGUR 5:1. Exempel på lexikonuppslagning som genererar loggfiler av typ a. Hämtad 2016-05-09.<sup>16</sup>

Med *navigeringsval* avses för det första klickningar i tre speciella fält i sökgränssnittet: uttalsknappen ”Lyssna”, länken till ”Bild svenska” och länken till ”Visa film” (se markeringar i figur 5:2).

<sup>16</sup> Motsvarar LSL4:s utseende och innehåll vid tidpunkten för insamling (se avsnitt 3.2.4).





FIGUR 5:2. *Typer av användaraktiviteter som genererar loggfiler av typ b. Hämtad 2016-05-09.*

För det andra avses de ändringar som användarna gör under fliken ”Inställningar” samt klickningar på frågetecknet ”?” (se markeringar till vänster i figur 5:2). Ett klick på frågetecknet genererar ett popup-fönster innehållande söktips av diverse slag. Under den vecka datainsamlingen genomfördes manipulerades inställningarna och följande informationskategorier klickades bort från ord-boksartiklarna: böjning, exempel, konstruktioner, sammansättningar och uttryck (se figur 5:3).



FIGUR 5:3. *Popup-fönster som visades vid klickning på ”Inställningar” vid undersökningstillfället. Hämtad 2016-08-09.*

I figur 5:4 återges hur artikeln för lemmat *fot* ser ut i standardläget.

**fot** [fo:t] LYSSNA subst.  
(foten, fötter, fötterna)

nedersta delen av benet (1)

Sammansättningar:

- fot|sula

Exempel:

- Kan du stödja på foten?

<även bildligt> nedersta delen av föremål

Exempel:

- huset låg vid foten av berget

Uttryck:

- **en black om foten** ett hinder, en belastning, en börda
- **försätter på fri fot** friger
- **går till fots** promenerar
- **på fri fot** inte fången
- **på stående fot** genast
- **står med båda fötterna på jorden** är realistisk
- **står på god fot med någon** har ett gott förhållande till någon

FIGUR 5:4. Komplet artikel för fot i LSL4. Hämtad 2016-05-25.

I figur 5:5 återges hur samma artikel ser ut när de nämnda informationskategorierna avlägsnats.

**fot** [fo:t] LYSSNA subst.

nedersta delen av benet (1)

<även bildligt> nedersta delen av föremål

FIGUR 5:5. Manipulerad artikel för fot i LSL4. Hämtad 2016-05-25.

Under datainsamlingsveckan informerades användarna om att det pågick en undersökning av LSL och att ordboksartiklarna för tillfället såg annorlunda ut. Meddelandet på Lexins startsida löd:

Just nu (under vecka 40) pågår en undersökning av användningen av *Lexins svenska lexikon*. Därför är inte all information om uppslagsorden synlig när du slår upp ett ord. För att få fram informationen, klicka på rutan ”Inställningar” och kryssa för de uppgifter du vill ska synas.

Syftet med att manipulera inställningarna var att få mer detaljerad kunskap om dels användarnas sökbeteende, dels vilka specifika informationskategorier användarna är intresserade av (jfr avsnitt 7.2.3 och kapitel 8).

Valet av tidpunkt för datainsamlingen påverkades av faktorer som användningsstatistik från Lexins webbsida och förekomst av allmänna helgdagar och skollov. Insamlingen avgränsades till en vecka i början av oktober 2008 respektive 2011 (se vidare avsnitt 5.4).

### 5.1.2 Metodologiska överväganden

De metodologiska övervägandena rör inte så mycket själva valet av metod utan snarare insamlingsförfarande och urvalsprocess. Avhandlingen innehåller som redan framgått tre delstudier. Syftet med loggfilsanalysen är att undersöka användarnas sökbeteende i form av uppslagningar och övriga aktiviteter på webbsidan. Med hjälp av olika frekvenstabeller jämförs lexikonuppslagningarna från 2008 och 2011. Resultaten ska sedan tolkas i förhållande till de data som framkommer av enkätundersökningen. Dessutom länkas de två typerna av data från 2011 samman och genererar ytterligare ett material som ligger till grund för delstudie 3 (se vidare avsnitt 5.3). Vid insamlingstillfället manipulerades som sagt inställningarna på LSL:s webbsida genom att vissa informationskategorier ”klickades ur”. Det innebar att ordboksartiklarna under denna period saknade väsentlig lexikografisk information, såvida inte användarna själva återställde inställningarna (se avsnitt 5.6). Detta är den huvudsakliga anledningen till att undersökningens genomförande begränsades till en vecka. Eftersom Lexin är en statligt finansierad ordbok ansågs det olämpligt att manipulera med webbsidan för mycket eller under för lång tid.

Loggfilsaterialet skulle i princip kunna bli hur stort som helst. Att samla in loggfiler i form av lexikonuppslagningar är nämligen en relativt enkel operation (se avsnitt 4.2.2). I det fall loggfilsanalys hade utgjort den huvudsakliga metoden skulle rimligen ett större urval ha gjorts, och framför allt från flera olika tidsperioder (se t.ex. Nygaard & Fjeld 2008). Analysen skulle då också ha kunnat göras ännu mer omfattande och djupgående. Loggfilerna utgör dock endast en del av materialet för denna avhandling och en mer genomgripande loggfilsanalys ligger därför utanför ramen för undersökningen.

Som framgår av avsnitt 4.2.2 kan man av flera olika skäl inte med säkerhet knyta en ip-adress till en unik användare (se också avsnitt 5.7 om etiska aspekter). Tillsammans med exakta tidsangivelser för uppslagningarna är emellertid sannolikheten många gånger stor att så är fallet (jfr Almind 2008). Den här möjligheten kommer emellertid bara i mycket liten utsträckning att utnyttjas i delstudie 1, medan den utgör en viktigare komponent i delstudie 3 (se avsnitt 5.3).

## 5.2 Metod i delstudie 2: Webbaserat frågeformulär

### 5.2.1 Insamlingsförfarande

Delstudie 2 syftar till att få en bild av användarna av LSL. Då jag (också) önskar nå så många faktiska användare som möjligt vänder jag mig lämpligen till användarna direkt via LSL:s webbsida. Detta sker med hjälp av ett webbaserat frågeformulär. Frågeformulär är som sagt en traditionell och frekvent använd undersökningsmetod, om än inte oproblematisk, vid lexikografiska användarstudier (Lew 2002, Tarp 2008a). Den har visat sig passande också vid forskning på internet (jfr Denscombe 2016:37–39), där stora delar av användningen av ordböcker numera sker.

Materialet utgörs av två uppsättningar enkätdata insamlade vid två skilda tillfällen. Det första materialet samlades in genom att en länk till en enkät lades upp på webbsidan för LSL3, dvs. LSL före uppdateringen, under fyra dagar i slutet av september 2007. Det andra materialet samlades in genom att motsvarande länk lades upp på webbsidan för LSL4, alltså LSL efter uppdateringen, under sju dagar i början av oktober 2011. En väsentlig skillnad mellan de två frågeformulären är att det förra formuläret innehåller 10 frågor och det senare 20 frågor. Frågeformulären återges i sin helhet i bilaga 1 (2007-enkäten) och bilaga 2 (2011-enkäten).

Båda versionerna av frågeformulär skapades i programverktyget ACE. ACE används för att göra kursutvärderingar vid Institutionen för numerisk analys och datalogi (NADA) vid KTH i Stockholm. Programmet har också utvecklats där. Länken till frågeformuläret uppenbarade sig i samband med användarnas inledande uppslagning. Användaren kunde välja att klicka på alternativet ”Hämta enkät” eller ”Nej tack”.

De potentiella undersökningsdeltagarna, undersökningspopulationen, utgjordes av samtliga LSL:s användare under de två insamlingsperioderna. Populationen var alltså helt okänd inför undersökningens genomförande. I efterhand kunde uppgift om ip-adress ha gett en indikation på hur många unika användare det rörde sig om. Uppgift om ip-adress fanns dock endast för materialet från 2011 (se nedan). Deltagandet i enkätundersökningen var frivilligt och urvalet gjordes genom självselektion (Denscombe 2016:38). Hur stort informanturvalet blev påverkades av användarnas villighet och möjlighet att svara. I så måtto blev urvalet positivt då informanterna utgjordes av särskilt intresserade LSL-användare. Dessutom ökade också chanserna att enkäten blev korrekt och fullständigt ifylld. Risker för såväl objektbortfall som partiellt bortfall torde därmed ha minskat (se vidare avsnitt 5.5.2).

I samband med det första undersökningstillfället inkom 370 enkäter. De utgör underlag för en studie som presenteras i Hult (2008a) (se avsnitt 4.3.3.1). Syftet var då att pröva metodens användbarhet och se om användarna över huvud taget fyllde i formuläret. Materialet föreligger inte i något statistikprogram eller liknande utan endast som sammanlagda data från samtliga ifyllda frågeformulär. Jag kan alltså inte se resultaten för varje individuellt frågeformulär vilket förstås begränsar analysmöjligheterna. I föreliggande undersökning fungerar därför det här materialet i första hand som ett betydelsefullt jämförelsematerial.

Vid det andra undersökningstillfället inkom 863 enkäter. De tillgängliggjordes i en komprimerad rådatafil som först laddades ned i kalkylprogrammet Microsoft Excel och därefter överfördes till statistikprogrammet SPSS. Eftersom SPSS kan hantera Excel-filer gick överföringen dessa program emellan smidigt. SPSS möjliggör förstås helt andra analysmöjligheter i jämförelse med vad som är fallet med enkätmaterial från 2007. Antalet ip-adresser under insamlingsveckan uppgår till 25 974 vilket innebär att svarsfrekvensen i relation till ip-adresserna är 3,3 % (se vidare avsnitt 6.1). Det rör sig alltså troligen om färre än 25 974 olika användare. Den egentliga svarsfrekvensen är troligen något högre. Användarna fick nämligen frågan om enkäten en gång per söksession. Det innebär att om de besvarade enkäten vid ett användningstillfälle dök länken till enkäten upp även vid nästa tillfälle de besökte webbsidan. I de fall en användare återbesöker webbsidan från samma ip-adress påverkas inte svarsfrekvensen. Men om en användare besöker webbsidan från flera ip-adresser blir den ip-baserade relativa svarsfrekvensen lägre än den egentliga. Och det omvända sker om flera olika användare fyller i var sin enkät från samma ip-adress. Det resulterar i en högre svarsfrekvens. Detta har dock skett i mycket liten utsträckning. Totalt 23 ip-adresser har fler än ett (nämligen två och i ett fall tre) enkät-id knutet till sig. Det är mindre än 3 % av det totala antalet enkäter. Antalet unika användare är alltså troligen mindre än 25 974 och den egentliga svarsfrekvensen är därför något högre än 3,3 %.

## 5.2.2 Metodologiska överväganden

Den insamlingsmetod som används i delstudie 2 ger som vi sett ett icke-sannolikhetsurval eftersom deltagarna i enkätstudien utser sig själva till informanter (de som har tid och lust att fylla i enkäten). När man, som i det här fallet, önskar undersöka en specifik grupp individer (de som använder LSL, och i synnerhet ordbokens målgrupp) är detta tillvägagångssätt ändamålsenligt (jfr Denscombe 2016:38). Ett icke-sannolikhetsurval gör det dock, som redan konstaterats, strikt tolkat omöjligt att generalisera resultaten. Men vi vet

också att resultaten likväl kan säga något om tänkbara tendenser som kan gälla även för större grupper (Sveningsson et al. 2003:66). Dessutom ger ett stort urval större tillförlitlighet när det gäller de slutsatser som dras (jfr Denscombe 2016:81). Och i föreliggande studie har vi att göra med ett jämförelsevis stort urval enkäter (jfr Welker 2006).

Ifråga om urvalsförfarandet hade ett alternativ förstås varit att distribuera en postenkät till ett slumpmässigt urval bestående av t.ex. 1 000 personer i Sverige. Det var emellertid aldrig aktuellt eftersom ett av avhandlingens syften var att länka informanternas enkätsvar med deras uppslagningar i LSL (se kommande avsnitt). Dessutom skulle antalet relevanta informanter antagligen ha blivit mycket litet. Det potentiella antalet individer som använder LSL, sett till Sveriges totala population, är nämligen mycket litet. Därtill kommer det faktum att enkäter i allmänhet brukar ha ett relativt stort bortfall (SCB Statistik 2015).

Länkningen förutsätter alltså att jag vänder mig direkt till dem som faktiskt använder ordboken. Teoretiskt sett skulle även ett slumpmässigt urval bestående av 1 000 Lexinanvändare kunna ha gjorts utifrån uppgifter om ip-adresser, under den aktuella insamlingsperioden. Det hade emellertid krävt någon slags registrering från användarnas sida i syfte att i nästa steg kunna distribuera enkäterna. Det här hade sannolikt också lett till ett mycket stort bortfall. Det helt avgörande är emellertid att en sådan registrering i vilket fall som helst inte lät sig göras. Lexin är som tidigare påpekats en statligt finansierad webbsida och utgör en gratistjänst och det är inte lämpligt att, om än mycket tillfälligt, kräva registrering för att använda webbsidan. Dels eftersom ordboken i första hand vänder sig till nya svenskar, dels eftersom den tidigare nämnda manipuleringen av innehållet (se avsnitt 5.1.1) faktiskt innebar att väsentlig lexikografisk information saknades. Frågan är också hur användarna hade ställt sig till detta. En del hade säkerligen motsatt sig en registrering och andra hade kanske uppfattat det som krångligt och avstått av det skälet.

Att uppskatta den sannolika svarsfrekvensen är svårt när populationen är okänd. Dessutom brukar svarsfrekvensen vid surveyundersökningar av det här slaget vara låg. Faktorer som visat sig ha betydelse för andelen svar är t.ex. att man riktar in sig på lämpliga människor för undersökningen, att man informerar om undersökningen i förväg och att man låter det framgå att ett deltagande gör skillnad (se Denscombe 2016:49–53). I föreliggande studie har de här faktorerna tillgodosetts. Dessutom var sannolikheten att uppnå ett stort urval hög med tanke på det stora antalet uppslagningar som sker i LSL varje månad. Det är emellertid inte samma sak som att uppnå en acceptabel svarsfrekvens. Den bästa indikatorn på ett tillfredsställande informanturval får man genom att jämföra med liknande studier (Denscombe 2016:54, se avsnitt 4.3.3.1).

Eftersom det inte fick ta alltför lång tid att fylla i enkäten valde jag att i huvudsak använda mig av fasta svarsalternativ. Däremot lade jag till ett textfält till flera av frågorna för den händelse informanterna ville kommentera sitt svar på något sätt. Jag kunde emellertid ha sagt något om ungefär hur lång tid enkäten skulle ta att fylla i, något som jag inte gjorde (se vidare 5.2.1 och 5.5.1). Man får inte glömma att man avbryter användarna i en pågående aktivitet och att denna aktivitet har till syfte att lösa ett språkligt problem som uppstått i en situation som ligger utanför själva ordbokskonsultationen. Den tillfredsställande svarsfrekvensen (i absoluta tal räknat) i 2007-enkäten tydde likväl på att jag kunde inkludera fler frågor utan att äventyra svarsfrekvensen i alltför stor utsträckning.

Som vi vet riktar sig LSL till inlärare som är relativt avancerade nybörjare och som sådana torde de ha tillräckliga svenskkunskaper för att kunna svara på enkätens frågor, i synnerhet med tanke på att jag försökte formulera frågorna och svarsalternativen på en så enkel svenska som möjligt. En inte oväsentlig brist med frågeformulären är dock trots allt att de enbart fanns tillgängliga på svenska. Det är en besvärande omständighet som ger en snedvridning i urvalet. Att översätta enkäten till de övriga Lexin-språken, så som t.ex. gjordes i användarundersökningen av norska Lexin (se avsnitt 4.3.3.1), fanns det dock helt enkelt inte resurser till. I den norska undersökningen översatte man i och för sig inte till alla Lexin-språken, vilket också kan diskuteras.

Man kan dock fråga sig varför enkäten inte fanns tillgänglig på engelska. Huruvida svarsfrekvensen hade varit en annan om enkäten även varit tillgänglig på det språket kan man bara spekulera i. Kanske hade en annan grupp av användare kommit med i urvalet. Det hade emellertid lett till snedvridning åt ett annat håll. Då hade enkäten i och för sig blivit tillgänglig för fler användare men jag hade också tagit hänsyn till en faktor som jag inte var intresserad av, nämligen språkkunskaper i engelska. Eftersom svenska är allas nya språk menar jag att den ingången är mer rättvis och lämplig med tanke på undersökningens syften.

## 5.3 Metod i delstudie 3: Datatriangulering

### 5.3.1 Insamlingsförfarande

Som redan konstaterats kan metodkombination, och i det här fallet datatriangulering, ge en mer fullständig och nyanserad bild av en företeelse (se Hult 2012). Delstudie 3 baserar sig på material från delstudie 1 och 2 och är me-

tomdämsigt av mer tentativ karaktär. Endast data från 2011-materialet ingår. Vid sidan av utprovning av metodens genomförbarhet är det övergripande syftet att närstudera enskilda informanternas sökbeteende. I det här fallet innebär datatrianguleringen att insamlingen av enkäter och loggfiler synkroniseras. Trianguleringen är alltså simultan (parallell) och inte sekventiell och samma vikt läggs vid båda materialen (Denscombe 2016:217–218).

Enkäterna länkas till respektive informants aktiviteter på LSL:s webbsida med hjälp av ip-adresser. Antalet länkningsar är alltså lika stort som antalet informanter i enkätstudien, dvs. 802. Försök att studera enskilda informanternas autentiska ordboksanvändning har t.ex. gjorts av Almind (2008) och studier som inbegriper både loggfilsanalys och enkätstudie har genomförts av de Schryver & Joffe (2004) och Lorentzen & Theilgaard (2012). I dessa undersökningar har dock inte delmaterialen kopplats ihop på det sätt som det görs här.

Fokus i delstudie 3 ligger på de informanter som 1) har uppgett att de har något av Lexin-språken som sitt modersmål och 2) har anlänt till Sverige någon gång under 2000-talet. Därutöver begränsas urvalet till de informanter som i enkäten uppger att de använder LSL varje dag. Detta kriterium utnyttjas för att öka sannolikheten att påträffa ordboksanvändare som använder LSL kontinuerligt och är välbekanta med ordboken. Sammanlagt 249 informanter uppfyller dessa tre kriterier. Dessa användare får här representera LSL:s målgrupp (jfr Beijer & Fre Woldu 1997:4 i avsnitt 3.2.1).

Delstudie 3 och tillhörande metod är som sagt av tentativ karaktär. Därför riktas uppmärksamheten mot några enskilda informanter och deras ordboksaktiviteter. Tillsammans utgör de ett diskussionsunderlag där gjorda iakttagelser syftar till att ge svar på den här metodens eventuella nytta.

### 5.3.2 Metodologiska överväganden

Genom att kräva någon form av registrering bland användarna, t.ex. via inloggning, hade jag på ett säkrare sätt kunnat knyta en ip-adress till en användare och hans aktiviteter på webbsidan. Detta gick emellertid inte att genomföra (jfr ovan när det gäller urvalet av informanter till enkätundersökningen).

För att få bättre kontroll över enskilda användares söksessioner hade man kunnat implementera så kallade *cookies* vid insamlingsprocessen. Eftersom Lexin använder sig av långvariga cookies för att komma ihåg inställningar är det möjligt att identifiera en användare med en cookie. På så sätt kan loggfiler kopplas till specifika användare mycket säkrare än med ip-adresser (Viggo Kann, mejlkonversation 2010-04-14). Varje lexikonuppslagning och ändring i inställningar hade då märkts med ett unikt nummer för sessionen, dvs. från att användaren går till webbplatsen till att hen lämnar den. Det fanns dock inte resurser till att genomföra detta.



Mot bakgrund av ovan nämnda förhållanden uppstod en osäkerhet kring metodens lämplighet. Delar av analysen skulle kunna kritiseras för att vara alltför spekulativ. Men som redan framgått är det trots allt möjligt att identifiera enskilda användare om loggfilen innehåller uppgifter om brukaridentitet, söksträng, sökmetodik och sökväg (se avsnitt 4.2.2). Dessutom är metoden i delstudie 3 av mer tentativ karaktär och spekulationer kan därför inom rimliga gränser tillåtas.

## 5.4 Material i delstudie 1: Loggfilerna

### 5.4.1 Lexikonuppslagningar

Som framgår av avsnitt 3.2.4.3 gjordes den första internetversionen av Lexin tillgänglig år 1995 och tio år senare tillkom den enspråkiga svenska delen, alltså LSL. Statistik på antal uppslagningar har insamlats sedan 1997. Det året gjordes totalt 3,3 miljoner uppslagningar i ordböckerna och knappt 20 år senare, 2015, utfördes 330 miljoner uppslagningar. Uppslagningarna i LSL utgör i genomsnitt cirka 4 % av det totala antalet uppslagningar i Lexin-lexikonet.

Materialet till delstudie 1 utgörs som redan framgått av en veckas lexikonuppslagningar, dvs. loggfiler av typ a, i början av oktober 2008 och 2011. De uppgår till 153 820 respektive 205 591 uppslagningar. Samtliga lexikonuppslagningar ingår i loggfilsanalysen och här har alltså inget urval gjorts. I tabell 5:1 redovisas antal uppslagningar per dag i LSL samt det totala antalet uppslagningar under de två aktuella insamlingsveckorna.

TABELL 5:1. *Antal uppslagningar per dag i LSL under datainsamlingsperioderna 2008 och 2011.*

Veckodag	Måndag	Tisdag	Onsdag	Torsdag	Fredag	Lördag	Söndag	Totalt
<b>Uppslagningar 2008</b>	26 533	26 968	27 162	24 860	15 607	12 170	20 520	153 820
<b>Uppslagningar 2011</b>	37 483	33 968	35 511	33 793	22 028	17 936	24 872	205 591

Det genomsnittliga antalet uppslagningar per dag under hela oktober 2008 är ca 20 400 och motsvarande siffra för oktober 2011 är 28 700. Genomsnittet per månad för hela 2008 och 2011 är något lägre. Tabellen kommenteras närmare

i avsnitt 6.1. Under datainsamlingsveckan 2011 fördelar sig de 205 591 uppslagningarna på 25 974 ip-adresser. Det motsvarar i genomsnitt 8 uppslagningar per ip-adress.

En uppslagning i LSL leder till något av följande utfall:

- 1) Träff (T)
- 2) Automatkorrigering (A)
- 3) Korrigering med förslag (F)
- 4) Icke-träff (?)

Inom parentestecken anges den beteckning som utfallen har i de kommande tabellerna. Typ 1 innebär att ordet som eftersöks finns i LSL, antingen som ett uppslagsord eller som ett morfologiskt exempel. T står för Träff, dvs. uppslagning (se figur 5:1 i avsnitt 5.1.1 för ett exempel). Typ 2 innebär att uppslagningen dirigeras till ett ord som finns i LSL och som slås upp. Det ordet är formmässigt mycket likt det sökta ordet. Typ 3 betyder att ordet inte ger träff men att det samtidigt presenteras en rad uppslagsord som användaren kan ha åsyftat. Typ 4 innebär att söksträngen inte återfinns i LSL och att användaren möts av meddelandet "Ordet x finns inte i lexikonet". Det betecknas med ett frågetecken (?) i tabellerna. Strax nedan exemplifieras de fyra typerna av utfall i korta utdrag från 2008- och 2011-materialen.

En loggfil från 2008-materialet innehåller information om utfallstyp, datum, tidpunkt, utfall av uppslagning samt antal ordförslag som ges ifall uppslagningen inte återfinns i LSL. I tabell 5:2 återges uppslagningar från en minut 2008, närmare bestämt måndagen 6 oktober mellan kl. 10.00.00 och 10.00.59. Tabellen återger materialet i bearbetat skick och är alltså inte ett exempel på hur loggfilen ser ut i obearbetat tillstånd.

Den första kolumnen anger radnummer. Den andra kolumnen anger den typ av utfall uppslagningen resulterat i, dvs. träff (T), automatkorrigering (A), korrigering med förslag (F) eller icke-träff (?). I kolumn 3 anges datum. Siffran 6 står för måndag 6 oktober 2008. Kolumn 4 anger klockslaget för uppslagningen och i kolumn 5 återges själva söksträngen. Kolumn 6 innehåller antingen det uppslagsord som en automatkorrigering dirigeras till, eller antalet förslag som ges vid en F-korrigering, följt av de enskilda ordförslagen.

Av de 23 uppslagningar som utförs under den aktuella minuten leder 16 till träff i LSL, t.ex. rad 1 och 3 som innehåller substantivet *bagatell* och adjektivet *ställföreträdande*. En uppslagning, *välkommen*, har automatkorrigerats (A) till ordboksartikeln *välkommen* (rad 11). I tabell 5:2 återfinns också två uppslagningar som har lett till korrigering med förslag (F), dels *habitat* på rad 2 som leder till sex förslag (*hatat, haltat, rabiät, habilt, habila, habilast*), dels *pitcha* på rad 23 som leder till två förslag (*picka, bitch*).

TABELL 5:2. Uppslagningar 2008, måndag 6 oktober mellan kl. 10.00.00 och 10.00.59.

Rad	Typ	Dag	Tidpunkt	Söksträng	Förslag
1	T	6	10:00:04	bagatell	
2	F	6	10:00:05	habitat	6: hatat, haltat, rabiat, habilt, habila, habilast
3	T	6	10:00:06	Ställföreträdande	
4	T	6	10:00:08	patriarkaliska	
5	T	6	10:00:10	Bödel	
6	?	6	10:00:11	kallibreras	
7	T	6	10:00:13	budget	
8	T	6	10:00:16	slutför	
9	T	6	10:00:16	budget	
10	?	6	10:00:16	kallibrera	
11	A	6	10:00:18	välkomen	välkommen
12	T	6	10:00:18	brer	
13	T	6	10:00:23	palm	
14	T	6	10:00:31	bedrift	
15	T	6	10:00:33	brer	
16	?	6	10:00:33	försluta	
17	T	6	10:00:37	hoppa	
18	T	6	10:00:40	tyvärr	
19	T	6	10:00:42	sol	
20	?	6	10:00:43	försluter	
21	T	6	10:00:49	etablera	
22	T	6	10:00:54	bragd	
23	F	6	10:00:54	pitcha	2: picka, bitch

Figur 5:6 på nästa sida visar utfallet av söksträngen *habitat* på LSL:s webbsida. Om användaren väljer att slå upp något av förslagen visar det sig i loggfilerna. (Jag kan dock inte hitta något av de listade orden i nära anslutning till uppslagningen *habitat* i mitt loggfilsmaterial.)



FIGUR 5:6. Resultatet av söksträngen habitat. Hämtad 2016-05-18.

Slutligen har vi fyra uppslagningar som resulterar i icke-träffar (se rad 6, 10, 16, 20). Det indikeras av frågetecknet i kolumn 2. De utgörs dock bara av två olika ord, *förluta/förluter* och de felstavade formerna *\*kallibreras/\*kallibrera*. I det här fallet är de ord som troligen åsyftas inte uppslagsord i LSL, men icke-träffar kan också utgöras av grovt felstavade ord som rättstavningshjälpen inte kan identifiera. Även om loggfilerna från 2008 inte innehåller ip-adresser kan man ändå med stor sannolikhet anta att uppslagningarna på t.ex. rad 16 och 20 (*förluta* och *förluter*) har utförts av samma användare eftersom de befinner sig så pass nära varandra i tid. Första sökningen görs på verbets infinitivform och inför den andra sökningen har söksträngen modifierats och användaren söker istället på verbets presensform.

I tabell 5:3 illustreras en loggfil från 2011-materialet som även den innehåller en minuts uppslagningar.

Av de 35 uppslagningarna under den aktuella minuten leder 22 till träffar i LSL. De tre första träffarna är verbet *skakar* på rad 3, substantivet *förtret* på rad 5 och adjektivet *partisk* på rad 7. Två uppslagningar leder till automatkorrigering, *tyna bort* dirigeras till *tonar bort* (rad 1) och *ubät* dirigeras till *ubåt* (rad 4). Vidare finns fem uppslagningar som leder till korrigering med förslag, t.ex. *våliga* på rad 6 som resulterar i fem ordförslag (vilket alltså siffran 5 i kolumn 6 står för). En sökning direkt i LSL4 avslöjar att de fem ordförslagen är *dåliga*, *tåliga*, *veliga*, *vågiga* och *vådliga*. I tabellen finner vi slutligen sex uppslagningar som leder till en icke-träff. De är det spanska uttrycket *Que bien* (rad 9, 12 och 15), den tvådelade söksträngen *gastrointestinala störningar* (rad 23) och *frukostbricka* (rad 35).

I tabell 5:3 har som synes en kolumn tillkommit i jämförelse med tabell 5:2. Den innehåller delvis anonymiserade ip-adresser. Som diskuterades i avsnitt 5.1.2 är en ip-adress inte automatiskt detsamma som en unik användare. Om samma ip-adress upprepade gånger återkommer under en begränsad tids-session är sannolikheten dock stor att så är fallet. Det kan man t.ex. misstänka vad gäller uppslagningarna i rad 7, 10, 18 och 27 (*partisk*, *därtill*, *Dessvärre* och *Desbättre*). Att det finns ett släktskap mellan sökorden *dessvärre* och *desbättre* ökar också sannolikheten för att det rör sig om en och samma användare.

TABELL 5:3. Uppslagningar 2011, måndag 3 oktober mellan kl. 10.01.00 och 10.01.59.

Rad	Typ	Datum	Tidpunkt	Söksträng	Förslag	Ip-adress
1	A	3	10:01:01	tyna bort	tonar bort	193.181.nn.nnn
2	F	3	10:01:02	rokad	7	82.99.nn.nnn
3	T	3	10:01:06	skakar		217.115.nn.nn
4	A	3	10:01:06	ubät	ubåt	78.72.nnn.nn
5	T	3	10:01:09	förtret		80.216.nn.nnn
6	F	3	10:01:11	våliga	5	80.217.nnn.nnn
7	T	3	10:01:13	partisk		81.230.nn.nnn
8	T	3	10:01:17	GRADERA		109.228.nnn.nn
9	?	3	10:01:17	Que bien		80.65.nnn.nnn
10	T	3	10:01:19	Därtill		81.230.nn.nnn
11	F	3	10:01:21	konkurere	6	83.137.n.n
12	?	3	10:01:21	Que bien		80.65.nnn.nnn
13	T	3	10:01:22	mer		109.228.nnn.nn
14	T	3	10:01:25	konkurrera		83.137.n.n
15	?	3	10:01:27	Que bien		80.65.nnn.nnn
16	T	3	10:01:27	tog		80.254.nnn.nn
17	T	3	10:01:31	hitta		109.228.nnn.nn
18	T	3	10:01:34	Dessvärre		81.230.nn.nnn
19	T	3	10:01:35	radio		213.113.nnn.nnn
20	T	3	10:01:37	skvallrar		130.243.nnn.nn
21	T	3	10:01:37	underlägsen		193.181.nn.nnn
22	T	3	10:01:41	tuff		46.230.nnn.nnn
23	?	3	10:01:43	gastrointestinala störningar		195.67.nnn.nnn
24	T	3	10:01:44	annat		85.188.nnn.nn
25	T	3	10:01:46	krossar		217.115.nn.nn
26	T	3	10:01:47	referens		109.58.nnn.nnn
27	T	3	10:01:48	Dessbättre		81.230.nn.nnn
28	T	3	10:01:49	känguru		178.174.nnn.nnn
29	T	3	10:01:49	räfsar		195.67.nnn.nn
30	F	3	10:01:51	födda	6	90.227.nnn.nnn
31	F	3	10:01:52	gastro	9	195.67.nnn.nnn
32	?	3	10:01:52	fraternity		193.178.nnn.n
33	T	3	10:01:53	klumpig		62.109.nn.nnn
34	T	3	10:01:55	ta		80.254.nnn.nn
35	?	3	10:01:58	frukostbricka		83.253.nnn.nnn

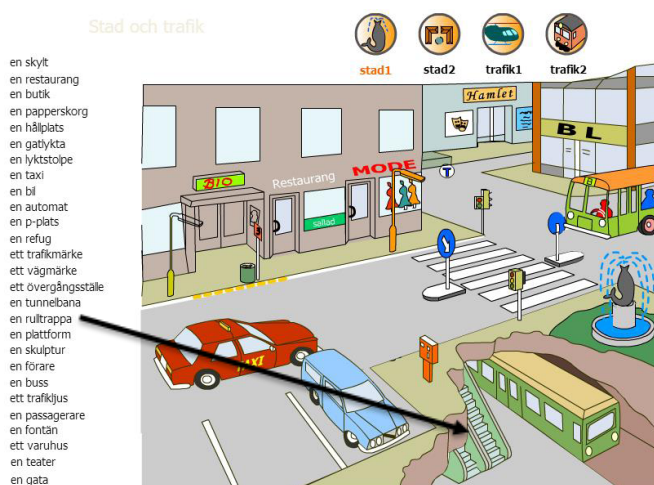
## 5.4.2 Navigeringsval

Navigeringsval, loggfiler av typ b, återfinns endast i 2011-materialet. För tydlighetens skull återges på nytt figuren som visar vilka typer av data som ingår i materialet med uppgifter om navigeringsval (jfr avsnitt 5.1.1).



FIGUR 5:7. *Typer av användaraktiviteter som genererar loggfiler av typ b. Hämtad 2016-05-09.*

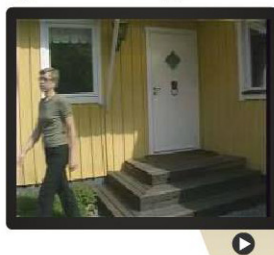
När användarna klickar på uttalsknappen dirigeras de om till en annan webbadress där uttalet läses upp. För att återgå till söksidan behöver man klicka på "Bakåt"-knappen längst upp till vänster på webbsidan. När man klickar på BILD SVENSKA som återfinns i artikeln *rulltrappa* länkas man på samma sätt till en bild, som återges i figur 5:8.



FIGUR 5:8. *BILD SVENSKA från uppslagsordet rulltrappa (pil infogad i efterhand). Hämtad 2016-05-09.*

Länken leder i det här fallet till en scenisk illustration under huvudrubriken ”Stad och trafik” (se avsnitt 3.1.2.8 för olika typer av illustrationer). Länken VISA FILM leder till en filmsekvens som realiserar verbfrasen *gå nedför trappan* (se figur 5:9).

**går nedför trappan**



FIGUR 5:9. VISA FILM från uppslagsordet trappa. Hämtad 2016-05-09.

När det gäller loggfiler av typ b fanns det vid insamlingsförfarandet inte möjlighet att skilja ut klickningar som görs från LSL från de klickningar som görs från något av de tvåspråkiga lexikonerna. Det är därför näst intill omöjligt att exakt få reda på i hur stor utsträckning länkarna till uppläst uttal, bilder och filmer utnyttjas av användare som söker i just LSL. Totalt har de tre länkarna klickats på vid 1 710 028 tillfällen under den vecka materialet samlades in.

TABELL 5:4. Antal klickningar som görs i rutan framför respektive informationskategori.

Informationskategori som klickas för	Antal klick
Avledning	82
Bilder	152
Böjning	9784
Definition	144
Exempel	9269
Förkortningar	121
Konstruktioner	8434
Motsatser	120
Ordklass	113
Sammansättningar	8972
Uppslagsord	60
Uttal	70
Uttryck	8117

Med navigeringsval avses också klickningar på frågetecknet ”?” och de ändringar som användarna gör under fliken ”Inställningar” (se figur 5:7). Antalet klickningar på ”Frågetecknet”-knappen uppgår till 7 599. Av tabell 5:4 framgår hur många klickningar som görs i inställningarna i rutorna för de möjliga informationskategorier en ordboksartikel i LSL kan innehålla. Här finns det alltså inte heller klara siffror på antal klickningar som görs från just LSL.

Av tabell 5:4 framgår tydligt att de informationskategorier som vid tidpunkten för insamlingen var urkryssade har kryssats för i långt högre utsträckning än de övriga kategorierna. Användarna har alltså aktivt valt att återställa inställningarna till standardläget. Navigeringsvalen behandlas som sagt endast kort i loggfilsanalysen men utgör ett betydelsefullt inslag i delstudie 3, se vidare avsnitt 5.6.

## 5.5 Material i delstudie 2: Frågeformulären

I delstudie 2 kartläggs sålunda användarna av LSL. Materialet består ursprungligen av två uppsättningar enkätmaterial, 370 frågeformulär från 2007 och 863 frågeformulär från 2011. Antalet ip-adresser från insamlingsperioden 2011 uppgår som sagt till 25 974. Förutsatt att adresserna representerar unika användare är svarsfrekvensen mycket låg (3,3 %). Antalet enkäter får ändå betraktas som mycket tillfredsställande i förhållande till liknande studier och med tanke på insamlingsperiodens längd (se avsnitt 4.3.3.1).

Materialen är inte helt jämförbara då 2007 års enkät består av tio frågor (se bilaga 1) och 2011 års enkät innehåller 20 frågor (se bilaga 2). Som redan framgått är 2011-undersökningen överordnad 2007-undersökningen (se avsnitt 5.2.1). Dessutom ingår de flesta frågor från enkäten 2007 i enkäten 2011 och därför fokuseras i det följande på frågeformuläret som används i 2011-undersökningen.

Innan frågeformuläret från 2011 lades upp på LSL:s webbsida kontrollerades det i tre forum. För det första granskades det av en grupp medarbetare, varav två med utländsk bakgrund, vid Svensk Nationell Datatjänst, SND. SND är en serviceorganisation för svensk forskning inom humaniora, samhällsvetenskap och medicin. Den tillhandahåller bl.a. stöd och rådgivning under hela forskningsprocessen (SND 2016). För det andra granskades formuläret av Magnus Bredberg, adjunkt i statistik vid Högskolan i Skövde. För det tredje testades frågeformuläret på ett tiotal studenter med utländsk bakgrund som läste en kurs i svenska för nybörjare på Institutionen för svenska språket vid Göteborgs uni-



versitet under hösten 2011.<sup>17</sup> Eftersom dessa studenter tillhör Lexins målgrupp kunde jag få en fingervisning om hur väl jag lyckats med formulärets språkliga utformning. Slutligen erhöll jag också konkreta förbättringsförslag av Viggo Kann. Med facit i hand kanske lärare i svenska som andraspråk borde ha rådfrågats i större utsträckning.

Kommentarerna från de tre granskningsforumen rörde såväl frågornas och svarsalternativens ordningsföljd som språkliga formuleringar, t.ex. konsekvensen av ett visst ordval. Två mycket värdefulla förslag gällde också frågeformulärets design, men de föreslagna förändringarna var tyvärr inte heller genomförbara i det program (ACE) som enkäten lades upp i (se avsnitt 5.2.1). Den ena kommentaren gällde att de informanter som hade svenska som förstaspråk automatiskt skulle ledas vidare till de två sista frågorna eftersom fråga 12–19 inte gällde dem. Nu fick de istället scrolla förbi dem och ned till fråga 20 för att fortsätta att fylla i enkäten. Det andra förslaget gällde att det inte skulle vara möjligt att kryssa för fler svarsalternativ än vad som efterfrågades. Det gick inte heller att implementera i ACE. Studenternas svar gjorde mig uppmärksam på mindre lyckade formuleringar av både frågor och svarsalternativ vilket ledde till förbättringar. Beträffande frågornas ordningsföljd försökte jag följa de rekommendationer som finns, t.ex. att okomplicerade och okontroversiella frågor placeras i början av formuläret.

Ifråga om etiska aspekter kunde det ha framgått tydligare att jag inte hade för avsikt att röja deltagarnas identitet eller låta materialet hamna i orätta händer. Även om det inte ställdes några känsliga frågor som kom i konflikt med personuppgiftslagen (se avsnitt 5.7), så kan möjligen vissa informanter vara av annan uppfattning. En lösning på detta problem hade varit att införa svarsalternativet ”Vill inte svara”, vilket också hade gett mig information om huruvida det uteblivna svaret just var ett uttryck för en ovilja att fylla i delar av enkäten.

### 5.5.1 Beskrivning av frågeformulären

Den utökade versionen av frågeformuläret som används i 2011 års undersökning innehåller 20 frågor. De frågor som kvarstår efter den första enkäten gäller: ålder, kön, sysselsättning, utbildning (i Sverige och hemlandet), modersmål, svenskkunskaper, sökfrequens, användningssituation och efterfrågade informationskategorier. Även data från enkätens kommentarsfält kan jämföras i de två materialen.

---

<sup>17</sup> Tack till Monica Nyberg, adjunkt vid Institutionen för svenska språket, och hennes studenter för möjligheten att pilota enkäten.

I det följande presenteras och diskuteras den slutgiltiga versionen av frågeformuläret. I internetversionen hade formuläret en ljusgul bakgrund. Formuläret inleddes med följande text:

Just nu pågår en undersökning av användningen av *Lexins svenska lexikon*, som nyligen har uppdaterats. Svara på några frågor kring dig och din användning av lexikonet så är du med och gör det ännu bättre!

Vid flervalsfrågorna klickar du i önskad ruta (till vänster om det valda alternativet). Om du vill skriva i en kommentarruta måste du först flytta markören in i rutan.

Skicka iväg din enkät genom att klicka på knappen märkt "Skicka enkät" längst ner.

Tack för hjälpen!

Texten i de tre inledande raderna ingick i sina huvuddrag också i det popup-fönster som genererades när en användare valde att söka i LSL och erbjöds att delta i undersökningen. De påföljande raderna utgörs av en standardtext för just denna enkät i ACE (se avsnitt 5.2.1). Här finns möjligtvis en del diskutabla ordval med tanke på svenskkunskaperna hos de användare jag i första hand riktade enkäten till. Förhoppningsvis inverkade dessa ordval inte negativt på informanternas förståelse av hur formuläret skulle fyllas i.

Sedan följer de 20 frågorna. På de frågor som har ringar framför svarsalternativen kan endast ett (1) svar ges. Om svarsalternativen föregås av fyrkanter kan alla alternativ kryssas för, om så önskas.

I fråga 1 till 4 efterfrågas information om födelseår, kön, högsta avslutade utbildning i Sverige och sysselsättning. Frågorna och svarsalternativen lyder:

1. Födelseår:

2. Kön:

- Kvinna
- Man

3. Högsta avslutade utbildning i Sverige:

- Ingen avslutad utbildning
- Grundskola
- Gymnasium
- Folkhögskola
- Komvux (eller motsvarande)
- Universitet/högskola
- Annan (ange nedan)

Annan utbildning:

4. Vad är din huvudsakliga sysselsättning?

- o Jag studerar
- o Jag arbetar
- o Jag söker arbete
- o Annan (ange nedan)

Annan huvudsaklig sysselsättning:

Svaren på de inledande frågorna ger bakgrundsinformation om vilka som väljer att svara på enkäten och ger också en uppfattning om urvalets sammansättning. Att efterfråga födelseår istället för ålder ger säkrare uppgifter om hur gammal informanten är. Genom uppgifter om födelseår kan jag också inordna informanternas ålder i intervaller, om jag så vill.

Svarsalternativen för fråga 2 ”Kön”, dvs. ”Kvinna” och ”Man”, skulle idag ha utökats med ett eller två svarsalternativ av typen ”Annat” och ”Vill inte svara”.

Frågan om utbildning i Sverige är placerad i början av enkäten medan frågan om utbildning i hemland är placerad i ett senare avsnitt som riktar sig speciellt till dem som har svenska som andraspråk. Jag önskade hålla isär frågorna till dessa användare från de frågor som riktade sig till alla informanter. I efterhand kan konstateras att frågorna om utbildning i Sverige respektive hemland lämpligen skulle ha placerats i direkt anslutning till varandra, såsom skedde i enkäten från 2007. Att placera dem efter varandra hade onekligen underlättat för informanterna när de fyllde i enkäten. Det är nämligen många kommentarer som handlar om utbildning utomlands men som tas upp under frågan om utbildning i Sverige. I övrigt ändrades frågeformuleringen från ”Utbildning i Sverige” till ”Högsta avslutade utbildning”.

Vad gäller formuleringen av fråga 4, som rör sysselsättning, övergav jag en formulering med enklare svenska (ungefär ”Vad gör du på dagarna?”) för den starkt etablerade men kanske språkligt mer utmanande formuleringen ”Vad är din huvudsakliga sysselsättning”. Det är svårt att avgöra vilken formulering som är lättast att förstå för dem som har bristande kunskaper i svenska språket. Jag utgick ifrån att många säkert känner igen den etablerade formuleringen medan den enklare formuleringen är relativt oprövad. Dessutom framgår det av svarsalternativen vad som efterfrågas. I enkätens elektroniska version kan man bara fylla i ett (1) alternativ, eftersom det är den huvudsakliga sysselsättningen som efterfrågas. Det finns dock ett kommentarsfält där informanterna kan utveckla sina svar, om så önskas.

Fråga 5 till 11 handlar om informanternas användning av LSL. I fråga 5 efterfrågas sökfrekvens:

5. Hur ofta brukar du söka i *Lexins svenska lexikon*?

- o Detta är första gången
- o Högst en gång i månaden
- o Flera gånger i månaden
- o Flera gånger i veckan
- o Flera gånger per dag

Svarsalternativen utökades från fyra till fem i förhållande till 2007 års enkät. Det sistnämnda alternativet blev tillagt. En mer träffsäker svarsformulering hade möjligen varit ”Minst en gång per dag”.

I fråga 6 ska informanterna ange i vilket syfte ordboken används:

6. Använder du mest *Lexins svenska lexikon* för att (ange max två alternativ):

- förstå ett ord i en text som du läser?
- få hjälp med ett ord när du skriver en text?
- få hjälp med ett ord när du ska översätta en text?
- få hjälp med ett ord som du hört?

Ev. annan anledning:

I lexikografiska termer motsvaras det första och det fjärde alternativet av funktionerna ”Reception”, i skrift respektive i tal. Det andra svarsalternativet motsvarar ”Produktion” och alternativ 3 ”Översättning” (se avsnitt 2.3.1). Vidare meddelas i fråga 6 att max två alternativ ska kryssas för. Denna begränsning är dock inte genomförd i teknisk bemärkelse vilket innebär att man kan kryssa för ett obegränsat antal alternativ. Detta är olyckligt eftersom vissa informanter verkar ha uppfattat instruktionen och andra inte (se vidare avsnitt 7.2).

Fråga 7 och 8 efterfrågar mer specifikt vilka informationstyper som informanterna *ofta* respektive *aldrig* söker uppgifter om. De angivna alternativen motsvarar de informationstyper som återfinns i LSL och anges också i samma termer som används i ordboken.

En mycket vanlig lexikografiskt relaterad språkfråga är hur ett ord stavas. Alternativet ”Stavning” kunde förstås ha funnits med som ett alternativ, vilket det faktiskt gjorde i enkäten från 2007 (se bilaga 1, fråga 8). När enkäten sedan skulle utvidgas utgick jag ifrån de informationskategorier som listades i den uppdaterade versionen av LSL. Där ingick inte explicit stavning vilket medförde att denna kategori förbisågs.

I fråga 7 ombeds informanterna kryssa för max tre alternativ. I likhet med i fråga 6 är denna begränsning inte genomförd i teknisk bemärkelse vilket innebär att man kan kryssa för ett obegränsat antal alternativ (se vidare avsnitt 7.2).

7. Vilka typer av information söker du mest efter i *Lexins svenska lexikon* (ange max tre alternativ)?

- Uttal
- Ordklass
- Böjning
- Förkortningar
- Betydelse
- Motsatser
- Sammansättningar
- Exempel
- Konstruktioner
- Uttryck
- Bild
- Film

8. Finns det information som du aldrig söker efter i *Lexins svenska lexikon* (kryssa för så många du behöver)?

- Uttal
- Ordklass
- Böjning
- Förkortningar
- Betydelse
- Motsatser
- Sammansättningar
- Exempel
- Konstruktioner
- Uttryck
- Bild
- Film

I de två påföljande frågorna, 9a och 9b, efterfrågas informanternas kännedom om att det går att se bilder och filmer i LSL. De fås fram genom att man klickar på länkarna "Bild svenska" och "Visa film". Länkarna är placerade i anknötning till det betydelsemoment av ett ord som de tillhör (se figurerna 5:8 och 5:9 i avsnitt 5.4.2).

9 a. Brukar du titta på bilderna som finns till vissa ord (dvs. klickar du på länken Bild svenska)?

- Ofta
- Ibland
- Aldrig

9 b. Brukar du titta på filmerna som finns till vissa ord (dvs. klickar du på länken Visa film)?

- Ofta
- Ibland
- Aldrig

Kategorierna ”Bild” och ”Film” återfinns som synes också bland svarsalternativen i de två föregående frågorna om informationstyper. Det finns tre anledningar till att jag även tar upp dem separat: 1) Jag är intresserad av i vilken utsträckning just dessa två kategorier efterfrågas. 2) Om man varken kryssat för dem i fråga 7 eller 8 så har man kanske inte uppfattat dem som egna informationskategorier över huvud taget och en mer tydlig frågeformulering kan få fler att förstå vad det är jag frågar om. 3) Svaren kan senare kontrolleras mot antalet klickningar på dessa länkar, vilka återfinns i loggfilerna.

Ordet *brukar* som används i frågeformuleringen kan möjligen orsaka huvudbry. Det är ett ganska vagt ord som kan ha olika betydelse för olika människor. Förhoppningsvis är svarsalternativen tillräckliga för att uppnå tillförlitliga svar.

Den näst sista frågan om användningen av LSL gäller huruvida informanterna har läst användarmanualen som finns på Lexins webbsida.

10. Har du någon gång läst instruktionerna till Lexin, där det står om hur man ska tolka och förstå den information som finns i Lexinordböckerna?

- Ja
- Nej
- Vet inte

Det slås ofta och gärna fast att användare i mycket liten utsträckning utnyttjar de användarmanualer som ordböcker i regel innehåller (Svensén 2004:542). Samtidigt lägger lexicografer ned mycket tid på att författa dessa texter. Det är av intresse att ha ett mer dagsaktuellt svar på frågan och därför är den med i enkäten.

Till sist vill jag veta vilka andra typer av språkliga hjälpmedel som användarna gör bruk av, vid sidan av LSL.

11. Hur löser du språkliga problem om du inte använder *Lexins svenska lexikon*? (Du får kryssa för flera.)

- Använder en annan ordbok på Internet.
- Använder en tryckt ordbok.
- Frågar någon.
- Använder Google eller någon annan sökmotor på Internet.
- Annat sätt (ange nedan).

Ev. annat sätt:

LSL är bara en av möjliga ordböcker, såväl tryckta som elektroniska, att vända sig till med lexikografiskt relevanta problem. Ett både spännande och angeläget spörsmål är naturligen vilka andra hjälpmedel informanterna utnyttjar.

Det tredje och näst sista avsnittet i enkäten, fråga 12–19, riktar sig till dem som har annat modersmål än svenska. Avsnittet inleds med följande text:

Fråga 12 till 19 riktar sig till dig som har annat modersmål (=det första språk du lärde dig) än svenska. Övriga kan gå direkt till fråga 20.

Syftet med dessa frågor är dels att ta reda på huruvida informanterna tillhör LSL:s målgrupp, dels att få en bild av dessa personers språkliga bakgrund och svenska språkets närvaro i deras rådande livssituation i Sverige. Ordet *modersmål* förklaras inom en parentes som 'det första språk du lärde dig'.

Först efterfrågas informanternas ankomstår till Sverige.

12. Vilket år kom du till Sverige? (Gäller dig som har flyttat till Sverige från ett annat land.)

Tillsammans med uppgiften om födelseår kan fråga 12 ge svar på hur gammal informanten var när hen kom till Sverige. Med fråga 12 kan jag också skilja ut de användare som har annat modersmål än svenska och som bor i Sverige, dvs. som tillhör LSL:s målgrupp, från dem som har ett annat modersmål än svenska men som inte befinner sig i Sverige (jfr distinktionen mellan *andraspråk* och *främmandespråk* i avsnitt 3.1.1).

Därefter efterfrågas vilken utbildning användarna (med annat modersmål än svenska) har i svenska språket.

13a. Har du genomgått SFI-utbildning? (SFI=svenska för invandrare.)

- Ja
- Nej

13b. Har du genomgått någon annan kurs eller utbildning i svenska språket?

- Ja (ange vilken nedan).
- Nej

Ev. annan kurs eller utbildning i svenska:

Genomförd sfi-utbildning och/eller annan utbildning i svenska språket ger en indikation på vilken formell nivå av svensk-kunskaper informanterna har uppnått. Sfi, dvs. Svenska för invandrare, är en kostnadsfri utbildning i svenska språket och har bedrivits sedan 1960-talet. För att få läsa sfi krävs att man är folkbokförd i en kommun i Sverige, är minst 16 år gammal och saknar grund-

läggande kunskaper i svenska språket (Lernia 2016). Folkbokförd kan man bli om man har uppehållstillstånd. Det kan ta flera år innan man får reda på om man beviljas uppehållstillstånd eller inte. Det kan således dröja många år från det att man anlänt till Sverige till dess att man börjar läsa kursen sfi. Om informanterna svarar ja på frågan om de genomgått en sfi-utbildning kan man dels anta att de har vissa grundläggande kunskaper i svenska språket, dels få indirekt information om att de har uppehållstillstånd i Sverige och således är folkbokförda här. Det här var alltså förutsättningarna vid tidpunkten för datainsamlingen. I övrigt kunde frågan om sfi-utbildning ha specificerats på så sätt att informanterna skulle uppge hur långt de kommit i sin svenskutbildning, dvs. vilka kunskapskrav (A–D) som hitintills uppnåtts.<sup>18</sup>

Därefter följer frågor om modersmål och svenska språkets närvaro i användarnas vardag.

14. Vilket är ditt modersmål?
15. Vilket språk är du bäst på?
16. Hur ofta talar eller använder du svenska?
  - o Nästan hela tiden.
  - o Varje dag.
  - o Flera gånger i veckan.
  - o Mer sällan.
17. Vilket/vilka språk talar du främst med...
  - med din familj?
  - med dina vänner?
  - med kompisarna i skolan eller på jobbet?
18. Hur bra tycker du att dina svenskkunskaper är?
  - o mycket bra
  - o bra
  - o ganska bra
  - o ganska dåliga
  - o dåliga

---

18 I skrivande stund är kursplanen för sfi-utbildningen under revidering (oktober 2016). De fakta som presenteras här kan alltså inom kort vara inaktuella. De var dock giltiga vid undersökningens genomförande.



I frågan om modersmål kunde det ha varit lämpligare att skriva "Vilket/vilka" istället för bara "Vilket". Det språk som nu anges är förmodligen informanternas starkaste språk. Ett tiotal informanter uppger mer än ett språk och frågan är hur många fler som velat ange fler språk men som begränsat sitt svar till bara ett språk på grund av frågans formulering.

Påföljande fyra frågor innehåller olika grad av självskattning. Vid tolkningen av svar på självskattningsfrågor manas till försiktighet (jfr citatet från Hatherall 1984:184 i avsnitt 4.2.1). Det är t.ex. oklart vilken måttstock enskilda informanter utgår ifrån när de bedömer nivån på sin språkförmåga i svenska. Tanken är att frågorna tillsammans ska kunna ge en upplysande bild av modersmålet och svenska språkets användning i användarens vardag. Formuleringen i fråga 16 byggdes från 2007 till 2011 ut från "...talar du svenska" till "...talar eller använder du svenska" eftersom "talar" inte inbegriper t.ex. att lyssna på radio och läsa tidningen. Fråga 18, självskattning av svenskkunskaper, säger givetvis inte så mycket om informantens formella språkkunskaper. Självskattningen kan dock bidra med intressanta iakttagelser satta i relation till andra frågor, t.ex. frågorna om sfi-utbildning, ankomst till Sverige, hur ofta och med vilka man använder svenska m.fl. I samband med omarbetningen av enkäten vände jag på ordningen på svarsalternativen och ett alternativ, "Ganska bra", lades till. Detta för att skillnaden på "Bra" och "Ganska dåliga" kan uppfattas som för stor.

Avsnittet avslutas med frågan om utbildning i hemlandet, vilken till viss del redan har kommenterats (se ovan).

19. Högsta avslutade utbildning i ditt hemland:

- o Ingen avslutad utbildning
- o Grundskola (eller motsvarande)
- o Gymnasium (eller motsvarande)
- o Folkhögskola (eller motsvarande)
- o Universitet/högskola (eller motsvarande)
- o Annan (ange nedan)

Annan utbildning:

Ev. kommentar:

Det var inte helt enkelt att formulera svarsalternativ till frågan om utbildning i hemlandet. Skolsystemen kan se mycket olika ut i olika länder och ha benämningar som inte alls stämmer överens med det svenska systemet. Det ansågs därför bättre att utgå från det svenska systemets namn. Många har säkert varit i kontakt med det svenska utbildningssystemet och känner igen de benämningar som används. Min förhoppning är också att användarna ska kunna översätta den utbildning de har med sig från hemlandet till den nivå som utbildningen motsvarar i svenska termer. Därav kommer att det finns två kommentarsfält till denna fråga.

Slutligen ombeds informanterna att ge respons på *Lexins svenska lexikon* (20) samt uppge namn och e-postadress om de kan tänka sig att vara med i fler undersökningar av Lexin.

20. Här kan du skriva ned dina åsikter eller kommentarer kring *Lexins svenska lexikon*:

Lämna gärna ditt namn och din e-postadress om du kan tänka dig att vara med i fler undersökningar av Lexin:

Frågeformuläret avslutas som synes med ett tack, information om att enkäten bara kan skickas iväg en gång samt kontaktuppgifter till undertecknad.

Stort tack för hjälpen!

OBS! Det ifyllda formuläret kan bara skickas iväg en gång!

ann-kristin.hult@svenska.gu.se, Lexikaliska institutet, Institutionen för svenska språket, Göteborgs universitet.

## 5.5.2 Redovisning av bortfall

Det externa bortfallet utgörs av de användare som avvisar enkäten (se avsnitt 5.2.1). Det interna bortfallet består dels av *objektbortfall* (t.ex. ett tomt frågeformulär som skickats in), dels av *partiellt bortfall* eller *variabelbortfall* (svar saknas på en eller flera frågor). Tabell 5:5 visar uteblivna svarsfrekvenser för respektive enkätmaterial. I uppställningen framgår också tydligt vilka frågor som återfinns i båda frågeformulärens och där det är möjligt att göra jämförelser.

I enkätmaterialen från 2007 är det externa bortfallet okänt. Det interna bortfallet är förhållandevis litet. Av de 370 inskickade enkäterna är 14 tomma. De återstående 356 enkäterna ingår i studien. Det externa bortfallet i enkätmaterialen från 2011 är som sagt svårt att uppskatta men antalet unika registrerade ip-adresser under insamlingsperioden (25 974) kan ge en uppfattning om hur många som nåtts av frågeformuläret. Det interna bortfallet har jag däremot i högre grad kontroll över. Av tabellen framgår att det partiella bortfallet varierar stort mellan frågorna både inom respektive enkät och mellan enkäterna. Ett högt bortfall kan ha helt naturliga förklaringar. Flera av frågorna i 2011-enkäten riktade sig t.ex. bara till informanter som har svenska som andraspråk. Ett högt bortfall behöver därför inte vara negativt på något sätt utan kan i detta fall ge en indikation om hur många informanter som har svenska som modersmål. I andra fall kan man däremot ställa sig frågande till antalet uteblivna svar, t.ex. varför 38 informanter inte har besvarat frågan om ålder.

Om det partiella bortfallet är alltför högt kan hela enkäten behöva avlägsnas från analysen, dvs. övergå till ett objektbortfall. Operationella regler kan sättas upp för när det ska ske. Det finns dock inga generella anvisningar för var gränsen går, utan det är olika för olika studier och upp till den enskilde forskaren att bestämma. I det följande fokuserar jag på de 863 enkäterna i 2011-undersökningen.

TABELL 5:5. *Antal uteblivna svar på enskilda frågor i enkätmaterial från 2007 och 2011.*

	Fråga	Antal uteblivna svar	
		2007	2011
1	Ålder	0	38
2	Kön	2	12
3	Sysselsättning	4	13
4	Modersmål	34	111
5	Svenskkunskaper	160	93
6	Utbildning i Sverige	28	28
7	Utbildning i hemland	226	105
8	Sökfrequens	13	5
9	Eftersökta informationskategorier	okänt	9
10	Ej eftersökta informationskategorier	–	125
11	Användningssituation	okänt	8
12	Hittas eftersökt information?	38	–
13	Klickar du på länk till "Bild svenska"?	–	17
14	Klickar du på länk till "Visa film"?	–	14
15	Läser du Lexin-instruktioner?	–	12
16	Används andra hjälpmedel?	–	15
17	Ankomst till Sverige	–	171
18	Sfi-utbildning	–	106
19	Övrig utbildning i svenska	–	116
20	Bästa språk	–	113
21	Bruk av svenska	–	93
22	Språk med familj	–	114
23	Språk med vänner	–	115
24	Språk i skola/på jobb	–	119
25	Avslutande kommentar	246	431

De operationella reglerna har tagits fram enligt följande. Jag har utgått ifrån hur många svar som avgetts på de frågor som riktar sig till alla, dvs. 1–11 och 20. Vid val mellan en sträng bedömning och ett mer tolerant förhållningssätt har jag stannat för det senare eftersom enkäten riktar sig till personer som har begränsade kunskaper i svenska. Att formuläret endast har funnits tillgängligt på svenska anser jag vara ett skäl till att tillåta ett större partiellt bortfall. Det finns informanter som helt klart har uppvisat en önskan att fylla i hela formuläret men som förmodligen på grund av begränsade svenskkunskaper missuppfattat vissa frågor. Jag har då låtit dessa formulär vara kvar.

För fråga 1–4, rörande bakgrundsuppgifter om informanterna, har jag begärt svar på minst två frågor. För fråga 5–11 om informanternas användning av LSL begärs svar på sex av åtta (en fråga är uppdelad i två, därav åtta). Tillhör informanten målgruppen och har fyllt i fråga 12–19 (dvs. frågorna som riktade sig till personer med annat modersmål än svenska) i tillfredsställande grad bortser jag ibland från det kravet. I tveksamma fall har jag alltid sett på svarsfrekvensen för hela enkäten och de operationella reglerna har då i vissa fall frångåtts. I något fall kan man också misstänka något tekniskt fel, t.ex. en enkät som endast innehöll text i ett kommentarsfält men inte något svar på någon fråga. Denna text är då med i analysen av kommentarerna trots att enkäten i sin helhet inte ingår i resultatredovisningen. Det visar sig att ett femtiotal informanter inte har gett något svar på de två första frågorna, dvs. födelseår och kön. Det kan vara så att man av olika skäl inte vill ange dessa uppgifter. Frågan om utbildning i Sverige har också ett anmärkningsvärt högt svarsbortfall (se rad 6 i tabell 5:5). En naturlig orsak till det kan vara att de som inte har någon utbildning i Sverige hoppar över frågan istället för att fylla i ”Ingen avslutad utbildning”.

Frågorna 9, 10 och 11 handlar om huruvida användarna har klickat på länkarna till bilder och filmer samt om de har läst instruktionerna till hur man använder Lexin. Det kan mycket väl vara så att man inte känner till att Lexin innehåller bilder och filmer eller inte förstår vilka instruktioner som åsyftas. Det finns en risk för att informanterna då väljer att helt avstå från att svara på frågan. Jag valde därför att inte ställa några svarskrav på dessa frågor.

Risken för språksvårigheter är som sagt en viktig anledning till att jag tillåter ett större partiellt bortfall än om ordbokens målgrupp hade varit en annan. Det finns också exempel i materialet där man kan misstänka att språksvårigheter ligger bakom det uteblivna svaret. Några informanter har t.ex. fyllt i att de informationskategorier som oftast eftersöks också är de som aldrig efterfrågas. Men bortfall och motstridiga svar kan förstås också bero på slarv. Informanter har en tendens att svara snabbt och inte sällan utan vidare engagemang. Den genomsnittliga informanten anstränger sig heller inte särskilt mycket för att förstå en oklar eller svår fråga. Är innehållet i en fråga inte entydigt är risken stor att svaren inte är tillförlitliga (Japiec et al. 1997).

De operationella reglerna har resulterat i att 61 enkäter avlägsnats från materialet. Underlaget för resultatkapitel 7, Användarna av LSL, omfattar således totalt 802 informanter.

## 5.6 Material i delstudie 3: Enskilda informanterns aktiviteter

Resultatet av länkningen av loggfiler och enkäter utgör underlag för delstudie 3. Antalet länkningar är lika stort som antalet informanter i enkätstudien, dvs. 802. Som framgår av avsnitt 5.3.1 avgränsas urvalet först till de informanter som uppfyller kriterierna för målgruppen och som dessutom svarar att de använder LSL frekvent. Totalt 249 informanter uppfyller dessa kriterier. Länkningarna tillhörande målgruppen beskrivs i inledningen till kapitel 8, Målgruppsanvändning. Därefter fokuserar jag på fem enskilda informanterns aktiviteter. De har valts ut med utgångspunkt i redan kända variabler som modersmål, kön och ålder. Därtill har jag försökt finna länkningar som kan generera belysande exempel och intressanta iakttagelser. En länkning kan vidare innehålla allt ifrån några enstaka aktiviteter till hundratals sådana. Det finns också informanter som bara gör en enda uppslagning, fyller i enkäten och sedan inte använder LSL mer under insamlingsperioden. De finns åter andra som gör en sökning, fyller i formuläret och sedan kryssar för de borttagna informationskategorierna och sedan lämnar ordboken. Det finns slutligen också länkningar som innehåller flera hundra aktiviteter men där fler än ett enkät-id återfinns. Då får man vara försiktig med vilka aktiviteter som kan kopplas ihop med den aktuella målgruppsinformanten. De länkningar som väljs ut får heller inte innehålla för många aktiviteter eftersom de skulle bli svårt att på ett läsvänligt sätt redovisa för dem.

Tabell 5:6 på nästa sida ger exempel på användaraktiviteter utförda av informanten bakom enkät-id 2011:287.<sup>19</sup> Av kolumn 1 framgår att tolv aktiviteter har utförts under den aktuella insamlingsveckan. Åtta aktiviteter infaller på måndag 3 oktober, tre på tisdag 4 oktober samt en sista aktivitet fredagen 7 oktober. Kolumn 2 och 3 innehåller uppgifter om datum och klockslag för användningen. Aktiviteten i rad 1 utförs t.ex. 3 oktober 2011 strax före halv tio på förmiddagen. Kolumn 4 ger information om aktivitetstyp (jfr sökväg, söksträng och sökmetodik hos Almind 2008:27). Med aktivitetstyp avses dels

---

19 Enkäternas egentliga identifikationsnummer har ersatts av året 2011 följt av kolon och ett nummer mellan 1 och 863 (=det ursprungliga antalet enkäter).

utfall av lexikonuppslagning i träff, icke-träff etc. (se avsnitt 5.4.1), dels sökmetodik, dvs. vilka övriga sökfunktioner som aktiveras. Ordet "Checking" innebär att en ruta kryssas för i inställningarna. I kolumn 5 ges sedan specifika uppgifter om respektive aktivitetstyp. Vi ser då att en informationskategori har kryssats för i rad 1, nämligen Exempel. Den första lexikonuppslagningen sker i rad 6 och resulterar i en träff på söksträngen åtaganden.

TABELL 5:6. *Aktiviteter knutna till informanten bakom enkät-id 2011:287.*

Rad	Datum	Tidpunkt	Aktivitetstyp	Aktivitet
1	3	09:29:00	Checking	Exempel
2	3	09:29:04	Checking	Konstruktioner
3	3	09:29:06	Checking	Sammansättningar
4	3	09:29:07	Checking	Uttryck
5	3	09:29:09	Checking	Böjning
6	3	09:29:46	Träff	åtaganden
7	3	09:29:55	Frågeformulär	2011:287
8	3	09:32:03	Träff	sammansättningar
9	4	11:39:08	Träff	instruktioner
10	4	12:13:37	Korr F 3	mail → mil, mal, mall
11	4	12:13:39	Träff	mall
12	7	11:06:42	Korr A	kontroversionell → kontroversiell

Om vi sedan tar fram den enkät som kan knytas till dessa aktiviteter får vi veta mer om användaren, t.ex. att hen troligen är en kurdisk man född 1977 som kom till Sverige 2004. På frågan om vilka informationskategorier han söker mest efter i LSL kryssar han för Betydelse, Motsatser och Exempel. Han har vidare skrivit en kommentar på 51 ord. Den lyder:

Jag tycker att det är jätte bra hemsida och det hjälper mycket till mig. Det funkar som min assistent på jobbet. Dessutom får jag hjälp att lära mig mitt modersspråk(Kurdiska) också som var förbjudet i mitt hemland med hjälp av sidan. Jag tackar jätte mycket till er. Med vänliga hälsningar.<sup>20</sup>

Av kommentaren framgår bl.a. att det svensk-kurdiska lexikonet används i syfte att lära sig modersmålet och alltså inte enbart som hjälp till att lära sig svenska språket, vilket är ordbokens primära syfte.

20 Kommentaren från materialet återges i oredigerat skick. Detsamma gäller för samtliga informantkommentarer i avhandlingen.

På liknande sätt presenteras i kapitel 8 de fem enskilda användarna och deras aktiviteter i LSL. Vid resultatredovisningen ges givetvis mer utförliga presentationer av länkningarna följt av näranalyser.

## 5.7 Etiska aspekter

All forskning måste ta hänsyn till följande fyra etiska aspekter: *informationskravet*, *samtyckeskravet*, *konfidentialitetskravet* och *nyttjandekravet*. Informationskravet innebär en skyldighet att informera potentiella deltagare om vad som sker. Enligt samtyckeskravet måste forskare ha tillstånd för att få samla in känsliga uppgifter. Därtill kommer kraven på sekretess, dvs. konfidentialitet, och till sist nyttjandevillkoret, att uppgifterna som samlas in endast används för forskningsändamål. Att uppfylla samtliga krav kan, även med den bästa vilja, vara en utmanande uppgift. Det kan t.ex. vara problematiskt att uppfylla informationskravet i de fall när forskaren vill kunna observera utan att ge sig till känna eftersom det skulle underminera hela undersökningen (jfr avsnitt 4.2.2 om observer's paradox). Det kan också vara svårt att komma i kontakt med användarna och informera dem om vad som pågår. Enligt Sveningsson et al. (2003:175–189) kan man dock modifiera kraven om man ser till att ingen enskild användare är identifierbar i det resultat som presenteras för forskningsvärlden och allmänheten, och kanske inte ens för forskaren själv.

I föreliggande undersökning har vissa etiska aspekter tagits i beaktande medan andra ha fått stå tillbaka. Jag uppfyller till viss del informationskravet då jag informerar om att det pågår en undersökning av användningen av webbsidan *Lexins svenska lexikon*. Jag meddelar dock inte explicit att användarnas uppslagningar och navigeringsval registreras. Samtyckeskravet kan vara besvärligt vad gäller forskning med hjälp av internet. Om man t.ex. skickar ut en postenkät så är ett svar lika med ett godkännande, dvs. samtycke. Detsamma borde gälla webbaserade frågeformulär. Förutsatt att undersökningen inte innehåller några personuppgifter eller känsliga uppgifter krävs inget formellt samtycke, oavsett vilken ålder informanten har. Materialet innehåller vidare ip-adresser. De räknas som personlig uppgift om de kan kopplas till en fysisk person. I mitt fall kan de inte med säkerhet knytas till en specifik person, speciellt inte så här långt i efterhand (såvida man inte har en fast ip-adress). I övrigt innehåller materialet enligt personuppgiftslagen inga känsliga personuppgifter och forskningen behöver således inte etikprövas. Informanterna ombads uppge namn och mejladress om de ville delta i fler undersökningar. Dessa uppgifter ingår dock inte i analysen.

Slutligen kommer materialet endast att användas för forskningsändamål och ingen informant kommer att identifieras, även om det i något eller några fall skulle vara möjligt.

## 5.8 Inför resultatredovisningen

I det följande redovisas resultaten från de tre delstudierna i tre separata kapitel. I tabell 5:7 ges en översikt över vilken metod och vilket material som ligger till grund för respektive delstudie.

TABELL 5:7. *Delstudie 1–3: Strategi, metod och material.*

	<b>Delstudie I: Användningen av LSL</b>	<b>Delstudie II: Användarna av LSL</b>	<b>Delstudie III: Målgrupps-användning</b>
<b>Strategi</b>	Observation	Enkät	Datatriangulering
<b>Metod</b>	Loggfilsanalys av a) lexikon-uppslagningar b) navigeringsval	Webbaserat frågeformulär	Länkning av loggfiler och enkätdata
<b>Material</b>	2008/2011: loggfiler av typ a, lexikonuppslagningar  2011: loggfiler av typ b, navigeringsval	2007: 10 frågor, 356 enkäter  2011: 20 frågor, 802 enkäter	5 enskilda informanternas ordboksaktiviteter och enkätuppgifter



## 6. Delstudie I: Användningen av LSL

I detta första resultatkapitel redogörs för den faktiska användningen av LSL. Som redan beskrivits i kapitel 5 utgörs underlaget till delstudie 1 av två typer av loggfiler. *Typ a* innehåller enkla lexikonuppslagningar i LSL, dvs. de ord eller ordkombinationer som användarna har skrivit in i sökfältet och slagit upp. *Typ b* utgörs av användarnas övriga aktiviteter på webbsidan, dels i form av klickningar på länkarna till uppläst uttal, bilder och filmer, dels ändringar i inställningarna samt klickningar på knappen med frågetecknet (?). Loggfiler av typ b återfinns endast i 2011-materialet och berörs, som redan konstaterats, endast kort i föreliggande kapitel men får en mer betydande roll i kapitel 8.

I det följande presenteras först statistiska uppgifter om uppslagningarna i LSL. Därefter tar en mer språkvetenskaplig analys av innehållet i loggfilerna vid, vilken inbegriper en jämförelse mellan uppslagningarna i LSL före och efter uppdateringen av ordboken. Avslutningsvis ges en sammanfattning.

### 6.1 Loggfiler av typ a: Lexikonuppslagningar

I tabell 6:1 på nästa sida återges ännu en gång lexikonuppslagningarnas fördelning under de två insamlingsperioderna (jfr tabell 5:1 i avsnitt 5.4.1).

Av tabell 6:1 kan man utläsa att i båda versionerna av LSL är söktrycket högst måndag till torsdag. Användarna är alltså som mest aktiva då. Under fredag och lördag är uppslagningarna markant färre för att åter öka under söndagen. Liknande resultat framkommer t.ex. i Bergenholtz & Johnsen (2005:125).

Uppslagningarna i samtliga Lexin-lexikon uppvisar under hela oktober månad 2011 samma mönster. Man kan tänka sig att detta användningsmönster speglar en genomsnittlig arbets- eller studievecka. Arbetsinsatsen, t.ex. i form av deltagande i lektionsundervisning, är som störst under veckans inledande dagar. Helgen är för många människor studie- och arbetsbefriad och söktrycket går naturligen ner, och det redan under fredagen. På söndagen ställs siktet in på en ny arbets- eller studievecka och förberedelser inför denna verkar generera en ökad användning av LSL. Vad gäller tidpunkt på dygnet har LSL det högsta söktrycket mellan cirka 09.00 och 15.00 samt mellan cirka 19.00 och 22.00, även det ett mönster som går igen om man ser till det totala antalet uppslagningar i hela Lexin (Lexinstatistik 2016).

TABELL 6:1. *Antal uppslagningar per dag under datainsamlingsperioderna 2008 och 2011.*

Veckodag	Måndag	Tisdag	Onsdag	Torsdag	Fredag	Lördag	Söndag	Totalt
<b>2008</b>	26 533	26 968	27 162	24 860	15 607	12 170	20 520	153 820
<b>2011</b>	37 483	33 968	35 511	33 793	22 028	17 936	24 872	205 591

Tidigare framhölls att så länge det svensk-engelska lexikonet fanns tillgängligt i Lexin utfördes majoriteten av uppslagningarna i detta lexikon. När det togs bort (2010) minskade användningen av Lexin drastiskt (se avsnitt 3.2.4). Det kan därför tyckas oväntat att loggfils materialet 2011 är så mycket större (en ökning med 33 %) än materialet från 2008. En rimlig förklaring är att användningen av internetordböcker sedan 2008 generellt sett har ökat i förhållande till användningen av tryckta ordböcker. Ytterligare en tänkbar förklaring till att antalet uppslagningar är så mycket större 2011 kan vara att användarna har upptäckt att LSL har reviderats och att uppslagningar nu i större utsträckning ger träff i ordboken (se tabell 6:2). När användarna av det borttagna svensk-engelska lexikonet, av vilka säkerligen många var svenska modersmålstalare, tvingades vända sig någon annanstans kan nätversionen av Lexin-ordböckerna också ha fått en ännu starkare profil som inlärningslexikon för icke svensktalande. Tillsammans med det faktum att Lexin-ordböckerna fortfarande är kostnadsfria internetordböcker – och (bättre) tvåspråkiga alternativ saknas – gör detta att användningen sannolikt ökar bland dem som tillhör Lexins målgrupp.

I föregående kapitel redogjordes för de fyra typer av utfall som en uppslagning i LSL kan resultera i, dvs. träff, automatkorrigering, korrigering med förslag och icke-träff (se avsnitt 5.4.1). Tabell 6:2 visar frekvenser för de fyra typerna av utfall i 2008 och 2011 års material.

TABELL 6:2. *Frekvenser för uppslagningar i materialet från 2008 och 2011 fördelade på träff, automatkorrigering (A), korrigering med förslag (F) och icke-träff.*

Typ	Frekvens 2008		Frekvens 2011	
	Absolut	Relativ (%)	Absolut	Relativ (%)
Träff	94 677	62	150 221	73
Korrigering A+F	28 083	18	37 068	18
- varav A	12 401	8	12 466	6
- varav F	15 682	10	24 602	12
Icke-träff	31 060	20	18 302	9
<b>Totalt</b>	<b>153 820</b>	<b>100</b>	<b>205 591</b>	<b>100</b>

Av tabell 6:2 framgår som redan nämnts att skillnaden i antalet uppslagningar mellan de två materialen är stor. Den motsvarar en ökning med 33 % i 2011-materialet. LSL4 (2011) har också en mycket bättre träffstatistik än LSL3 (2008), 73 % jämfört med 62 %. Andelen träffar har ökat med 11 procentenheter och andelen icke-träffar har minskat på motsvarande sätt. De korrigerade uppslagningarna (A+F) ligger på samma procenttal, 18 %. De automatkorrigerade uppslagningarna (A) har minskat procentuellt men är något fler 2011. Korrigeringarna med förslag (F) har däremot ökat med två procentenheter (från 10 till 12 %) vilket motsvarar en ökning med nästan 9 000 uppslagningar i 2011-materialet. Eftersom de korrigerade sökningarna inte är rena träffar i ordboken skulle de kunna redovisas tillsammans med icke-träffarna. Andelen misslyckade uppslagningar skulle i så fall ligga på 38 % 2008 och 27 % 2011.

I jämförelse med andra användarstudier ligger andelen rena träffar på en lägre nivå i LSL3. I den uppdaterade versionen från 2011, dvs. LSL4, närmar sig andelen rena träffar däremot nivåer som redovisas i andra loggfilsstudier. I Lorentzens & Theilgaards (2012:656) studie förefaller andelen träffar vara inemot 80 % (de redovisar endast andelen uppslagningar som *inte* gav något resultat). Nygaard & Fjeld (2008:59) redovisar också andelen misslyckade sökningar, 30 %, som innebär att 70 % av uppslagningarna ger träff. Här handlar det dock om två nordiska enspråkiga ordböcker som är mycket större än Lexin och det är naturligt om de har högre träffprocent. Hos de Schryver & Joffe (2004:191) vars studie inbegriper en mindre tvåspråkig ordbok är andelen lyckade uppslagningar 67 % före en uppdatering och 75 % efter denna uppdatering.

## 6.1.1 Träffar 2008 och 2011

### 6.1.1.1 Söksträngar

Enligt tabell 6:2 ger 62 % av uppslagningarna i hela 2008-materialet träff i LSL medan motsvarande andel i 2011-materialet är 73 %. Den absolut vanligaste träffen i båda materialen utgörs av substantiv i obestämd form singular. Därefter kommer verb i infinitivform och adjektiv i grundform. De här formerna sammanfaller vanligen med lemmaformen i de allra flesta ordböcker, dock inte vad gäller verben i Lexin-ordböckerna (se vidare Svensén 2004:128–130). I Lexin får man som bekant träff även på böjningsformer men användarna söker alltså trots det främst på ordens grundform (se avsnitt 3.2.4). Det kan bero på att användarna inte känner till att det går att söka på böjningsformer eller att de är så pass vana ordboksanvändare att de vet att uppslagsord normalt står i grundform.

Om man ser till verbformer i materialet sker dubbelt så många sökningar på verb 2011 jämfört med 2008. Söksträngen utgörs dessutom mycket oftare av presensform i det senare materialet. Det är intressant med tanke på att presensformen som sagt är uppslagsform i Lexin-ordböckerna, till skillnad från vad som vanligen är praxis i (inlärnings)ordböcker (se avsnitt 3.2.2). Det är därför en kraftig indikation på att användarna i 2011-materialet i större utsträckning utgörs av LSL:s målgrupp. Presensformens ökade frekvens kan också förklaras i termer av att, som redan påpekats, många svenska modersmålstalare har lämnat Lexin sedan det svensk-engelska lexikonet avlägsnades.

Även sökningar på adverb är relativt vanliga sett till hela 2008- och 2011-materialen. Därefter blir frekvenserna per ordklass mycket lägre.

I tabell 6:3 återges de 50 vanligaste söksträngarna som resulterat i en träff i LSL, i respektive material.

Av tabell 6:3 framgår tydligt att söksträngarna i de två materialen är påfallande olika. Det märks redan vid en jämförelse av de fem första sökorden, undantaget interjektionen *hej* som förekommer i båda materialen (plats 1 respektive 3). I 2008-materialet återfinns därutöver *moral*, *abstrakt*, *fitta* och *kuk*. Söksträngarna utgörs av tre substantiv och ett adjektiv i grundform. Dels har vi att göra med mer skriftspråkliga ord, dels könsbeteckningar. I 2011-materialet finner vi *reser*, *lägenhet*, *preskriberat* och *jag*. Söksträngarna utgörs av substantiv, verb och pronomen i både grundform och böjningsform. Dessa ord är av mer vardagsnära slag, med undantag av verbet *preskriberat*.

TABELL 6:3. De 50 mest frekventa träffarna i 2008 och 2011 års material.

Plats	2008		2011	
	Söksträng	Frekvens	Söksträng	Frekvens
1	hej	155	reser	585
2	moral	92	lägenhet	527
3	abstrakt	85	hej	438
4	fitta	85	preskriberat	264
5	kuk	84	jag	210
6	sympati	81	irriterad	195
7	naiv	75	har	186
8	etik	72	vår	169
9	bajs	72	z	144
10	konkret	68	bil	144
11	budget	66	gudfar	143
12	demokrati	65	hus	141
13	integritet	65	gör	135
14	objektiv	65	bör	134
15	ideal	64	tar	133
16	perspektiv	60	gå	132
17	relevant	60	supinum	125
18	solidaritet	60	är	120
19	satir	59	gudfadern	119
20	knulla	58	går	119
21	definiera	57	far	118
22	genom	56	bara	116
23	bedrift	55	bord	107
24	prägl	55	för	106
25	jag	54	barn	106
26	hora	54	hålla	105
27	främja	54	vill	105
28	lojal	54	nedanför	101
29	dock	52	se	100
30	subjektiv	52	a	99
31	dynamisk	52	man	99
32	integrera	51	står	96
33	ö	51	vara	95
34	komplex	51	ligger	93
35	aspekt	50	förmåga	92
36	radikal	50	bar	91
37	enda	49	stor	90

38	gammal	49	sa	90
39	rationell	48	bok	89
40	inflytande	48	bra	89
41	å	48	sårad	89
42	är	48	sträcka	88
43	konventionell	48	pågår	88
44	faktiskt	47	byta	87
45	bistånd	46	själ	86
46	spontan	46	ta	86
47	vision	46	skäl	85
48	ödmjuk	45	om	84
49	karaktär	45	brist	82
50	konvention	45	låt	82

En jämförelse av samtliga ord i tabellen uppvisar liknande resultat. Exempelvis är det få ord som är gemensamma i de två materialen; utöver *hej* påträffas *jag* (plats 25 respektive 5) och *är* (plats 42 respektive 18).<sup>21</sup> Att ordet *hej* hamnar i toppen i båda materialen kan tyckas märkligt. Ur ett andraspråksperspektiv ter det sig emellertid ganska naturligt att man som inlärare på nybörjarnivå har ett behov av att få veta i vilka situationer just den här hälsningsfrasen används i svenskan. När det gäller dynamiska webbsidor, där man i realtid kan interagera med (personal på) en myndighet eller ett företag, är det heller inte ovanligt att kommunikationen mellan dator och användare inleds med en hälsningsfras. Det skulle också kunna vara en bidragande orsak till ordets prominenta placering i tabellen.

Av tabell 6:3 kan man vidare utläsa att nästan hälften av söksträngarna i 2008-kolumnen, 23 till antalet, utgörs av substantiv, t.ex. *sympati* (plats 6), *budget* (11) och *perspektiv* (16). Näst vanligast är adjektiven som är 15 stycken, däribland *naiv* (7), *relevant* (17) och *lojal* (28). Verben uppgår till 6 stycken, t.ex. *prägl*a (24), *främj*a (27) och *integr*era (32). De resterande 5 söksträngarna utgörs av en preposition, ett pronomen, ett adverb, en interjektion och slutligen en söksträng som antingen är ett adverb eller ett neutrumböjt adjektiv (*faktiskt*, 44). Som redan antytts kan man se att söksträngarna 2008 i de allra flesta fall utgörs av det aktuella ordets grundform, t.ex. *moral* (2), *definiera* (21) och *rationell* (39). De enda undantagen är söksträngen *är*, dvs. presensformen av verbet *vara*, tillika uppslagsformen i Lexin, och kanske ovan nämnda *faktiskt*. Slutligen har vi en söksträng som antingen är ett substantiv eller ett adjektiv, nämligen ordet *objektiv* (14).

---

21 I det följande meddelas inom parentes exempelordens placering i aktuell tabell, om inget annat anges.

Söksträngarna från 2011 har till att börja med en annan ordklassfördelning. Substantiven är vid första anblicken färre 2011 än 2008 och uppgår till 16 stycken, t.ex. *lägenhet* (plats 2), *bord* (23) och *barn* (25). Det förekommer bara ett ord som otvetydigt är ett adjektiv (*stor*, 37). Verben är dock många fler 2011 än 2008, 17 stycken, t.ex. *har* (7), *gör* (13) och *står* (32). Vidare återfinns en preposition, ett pronomen och en interjektion. De resterande 13 söksträngarna utgörs av homografa ordformer och kan föras till mer än ordklass, t.ex. *bra* (40), *sträcka* (42) och *om* (48).

Homografi är en vanlig företeelse i svenskan (se t.ex. Berg 1978). Med homografi avses dels *extern homografi*, dvs. ”den relation som råder mellan i skrift identiska former tillhörande olika lemman”, dels *intern homografi*, dvs. ”den relation som råder mellan i skrift identiska flexionsformer (grundformen inräknad) inom ett och samma lemma” (a.a. s. XVI). Som externa homografer räknas exempelvis *cyklar*, *fiskar* och *spolar* och exempel på interna homografa ordformer är *barn*, *flyg* och *kasta*. Vad gäller de homografer som förekommer i mitt material kan man i en del fall sluta sig till vilken ordklass (och därmed ofta betydelse) som användaren har åsyftat. I många fall är det däremot omöjligt att veta. När en homograf ordform slås upp i LSL4 genereras därför alla de ordboksartiklar som har en uppslags- eller böjningsform som berörs av den homografa söksträngen (se avsnitt 3.2.4 för ett exempel). I det följande återkommer jag till diskussionen om homografi endast i de fall det anses relevant för undersökningen.

Verben verkar av innehållet i tabell 6:3 att döma spela en viktigare roll för användarna 2011 än 2008. Förutom att de är många fler 2011 innehas första platsen av ett verb (*reser*) medan det första verbet 2008 återfinns först på plats 20 (*knulla*). De flesta verben från 2011 står som redan antytts dessutom i presens. Ytterligare exempel är *tar* (15), *ligger* (34) och *pågår* (43). Andra böjda verbformer förekommer också såsom *preskriberat* (4) och *sa* (38).

Även betydelsemässigt uppvisar söksträngarna i de två topp 50-listorna uppenbara skillnader. För det första innehåller söksträngarna i kolumnen för 2008 mer abstrakta substantiv, t.ex. substantiven *etik* (plats 8), *demokrati* (12) och *solidaritet* (18) och mindre centrala verb och adjektiv såsom *präglade* (24), *främja* (27) och *integrera* (32) respektive *konkret* (10), *objektiv* (14) och *relevant* (17). I kolumnen för 2011 finner vi däremot mer konkreta substantiv som exempelvis *lägenhet* (2), *bil* (10), *hus* (12) och *barn* (25) och mer centrala verb och adjektiv som t.ex. *är* (18), *går* (20) och *vill* (27) respektive *stor* (37) och *bra* (40). I samma kolumn finns också två ord som betydelsemässigt ligger långt ifrån varandra men som är homofona och därför kan vara svåra att hålla isär för en inlärare av svenska, nämligen *själ* (45) och *skäl* (47).

Ytterligare två typer av sökningar bör uppmärksammas. För det första kan det tyckas anmärkningsvärt att söksträngarna *gudfader* (11) och *gudfadern* (19)

hamnar på topp 50-listan år 2011 och det går bara att spekulera i bakgrunden till det. Det är svårt att inte tänka på filmtrilogin *Gudfadern* men heller inte osannolikt att det beror på något helt annat. Sökordet har av någon anledning varit eftersökt under den vecka som datainsamlingen gjordes för denna undersökning. De här söksträngarna kan tänkas falla inom ramen för den typ av frekventa sökningar som uppvisar en tidsbunden social relevans (Wolfer et al. 2014:284–286) där specifika händelser genererar tillfälliga frekvenstoppar för ord som i vanliga fall slås upp mycket mer sällan (se vidare avsnitt 4.3.3.2). För det andra påträffas i topp 50-listan för 2008 års material fyra ord med sexuell innebörd samt ytterligare ett starkt vardagligt ord: *fitta* (4), *kuk* (5), *bajs* (9), *knulla* (20), *hora* (26). Man kan tänka sig att vissa användare, kanske i första hand barn och ungdomar i skolåldern, ägnar sig åt en slags okynnessökning på könsord och dylikt. Den här typen av uppslagningar identifieras också av de Schryver & Joffe (2004) som menar att det kan handla om en generell företeelse vid användning av alla (internet)ordböcker. Författarna hänvisar till internetordboken *Cambridge Dictionaries Online* där topp 50-listan över mest frekventa sökord redigerades om på grund av sitt oanständiga innehåll, innan listan publicerades på webbsidan (de Schryver & Joffe 2004:190). I 2011-materialet dyker däremot dessa ord upp mycket längre ned i frekvenslistorna. Det här är ytterligare en indikation på att användargrupperna före och efter upp-dateringen skiljer sig mycket åt.

Materialen skiljer sig dessutom markant åt frekvensmässigt. Den högsta frekvensen 2008 har ordet *hej* som har slagits upp vid 155 tillfällen. 2011 toppar ordet *reser* listan med 585 uppslagningstillfällen. Vi vet sedan tidigare att 2011-materialet är 33 % större än 2008-materialet. En motsvarande ökning är alltså att vänta för orden från 2011. Frekvensskillnaderna för de vanligaste söksträngarna är emellertid mycket högre än så. De allra mest frekventa söksträngarna 2011 eftersöks nästan tre gånger så ofta som 2008-orden. Orden i den nedre delen av tabellen slås upp dubbelt så många gånger och så småningom jämnar frekvensskillnaderna mellan materialen ut sig. De mest frekventa orden slås alltså upp mycket oftare 2011 än 2008.

När ord beskrivs som centrala och vanliga menar man i regel att de är mycket frekventa i allmänspråket. Detta mäts i sin tur genom olika frekvensöversikter över ords förekomster i t.ex. tidningstexter (jfr Allén 1972). För att få perspektiv på uppslagningarna i ordboken kan man därför undersöka hur vanliga de mest frekventa söksträngarna är i förhållande till just allmänspråket. I det följande jämförs orden i tabell 6:3 med deras förekomst i en referenskorpus. Tidigare nämnda Wolfer et al. (2014) är en av flera användarstudier där frekvenser från lexikonuppslagningar jämförs med frekvenser i referenskorpusar. Avsikten med en sådan jämförelse kan vara att få ett mer gediget underlag till lemmaurvalsprocessen. I det här fallet är syftet emellertid att verifiera eller ny-



ansera de slutsatser som hittills har kunnat dras om de mest frekventa uppslagningarna i LSL3 och LSL4. En lämplig referenskorpus kan t.ex. bestå av dagspress. Det är en typ av text som utkommer regelbundet och i stor upplaga och som avhandlar ett stort antal ämnen. De största dagstidningarna har dessutom en gratis internetversion av pappersförlagan. Många människor kommer således i kontakt med dagstidningar i sin vardag och de läses av en stor heterogen grupp människor (jfr Sköldberg 2004:43).

TABELL 6:4. *Frekvenser i gp08 och gp11 för de 25 vanligaste söksträngarna i 2008 och 2011 års material.*

	LSL3	gp08	LSL4	gp11
Plats	Söksträngar 2008	Frekvens (antal per miljon token)	Söksträngar 2011	Frekvens (antal per miljon token)
1	hej	6	reser	36
2	moral	7	lägenhet	60
3	abstrakt	3	hej	7
4	fitta	0,2	preskriberat	0,9
5	kuk	0,3	jag	3 784
6	sympati	3	irriterad	7
7	naiv	4	har	8 981
8	etik	5	vår	304
9	bajs	1	z	0,2
10	konkret	18	bil	133
11	budget	32	gudfar	0,3
12	demokrati	25	hus	143
13	integritet	11	gör	876
14	objektiv	3	bör	175
15	ideal	7	tar	571
16	perspektiv	39	gå	573
17	relevant	5	supinum	0
18	solidaritet	6	är	14 667
19	satir	3	gudfadern	0,9
20	knulla	0,4	går	1 747
21	definiera	3	far	50
22	genom	494	bara	1 348
23	bedrift	2	bord	37
24	prägl	2	för	10 370
25	jag	3 116	barn	559

I det här sammanhanget är det till att börja med passande att använda en korpus innehållande texter från 2008 och 2011, dvs. då datainsamlingen genomfördes. Den största korpusen av svenska texter finns i Korp, ett konkordansverktyg innehållande över 200 korpusar som tillhandahålls av Språkbanken vid Göteborgs universitet (Borin et al. 2012). De korpusar som passar mitt syfte utgörs av dagstidningstext från Göteborgs-Posten 2008 (16 927 185 token) respektive 2011 (19 935 692 token) (hädanefter gp08 och gp11). I tabell 6:4 har frekvenser tagits fram för de 25 vanligaste söksträngarna i loggfils-materialen från 2008 och 2011 (relativ frekvens genom antal ord per miljon token) (Korp 2016). Vissa av ordformerna är dessutom homografa, men vi bortser från det.

Av tabell 6:4 framgår för det första att de vanligaste söksträngarna i LSL inte kommer i samma frekvensordning i någon av referenskorpusarna. I gp08 får istället *budget* (11), *demokrati* (12), *perspektiv* (16), *genom* (22) och *jag* (25) de högsta frekvenserna. Vanligast i gp11 är *jag* (5), *har* (7) *är* (18), *går* (20) och *för* (24). Intressantare att lägga märke till är kanske att frekvenserna för gp11-orden är avsevärt högre än för gp08-orden. Om man till exempel räknar antalet söksträngar som har en frekvens på 30 eller mer i presskorpusarna är de 4 till antalet i mitt 2008-material och 18 i mitt 2011-material. Om man å andra sidan ser till de ord som har en frekvens på 10 eller lägre i presskorpusarna är de 17 till antalet i mitt 2008-material och 7 i mitt 2011-material. En sådan jämförelse bekräftar det som noterades i jämförelsen av söksträngarna i loggfils-materialet (se tabell 6:3), dvs. att 2011-materialet består av vanligare och mer centrala svenska ord än 2008-materialet. Som redan antytts indikerar uppslagningar av sådana ord att användarna i 2011-materialet är mindre avancerade inlärare än användarna i 2008-materialet.

#### 6.1.1.2 Mest sökta lemman

Majoriteten av uppslagningarna i LSL ger som vi vet en träff i ordboken i såväl 2008 som i 2011 års loggfiler. Vi vet också vilka söksträngar som är vanligast. I tabell 6:5 har samtliga söksträngar lemmatiserats. Det innebär att uppslagningar som tillhör samma ord, dvs. ett grundord och dess böjningsformer, räknas samman och inordnas under en och samma träff/lemma (NLO 1997:173–174). Eftersom söksträngarna existerar i isolation kan vi dock inte veta om en homograf söksträng som *fiskar* avser presensformen av verbet *fiska* eller pluralformen av substantivet *fisk*. I sådana fall räknas söksträngen till båda lemmarna. Söksträngen *bar* (rad 15 i tabell 6:5 nedan) räknas alltså till tre lemmarna: substantivet *bar*, adjektivet *bar* och verbet *bära*. I vissa fall kan man med hjälp av söksträngarna avgöra vilken ordklass det troligen är fråga om. I

fallet med *resa* vet vi t.ex. att presensformen, dvs. *reser*, är den mest frekventa söksträngen i 2011-materialet. Av de 672 förekomsterna i tabell 6:5 utgörs 585 av presensformen. Det är alltså verbet *resa*, och inte substantivet, som intar förstaplaceringen bland mest sökta lemmarna 2011. Tack vare lemmatiseringen får man en bättre helhetsbild (om än i vid bemärkelse) över samtliga ord som faktiskt slås upp oftast i LSL.

Tabell 6:5 återger de 50 mest sökta lemnarna som ger träff i LSL i 2008- och 2011-materialen.

TABELL 6:5. De 50 mest sökta lemnarna som ger träff i 2008 och 2011 års loggfilsmaterial.

Plats	2008		2011	
	Lemma	Frekvens	Lemma	Frekvens
1	<b>hej</b>	155	resa	672
2	<b>vara</b>	141	lägenhet	531
3	<b>den</b>	125	<b>hej</b>	438
4	anse	108	<b>gå</b>	320
5	<b>se</b>	103	<b>ha</b>	304
6	<b>känna</b>	101	<b>ta</b>	295
7	moral	97	preskribera	272
8	abstrakt	96	<b>jag</b>	271
9	<b>få</b>	95	<b>vara</b>	269
10	<b>jag</b>	95	<b>göra</b>	264
11	<b>göra</b>	94	gudfar	262
12	prägl	91	<b>lägga</b>	253
13	fitta	90	<b>se</b>	226
14	god	89	ligga	224
15	kuk	88	bar	214
16	sympati	88	irritera	212
17	objektiv	83	säga	206
18	omfatta	81	skriva	205
19	<b>gå</b>	80	sätta	202
20	<b>ha</b>	80	hålla	201
21	integrera	80	vilja	197
22	<b>ge</b>	79	böra	189
23	integritet	78	<b>den</b>	186
24	naiv	78	<b>få</b>	183
25	<b>ta</b>	78	skola	176
26	gammal	77	vår	176
27	konkret	77	<b>känna</b>	175
28	bajs	74	fara	173
29	etablera	74	äta	173

30	etik	73	för	168
31	<b>komma</b>	73	hus	165
32	relevant	73	be	163
33	radikal	72	<b>ge</b>	163
34	<b>råda</b>	72	<b>komma</b>	161
35	aspekt	71	föra	157
36	definiera	71	byta	155
37	dynamisk	71	<b>finna</b>	155
38	<b>finna</b>	71	bil	153
39	budget	70	låta	152
40	en	70	förstå	148
41	initiativ	70	sitta	148
42	komplex	70	z	146
43	<b>lägga</b>	70	god	144
44	solidaritet	69	stå	144
45	du	68	ko	142
46	främja	68	<b>råda</b>	141
47	ideal	68	far	140
48	led	68	falla	135
49	norm	68	man	135
50	ambitiös	67	slå	134

Av tabell 6:5 kan för det första utläsas att fler ord är gemensamma för de två materialen än vad som var fallet med de olemmatiserade söksträngarna. De fetstilta orden i tabell 6:5 markerar de 16 gemensamma lemmarna i de två materialen. Det som förenar dessa ord är att de flesta av dem är verb och att de är mycket centrala i svenskan. Verben är dessutom ofta flertydiga och förekommer gärna i fraser av olika slag, t.ex. *se, få, gå, lägga, ge* och *komma* (jfr Viberg 2004). Utöver verben återfinns i båda materialen *hej, jag* och *den*. De två förstnämnda känner vi igen från frekvenstabellen över söksträngarna.

Om fokus istället läggs på de 68 lemmarna som särskiljer materialen framträder liknande resultat som för söksträngarna, även om det rör sig om andra ord. I 2008-materialet är substantiven vanligast. Lemmarna utgörs dessutom ofta av abstrakta substantiv, t.ex. *sympati* (16), *integritet* (23) och *aspekt* (35) och relativt avancerade verb som *omfatta* (18), *integrera* (21) och *etablera* (29) samt adjektiv som *relevant* (32) och *radikal* (33). I 2011-materialet är verben klart vanligast. Förutom det mer sällsynta verbet *preskribera* (7) tillhör verben här ett centralt svenskt ordförråd, t.ex. *ligga* (14), *skriva* (18), *äta* (29) och *förstå* (40). Substantiven är konkreta, t.ex. *lägenhet* (2), *hus* (31), *bil* (38) och *ko* (45). Endast ett adjektiv förekommer 2011, *god* (43). Noteras bör slutligen att bland de vanligaste uppslagningarna i de två materialen återfinns mycket få sammansättningar och avledningar.

Vidare innehåller båda materialen flertydiga ord men de starkt polysema orden förefaller vara fler i 2011-materialet. I tabell 6:6 upprepas de 25 vanligaste lemmarna från tabell 6:5 med antal tillhörande numrerade betydelsemoment (lexem i tabellen) samt flerordsuttryck i LSL4. Tabellen ska läsas så att vid en uppslagning av ordet *resa* (rad 1 i 2011-kolumnen) i LSL4 visas två lemmarna som tillsammans har fyra betydelsemoment och ett flerordsuttryck.

TABELL 6:6. *Antal lemmarna, betydelsemoment (lexem) och uttryck i LSL4 för de 25 mest frekventa lemmarna i 2008 och 2011 års material.*

Rad	LSL4							
	2008	Lemna	Lexem	Uttryck	2011	Lemna	Lexem	Uttryck
1	hej	1	1	2	resa	2	4	1
2	vara	5	6	4	lägenhet	1	1	0
3	den	2	3	2	hej	1	1	2
4	anse	1	1	0	gå	1	5	53
5	se	1	3	7	ha	3	4	8
6	känna	1	3	1	ta	1	6	34
7	moral	1	2	0	preskribera	1	1	0
8	abstrakt	1	1	1	jag	2	2	1
9	få	2	4	30	vara	5	6	4
10	jag	2	2	1	göra	1	4	14
11	göra	1	4	14	gudfar	1	1	0
12	prägl	1	1	0	lägga	1	3	13
13	fitta	1	1	0	se	1	3	7
14	god	1	3	26	ligga	1	1	16
15	kuk	1	1	0	bar	4	4	4
16	sympati	1	1	0	irritera	1	1	0
17	objektiv	2	2	0	säga	1	1	14
18	omfatta	1	1	0	skriva	1	2	1
19	gå	1	5	53	sätta	1	3	23
20	ha	3	4	8	hålla	1	5	31
21	integrera	1	1	0	vilja	3	3	9
22	ge	1	1	8	böra	1	2	1
23	integritet	1	1	0	den	2	3	2
24	naiv	1	1	0	få	2	4	30
25	ta	1	6	34	skola	4	6	0
<b>Totalt</b>		<b>35</b>	<b>59</b>	<b>191</b>		<b>42</b>	<b>76</b>	<b>268</b>

Tabell 6:6 åskådliggör att såväl lemman och betydelsemoment som uttryck är fler för de lemmatiserade söksträngarna 2011 än 2008. Antalet lemman ökar med 10 %, lexemen med 29 % och uttrycken med 40 %. Skillnaden är ännu större om man ser till samtliga 50 ord från tabell 6:5. Antal lexem/lemma blir då 1,6 för 2008-orden och 1,9 för 2011-orden. Antal uttryck/lemma är 3,4 för 2008-orden och 5,7 för 2011-orden. En trolig anledning till att de här orden slås upp, och kanske i ännu högre grad av andraspråkstalare än förstaspråkstalare, är just deras flertydighet. Lemmana i 2011-kolumnen utgörs som vi tidigare sett av mera vanliga och centrala svenska ord än i 2008-kolumnen. Som tabellen visar är alltså de vanligare och centralare orden i högre grad flertydiga (jfr avsnitt 3.1.1 om svåra ord ur ett ordinlärningsperspektiv). Detta bekräftar ännu en gång diskrepansen användargrupperna emellan.

Vad gäller de obskena orden är förhållandet detsamma för de lemmatiserade uppslagningarna som för söksträngarna (se tabell 6:3). Dessa ord återfinns fortfarande endast i 2008 års kolumn. Det första ordet som kan föras till denna kategori i 2011-materialet återfinns på plats 624, med 44 uppslagningar (*hora*).

I avsnittet om de mest frekventa söksträngarna undersöktes hur vanliga de är i förhållande till allmänspråket. Det var ett sätt att få perspektiv på uppslagningarna i ordboken. När det gäller de mest frekventa lemmana i mitt material kan det istället vara intressant att jämföra dem med någon form av basordförråd i svenskan. Termen *ordförråd* kan som redan beskrivits definieras utifrån ett receptivt eller ett produktivt perspektiv (se avsnitt 3.1.1). Enligt Enström (2016:39) ger ett ordförråd på cirka 5 000 ord en nittioprocentig täckning av de flesta texter. En ännu högre täckningsgrad uppnås framför allt genom ytterligare inläring av ord som är typiska för olika ämnesområden och texttyper. För att kunna tillgodogöra sig innehållet i en svensk dagstidning bör man kunna förstå ca 40 000 ord (a.a. s. 41).

När ett ordförråd på omkring 18 000 ord skulle väljas ut till den första upplagan av LSL använde man sig av i huvudsak fyra olika källor, bl.a. ett frekvensbaserat skriftspråkligt ordförråd och ett praktiskt ordförråd hämtat ur ordindex till ett antal läromedel för invandrarundervisning (Gellerstam 1999b:6). Ett senare försök att ta fram ett frekvensbaserat svenskt basordförråd gjordes inom ramen för Kelly-projektet (Volodina & Johansson Kokkinakkis 2012). Projektet gick ut på att skapa ett verktyg för språkinläring, och ett basordförråd på nio olika språk togs fram. De svenska orden är excerperade från ett omfattande textmaterial från webben på 114 miljoner tokens och ska representera ett samtida språkbruk. Webtexterna täcker dels in olika genretyper, dels både skriftspråkliga och mera talspråkliga texttyper (t.ex. bloggtexter). Den så kallade Kelly-listan är lemmatiserad och omfattar drygt 8 400 lemman. Orden är frekvenssorterade vilket innebär att ju högre upp i listan de står, desto mer centrala är de. Överst i listan återfinns lemmana *och*, *vara*, *i*, *ha* och *dess*.

I tabell 6:7 ges frekvenser för de 25 mest frekventa lemmarna i 2008 och 2011 års material tillsammans med information om deras placering i Kelly-listan. För tydlighets skull har de ord som är gemensamma i de två kolumnerna markerats med fetstil.

TABELL 6:7. De 25 mest frekventa lemmarna från 2008 och 2011 års material och deras placering i Kelly-listan.

Rad	2008		2011	
	Lemma	Placering i Kelly-listan	Lemma	Placering i Kelly-listan
1	<b>hej</b>	1076	resa	914
2	<b>vara</b>	2	lägenhet	1407
3	<b>den</b>	21	<b>hej</b>	1076
4	anse	183	<b>gå</b>	58
5	<b>se</b>	53	<b>ha</b>	4
6	känna	121	<b>ta</b>	51
7	moral	2119	preskribera	-
8	abstrakt	4324	<b>jag</b>	16
9	<b>få</b>	30	<b>vara</b>	2
10	<b>jag</b>	16	<b>göra</b>	45
11	<b>göra</b>	45	gudfar	-
12	prägl	1945	lägga	172
13	fitta	-	<b>se</b>	53
14	god	200	ligga	168
15	kuk	-	bar	1355
16	sympati	4082	irritera	6218
17	objektiv	2848	säga	59
18	omfatta	853	skriva	96
19	<b>gå</b>	58	sätta	250
20	<b>ha</b>	4	hålla	138
21	integrera	3266	vilja	43
22	ge	91	böra	70
23	integritet	1706	<b>den</b>	21
24	naiv	3821	<b>få</b>	30
25	<b>ta</b>	51	skola	15

Av tabellen kan man för det första utläsa att fyra ord från loggfilsmaterialiet saknas helt i Kelly-listan, dels två obscena ord i 2008-kolumnen (rad 13, 15), dels två ovanligare ord i 2011-kolumnen (rad 7, 11). Vidare har verbet *vara* (rad 2 och 9) högst placering i Kelly-listan av samtliga ord i de två loggfilsmaterialen. Lägst placering har verbet *irritera* (rad 16 i 2011-kolumnen). Fyra av

de tolv unika orden i 2008-kolumnen har en placering på högre än 1 000, t.ex. *god* på rad 14 med en placering i Kelly-listan på 200. I 2011-kolumnen har 9 av de tolv unika orden en placering på 1 000 eller högre såsom *skriva* på rad 18 med placeringen 96. Lemmana i 2011-kolumnen befinner sig alltså närmare kärnan i basordförrådet (enligt Kelly-listan) än lemmarna i 2008-kolumnen.

Sammanfattningsvis kan sägas att loggfilerna med användarnas uppslagningar som leder till träff, framför allt i 2011-materialet, visar att vanliga svenska ord slås upp i stor utsträckning. Tillsammans med det faktum att träffprocenten är mycket högre i den uppdaterade versionen av LSL bekräftas att lemmaurvalet är väl anpassat till LSL:s faktiska användare. Det totala antalet lemman i 2008-loggfils materialet uppgår till ca 15 000 och i 2011-loggfils materialet är motsvarande siffra drygt 16 000. LSL innehöll före uppdateringen ca 28 500 uppslagsord och efter uppdateringen uppgår lemmaantalet till ca 31 500. Av detta kan man sluta sig till att under de aktuella datainsamlingsveckorna slogs omkring hälften av LSL:s uppslagsord upp mellan 1 och 155 gånger år 2008 och mellan 1 och 672 gånger 2011 (jfr tabell 6:5). Med tanke på att materialen endast innehåller en veckas uppslagningar får man nog säga att det är relativt många av ordbokens lemman som slås upp. Det hade varit intressant att veta hur det förhåller sig över ett helt år.

Hitills har fokus legat på de ord som faktiskt återfinns bland de mest frekventa träffarna i materialet. Det kunde också vara intressant att uppmärksamma ord som *inte* påträffas i loggfilerna, men det faller utanför ramen för den här undersökningen.

## 6.1.2 Korrigerade uppslagningar 2008 och 2011

Det finns som redan framgått två typer av korrigerings som en uppslagning i LSL kan leda till (se avsnitt 5.4.1). Antingen automatkorrigeras uppslagningen till en annan ordboksartikel eller så genererar uppslagningen en korrigeringsförslag (minst två) som innebär att ett visst antal existerande uppslagsord presenteras. Dessa förslag genereras utifrån editeringsavstånd (se avsnitt 3.2.4.2).

### 6.1.2.1 Automatkorrigering

I tabell 6:8 listas de 25 vanligaste söksträngar i materialen från 2008 och 2011 som leder till automatkorrigering. Kolumnen ”Utfall” innehåller det uppslagsord som den sökta strängen automatkorrigeras till. Det kan vara på plats att påminna om att i LSL4 är de allra flesta böjningsformer sökbara och



ger alltså direkt träff i ordboken, till skillnad från i LSL3. En böjningsform som ännu inte är sökbar är däremot presens particip.

Av tabell 6:8 framgår att 20 av de vanligaste uppslagningarna som automatkorrigeras i 2008-materialet är rättstavade svenska ord, t.ex. *estetik* (plats 1), *benägenhet* (5), *dekadens* (10), *inventering* (15) och *mytoman* (25). De fem felstavade orden är *\*konjunktur* (6), *\*igentligen* (12), *\*orginal* (16), *\*originell* (17) och *\*assitten* (20). I alla fem fallen handlar det om en enda bokstav som ska läggas till eller bytas ut för att ordet ska bli rättstavat.

TABELL 6:8. De 25 vanligaste uppslagningarna som leder till automatkorrigering 2008 och 2011, i frekvensordning.

Plats	2008			2011		
	Söksträng	Utfall	Frekvens	Söksträng	Utfall	Frekvens
1	estetik	estetisk	30	tragidin	tragedin	141
2	paranoid	paranoia	27	delegerings	legerings	26
3	kontext	konvext	26	återstående	återseende	22
4	symtom	symptom	23	sovjetunionen	Sovjetunionen	22
5	benägenhet	belägenhet	20	ont om	vet om	22
6	konjunktur	konjunktur	18	Upp satt	uppsatt	18
7	hybris	hybrid	17	vore	vare	17
8	migration	emigration	17	akompanjerad	ackompanjerad	17
9	avsedd	ansedd	16	argsnit	argsint	17
10	dekadens	dekadent	16	defenition	definition	17
11	frekvent	frekvens	16	fajitas	fajtas	17
12	igentligen	egentligen	16	förespråk	förespråka	17
13	pömsig	pösigt	16	fasanfullt	fasansfullt	17
14	diskurs	diskus	15	flatig	flitig	17
15	inventering	investering	15	agater	gater	17
16	orginal	original	15	genorositet	generositet	17
17	originell	originell	15	gallande	gillande	17
18	rabbi	rabbin	15	häradad	härddad	17
19	rationalitet	nationalitet	15	hörsam	hörsal	17
20	assitten	assietten	14	inhumana	humana	17
21	beteckning	betäckning	14	interagera	integrera	17
22	foajen	foajén	14	kirug	kirurg	17
23	inneha	innehav	14	majonäs	majonnäs	17
24	retorik	retorisk	14	nedgrävd	nedärvd	17
25	mytoman	pyroman	13	nerrivet	nerringt	17

Totalt sett korrigeras 15 av de 25 uppslagningarna i 2008 års material till (i varierande grad) relevanta, existerande uppslagsord. Goda exempel är *konjunktur* till *konjunktur*, *orginal* till *original* och *foajen* till *foajén*. Automatkorrigeringarna kan också vara mindre lyckade. Av de misslyckade slaget ser vi *kontext* till *konvext*, *diskurs* till *diskus*, *inventering* till *investering*, *rationalitet* till *nationalitet* och *mytoman* till *pyroman*. De här korrigeringarna leder till uppslagsord som visserligen har likheter rent ortografiskt med sökordet, men som troligen inte är relevanta.

I 2011-materialet är 12 av de 25 uppslagningarna som leder till automatkorrigering rättstavade svenska ord, t.ex. *delegerings* (2), *fajitas* (11), *agater* (15), *hörsam* (19) och *nerrivet* (25). Det rör sig ofta om böjda former. Exempel på rent felstavade ord är *\*tragidin* (1), *\*akompanjerad* (8), *\*argsnit* (9), *\*definition* (10), *\*fasanfullt* (13), *\*genorositet* (16), *\*gallande* (17) och *\*majonäs* (23). Även här har vi alltså att göra med smärre felstavningar. Korrigeringen *sovjetunionen* till *Sovjetunionen* visar att rättstavningshjälpen är känslig för versaler. Korrigeringen från *återstående* till *återseende* påvisar dessutom att presens participformen inte är sökbar.

Totalt sett korrigeras 13 av de 25 uppslagningarna i 2011 års material till relevanta existerande uppslagsord. Det är i det närmaste lika många som i 2008 års material. Relevanta korrigeringar är t.ex. *tragidin* till *tragedin*, *Upp satt* till *uppsatt*, *argsnit* till *argsint*, *fasanfullt* till *fasansfullt* och *kirug* till *kirurg*. Bland de (högst troligen) misslyckade automatkorrigeringarna återfinns *ont om* till *vet om*, *fajitas* till *fajtas*, *hörsam* till *hörsal* och *nerrivet* till *nerringt*.

De automatkorrigerade söksträngarna uppvisar i de två materialen liknande skillnader som analysen av träffarna i LSL visar. I 2008-materialet står söksträngen t.ex. oftare i grundform än i 2011-materialet. Orden skiljer sig dock inte lika mycket åt vad gäller graden av abstraktion. Uppslagningen *tragidin* har vidare en frekvens (141) som avviker stort ifrån övriga frekvenser i båda materialen. Söksträngen är kanske ett av exemplen på ett ord med tidsbunden social relevans (jfr tabell 6:3 i avsnitt 6.1.1 och exemplet *gudfadern*).

Av de automatkorrigerade uppslagningarna från 2008 ger 10 direkt träff år 2011, dvs. i LSL4. Det beror antingen på att de orden numera ingår i lemmalistan, t.ex. *paranoid*, *benägenhet* och *dekadens*, eller på att fler böjningsformer nu är sökbara, t.ex. *ansedd*. I absoluta tal är det inte så många fler uppslagningar som leder till en lyckad automatkorrigering 2011 i jämförelse med 2008 och procentuellt sett har de minskat med två enheter (se tabell 6:2). Att fler böjningsformer blev sökbara i LSL4 har alltså inte haft någon större effekt på andelen automatkorrigeringar, utan det är helt enkelt andra ord som automatkorrigeras eftersom de flesta böjningsformer numera ger direkt träff i ordboken.

### 6.1.2.2 Korrigering med förslag

Den andra typen av korrigering är korrigering med förslag. I tabell 6:9 och 6:10 listas de 25 vanligaste uppslagningarna av detta slag i respektive material. Där anges också antalet ordförslag som korrigeringen genererar samt exempel på föreslagna uppslagsord.

TABELL 6:9. *De 25 vanligaste uppslagningarna som leder till korrigering med förslag 2008.*

2008				
Plats	Antal	Söksträng	Utfall	Förslag
1	88	empati	3	apati, ympat, sympati
2	36	gamla	5	–
3	35	cv	8	ev, av, tv, ca, co, cl, cm, cp
4	35	gammla	16	–
5	30	präglas	3	präгла, präglad, präglat
6	26	engagerad	4	engagera, engagerat, engagerar, engagerade
7	26	sarkasm	2	parkas, barkass
8	24	neger	2	seger, niger
9	21	autonomi	3	–
10	19	definiera	2	–
11	19	pedant	15	dant, pant, plant, pekat, pedal, nedan, ledan, medan
12	18	håglös	3	–
13	17	anal	6	ana, anar, banal, kanal, antal, anal-
14	17	fingerad	4	fingerar, fingerat, fingerar, fingerade
15	17	livegna	3	liven, literna, livlina
16	17	vasall	7	svall, vall, nasal, nasal-, nasala, avfall, avkall
17	16	donator	7	–
18	16	epitet	15	pietet, pitt, bitet, nitet, litet, vitet, piket, opiet, slitet
19	16	karma	11	krama, arma, karm, larma, varma, kamma, karga
20	16	sverige	2	Sverige, sverige-
21	16	åäö	72	-ö, dö, mö, kö, hö, tö, än, ät
22	15	dement	2	gement, cement
23	15	dynamik	2	dynamit, dynamisk
24	15	järv	5	jäv, kärv, värv, järn, djärv
25	15	milis	18	–

Vad gäller ordförslag som ges kan vi för uppslagningarna 2008 inte alltid få helt tillförlitliga uppgifter om vilka konkreta ord som presenteras eftersom LSL3 inte längre finns tillgänglig. Därför saknar tabell 6:9 ordförslag för vissa uppslagningar. Genom att slå upp orden i något av de tvåspråkiga lexikonerna – som ju ännu inte uppdaterats – kan man emellertid se vilka svenska ordförslag som då ges. De motsvarar delvis de förslag som gavs vid en sökning i 2008 års version av LSL. Något som dock måste beaktas är att rättstavningshjälpen har förbättrats sedan 2008 och fungerar på alla Lexin-lexikonerna – uppdaterade eller inte. Detta kan alltså påverka vilka förslag som ges. Ordförslagen anges vidare i den ordning de presenteras i LSL. Om förslagen är många anges så många som utrymmet i tabellen tillåter.

Samtliga söksträngar förutom *åö* i tabell 6:9 utgör identifierbara svenska ord. Alla ord utom ett, *gammla* (plats 4), är dessutom rättstavade. Söksträngarna utgörs av grundformen i 20 fall av 25. Även här tillhör många av orden ett mer avancerat ordförråd än vad som kan förväntas av en nybörjarinlärare i svenska. Exempel på sådana substantiv är *empati* (1), *sarkasm* (7), *autonomi* (9) och *vasall* (16). Bland adjektiven finner vi t.ex. *håglös* (12) och *dement* (22). Bland de få verben är *fingerad* (14) ett lämpligt exempel.

I LSL4, dvs. LSL av idag, leder 14 av de aktuella söksträngarna till träff, däribland *empati*, *håglös* och *dement*. En av söksträngarna, *milis* (25), automatkorrigeras numera olyckligtvis till uppslagsordet *mil*. Övriga 10 söksträngar utgörs följaktligen fortfarande av lemmaluckor, t.ex. *sarkasm* (7), *pedant* (11), *epitet* (18), *karma* (19) och *järv* (24). När det gäller ordformer som *sarkasm* och *pedant* kunde mycket lämpliga ordförslag ha varit de befintliga lemmarna *sarkastisk* och *pedantisk*, istället för de givna förslagen *parkas* och *barkass* respektive *dant*, *pant* och *plant* m.fl. Vi återkommer till rättstavningshjälpens funktionalitet längre fram.

Uppslagningarna från 2011 som leder till korrigering med förslag presenteras i tabell 6:10. Söksträngarna 2011 utgör till att börja med en mer blandad ordsamling än de från 2008. Utländska ord varvas med oidentifierbara söksträngar och mer svårbedömda felstavningar av förmodade svenska ord. De rättstavade svenska orden är 10 stycken, t.ex. *generera* (5), *hejdå* (11), *löpande* (20) och *heroisk* (21). De felstavade svenska orden verkar vara 7 till antalet, såsom *kanalle* (6), *tyvär* (16) och *sveder* (17). De utländska söksträngarna uppskattas till 5 stycken, exempelvis *cat* (2), *tikanis* (4) och *hog* (24). Söksträngen *tikanis* tolkas här som den grekiska frasen *ti kanis* med den ungefärliga betydelsen 'Hur mår du?'. Till sist har vi 3 stycken enstaviga söksträngar, *za* (10), *ze* (15) och *zi* (22) som är mer svåridentifierade. Det är tänkbart att de två förstnämnda strängarna är polska ord som betyder 'för, istället för' respektive 'från'.

TABELL 6:10. De 25 vanligaste uppslagningarna som leder till korrigering med förslag 2011.

2011				
Plats	Antal	Söksträng	Utfall	Förslag
1	36	lexin	31	exen, exil, lin, lien, len, kemin, kexen, taxin
2	34	cat	6	at, ca, fat, lat, mat, hat
3	31	hårdigt	5	färdigt, värdigt, härdigt, härligt, ihärdigt
4	29	tikanis	6	tians, tikars, fikans, tikens, bikinis, tikarnas
5	25	generera	13	genera, genrer, genrers, genererna, generar
6	24	kanalle	15	nalle, kalle, knall, knall-, knalla, knalls, kanal
7	23	bevar	8	vevar, revar, bovar, bävar, benar, betar, bevara
8	22	hi	13	bi, ni, vi, si, hd, ha, ho, hs, hr, hö, hy, hiv, hit
9	22	fafarr	10	farfar, fasar, faxar, katarr, safari, faxars, fanfar
10	19	za	8	ca, la, ja, ha, va, ta, sa, zs
11	19	hejdå	2	hejd, hejda
12	18	leventet	3	leendet, adventet, elementet
13	18	fate	12	fat, gate, late, sate, fete, fat-, fatt, fats, flate
14	18	bräserad	10	raserad, baserad, brisera, bogserad, blaserad
15	17	ze	9	fe, ge, de, be, le, ve, te, se, zs
16	17	tyvär	2	tvär, tyvärr
17	17	sveder	5	seder, smeder, svider, sveper, -sveder
18	17	haye	6	hage, hade, hane, hale, hake, hare
19	17	fez	7	fe, feg, fen, fel, fem, fet, fes
20	16	löpande	2	lönande, -löpande
21	16	heroisk	8	heroin, serbisk, persisk, hednisk, hektisk
22	15	zi	5	bi, ni, vi, si, zs
23	15	sverige	2	Sverige, sverige-
24	15	hog	15	ho, fog, dog, bog, nog, log, tog
25	14	zo	13	Co, go, do, bo, lo, jo, ko, ho, ro, po

När det gäller antalet ordförslag som ges varierar detta stort i båda materialen. I de flesta fall är de färre än tio. Två undantag är söksträngen *ääö* (21) i tabell 6:9 som ger hela 72 ordförslag och (lustigt nog) *lexin* (1) i tabell 6:10 som ger 31 förslag. Vid första anblicken är kanske inte antalet ordförslag så besvärande. I tabell 6:9 genereras relevanta ordförslag vid uppslagningarna på passivformen *präglas* (rad 5) och perfekt participformen *fingerad* (rad 14) i form av böjningsformer av dessa verb. (De angivna böjningsformerna blir som vi redan vet sökbara efter uppdateringen av LSL3 och ger alltså träff i LSL4.) Men när

man tittar närmare på vilka konkreta förslag som presenteras och framför allt i vilken ordning de kommer blir bedömningen en annan. Det visar sig då att editeringsavstånd inte är en tillräckligt vass urvalsmetod att basera ordförslagen på. För det första är ordförslagen, utifrån ett semantiskt perspektiv, många gånger väsensskilda från sökordets betydelse, och därmed helt irrelevanta, se t.ex. *generera* (rad 5, 13 förslag) i tabell 6:10 med ordförslag som *genera*, *genrer* och *generna*. Ett annat exempel från samma tabell är *bräserad* (rad 14, 10 förslag) som bland annat resulterar i förslagen *raserad*, *baserad* och *bogserad*. För det andra hamnar förslag som är tämligen goda ofta först i slutet av förslagslistan, se t.ex. *hårdigt* (3) i tabell 6:10 där de fem ordförslagen ges i följande ordning: *färdigt*, *värdigt*, *händigt*, *härligt*, *ihårdigt*. I samma tabell finner vi också redan nämnda *lexin* (1) där de mest rimliga förslagen, *lexika* och *lexikon*, placeras allra sist bland de 31 ordförslagen. Ytterligare två exempel är uppslagningarna *bevar* (7) och *zo* (25) i tabell 6:10 där de relevanta ordförslagen, *bevara* och *bevares* respektive *zoo*, anges sist respektive näst sist.

Rättstavningshjälpen är som sagt baserad på editeringsavstånd (se avsnitt 3.2.4) och inte på uttal eller utifrån några semantiska kriterier. Att basera rättstavningshjälpen endast på editeringsavstånd verkar fungera vid mindre felstavningar förutsatt att motsvarande rättstavade ord återfinns i ordboken. Men i de fall det sökta ordet inte ingår i LSL blir utfallet av en korrigering många gånger helt missvisande. Söksträngen *rationalitet* skulle t.ex. i bästa fall leda användaren till adjektivet *rationell* istället för substantivet *nationalitet*, såsom sker här.

SAOL Plus (2007), cd-versionen av SAOL13, är ett exempel på en ordlista som har en rättstavningshjälp baserad på både stavning och uttal (Berg et al. 2008a, jfr Lew & Mitton 2011). Den rättstavningshjälpen är utvecklad av företaget Oribi AB och har till viss grad också anpassats specifikt för uppslagsorden i SAOL. Söksträngen *sjonalister* ger t.ex. som enda förslag uppslagsordet *journalist*. I LSL4 leder motsvarande söksträng till en icke-träff. Söksträngen *haye* (18) från tabell 6:10 resulterar vidare i 10 ordförslag i SAOL Plus. Bland förslagen återfinns bland andra *hagen*, *haj* och *hojen* som visar på en uttalsmässig rättstavningskorrektion, att jämföra med de editeringsbaserade ordförslagen i LSL4 vilka är *hage*, *hade*, *hane*, *hale*, *hake* och *hare*. Vidare leder söksträngen *zo* (25) från tabell 6:10 till de två uppslagsorden *zoo* och *zoo-* i SAOL Plus. I LSL4 hamnar *zoo* näst sist bland de 13 ordförslagen.

SAOL Plus ger inte alltid färre ordförslag än LSL4. Det verkar dock som att rättstavningshjälpen i ordlistan maximalt genererar 10 ordförslag. Förslagen är alltså många gånger vettigare tack vare att rättstavningshjälpen är anpassad efter både editeringsavstånd och uttalsmässig stavning.

När ordförslagen är få är den editeringsbaserade ordningsföljden en mindre allvarlig defekt. Men om användaren möts av ett större antal ordförslag blir det

långt besvärligare och onödigt tidsödande att finna det mest relevanta ordet – i den mån det över huvud taget existerar ett sådant. Att värdera vilka förslag som är relevanta är dessutom en helt annan sak för inlärare än för modersmåls-talare. Inlärare förutsätts följaktligen att på basis av sina begränsade svensk-kunskaper kunna avfärda irrelevanta ordförslag som i sin tur har baserats på editeringsavstånd.

Avslutningsvis kan konstateras att *få* ordförslag å ena sidan är mer användarvänligt än *många* förslag. Å andra sidan är inget av dessa scenarion önskvärt så länge ordförslagen inte är relevanta i relation till den sökta strängen. I sådana fall är det måhända mer lämpligt att låta användaren få veta att sökordet helt enkelt inte finns i ordboken. I annat fall skulle det vara funktionellt att begränsa antalet ordförslag som kan ges. Det bör man dock inte göra utan att en uttalsmässig rättstavningskontroll har implementerats. En reduktion av antalet ordförslag skulle nämligen i nuläget innebära att de mest relevanta ordförslagen många gånger uteblir.

### 6.1.3 Icke-träffar 2008 och 2011

I de föregående avsnitten redogjordes för uppslagningar som antingen genererar en träff i LSL eller resulterar i någon typ av korrigeringsord som i bästa fall leder fram till ett relevant uppslagsord. I det här avsnittet redovisas de allra vanligaste uppslagningar som resulterar i en icke-träff och där användare alltså möts av meddelandet "Ordet x finns inte i lexikonet". Tabell 6:11 innehåller de 25 vanligaste söksträngarna som leder till icke-träff i 2008- och 2011-materialet.

Andelen rena icke-träffar är som vi vet 20 % 2008 och 9 % 2011 (se tabell 6:2 i början av kapitlet). Risken att en uppslagning resulterar i en icke-träff är alltså avsevärt lägre i LSL4. Listor med icke-träffar kan som redan framgått vara en mycket bra utgångspunkt när det gäller att uppdatera lemmaurvalet i en ordbok (se t.ex. de Schryver & Joffe 2004, Nygaard & Fjeld 2008). Sådana listor användes också vid uppdateringen av LSL (se avsnitt 3.2.3). Bland orden i 2008-kolumnen i tabell 6:11 har 14 inkluderats, t.ex. *kontroversiell* (2), *kognitiv* (3), *integration* (8), *melankolisk* (19) och *relevant* (24). Söksträngarna *tillskärning* (4), *besserwisser* (11) och *etnicitet* (13) ger däremot fortfarande ingen träff i LSL4. Uppslagningen *holistisk* (22) automatkorrigeras i LSL4 till *politisk* och sökningen på ordet *överkurs* (9) genererar en lista med nio ordförslag. Orden *tillskärning*, *besserwisser*, *etnicitet*, *holistisk* och *överkurs* utgör alltså nyordskandidater till en framtida uppdatering. Från ett användarperspektiv skulle det förmodligen vara mycket värdefullt att få reda på att dessa ord utgör fullgoda svenska lexikala enheter, även om de inte återfinns i LSL.

TABELL 6:II. De 25 vanligaste icke-träffarna under insamlingsperioderna 2008 och 2011.

Plats	2008		2011	
	Icke-träff	Antal	Icke-träff	Antal
1	bambarder	61	what r u doing	193
2	kontroversiell	57	pennfodral	81
3	kognitiv	55	farsi	57
4	tillskränning	53	ekipera	32
5	benämning	43	buffelhudar	30
6	empatisk	41	framförallt	29
7	interaktion	40	punktform	25
8	integration	36	sondmatning	19
9	överkurs	34	incident	19
10	intention	34	en mening	16
11	besserwisser	34	globalisering	16
12	kontinuitet	33	disktrasa	14
13	etnicitet	33	gap hacker	13
14	innovativ	32	golvmopp	13
15	rekrytering	30	triangelhandeln	13
16	ä	27	drivenmotor	13
17	konsumera	26	snirkliga utsmyckningar	12
18	intervention	26	encumber	12
19	melankolisk	24	all betala	12
20	optimal	24	population	11
21	nyckfull	23	krig mot azteker och indianer	11
22	holistisk	23	pos selene	11
23	skildring	22	spindelled	10
24	relevant	22	nuförtiden	10
25	logistik	21	involvera	10

Typiskt för orden från 2008 är att de är rättstavade ord i grundform förutom *bambarder* på första plats och *ä* på sextonde plats. Tidigare påpekades att avledningar och framför allt sammansättningar är få bland de mest frekventa uppslagningarna. Bland icke-träffarna återfinns däremot både avledningar och sammansättningar, t.ex. *empatisk* (6), *överkurs* (9) och *rekrytering* (15).

De vanligaste icke-träffarna från 2011 utgör en mer brokig skara av uppslagningar. Många av orden står även här i grundform men fler av dem står i böjningsform än i 2008-kolumnen. Några av söksträngarna utgörs dessutom av flera ord (*en mening* 10, *snirkliga utsmyckningar* 17) och av utländska ord och fraser (t.ex. *encumber* 18, *what r u doing* 1). Många av orden i 2011-kolumnen används oftare i vardagliga sammanhang än orden från 2008-kolumnen som



upplevs som mer avancerade, jämför t.ex. *kognitiv* (3), *benämning* (5) och *holistisk* (22) från 2008-materialet med *pennfodral* (2), *disktrasa* (12) och *golvmopp* (14) från 2011-materialet. Mönstret känns igen från de tidigare analyserna av träffar och korrigerade uppslagningar i LSL.

I avsnitt 4.3.3.2 redogjordes för en pilotstudie av Hult (2008b) som inriktar sig på icke-träffarna från en dags sökningar i LSL3 (i oktober 2006). Eftersom urvalet var såpass litet kunde en noggrannare granskning av hela 2006-materialet göras. De iakttagelser som görs i 2006-materialet visar sig överensstämma med flera av de observationer som görs beträffande orden i både 2008- och 2011-kolumnen i tabellen ovan. Användarna söker t.ex. på såväl grundform som pluralform och bestämda former. I likhet med 2008-orden utgörs icke-träffarna 2006 vanligen av rättstavade svenska ord. I likhet med 2011-orden återfinns bland 2006-orden också många avledningar och sammansättningar. Dessutom är de utländska söksträngarna oftast på engelska. I likhet med de material som undersöks här förekommer i 2006-materialet så kallade encyklopediska/googleinspirerade uppslagningar (a.a. s. 250, jfr de Schryver et al. 2006:71). Slutligen är flerordsuppslagningar en relativt vanlig typ av icke-träff i 2006-materialet. Framför allt slår användarna upp reflexiva verb och partikelverb men även kollokationer och mer eller mindre fasta konstruktioner samt tillfälliga sådana. Om dessa uppslagningar handlar nästa avsnitt.

#### 6.1.4 Flerordsuppslagningar

De allra flesta uppslagningarna i LSL, 98 % 2008 och 96 % 2011, utgörs av en enkel söksträng, vanligen ett ord (jfr Lorentzen & Theilgaard 2012 där 96 % av uppslagningarna består av ett ord, se avsnitt 4.3.3.2). Flerordssökningarna är således mycket få procentuellt sett. I absoluta tal uppgår de däremot till ett icke oansenligt antal, omkring 2 600 i 2008-materialet och drygt 7 700 i 2011-materialet. En stor majoritet, omkring 80 %, av flerordssökningarna i båda materialen utgörs av två ord. Ytterligare ca 13 % består av tre ord och den resterande andelen flerordsuppslagningar innehåller fyra ord eller fler.

En förklaring till att andelen flerordsuppslagningar inte är högre än ett par procent kan vara att användarna generellt sett är högst medvetna om hur (tryckta) ordböcker är strukturerade på makronivå (se avsnitt 2.3.1), dvs. för att få åtkomst till ordboksartiklarna måste först ett enskilt ord identifieras och eventuellt också dess grundform. Det är t.ex. tänkbart att en användare som slår upp verbet *komma* är ute efter partikelverbet *komma bort*. Det kan vi omöjligt veta. Däremot vet vi att de vanligaste lemmarna i såväl 2008- som 2011-materialet i genomsnitt har ca 3–5 uttryck/lemma (se tabell 6:6 i avsnitt 6.1.1). Det kan därför tänkas att användarnas faktiska efterfrågan på (förklaringar till) fler-

ordsuttryck är större än vad den låga andelen flerordsuppslagningar antyder.

Vid en genomgång av ett urval av 100 flerordssökningar i respektive material kan olika typer av flerordskombinationer identifieras (jfr termen *ordförbindelse*, NLO 1997:206).<sup>22</sup> Det finns för det första ordkombinationer som utgörs av etablerade och mer eller mindre fasta fraser, t.ex. *du är en klippa, heder och samvete, tårta på tårta* och *vända på slantarna*. För det andra återfinns ordkombinationer som är av den mer tillfälliga typen, t.ex. *den lilla, tänker borra* och *jorden saltades*. För det tredje förekommer ofta verb följt av partikel eller dylikt, t.ex. *måna om, ebba ut, känner sig, flaggat för* och *hålla på och*. För det fjärde påträffas flerordssökningar som utgörs av särskrivningar såsom *tusch penna*. De är emellertid mycket få. Slutligen finns exempel på ordkombinationer som utgörs av googleinspirerade sökningar (se ovan). Denna typ av flerordssökning är också ovanlig. Ett exempel från materialet är *Astrid Lindgren*.

Ordkombinationer av de typer som exemplifieras här verkar alltså av användarna uppfattas som mer eller mindre möjliga enheter i ordböcker. Men det är i många fall inte den typen av flerordslemman som faktiskt ingår i LSL. Däremot kan fraserna mycket väl ingå i ordboksartiklarnas språkexempel. En bildlig användning av ordet *klippa* illustreras t.ex. med språkprovet 'målvakten är en klippa i försvaret'.

I LSL3 ingick 642 flerordslemman och inom ramen för uppdateringsprojektet utökades de till 1 118. Utökningen består till största del av reflexiva verb och partikelverb (se Malmgren 2012). Dessa är således sökbara. Revideringen av LSL inbegrep också en uppgradering av idiomen till sublemman (se avsnitt 3.2.3). I egenskap av sublemman kunde dessa flerordsuttryck, i likhet med flerordslemman, vara sökbara. Det skulle exempelvis innebära att en uppslagning på den ovan nämnda frasen *heder och samvete* – tillika sublemma i LSL – som i nuläget leder till en ren icke-träff, istället skulle resultera i en lyckad uppslagning.

Analysen av flerordsuppslagningarna skulle förstås kunna utvidgas och fördjupas.<sup>23</sup> Det ligger emellertid utanför ramen för denna undersökning. Uppslagningarna utgör trots allt en mycket låg andel av det totala antalet uppslagningar i de två materialen.

---

22 En djupgående klassificering av ordförbindelser i ordböcker presenteras av Farø & Lorentzen (2009).

23 Se Prentice (2010:1–145) för en studie av första- och andraspråksinlärares behärskning av kollokationer och idiom.

## 6.2 Loggfiler av typ b: Navigeringsval

Som tidigare konstaterats fanns det vid insamlingsförfarandet inte möjlighet att skilja ut klickningar som görs från LSL från de klickningar som görs från något av de tvåspråkiga lexikonerna. Det är därför näst intill omöjligt att exakt få reda på i hur stor utsträckning länkarna till uppläst uttal, bilder och filmer utnyttjas av användare som söker i just LSL. Totalt har de tre länkarna klickats på vid 1 710 028 tillfällen under den vecka 2011-materialet samlades in. I avsnitt 3.2.4 framkom att antalet lexikonuppslagningar i LSL uppgår till omkring 4 % av det totala antalet uppslagningar i Lexin. Det är rimligt att anta att andelen klickningar på länkarna ligger på samma, eller möjligtvis lägre, nivå. Vi ska nämligen senare se att det är vanligare att länkarna klickas på från de tvåspråkiga lexikonerna än från LSL (se vidare kapitel 8). Antalet klickningar per dag uppvisar samma mönster som uppslagningarna, dvs. att länkarna utnyttjas mest under veckans inledande fyra dagar. Därefter går användningen markant ned under fredag och lördag för att åter öka under söndagen (jfr avsnitt 6.1).

När det gäller klickningar på knappen med ett frågetecken och ändringar i inställningarna kan vi alltså inte heller veta från vilket lexikon aktiviteterna utförs. Antalet klickningar på frågetecknenknappen uppgår som tidigare konstaterats till 7 599. I vilken utsträckning knappen genomsnittligen aktiveras finns det dock inga tillgängliga uppgifter om. Man kan däremot tänka sig att användarna klickar på knappen i större utsträckning under datainsamlingsveckan, eftersom många av dem har reagerat på ordboksartiklarnas avvikande utseende.

Av tabell 6:12 på nästa sida framgår hur många klickningar som görs i inställningarna för de möjliga informationskategorier en ordboksartikel i LSL kan innehålla. Av tabellen framgår att de urkryssade kategorierna åter har kryssats för mellan 8 117 (Uttryck) och 9784 (Böjning) gånger. De övriga kategorierna har endast kryssats för ca 50–200 tillfällen. Det skulle emellertid behövas bidragande upplysningar om hur ofta användarna ändrar i inställningarna i vanliga fall för att kunna uttala sig om mer om uppgifterna i tabellen. Av loggfilerna kan man dock utläsa att vanligen kryssas samtliga fem informationskategorier för, och att det dessutom oftast sker inom loppet av några sekunder (se tabell 8:3, rad 3–7 för ett exempel). Användarna ger på så sätt sken av att genomföra denna aktivitet utan någon större reflektion.

Antalet förkryssningar kan tillsammans med uppgift om det totala antalet ip-adresser (dvs. 25 974) som registrerades under insamlingsveckan ge en grov uppskattning av hur stor andel av informanterna som återställde inställningarna, och därmed ordboksartiklarnas utseende, till standardläget (se avsnitt 5.1.1). Det skulle i så fall betyda att åtminstone omkring en tredjedel av användarna uppmärksammat den tillfälliga förändringen och gått till in-

ställningarna och kryssat för de urkryssade rutorna. Kanske har de också uppmärksammat den instruktion som fanns på webbsidan som informerade om den pågående användarundersökningen av LSL, att inställningarna för tillfället hade ändrats och hur man ändrade tillbaka till det ursprungliga utseendet på ordboksartiklarna.

TABELL 6:12. Antal klickningar som görs i rutan framför respektive informationskategori.

Informationskategori	Antal klick
Avledning	82
Bilder	152
Böjning	9784
Definition	144
Exempel	9269
Förkortningar	121
Konstruktioner	8434
Motsatser	120
Ordklass	113
Sammansättningar	8972
Uppslagsord	60
Uttal	70
Uttryck	8117

Syftet med att manipulera inställningarna var att få mer detaljerad kunskap om dels användarnas sökbeteende, dels vilka specifika informationskategorier användarna är intresserade av. Av resultatet att döma kan man dra vissa slutsatser om användarnas sökbeteende, t.ex. att de som aktivt ändrar i inställningarna väljer att kryssa för samtliga fem avaktiverade informationskategorier, kanske en naturlig önskan att få så mycket information som möjligt. Några intressanta slutledningar om användarnas eventuella specifika behov av enskilda kategorier kan emellertid inte göras. Böjning, Exempel och Sammansättningar har visserligen kryssats för i större utsträckning än Konstruktioner och Uttryck, men det är vanskligt att göra några tolkningar av det (jfr avsnitt 7.2.3 om informanterns förmodade svårighet att skilja på t.ex. Konstruktion och Exempel).

Som tidigare konstaterats spelar navigeringsvalen en mer betydelsefull roll i delstudie 3 som redovisas i kapitel 8.

## 6.3 Sammanfattning

Uppslagningarna i LSL3 och LSL4 skiljer sig åt i olika avseenden men uppvisar också likheter. I båda versionerna av LSL är t.ex. söktrycket högst måndag till torsdag. Under fredag och lördag är uppslagningarna markant färre för att sedan åter öka under söndagen. Man kan tänka sig att detta användningsmönster speglar en genomsnittlig arbets- eller studievecka. Den absolut vanligaste träffen i båda materialen är substantiv i obestämd form singular. Därefter kommer verb i infinitivform och adjektiv i grundform. Flerordssökningar är procentuellt sett få i båda materialen.

Skillnaderna mellan de två materialen är dock fler än likheterna. LSL4 uppvisar för det första en bättre träffstatistik än LSL3. Det beror framför allt på att fler böjningsformer har blivit sökbara i den uppdaterade versionen av LSL. Sett till de mest frekventa söksträngarna är de två materialen påfallande olika. För det första verkar verben spela en viktigare roll för användarna 2011 än 2008. I 2011-materialet är verben nämligen fler till antalet och de har dessutom högre frekvenser än i 2008-materialet. I 2011-materialet utgörs söksträngarna dessutom oftare av böjda verbformer, i synnerhet presens.

De uppslagningar som genererar en träff i LSL uppvisar betydelsemässigt uppenbara skillnader. För det första utgörs många av 2008-orden av abstrakta substantiv, t.ex. *etik*, och av mindre centrala verb och adjektiv såsom och *integrera* respektive *relevant*. Bland 2011-orden finner vi däremot oftare konkreta och mycket centrala substantiv som exempelvis *bil*, och mer centrala verb och adjektiv som t.ex. *går* respektive *stor*. En jämförelse mellan de mest frekventa söksträngarna i de båda materialen med deras förekomst i dagspress styrker slutsatsen att uppslagningarna i 2011-materialet i högre utsträckning består av vanligare och mer centrala svenska ord än uppslagningarna i 2008-materialet.

LSL har en rättstavningsfunktion som är baserad på editeringsavstånd och inte på uttalsmässig stavning. En uppslagning som inte ger träff kan leda till en automatkorrigering eller en korrigering med förslag. Automatkorrigeringen tycks fungera optimalt när det handlar om smärre felstavningar, dock inte initialt i söksträngen, förutsatt att motsvarande rättstavade ord finns i ordboken. För de mest frekventa uppslagningarna som leder till korrigering med ordförslag varierar antalet förslag mellan 2 och 72. Vid en närmare granskning av vilka konkreta förslag som presenteras, och framför allt i vilken ordning de kommer, visar det sig att editeringsavstånd inte är en tillräckligt vass urvalsmetod att basera ordförslagen på. De föreslagna lemmarna är t.ex. många gånger utifrån ett semantiskt perspektiv väsensskilda från sökordets betydelse.

Rättstavningshjälpens funktionalitet kan således diskuteras. Fler sökbara böjningsformer har trots allt lett till en förbättrad träffstatistik i LSL4. Färre söksträngar genererar rena icke-träffar och fler uppslagningar resulterar istället i

någon typ av korrigerig, oftast korrigerig med förslag. Förslagen verkar dock alltför ofta vara till mer skada än nytta. Från ett användarperspektiv skulle det förmodligen vara mycket värdefullt att få reda på att dessa ord utgör fullgoda svenska lexikala enheter, även om de inte återfinns i LSL.

Sammanfattningsvis kan sägas att söksträngarna i 2011-materialet i större utsträckning består av centrala svenska ord. Där finner vi mer basala och vanliga ord i betydelsen ord som man lär sig tidigt i sin svenskinläring och som är mycket frekventa i svenskan. Det rör sig om mycket vanliga verb med flera grundbetydelser och vanligen många överförda användningar, inte sällan förekommande i olika typer av fasta fraser. Skillnaderna mellan de två materialen kan förmodligen kopplas till avlägsnandet av det svensk-engelska lexikonet.

## 7. Delstudie II: Användarna av LSL

I detta andra resultatkapitel redogörs i första hand för enkätundersökningen från 2011, här Studie<sup>11</sup>. Den genomfördes i syfte att kartlägga användarna av LSL4. Analysen omfattar 802 frågeformulär. För ungefär hälften av frågorna finns som redan framgått ett jämförelsematerial i form av en enkätundersökning från 2007 av LSL3, här Studie<sup>07</sup>. Studie<sup>07</sup> omfattar 356 enkäter. En närmare beskrivning av enkätmaterialen ges i avsnitt 5.5.

I det följande redovisas enkätmaterialen i fyra avsnitt. Först presenteras bakgrundsinformation om informanterna. Sedan redogörs för deras uppfattning om sin användning av LSL. Därefter ges en närmare beskrivning av de informanter som har annat modersmål än svenska. Det sista avsnittet behandlar innehållet i det som utgörs av enkätens avslutande öppna kommentarsfält. Jämförelsen mellan de två enkätmaterialen görs löpande, i de fall det är aktuellt.

### 7.1 Bakgrundsinformation

Informanternas bakgrundsdata står som sagt i fokus i detta första avsnitt. Det inkluderar variablerna ålder, kön, sysselsättning och utbildning i Sverige samt utbildning i eventuellt annat hemland.

### 7.1.1 Födelseår

Informanterna i Studie<sup>11</sup> uppvisar en stor spridning vad gäller födelseår. Den äldsta informanten är född 1940 och fyller alltså 71 år vid tidpunkten för materialinsamlingen, dvs. år 2011. Den yngsta informanten är född år 2000 och fyller således 11 år under insamlingsåret. Mellan åren 1940 och 2000, dvs. det tidigaste och senaste födelseåret, är samtliga årtal representerade förutom tre (1942, 1943 och 1947). Medelåldern är 32 år och medianåldern är 31 år.

I tabell 7:1 redovisas information om antalet informanter per åldersgrupp i Studie<sup>11</sup> samt motsvarande uppgifter från Studie<sup>07</sup>.

TABELL 7:1. Informanter per åldersgrupp i Studie<sup>11</sup> och Studie<sup>07</sup>, i absoluta och relativa tal.

Åldersgrupp	Studie <sup>11</sup>		Studie <sup>07</sup>	
	Absolut frekvens	Relativ frekvens (%)	Absolut frekvens	Relativ frekvens (%)
-14	13	1,6	39	11
15-19	105	13,1	153	43
20-29	231	28,8	86	24,2
30-39	219	27,3	36	10,1
40-49	111	13,8	31	8,7
50-59	67	8,4	8	2,2
60-69/60- <sup>24</sup>	15	1,9	3	0,8
70-	3	0,4	-	-
Obesvarad	38	4,7	0	0
<b>Totalt</b>	<b>802</b>	<b>100</b>	<b>356</b>	<b>100</b>

Av tabell 7:1 framgår att åldersfördelningen i stort sett följer samma mönster i de två studierna på så sätt att det lägsta antalet informanter återfinns i den yngsta och den äldsta åldersgruppen. Över hälften av informanterna i Studie<sup>11</sup>, ca 56 %, är mellan 20 och 39 år. I Studie<sup>07</sup> är istället majoriteten informanter, drygt 67 %, mellan 15 och 29 år. Informanterna i Studie<sup>11</sup> är alltså genomsnittligt äldre än de i Studie<sup>07</sup>. Alternativet 70- som återges i tabellen fanns inte med som alternativ i Studie<sup>07</sup>, vilket i tabellen indikeras av minustecknen (-).

Av uppgifterna i tabellen kan man också sluta sig till att i Studie<sup>11</sup> är andelen informanter i grundskole- eller gymnasieåldern, dvs. informanter upp till 19 år,

---

24 I enkäten för Studie<sup>07</sup> efterfrågades ålder, inte födelseår, och informanterna ombads kryssa för ett av sju givna åldersintervall. Alternativet "60 eller mer" var det högsta åldersintervallet.



mycket lägre än i Studie<sup>07</sup>, nämligen 15 % mot 54 %. Att färre skolungdomar använder LSL 2011 är med största sannolikhet en konsekvens av att det svensk-engelska lexikonet togs bort i och med uppdateringen av Lexin (se vidare avsnitt 3.2.4.3). I oktober 2007 stod t.ex. uppslagningarna i det svensk-engelska lexikonet för 70 % av det totala antalet sökningar i Lexin-lexikonerna. Man vet också att det svensk-engelska lexikonet var mycket populärt och frekvent använt bland ungdomar i skolåldern, vilket sannolikt genererade användning av LSL av andra användare än den primära målgruppen (f.d. projektledaren för Lexin Christian Mattsson, muntlig kommunikation 2009-05-06).

Huruvida födelseårens spridning i materialet speglar hur det verkligen förhåller sig med Lexin-användarnas ålder är svårt att uttala sig om. Ålder kan vara en faktor som spelar in när det gäller benägenheten att delta i en (webb-)enkätundersökning (Denscombe 2016:55). Det man kan utläsa ur de redovisade uppgifterna är att LSL används av alla åldersgrupper även om de yngsta och de äldsta informanterna är klart färre. Lexin når således ut till användare i alla åldrar.

Ålder är en variabel som användarforskare traditionellt sett inte verkar ha lagt så stor vikt vid. Däremot beskrivs på vilken utbildningsnivå informanterna befinner sig och man kan då i viss mån dra slutsatser om hur gamla de är. Historiskt sett har enkätstudier i stor utsträckning varit inriktade på att undersöka ordboksanvändning bland studenter (se avsnitt 4.3.1). Det är förmodligen anledningen till att ålder sällan har redovisats explicit. När det gäller webb-enkätstudier är situationen en annan. Där vänder man sig i regel till en mycket mer heterogen användarskara och informanterna ombeds ofta uppge ålder. I t.ex. Lorentzen & Theilgaards (2012) enkätstudie är medelåldern drygt 37 år. Informanterna i Müller-Spitzers (2013) enkätstudie uppvisar en relativt jämn åldersfördelning förutom när det gäller de informanter som är yngre än 21 år, vilka är få (se avsnitt 4.3.3.1). I en användarundersökning av appversionen av den förra upplagan av *Svenska Akademiens ordlista*, SAOL13, är informanternas medelålder 43 år och medianåldern är 41 år (Holmer et al. 2015:362).

### 7.1.2 Kön

Könsfördelningen är i det närmaste identisk i de två enkätmaterialen. Något fler kvinnor än män, 54 % mot 46 %, ingår i Studie<sup>11</sup> och i Studie<sup>07</sup> är fördelningen nästan densamma, 53 % kvinnor och 47 % män. Med avseende på ålder är det framför allt i de yngre åldersgrupperna som kvinnorna är fler. Det är bara i åldersgruppen 50–59 som männen är markant fler.

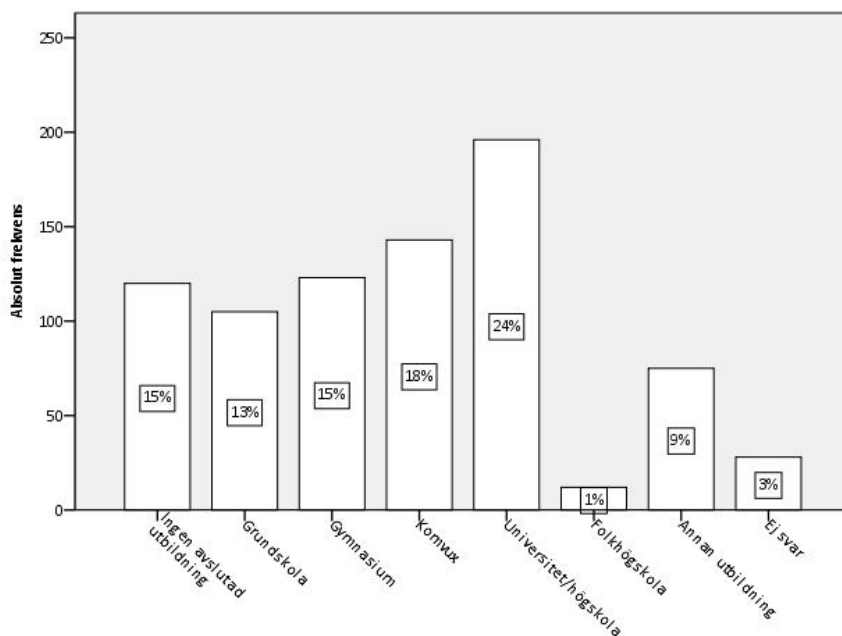
Inte heller kön är en variabel som användarforskare verkar ha lagt så stor vikt vid i traditionella enkätstudier av tryckta ordböcker. Det är däremot nå-

gonting som uppmärksammas i många webbenkätstudier (se avsnitt 4.3.3.1). Lorentzen & Theilgaard (2012) uppmärksammar t.ex. att könsfördelningen i deras enkätstudie är helt jämn. I undersökningen av norska Lexin deltar något fler män än kvinnor. I Müller-Spitzer (2013) har kvinnor besvarat enkäten i mycket större utsträckning än män. I Holmer et al. (2015:362) visar det sig också att huvuddelen av informanterna, 60 %, är kvinnor. De båda senare studierna pekar alltså på samma tendens som min studie. Det kan tilläggas att studier har visat att kvinnor är mer benägna att svara på enkäter än män (Holtti 2003:510). Det har också visat sig att kvinnor är mer benägna än män att svara på enkäter via mobiltelefon, medan män istället i högre grad än kvinnor svarar på enkäter via datorn (Userneeds 2016).

## 7.1.3 Högsta avslutade utbildning

### 7.1.3.1 Utbildning i Sverige

I figur 7:1 presenteras informanternas svar på frågan om högsta avslutade utbildning i Sverige i Studie<sup>11</sup>.



FIGUR 7:1. Studie<sup>11</sup>: Högsta avslutade utbildning i Sverige.

I figur 7:1 kan man för det första notera att 15 %, ca 110-tal informanter, inte har någon avslutad utbildning alls. Man kan tänka sig att många av dessa informanter är relativt nyanlända i Sverige och att de därför inte har haft möjlighet att skaffa sig en svensk utbildning än. Vidare uppger 13 %, drygt 100 informanter, att de har grundskoleutbildning och ca 15 %, uppemot 125 informanter, att de har uppnått gymnasiekompetens. Vidare har närmare 150 informanter, 18 %, en Komvux-utbildning (eller motsvarande). Komvux erbjuder framförallt möjlighet att slutföra grundskole- och gymnasiestudier samt sfi-utbildning. En del av dessa informanter skulle sannolikt ha återfunnits bland dem som anger grundskole- eller gymnasienivå om inte Komvux-alternativet funnits.

Det största antalet informanter finner vi i stapeln för universitetsutbildning. Där återfinns 200 informanter vilket motsvarar 24 %. Den minsta andelen informanter, 1 %, har kryssat för att de har en folkhögskoleexamen. Slutligen har 9 %, drygt 70 informanter, angett att de har en annan utbildning. Ett trettioital informanter, 3 %, har lämnat frågan obesvarad.

Sammanfattningsvis kan konstateras att informanterna i Studie<sup>11</sup> har en tämligen hög utbildningsnivå. Vilka typer av utbildningar det rör sig om framgår däremot inte. Det kan man dock i vissa fall få uppgifter om i informanternas eventuella kommentar till frågan om utbildning. De informanter som kryssat för ”Annan utbildning” ombads nämligen att precisera sitt svar. Det har 66 informanter gjort. Det är emellertid inte bara de informanter som kryssar för ”Annan utbildning” som väljer att kommentera sitt svar. Det gör även 88 informanter som valt något av de andra svarsalternativen, eller inget alls. Behovet av att kommentera sitt svar tycks alltså vara oberoende av vilket svarsalternativ som kryssas för. Av de totalt 154 kommentarerna till utbildning i Sverige är det vanligast att informanterna antingen preciserar sin utbildning eller utbildningsnivå eller att de uppger att de har en utbildning på gymnasie- eller högskolenivå, fast utomlands. Uppemot 100 kommentarer innehåller, i ungefär lika delar, något av dessa svar. I det följande ges några exempel på dessa.<sup>25</sup>

- (7:1) Statsvetenskap
- (7:2) Master
- (7:3) Har börjat sjuan.
- (7:4) Går på gymnasiet nu
- (7:5) Bussutbildning
- (7:6) Tandtekniker
- (7:7) Jag är läkare
- (7:8) Svenska som andraspråk SAS B

---

25 Som redan nämnts återges samtliga exempel från materialet i oredigerat skick.

- (7:9) Jag är utbildad jurist i Tyskland
- (7:10) Universitet i Ukraina
- (7:11) Jag studerade musik i Iran
- (7:12) Universitet i andra land

Behovet av att kommentera svaren kan bero på att informanterna vill lyfta fram ett de faktiskt har en gedigen utbildning eller ett kvalificerat yrke. Det är först och främst i rollen som svenskinlärare de är noviser.

Vidare har ett trettiotal informanter uppgett utbildningen ”sf” vilket förstås också är en typ av precisering. Som tidigare antytts återfinns flera av dessa informanter bland dem som kryssat för att de inte har någon avslutad utbildning.

Ett dussin informanter anger just annan typ av utbildning, exempelvis folkuniversitet, kvalificerad yrkesutbildning (KY) eller yrkeshögskola (YH) samt internutbildning. De resterande kommentarerna innehåller t.ex. information om att man inte bor i Sverige. Flera av de här informanterna studerar sannolikt svenska som främmandespråk.

Det finns förstås en risk att man kryssar för en utbildningsnivå som faktiskt inte har avslutats, vilket bevisligen är fallet för åtminstone en informant. Informanten bakom kommentaren ”Har börjat sjuan” har kryssat för svarsalternativet ”Grundskola”. Såsom frågan formulerades skulle hen egentligen ha kryssat för ”Ingen avslutad utbildning”. Informanten bakom kommentaren ”Går på gymnasiet just nu” kryssade också för alternativet ”Grundskola” men har till skillnad från den förre informanten svarat på det som faktiskt efterfrågas.

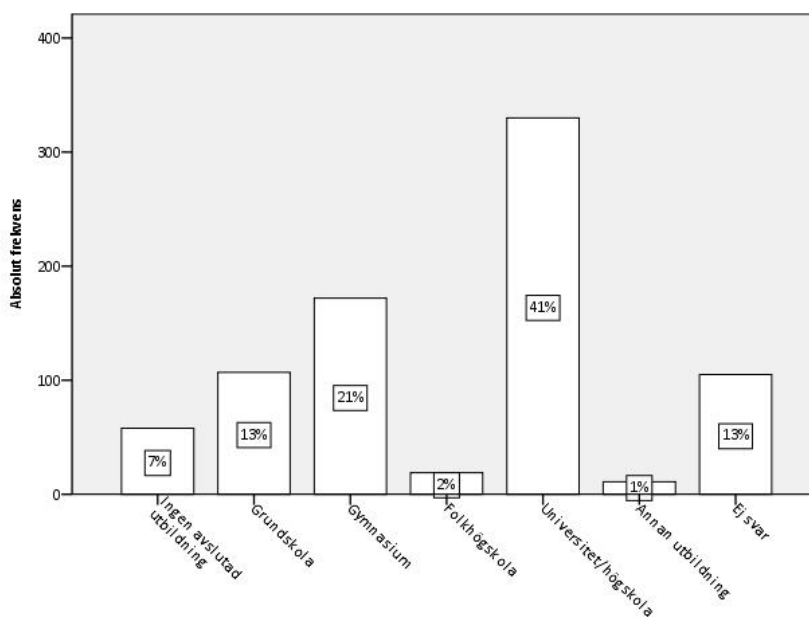
Informanterna i Studie<sup>11</sup> har vidare högre utbildning än informanterna i Studie<sup>07</sup>. Det stämmer väl överens med vad vi vet om informanternas ålder i de två studierna (se tabell 7:1), nämligen att Studie<sup>11</sup> utgörs av en äldre skara informanter. I Studie<sup>11</sup> är 24 % universitetsutbildade jämfört med 18 % i Studie<sup>07</sup>. I Studie<sup>07</sup> uppger dessutom nästan 80 % att de antingen har en grundskole- eller en gymnasieutbildning. Motsvarande andel i Studie<sup>11</sup> är avsevärt lägre, 28 %.

Tre svarsalternativ, ”Ingen avslutad utbildning”, ”Komvux” och ”Annan” ingick endast i frågeformuläret för Studie<sup>11</sup>. Det är förstås omöjligt att veta hur många av informanterna i Studie<sup>07</sup> som skulle ha kryssat för något av dessa alternativ, om det hade varit möjligt. I och med att så många som 42 % av informanterna i Studie<sup>11</sup> väljer något av dem skulle troligen också ett antal av deltagarna i Studie<sup>07</sup> ha gjort det. Å andra sidan kan man tänka sig att de informanter som uppnått grundskoleexamen på Komvux har kryssat för svarsalternativet ”Grundskola”.

När det gäller utbildning i Sverige visar det sig alltså att informanterna i Studie<sup>11</sup> har högre utbildning än de i Studie<sup>07</sup>. I följande avsnitt redogörs för informanternas utbildning i hemlandet, om annat än Sverige.

### 7.1.3.2 Utbildning i hemlandet

Totalt 697 informanter i Studie<sup>11</sup> har besvarat frågan om utbildning i hemlandet, vilket motsvarar 87 %. Detta talar starkt för att en stor majoritet av informanterna tillhör LSL:s målgrupp. Av figur 7:2 framgår hur svaren fördelar sig på de olika alternativen.



FIGUR 7:2. Studie<sup>11</sup>: Högsta avslutade utbildning i hemlandet.

Av figur 7:2 kan man utläsa att 7 %, omkring 60 informanter, inte har någon avslutad utbildning alls utomlands. Vidare har 13 %, drygt 100 informanter, en grundskoleutbildning från sitt hemland och 21 %, över 170 informanter, har en utländsk gymnasieutbildning. Ett tjugotal informanter, 2 %, anger att de utbildat sig på folkhögskolenivå (eller motsvarande). Även här finner vi att den vanligaste utbildningsnivån bland informanterna är universitet. Över 300 informanter, 41 %, har kryssat för detta alternativ. Endast ett tiotal informanter har kryssat för ”Annan utbildning”. Det dryga hundratalet informanter som avstår från att svara på frågan om utbildning utomlands utgörs rimligtvis till stor del av informanter med svenska som modersmål (se avsnitt 7.3.3 som redovisar informanternas svar på frågan om modersmål).

I Studie<sup>07</sup> svarade 130 informanter på frågan om utbildning i hemlandet vilket motsvarar 37 % av det totala antalet informanter, att jämföra med 87 % i Studie<sup>11</sup>. Antalet uteblivna svar ger som sagt en indikation på hur många av in-

formanterna som är presumtiva svenska modersmålstalare. I Studie<sup>07</sup> är de här informanterna alltså många fler, drygt 60 %, än i Studie<sup>11</sup> där de endast uppgår till 13 % av informanterna.

Den största andelen informanter i Studie<sup>07</sup>, 42 %, har en grundskoleutbildning från sitt hemland. Motsvarande andel i Studie<sup>11</sup> var 13 %. Ungefär lika stora andelar av informanterna i de två materialen har angett att de har gymnasiekompetens. Däremot är skillnaden stor när de gäller universitetsutbildning, 25 % i Studie<sup>07</sup> jämfört med 41 % i Studie<sup>11</sup>.

Svaren på frågorna om utbildning i Sverige och hemland visar alltså att informanterna i Studie<sup>11</sup> har hunnit längre i sin utbildning i hemlandet innan de kommit till Sverige än informanterna i Studie<sup>07</sup>.

#### 7.1.4 Sysselsättning

Sammanlagt 789 informanter i Studie<sup>11</sup> har svarat på frågan om sysselsättning. Deras svar fördelar sig såsom presenteras i tabell 7:2.

TABELL 7:2. *Informanternas sysselsättning i Studie<sup>11</sup>, i frekvens och procent.*

Sysselsättning	Absolut frekvens	Relativ frekvens (%)
1. Jag studerar	457	58
2. Jag arbetar	240	30
3. Jag söker arbete	70	9
4. Annan	22	3
<b>Totalt</b>	<b>789</b>	<b>100</b>

Uppemot 60 % av informanterna uppger att de studerar, 30 % att de arbetar och ca 9 % att de söker arbete. Tre procent kryssar för alternativet för annan sysselsättning. Totalt 67 informanter har kommenterat sitt svar på något sätt. Sett till det totala antalet informanter har en låg andel kommenterat sitt svar, cirka 8 %. Här är alltså behovet av att vidareutveckla sitt svar litet (jfr avsnitt 7.1.3.1 där detta behov var mera påtagligt). De informanter som valt att lämna en kommentar vill huvudsakligen framföra att de både arbetar och studerar eller specificera vilken sysselsättning de har. I tabell 7:3 har kommentarerna kategoriserats efter sitt huvudsakliga innehåll och presenteras där tillsammans med uppgift om frekvens samt exempel.

TABELL 7:3. Typ av kommentar, frekvens och exempel, Studie<sup>11</sup>.

Typ av kommentar	Antal	Exempel
Studerar och arbetar	22	<i>Jobbar också</i>
		<i>Arbetar och studerar</i>
		<i>Jag studerar o försöker att hitta jobb</i>
Precisering av svarsalt. 1–3, dvs. "Jag studerar", "Jag söker arbete" och "Jag arbetar"	16	<i>Språk</i>
		<i>Bussförare</i>
		<i>Tolkning</i>
		<i>Försvarsmakten</i>
		<i>Frilans översättare</i>
		<i>IT/affarsanalys</i>
		<i>Komvux</i>
"Annan", dvs. annan sysselsättning än något av svarsalternativ 1–3.	15	<i>Gör ingenting</i>
		<i>Husmor</i>
		<i>Föräldrarledig</i>
		<i>Pensionär</i>
		<i>Sjukpensionär</i>
		<i>Sitter hemma och softar</i>
		<i>Kroniskt sjuk, inte arbetsför</i>
Övrig kommentar	13	<i>hemmafru</i>
		<i>Jag vill jobba i framtiden</i>
Nonsens eller ej uttolkbart	1	<i>Rätt uttala</i>

I tabell 7:4 jämförs svar på frågan om sysselsättning från Studie<sup>11</sup> med dem i Studie<sup>07</sup>. I Studie<sup>07</sup> formulerades svarsalternativen annorlunda och mer formellt: "Studerande", "Yrkesarbetande" respektive "Arbetsökande". De här orden är etablerade i samband med olika typer av blanketter och frågeformulär och känns säkert igen och begrips av både modersmålstalande svenskar och svenskinlärare. Trots detta ändrades svarsalternativens utformning i enkäten från Studie<sup>11</sup> till enklare svenska – svarsalternativet "Yrkesarbetande" ändrades t.ex. till "Jag arbetar" (se avsnitt 5.5.1).

TABELL 7:4. Sysselsättning Studie<sup>11</sup> och Studie<sup>07</sup>, i procent.

Sysselsättning	2011 (%)	2007 (%)
Studerar	58	75
Arbetar	30	17
Söker arbete	9	3
Annan	3	5
<b>Totalt</b>	<b>100</b>	<b>100</b>

Av tabell 7:4 framgår för det första att andelen studerande är klart högre i Studie<sup>07</sup> där 75 % av informanterna uppger att de studerar. Andelen som arbetar är däremot mycket lägre, 17 % mot 30 % i Studie<sup>11</sup>. Majoriteten av informanterna i Studie<sup>11</sup> är ändå studerande. I övrigt är det fler informanter som är arbetssökande i det senare materialet.

Siffrorna i tabell 7:4 är förväntade med tanke på att det är fler informanter i Studie<sup>11</sup> än i Studie<sup>07</sup> som uppger att de har en universitetsutbildning (se avsnitt 7.1.3). I det senare materialet är det alltså fler informanter som är färdiga med sin utbildning och som söker arbete eller som har börjat arbeta.

När det gäller sysselsättning är kommentarerna få även i Studie<sup>07</sup>, och de består av samma typer som i Studie<sup>11</sup>.

### 7.1.5 Sammanfattning

I detta avsnitt har bakgrundsuppgifter om informanterna i Studie<sup>11</sup> presenterats, med jämförande uppgifter från Studie<sup>07</sup>. Födelseåren visar på en stor spridning i båda materialen och LSL verkar alltså nå ut till användare i alla åldrar. Informanterna i Studie<sup>11</sup> är däremot genomsnittligt äldre än informanterna i Studie<sup>07</sup>. Könsfördelningen är jämn i båda materialen.

Informanterna täcker dessutom in alla utbildningsnivåer i Sverige. I Studie<sup>11</sup> har emellertid en större andel av informanterna universitetsutbildning. I Studie<sup>07</sup> är istället andelen med enbart grundskoleutbildning störst. I Studie<sup>11</sup> uppger vidare en mycket stor andel av informanterna att de har en utbildning i sitt hemland (om annat än Sverige) och hela 41 % uppger att de har en universitetsutbildning. I Studie<sup>07</sup> är denna andel mycket lägre. Informanterna i 2011-materialet har alltså hunnit längre i sin utbildning i hemlandet än informanterna i 2007-materialet.

Slutligen är majoriteten av informanterna i Studie<sup>11</sup> studerande, men de studerande är dock relativt fler i Studie<sup>07</sup>. Däremot är andelen som har eller söker arbete större i Studie<sup>11</sup> än i Studie<sup>07</sup>. Fler informanter är alltså färdiga med sin utbildning, söker arbete eller har börjat arbeta.

Det står också klart att man med en webbenkät kan nå ut till en mer heterogen grupp användare än i en traditionell enkätstudie, där man ofta vänder sig direkt till en specifik typ av användare, i regel studenter (se avsnitt 4.3.1 för exempel på traditionella enkätstudier).

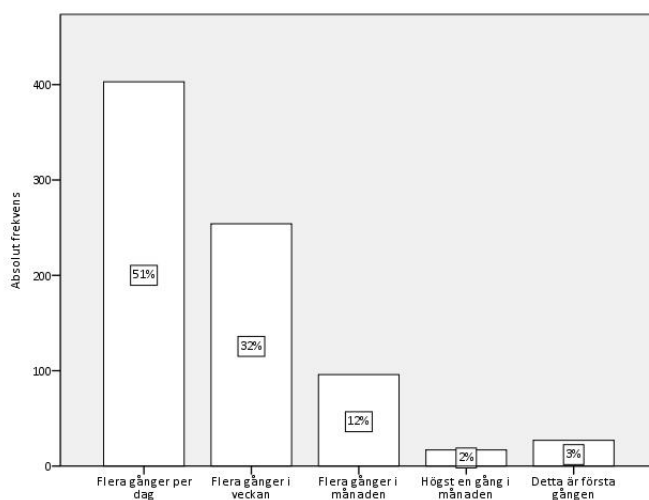


## 7.2 Informanternas utsagor om användningen av LSL

Efter de inledande frågorna om informanternas bakgrund följer frågor om deras användning av LSL. Syftet är att få svar på hur ofta LSL används, av vilken anledning ordboken används och vilka lexikografiska behov informanterna anser sig ha.

### 7.2.1 Sökfrequens

På frågan hur ofta LSL används svarar deltagarna i Studie<sup>11</sup> såsom presenteras i figur 7:3.



FIGUR 7:3. Studie<sup>11</sup>: Informanternas uppgifter om hur ofta LSL används.<sup>26</sup>

Av figur 7:3 framgår att hälften av informanterna uppger att de söker i LSL4 flera gånger varje dag och att ytterligare drygt 30 % använder lexikonet flera gånger i veckan. Över 80 % av informanterna är alltså relativt sett flitiga användare av LSL4. Som jämförelse kan noteras att ca 75 % av informanterna i Lorentzen & Theilgaards (2012:658) användarstudie av DDO uppger att de använder den undersökta ordboken dagligen eller flera gånger i veckan (se avsnitt 4.3.3.1). En mycket liten andel av informanterna uppger antingen att de söker i LSL någon gång i månaden eller att det är första gången. Det är ett väntat resultat med tanke på att många av de användare som sällan använder

26 Det hade varit bättre att ha alternativet ”Minst en gång per dag” än ”Flera gånger per dag”.

LSL troligen inte känner sig uppmanade att svara på en enkät om användning av just den ordboken.

I Studie<sup>07</sup> anger 44 % att de använder LSL3 flera gånger i veckan. Trettio procent använder lexikonet flera gånger i månaden och 13 % högst en gång i månaden. Resterande 13 % använder LSL för första gången. Enkäten i Studie<sup>07</sup> innehöll inte svarsalternativet ”Flera gånger per dag”. En jämförelse mellan de informanter i Studie<sup>07</sup> som använder LSL flera gånger per vecka med dem i Studie<sup>11</sup> som använder LSL minst flera gånger i veckan visar emellertid att informanterna i Studie<sup>11</sup> är frekventare ordboksanvändare. Andelen förstagångs-användare är dock klart högre i Studie<sup>07</sup> än i Studie<sup>11</sup>.

Noteras bör att det är informanternas egen uppskattning om hur ofta de använder ordboken som redovisas. Vi vet att data av detta slag kan brista i reliabilitet (se avsnitt 4.2.1). Reliabiliteten kan t.ex. äventyras av att människor i allmänhet kan ha svårt att minnas hur saker och ting förhåller sig. Man torde dock inte ha svårt att veta huruvida man använder en ordbok varje dag eller inte. Reliabiliteten kan också äventyras av att informanter vill imponera på dem som utför undersökningen och därför utmålar sig som mer kompetenta och ”duktigare” användare än vad de verkligen är. Den risken är kanske inte så stor i just det här fallet, inte minst för att informanterna är anonyma. Som vi senare ska se kan svarens sanningshalt i viss mån verifieras när enskilda informanternas aktiviteter undersöks närmare (se kapitel 8).

## 7.2.2 Användningssituation

Ett relevant och ofta förekommande spörsmål i frågeformulär om ordboksanvändning är i vilket syfte ordboken används. Det framkom inte minst i avsnittet om tidigare forskning (se avsnitt 4.3). I samband med frågan om anledningen till att man använder LSL uppmanades informanterna i Studie<sup>11</sup> att kryssa för max två av fem användningssituationer (se fråga 6 i bilaga 2). Det visar sig dock att en fjärdedel av informanterna har kryssat för fler än två svarsalternativ.

I tabell 7:5 redogörs för hur många kryss respektive svarsalternativ erhöles. I alternativ 1 och 4 beskrivs receptionssituationer och i alternativ 2 beskrivs en produktionssituation. Att reception har delats upp i kategorierna skrift och tal motiveras av att det annars saknas ett svarsalternativ som handlar om en talspråklig situation. Olyckligtvis fanns det emellertid inget alternativ som rör en talspråklig produktionssituation. Möjligen lindras misstaget av att LSL4 i första hand är en receptionsordbok och att det är mer relevant att få reda på användningen i de två receptionssituationerna än i motsvarande produktionssituationer.

TABELL 7:5. *Studie<sup>11</sup>: Informanternas uppgifter om användningssituation.*

Använder du mest <i>Lexins svenska lexikon</i> för att:	Absolut frekvens
1. ... förstå ett ord i en text som du läser?	513
2. ... få hjälp med ett ord när du skriver en text?	468
3. ... få hjälp med ett ord när du ska översätta en text?	402
4. ... få hjälp med ett ord som du hört?	249
5. Något annat	44
<b>Totalt</b>	<b>1 676</b>

Av tabell 7:5 framgår att LSL4 i synnerhet används i situationer av reception (alternativ 1+4). På andra plats hamnar situationer av produktion. Man kan således göra tolkningen att LSL används för de syften som avses – i första hand som receptionsordbok och i andra hand som produktionsordbok (se vidare avsnitt 3.2.3 om LSL4).

Hälften av informanterna, 402 av totalt 802, uppger att de använder LSL4 i översättningssituationer. Att ägna sig åt aktiviteten att översätta en text inbegriper både reception och produktion (jfr Hausmann 1977). Man kan hävda att när två eller fler språk är inblandade så handlar det alltid om översättning. Ur det perspektivet skulle det inte finnas ett behov av svarsalternativ 3. Man kan likväl tänka sig att översättning tolkas som en mycket konkret aktivitet där en specifik text ska översättas från ett språk till ett annat. Det sker normalt dels i undervisningssituationer, dels i olika yrkesmässiga sammanhang. I sådana fall förtjänar svarsalternativet en mer självklar plats. Inom såväl allmän tvåspråkig lexikografi som specifikt i funktionsläran skiljs också översättning ut från reception och produktion (se vidare avsnitt 2.3.2 och 2.3.3).

Vidare kryssar 44 informanter för alternativet "Något annat" och 20 av dem specificerar sitt svar i kommentarsfältet. Nedan följer några exempel:

- (7:13) få hjälp med att uttala ord
- (7:14) att veta olika form av ett ord
- (7:15) oftast kontrollerar jag stavningen
- (7:16) se hur ett ord används
- (7:17) hitta synonymer eller fler betydelser

I de flesta fall har informanterna på olika sätt preciserat vilken informationskategori de huvudsakligen vill veta mer om. Anmärkningsvärt nog beskriver de ofta situationer av produktion, bl.a. i ovan nämnda exempel. Därutöver har ett fyrtiotal informanter som inte kryssat för "Något annat" skrivit en kommentar. De flesta av dessa kommentarer handlar mer generellt om i vilken situation eller av vilken anledning LSL används. Några av dem citeras nedan.

- (7:18) Jag har dyslexi! Är mycket hjälpt av Lexin! Mina tonåringar-ungdommar använder också er! Tack! Jag vill översätta Engelska också, men det går inte! allt är bra!
- (7:19) som en elev som behöver förstå svenska orden på sina lektioner.
- (7:20) när jag skriver en text också, när jag smsas med någon
- (7:21) få hjälp med ett ord som jag vill använda dagligt
- (7:22) öva och kommunicera
- (7:23) öka mitt ordförråd
- (7:24) när jag inte vet vad ett ord betyder på mitt språk
- (7:25) för att leka med språket

Några informanter uppger att de använder LSL i alla de situationer som svarsalternativen beskriver. Kommentarererna kan också röra böjning, stavning och översättning. I det allra flesta fall beskrivs kommunikativa användarsituationer som alltså innefattar reception, produktion och översättning. En kognitiv användarsituation uppstår istället när man är i behov av och har en önskan om (mer) kunskap om något. Jag har inte funnit några kommentarer som tydligt ger uttryck för sådana behov.

I Studie<sup>07</sup> lyder motsvarande fråga om användningssituation som följer: ”I vilka situationer brukar du använda Svenska ord (dvs. LSL3)?”. Svartalternativen formulerades i det närmaste likadant som i Studie<sup>11</sup>, med ett undantag. Här var det istället svarsalternativet som rör produktion som delades upp i alternativen skrift och tal. Totalt uppger nästan hälften av informanterna att de använder LSL för att förstå ett ord i en text som de läser, dvs. för reception. En stor andel svarar också att de använder LSL för att få hjälp med ett ord när de skriver eller för att få veta hur ett ord uttalas, dvs. för situationer av produktion (i samband med skrift eller tal). En liten andel kryssar för ”Annan situation”. Några av informanternas kommentarer, som totalt uppgår till 27 stycken, återges nedan:

- (7:26) För att hitta synonymer när jag skriver på skoluppgifter.
- (7:27) för att hitta ordets EXAKTA betydelse
- (7:28) För att hitta synonymer i svenska och engelska, och ibland översättningar
- (7:29) jag ska lära mig svenska jag använder på alla situationer
- (7:30) Om man chattar med någon från ett annat land är det bra att snabbt och enkelt kunna slå upp ord man inte kan
- (7:31) Använder det i min undervisning
- (7:32) Förberedelser inför föreläsningar etc

I kommentarererna beskrivs eller preciseras såväl situationer av reception (t.ex. 7:30) som produktion (t.ex. 7:26). Av exemplet (7:28) antyds också att det svensk-engelska lexikonet används. Det hade ju vid tidpunkten för Studie<sup>07</sup> ännu inte hade tagits bort (se avsnitt 3.2.1).

LSL4 verkar alltså i större utsträckning användas som receptionsordbok än sin föregångare. Det är ett intressant resultat med tanke på att det övergripande syftet med uppdateringen av LSL var att stärka ordbokens karaktär av just receptionsordbok. I enkätundersökningen av norska Lexin (Rambøll 2012) framkom att ordboken framför allt används i samband med undervisningssituationer, vilka sannolikt inkluderar både reception och produktion.

### 7.2.3 Mest eftersökta informationskategorier

Syftet med föregående fråga var att få veta i vilken eller vilka övergripande situation(er) LSL används. Det aktuella spørsmålet rör vilka konkreta informationskategorier som användarna säger sig vara mest intresserade av att få uppgifter om. Frågan lyder: ”Vilka typer av information söker du mest efter i *Lexins svenska lexikon* (ange max tre alternativ)?”. Informanterna skulle kryssa för max tre alternativ, men även här har vissa informanter kryssat för fler alternativ än så (jfr frågan om användningssituation ovan samt se avsnitt 5.5.1). Informanternas kryss fördelar sig så som redovisas i tabell 7:6.

TABELL 7:6. *Informanternas uppgifter om mest eftersökta informationskategorier i Studie<sup>11</sup>.*

Informationskategori	Antal kryss
Betydelse	672
Exempel	376
Böjning	339
Uttal	284
Ordklass	258
Uttryck	234
Motsats	204
Sammansättning	180
Konstruktion	178
Förkortning	109
Bild	65
Film	40
<b>Totalt</b>	<b>2 939</b>

Av tabell 7:6 kan man utläsa att den mest eftertraktade informationskategorin i Studie<sup>11</sup> är Betydelse med 672 kryss. Det är ett förväntat resultat med tanke på att LSL i första hand är en receptionsordbok.<sup>27</sup> Betydelse värderas ofta mycket

27 Betydelse kan förstås också ha att göra med produktion, t.ex. i situationer där man vill försäkra sig om att man använder ett ord på ett semantiskt riktigt sätt.

högt när den här typen av frågor ställs, oberoende av vilken typ av ordbok som undersöks (Svensén 2004:549). Men det gäller egentligen samtliga kategorier som hamnar i den översta delen av tabell 7:6, dvs. även Exempel (376 kryss) och Bøjning (339 kryss). Liknande resultat framkommer t.ex. i Hartmann (1983) som i första hand undersöker andraspråksinlärares användning av en- och tvåspråkiga tryckta ordböcker och i Lorentzen & Theilgaard (2012), där en mer heterogen informantgrupps användning av en enspråkig internetordbok undersöks. Det verkar inte heller vara någon skillnad mellan användare av tryckta och elektroniska ordböcker.

Vidare har informationskategorierna Uttal, Ordklass, Uttryck och Motsats kryssats för av mellan 200 och 300 informanter. I något mindre utsträckning eftersöks uppgifter om Sammansättning, Konstruktion och Förkortning. Anledningen till att dessa informationskategorier får färre kryss kan vara att innebörden av kategorierna kan uppfattas som oklar. En informant skriver t.ex. att hen inte kan skilja på Exempel och Konstruktion. Konstruktionsuppgifter kan förmedlas på olika sätt och i LSL används döda språkprov. Vi har tidigare berört utmaningen som praktiska lexikografer ställs inför, nämligen att ta fram en korrekt och rättvisande beskrivning av ett ords konstruktionsmönster som samtidigt är användarvänlig i sitt presentationssätt (se avsnitt 3.1.2.6). Enligt andra användarstudier har konstruktionsuppgifter också visat sig vara besvärliga för användarna att förstå och göra bruk av (jfr Bergenholtz 1994). Detta är kanske den naturligaste förklaringen till att Konstruktion hamnar i tabellens nedre del.

I mycket liten omfattning önskas information genom Bild och Film (2 respektive 1 %). I en påföljande kontrollfråga efterfrågas specifikt om informanterna brukar klicka på länkarna till bilder och filmer. Det visar sig att drygt 60 % av informanterna uppger att de ibland eller ofta klickar på länken ”Bild svenska”. För länken ”Visa film” är motsvarande siffra 34 %. Det ger stöd åt de uppgifter som redovisades om navigeringsvalen, nämligen att klickningar mycket väl kan göras på de här länkarna men att det i större utsträckning verkar ske i samband med sökningar i tvåspråkiga lexikon (se avsnitt 6.2). I kapitel 8 kommer vi att få se ytterligare belägg på att det sannolikt förhåller sig på det sättet. Dessutom förefaller informanter som har uppgett ett sent ankomstår klicka på länkarna till bilder och filmer i större utsträckning än de som har bott i Sverige en längre tid (se kapitel 8).

I tabell 7:7 redovisas vilka informationskategorier som efterfrågas mest av de informanter som ingår i Studie<sup>07</sup>.

TABELL 7:7. Informanternas uppgifter om mest eftersökta informationskategorier i Studie<sup>07</sup>.

Informationskategori	Antal kryss
Betydelse	250
Stavning	166
Exempel	91
Böjning	54
Uttal	45
Stående uttryck	45
Sammansättning/avledning	44
Ordklass	41
Uttal via ljudfunktion	31
Bildtema	22
Annat: översättning, synonym	22
<b>Totalt</b>	<b>811</b>

I Studie<sup>07</sup> inleddes frågan med ”*Vilken typ av information ...*”, men informanterna upplystes om att de kunde kryssa för flera alternativ. Antalet informationskategorier i LSL och deras benämning skiljer sig dock åt före och efter uppdateringen av ordboken, vilket får till följd att svarsalternativen i de två frågeformulären avviker något från varandra.

Även i Studie<sup>07</sup> får informationskategorin Betydelse flest kryss. Det är kanske inte ett lika förväntat resultat med tanke på att majoriteten av informanterna här förefaller ha svenska som modersmål (jfr avsnitt 7.1.3.2). Vi vet emellertid att LSL också kan användas som skolordbok och måhända är det informanter i grundskoleåldern som använder ordboken för att få reda på ords betydelse. Vi vet sedan tidigare att en stor del av informanterna i Studie<sup>07</sup> är i just den åldern (se avsnitt 7.1.1). På andra till femte plats hamnar i fallande ordning Stavning, Exempel, Böjning, Uttal. Svarsalternativet Stavning ingick inte i Studie<sup>11</sup> (se avsnitt 5.5.1) men de övriga fyra informationskategorierna efterfrågas, relativt sett, i samma utsträckning i de båda studierna. Även här återfinns kategorin Bild i tabellens nedre del. Att Stavning intar andraplatsen i Studie<sup>07</sup> rimmar väl med ordbokens sekundära uppgift som produktionsordbok. Det gäller även Böjning och Uttal som återfinns på fjärde och femte plats, men med mycket lägre antal kryss.

Uppgifterna om eftersökta informationskategorier kan till viss del jämföras med de resultat som framkom av analysen av lexikonuppslagningarna. Exempelvis återfinns nästan inga sammansättningar bland de mest frekventa

uppslagningarna i LSL (se avsnitt 6.1.1). De efterfrågas också i begränsad omfattning av informanterna. I Studie<sup>11</sup> har 180 av totalt 802 informanter kryssat för informationskategorin Sammansättningar och i Studie<sup>07</sup> har 44 av totalt 356 informanter kryssat för samma kategori. Även om informanterna inte explicit ger uttryck för behov av information om specifika informationskategorier, kan sådana behov ändå vara ytterst verkliga (jfr avsnitt 2.3.3 om funktionslärans åtskillnad mellan subjektiva och objektiva behov).

Det hade också varit intressant att få reda på om det finns någon korrelation mellan ankomstår och efterfrågan på specifika informationskategorier, t.ex. ju senare ankomstår desto viktigare verkar betydelse eller illustrationer (bilder) vara för informanterna. Eftersom en absolut majoritet av informanterna har ett ankomstår mellan 2005 och 2011 och så få informanter har ett tidigare ankomstår, blir det dock svårt att göra några jämförelser.

Betydelse värderas som sagt allra högst i båda studierna. I avsnitt 3.2.1 (tabell 3:1) presenterades en uppställning av relationen mellan de informationskategorier som brukar ingå i en ordbok och de delar som innefattas i kunskapen av ett ord (efter Nation 2013:428). Där betraktas Betydelse som tillhörande receptiv ordkunskap. Exempel är den enda informationskategorin som ingår i både receptiv och produktiv ordkunskap. Den har också lyfts fram som en viktig informationstyp av informanterna. I själva verket är det nog fler kategorier som kan kopplas till situationer av både reception och produktion (jfr Malmgren 2009). Betydelse, som rankas högst av informanterna i båda materialen, är just en sådan kategori. LSL<sup>4</sup> är som tidigare beskrivits mer receptionsorienterad än sin föregångare (se avsnitt 3.2.4). Informanterna ger däremot uppenbarligen uttryck för behov av produktion. Från informanternas perspektiv är troligen dikotomin reception-produktion helt irrelevant. I den konkreta användningssituationen reflekterar de troligen inte så mycket över vilken av de två aktiviteterna som utförs. Inom såväl metalexikografen som i det praktiska lexikografiska arbetet har uppdelningen däremot större relevans, vilket inte minst framkommer i kapitel 2.

#### 7.2.4 Instruktioner till Lexin

Enligt Svensén (2004:542f.) läses användarinstruktioner i regel i mycket liten utsträckning. Ändå är användarvägledningar av olika slag ett stående inslag i de allra flesta ordböcker (Atkins & Rundell 2008:177). På frågan om informanten någon gång läst instruktionerna till Lexin, där det står om hur man ska tolka den information som finns i Lexin-ordböckerna, svarar en fjärdedel jakande och en fjärdedel att de inte vet. Den återstående hälften har inte läst instruktionerna. Att en fjärdedel av informanterna uppger att de har läst instruk-



tionerna till Lexin får nog anses vara en ganska hög siffra, och texterna fyller uppenbarligen en viktig funktion för vissa användare (liknande resultat redovisas i Holmer et al. 2015:363). Det är möjligt att informanterna har fått undervisning i hur man använder Lexin i samband med sina svenskstudier, men det vi inget om.

### 7.2.5 Hjälpmedel vid sidan av Lexin

Lexins webbsida är som vi vet välbesökt (se avsnitt 3.2.4). Det allmänna utbudet av såväl tryckta som elektroniska ordböcker är emellertid stort och man kan fråga sig till vilka andra hjälpmedel informanterna vänder sig för att lösa språkliga problem, vid sidan av LSL. I tabell 7:8 redovisas hur många informanter i Studie<sup>11</sup> som kryssat för respektive svarsalternativ. De kunde kryssa för alla fem alternativen om så önskades.

TABELL 7:8. Studie<sup>11</sup>: Användning av andra språkliga hjälpmedel vid sidan av LSL.

Alternativa hjälpmedel	Antal kryss	Relativ frekvens (%)
Google eller motsvarande	554	70
Annan ordbok på internet	420	52
Frågar någon	337	42
Annan tryckt ordbok	327	41
Annat sätt	46	6

Av tabellen framgår att 70 % använder Google eller en annan sökmotor och att över hälften av informanterna dessutom vänder sig till en annan ordbok på internet. Det är ett förväntat resultat i 2000-talets alltmer uppkopplade tillvaro. Omkring 40 % av informanterna uppger emellertid att de använder en tryckt ordbok. Man skulle kunna tänka sig att dessa ordboksanvändare tillhör en äldre generation. Det finns dock inga stöd för denna slutsats i mitt material. En rimlig förklaring är att många informanter säkerligen har tryckta ordböcker i hemmet och att dessa kommer till användning när de elektroniska versionerna inte är till någon hjälp. Det är heller inte säkert att man alltid har tillgång till elektroniska varianter.

En nästan lika stor andel av informanterna frågar någon om råd. Det är en situation man lätt kan föreställa sig eftersom de flesta informanterna uppger att de arbetar eller studerar. Då är man vanligen omgiven av människor som man kan fråga. Att endast 5,7 %, dvs. 46 informanter, kryssar för ”Annat sätt”

indikerar att de svarsalternativ som erbjöds täcker in alternativa hjälpmedel på ett tillfredsställande sätt.

Ett femtiotal informanter har valt att lägga till en kommentar. Bland de kommentarer som återkommer finns "Google translate", namngivning av en specifik ordbok (t.ex. "SAOL", "tyda.se", "babylon") eller det att man frågar en lärare. Ordbehandlingsprogrammet Words stavnings- och grammatikkontroll nämns en gång. Resultatet ligger i linje med Nilzon (2016) som undersöker språkintrouktionselevens användning av elektroniska ordböcker och översättningsverktyg. Han kommer fram till att jämte Lexin är Google translate det vanligaste språkliga hjälpmedlet. Enligt Nilzon verkar eleverna ha dålig kännedom om skillnaden mellan dessa helt olika typer av språkliga resurser, och deras främsta förtjänster (jfr Bergenholtz 2007). Eleverna tycks heller inte uppfatta detta som något problematiskt.

## 7.2.6 Sammanfattning

Informanterna i Studie<sup>11</sup> förefaller vara flitiga användare av LSL4. Sammanlagt uppger över 80 % att de använder ordboken flera gånger varje dag eller flera gånger i veckan. Informanterna i Studie<sup>11</sup> är frekventare ordboksanvändare än de i Studie<sup>07</sup>.

LSL4 verkar i större utsträckning användas som receptionsordbok än sin föregångare. Det är ett intressant resultat med tanke på att det övergripande syftet med uppdateringen av LSL var att stärka ordbokens karaktär av just receptionsordbok. Många informanter uppger emellertid att de använder LSL i både situationer av reception och av produktion. Betydelse är den specifika informationskategori som efterfrågas allra mest. Denna kategori får överlägset flest kryss i såväl Studie<sup>11</sup> som Studie<sup>07</sup>. Betydelse betraktas som en receptiv kategori. I den konkreta användningssituationen reflekterar informanterna troligen inte så mycket över vilken sorts aktivitet det är som utförs. Uppdelningen i reception och produktion är sannolikt av begränsad relevans för dem.

En fjärdedel av informanterna uppger att de har läst instruktionerna till Lexin och det får nog anses vara en ganska hög siffra. De fyller uppenbarligen en viktig funktion för vissa användare.

Slutligen uppger närmare 70 % av informanterna att de använder Google eller en annan sökmotor, vid sidan av LSL. Hälften av informanterna svarar också att de använder en annan ordbok på internet. Omkring 40 % använder en tryckt ordbok.

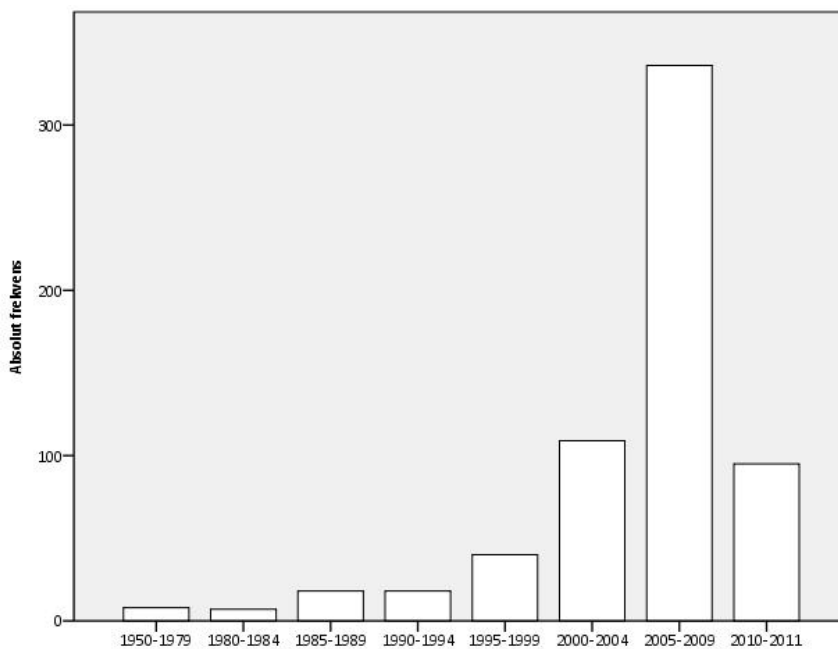
## 7.3 Användare med annat modersmål än svenska

En övergripande frågeställning är huruvida LSL används av den åsyftade målgruppen, dvs. personer som har ett annat modersmål än svenska och befinner sig på en relativt avancerad nybörjarnivå i sin svenskinläring (se avsnitt 3.2.1). Jämförelsen mellan uppslagningarna i LSL3 och LSL4 pekade t.ex. på att användarna av LSL4 i större utsträckning utgörs av andraspråksinlärare (se avsnitt 6.1.1).

Frågorna i formulärets tredje avsnitt fokuserar på LSL:s målgrupp. I detta avsnitt sammanfaller tre frågor från Studie<sup>11</sup> och Studie<sup>07</sup>, närmare bestämt de som rör modersmål, kunskaper i svenska språket och utbildning i hemland. Det sistnämnda spørsmålet har redan avhandlats i avsnitt 7.1.3.2.

### 7.3.1 Ankomst till Sverige

Sammanlagt 632 informanter i Studie<sup>11</sup> svarar på frågan om vilket år de anlände till Sverige. Det motsvarar 79 % av det totala antalet. De angivna ankomststären åskådliggörs i figur 7:4.



FIGUR 7:4. Studie<sup>11</sup>: Informanternas ankomstår till Sverige.

Av figuren kan man utläsa att majoriteten informanter anlände till Sverige under 2000-talet, närmare bestämt 86 %. De allra flesta av dem kom hit mellan 2005 och 2009 och omkring ett hundratal anlände så sent som 2010–2011. Vid tidpunkten för datainsamlingen hade dessa informanter alltså befunnit sig i Sverige i högst sex år. Majoriteten av informanterna har således kommit till Sverige så pass nyligen att de faller inom ramen för den typ av ordboksanvändare Lexin-ordböckerna vänder sig till.

I det följande redovisas uppgifter om informanternas utbildning i svenska språket, användning av svenska i olika sammanhang samt deras modersmål. Därefter försöker jag göra en uppskattning av hur många av informanterna som tillhör målgruppen för LSL, dvs. behärskar svenska på en relativt avancerad nybörjarnivå.

### 7.3.2 Sfi och/eller annan kurs i svenska språket

Sammanlagt 445 informanter i Studie<sup>11</sup> har genomgått sfi-utbildning. De allra flesta av dem har kommit till Sverige under 2000-talet och i synnerhet under perioden 2005–2009. I avsnitt 5.5.1 framkom att för att få läsa sfi krävs att man har uppehållstillstånd i Sverige. Det kan emellertid ta lång tid innan man får reda på huruvida man beviljas uppehållstillstånd eller inte. Det kan således dröja flera år från det att man anlant till Sverige till dess att man börjar läsa sfi.

Enligt statistik från Skolverket har antalet individer som deltar i sfi ökat sedan slutet av 1990-talet (Skolverket 2016). Omkring år 2007 var antalet sfi-studerande ca 70 000 och 2011 hade det ökat till ca 100 000. År 2015 uppgick det till 138 000. Ett hundratal språk finns representerade bland dem som deltar i sfi-utbildning. Arabiska är det vanligaste modersmålet. Vidare är majoriteten av sfi-studenterna 25–39 år.

De 445 informanter som anger att de har genomgått en sfi-utbildning har, får man anta, vissa grundläggande kunskaper i svenska språket. Som redan påpekats är det dock stor skillnad på att ha uppfyllt kunskapskraven för t.ex. kurs A, C eller D (se avsnitt 5.5.1). Indirekt får vi också veta att de informanterna har uppehållstillstånd i Sverige och att de således är folkbokförda här.<sup>28</sup>

Frågan om sfi-utbildning följdes upp med en fråga om annan kurs eller utbildning i svenska språket. Ett antal av 433 informanter uppger att de har en sådan utbildning och även dessa informanter har till största delen kommit till Sverige under 2000-talet, också här framför allt under perioden 2005–2009. Nästan 400 av dem specificerar också vilken utbildning det rör sig om. Ofta

---

28 Det här var alltså förutsättningarna vid tidpunkten för datainsamlingen 2011.

anges i förkortad form olika typer av grundläggande kurser i svenska språket (i vissa fall verkar det röra sig om ovan nämnda kunskapskrav för sfi, dvs. informanterna har angett någon av bokstäverna A, B, C eller D). Kommentarererna kan också innehålla andra typer av svenska utbildningar. Några exempel från materialet återges nedan.

- (7:32) styr och regler mekaniker
- (7:33) lastbil kurs taxi kurs
- (7:34) Sjukvårdssvenska
- (7:35) omvårdnad kurs packet
- (7:36) tolkutbildning

De anförda exemplen är ytterligare belägg för att informanterna tycks ha ett behov av att lyfta fram att de har en gedigen utbildning eller ett kvalificerat yrke (jfr avsnitt 7.1.3.1 om informanternas kommentarer i anslutning till frågan om utbildning i Sverige).

### 7.3.3 Modersmål

På frågan om modersmål svarar 690 informanter i Studie<sup>11</sup>, dvs. 86 %, att de har ett annat modersmål än svenska. Så många som 64 olika språk är representerade bland svaren. Ett tjugotal av informanterna anger fler än ett modersmål. I Studie<sup>07</sup> uppges 50 olika språk (inklusive svenska och teckenspråk) fördelat på 322 informanter (jfr tabell 5:5 i avsnitt 5.5.2). Här uppges ett tiotal informanter två eller tre språk. Det som informanterna uppfattar som sitt modersmål räknar jag som ett modersmål (jfr termen *förstaspråk*, se avsnitt 3.1.1).

I tabell 7:9 redovisas i fallande ordning hur många informanter i Studie<sup>11</sup> som har något av Lexin-språken som modersmål. Motsvarande uppgifter ges också för Studie<sup>07</sup>. Nordkurdiska och sydkurdiska redovisas gemensamt eftersom majoriteten av informanterna inte anger vilket av språken de talar. Icke-Lexinspråk redovisas tillsammans under rubriken ”Övriga språk”.

Av tabell 7:9 framgår att samtliga språk som företräds av ett lexikon på Lexins webbsida återfinns bland de angivna modersmålen i såväl Studie<sup>11</sup> som Studie<sup>07</sup>. Totalt sett har 63 % av informanterna i Studie<sup>11</sup> något av Lexin-språken som modersmål, medan motsvarande andel i Studie<sup>07</sup> endast är 29 %. Arabiska, persiska, spanska och kurdiska är de språk som har störst antal informanter i båda studierna. Ryska och somaliska har också många informanter i Studie<sup>11</sup> men få i Studie<sup>07</sup>.

TABELL 7:9. Antal informanter per Lexin-språk och övriga språk i Studie<sup>11</sup> respektive Studie<sup>07</sup>.

Studie <sup>11</sup>		Studie <sup>07</sup>	
Språk i Lexin	Informanter/ modersmål	Språk i Lexin	Informanter/ modersmål
Arabiska	105	Arabiska	23
Persiska	89	Persiska	18
Spanska	79	Kurdiska	15
Kurdiska (nord+syd)	49	Spanska	10
Somaliska	46	Engelska	5
Ryska	42	Finska	4
Finska	21	Bosniska	4
Serbiska (latinskt)	16	Turkiska	3
Grekiska	15	Albanska	3
Bosniska	14	Serbiska	2
Turkiska	13	Kroatiska	2
Kroatiska	6	Grekiska	2
Albanska	4	Ryska	1
Amhariska	4	Somaliska	1
Azerbajdzjanska	3		
Pashto	3		
Övriga 47 språk	181	Övriga 35 språk	47
Ej svar	103	Ej svar	34
Nonsens, ej uttolkbart	8		
Svenska	1	Svenska	182
<b>Totalt</b>	<b>802</b>	<b>Totalt</b>	<b>356</b>

Av tabellen framgår vidare att 181 informanter i Studie<sup>11</sup> anger ett annat modersmål än de som finns representerade i Lexin. Sammanlagt är det fråga om 47 icke-Lexinspråk. De icke-Lexinspråk som har flest representanter är, i fallande ordning (med antal informanter inom parentes): kinesiska (24), tyska (19), thailändska (16), polska (13) och assyriska/syrianska (10). Den resterande gruppen språk har mellan två och nio representanter, t.ex. engelska (9), bulgariska (5), danska (3), japanska (3), urdu (2). De resterande 25 angivna språken har endast en representant, t.ex. armeniska, isländska, kachin (talas i norra Burma), mongoliska, romani, slovenska, tagalog och uzbekiska.

Enligt Lexins webbsida pågår för närvarande (oktober 2016) arbete med lexikon på romska (arli, kelderash och lovari), thailändska och tigrinska. På andra ställen i Lexin-portalen framgår dock att ett svensk-tigrinskt lexikon

finns i både tryckt format och på internet. Dessutom är det möjligt att beställa svenskt-romskt lexikon (arli). Uppgifterna på webbsidan verkar helt enkelt vara i behov av en uppdatering.

I Studie<sup>07</sup> återfinns 47 informanter fördelade på 35 icke-Lexinspråk som modersmål. De språk som har flest representanter är, i fallande ordning (med antal informanter inom parentes): engelska (5), polska (5), thailändska (4), franska (3), dari (2), syrianska (2), ungerska (2). De resterande 25 angivna språken har endast en representant, t.ex. armeniska, estniska, filippinska, hebreiska, isländska, kinesiska, nederländska, norska, rumänska, pashto, somaliska och vietnamesiska. Somaliska fanns faktiskt redan tillgängligt vid tidpunkten för undersökningen men hade troligen lagts upp ganska nyligen. Detsamma gäller en tryckt svensk-vietnamesisk ordbok som gavs ut 2007. Däremot har pashto tillkommit i internetversionen (se vidare avsnitt 3.3.1 om Lexin-projektets kriterier för urval av språk).

Som framgår av tabell 7:9 är antalet informanter som har angivit att de har svenska som modersmål litet i Studie<sup>11</sup> men stort i Studie<sup>07</sup>. En av anledningarna till att detta är att i studien från 2011 skulle endast de med annat modersmål än svenska besvara frågan. I Studie<sup>07</sup> ombads däremot alla informanter att ange sitt modersmål. En stor andel av de 103 informanter som inte har besvarat frågan i Studie<sup>11</sup> har sannolikt svenska som sitt förstaspråk. Denna slutsats bekräftas inte minst av att antalet uteblivna svar på de flesta av frågorna som riktar sig till just andraspråkstalare är omkring 100 (se tabell 5:4 i avsnitt 5.5.2).

Genom att jämföra informanternas modersmål med hur många uppslagningar som görs i de tvåspråkiga Lexin-lexikonerna kan man få en uppfattning om hur representativt urvalet av informanterna är. Tabell 7:10 innehåller uppgifter om det totala antalet uppslagningar i respektive lexikon under oktober 2011 och september 2007, dvs. samma månader som datainsamlingen gjordes.

Av tabell 7:10 framgår att de språk som har många representanter i Studie<sup>11</sup> och Studie<sup>07</sup> också har höga antal uppslagningar i motsvarande lexikon i Lexin. Om man bortser från det svensk-engelska lexikonet i Studie<sup>07</sup> är det lexikonet från svenska till arabiska, spanska och ryska som används mest de undersökta månaderna både år 2011 och 2007. Användningen av det svensk-finska och även det svensk-spanska har kraftigt minskat mellan 2007 och 2011 medan bruket av det svensk-persiska och svensk-somaliska står för en dramatisk ökning. Det svensk-somaliska lexikonet hade nyligen publicerats vid tidpunkten för Studie<sup>07</sup> och detsamma gäller veterligen för den svensk-persiska ordboken. En tänkbar förklaring till den kraftigt ökade användningen är att brukarna ännu inte noterat tillkomsten av de nya lexikonerna. I mars 2008, dvs. ett halvår senare, görs t.ex. drygt 14 000 uppslagningar i det svensk-persiska lexikonet och i september samma år är antalet uppe i över en miljon.

TABELL 7:10. *Antal uppslagningar i de tvåspråkiga Lexin-lexikonerna september 2011 och oktober 2007, efter frekvens.*

Lexikon	Oktober 2011, antal uppslagningar per ordbok	Lexikon	September 2007, antal uppslagningar per ordbok
Svenskt-Arabiskt	4 383 317	Svenskt-Engelskt	20 936 370
Svenskt-Persiskt	3 368 131	Svenskt-Arabiskt	2 282 278
Svenskt-Somaliskt	2 694 017	Svenskt-Spanskt	2 081 881
Svenskt-Ryskt	1 164 713	Svenskt-Finskt	1 050 319
Svenskt-Spanskt	961 800	Svenskt-Ryskt	792 837
Svenskt-Kurdiskt (nord+syd)	580 447	Svenskt-Turkiskt	440 478
Svenskt-Turkiskt	470 981	Svenskt-Somaliskt	326 838
Svenskt-Grekiskt	255 633	Svenskt-Serbiskt	292 147
Svenskt-Serbiskt	253 620	Svenskt-Grekiskt	214 374
Svenskt-Finskt	230 700	Svenskt-Bosniskt	200 842
Svenskt-Albanskt	175 714	Svenskt-Albanskt	197 687
Svenskt-Bosniskt	139 087	Svenskt-Kurdiskt (nord+syd)	131 601
Svenskt-Amhariskt	89 565	Svenskt-Kroatiskt	105 008
Svenskt-Kroatiskt	75 290	Svenskt-Persiskt	87
Svenskt-Pashto	19 754		
Svenskt-Azerbajdzjanskt	12 753		
<b>Totalt</b>	<b>14 875 522</b>	<b>Totalt</b>	<b>29 054 747</b>
		Exkl. svenskt-engelskt	8 118 377

Vidare finns det ett starkt samband mellan rangordningen av språk efter antal uppslagningar på nätet (enligt tabell 7:10) och rangordningen av informanternas språk i mitt material (enligt tabell 7:9) (jfr de Schryver et al 2006). För Studie<sup>11</sup> ger Spearmans rangkorrelationstest det höga värdet 0,91.<sup>29</sup> Det innebär att mitt urval av informanter åtminstone är representativt vad gäller fördelningen på informanternas modersmål med tanke på hur den verkliga användningen ser ut.

29 Spearmans rangkorrelation beräknas enligt:  $p = 1 - \frac{6\sum d_i^2}{n(n^2-1)}$ , där  $d_i = x_i - y_i$  är skillnaden i rang och  $n$  är antalet observationer. Värdet 1 motsvarar högsta korrelation och värde 0 minsta korrelation (Wikipedia, *Spearmans rangkorrelation*).



Så länge det svensk-engelska lexikonet fanns tillgängligt utfördes som vi vet majoriteten av uppslagningarna i detta lexikon. Det fanns endast tillgängligt vid tidpunkten för Studie<sup>07</sup> men inte då Studie<sup>11</sup> genomfördes. Det är förklaringen till att det totala antalet uppslagningar är så mycket större i september 2007 än i oktober 2011. Om man räknar bort uppslagningarna i det svensk-engelska lexikonet kan man sluta sig till att de flesta andra tvåspråkiga lexikonen har ökat i antal uppslagningar. I undersökningen av norska Lexin (Rambøll 2012) framgick att undantaget det norsk-engelska lexikonet används ordböckerna från norska till arabiska, polska, somaliska, tigrinska och ryska mest (se avsnitt 4.3.3.1).

### 7.3.4 Bästa språk

Som tidigare konstaterats (se avsnitt 3.1.1) är det språk som man har som modersmål inte nödvändigtvis det språk som man behärskar bäst. Personer som har immigrerat till ett nytt land som barn behärskar i vuxen ålder ofta sitt andraspråk bättre än sitt förstaspråk. Genom svaret på frågan om vilket språk man anser sig vara bäst på kan man få en indikation på hur situationen ser ut för informanterna i just denna undersökning. Av svaren att döma är det otydligt så att majoriteten av informanterna uppger sitt modersmål som sitt bästa språk. Uppemot 80 % gör detta. Den återstående andelen, 20 %, består av informanter som uppger att deras bästa språk är engelska, svenska eller ett annat språk. Uppgifterna är på intet sätt förvånande eftersom den stora majoriteten av informanterna har kommit till Sverige relativt nyligen och dessutom många gånger som (unga) vuxna.

### 7.3.5 Bruk av svenska och svenskkunskaper

Drygt 70 % av informanterna med annat modersmål än svenska uppger att de talar eller använder svenska språket varje dag eller nästan hela tiden. Av svaren framgår också att majoriteten av informanterna använder modersmålet när de talar med sin familj men att de talar svenska i skolan eller på jobbet. I samtal med vänner ser bilden emellertid annorlunda ut. Med viss övervikt för svenska så talar informanterna i nära nog lika hög grad antingen svenska *eller* modersmål eller svenska *och* modersmål med sina vänner.

Som redan diskuterats är svaren på självskattningsfrågor många gånger vanskliga att tolka, inte minst eftersom det (i det här fallet) är omöjligt att veta efter vilken måttstock svenskkunskaperna bedöms (se avsnitt 4.2.1). På frågan om hur bra informanterna bedömer att deras kunskaper i svenska är kryssar

61 % av informanterna i Studie<sup>07</sup> och 57 % av informanterna i Studie<sup>11</sup> för alternativen ”Bra” eller ”Mycket bra”. Det är intressant med tanke på att majoriteten av 2007-informanterna är modersmålstalare av svenska medan majoriteten av 2011-informanterna har ett annat modersmål än svenska. Detta är ett lysande exempel på svårigheten med att tolka svar på frågor av det här slaget (jfr Lew 2004:71).

Det kan dock vara intressant att ta reda på i vilken utsträckning informanterna själva anser att de talar eller använder sig av svenska språket eller vilken uppfattning de har om nivån på sina språkkunskaper. I sig själv har svaret på frågan kanske begränsat intresse, men tillsammans med frågorna om modersmål, bästa språk och med vem man använder vilket språk, kan den som sagt bidra med relevant information om informanternas vardagliga språkanvändning.

### 7.3.6 Sammanfattning

Svaren på frågorna i detta avsnitt visar för det första att majoriteten av informanterna i Studie<sup>11</sup> har svenska som sitt andraspråk, till skillnad från informanterna i Studie<sup>07</sup> där svenska är det vanligaste modersmålet. Totalt sett uppger 86 % av informanterna att de har ett annat modersmål än svenska och 64 % anger något av Lexin-språken. De vanligaste modersmålen i båda materialen (dvs. fränsett svenska i Studie<sup>07</sup>) är arabiska, persiska, kurdiska och somaliska. Det innebär att mitt urval av informanter förefaller vara representativt vad gäller fördelningen på informanternas modersmål med tanke på hur den verkliga användningen ser ut.

Vidare har majoriteten av informanterna i Studie<sup>11</sup> kommit till Sverige under 2000-talet. Många av dem har läst sfi vilket innebär att de har uppehållstillstånd i Sverige. Naturligt nog uppger de flesta att modersmålet är deras bästa språk. Många informanter har svarat att de dagligen kommunicerar på svenska, men att det framför allt sker i skolan eller på jobbet och inte i hemmet.

Min samlade bedömning av enkätsvaren är att en stor del av informanterna verkar tillhöra den primära målgruppen för LSL, dvs. inlärare på relativt avancerad nybörjarnivå. Utan djupare kännedom om informanternas reella svensk-kunskaper är det dock omöjligt att veta hur det verkligen förhåller sig. En fingervisning om detta kan däremot informanternas fritt skrivna kommentarer ge. Dessa behandlas i följande avsnitt.

## 7.4 Kommentarer om LSL

Båda versionerna av frågeformulären avslutades med ett kommentarsfält där informanterna kunde skriva ned eventuella synpunkter på LSL. Informanternas kommentarer kan t.ex. ge värdefull information om vilka lexikografiska behov de anser sig ha. Som forskare kan man därtill få reda på saker man inte tänkt på, t.ex. ifråga om utformningen av enkäten eller när det gäller det ämne som undersöks (jfr Müller-Spitzer 2013 som tar ett lite annorlunda grepp i utformningen av sin enkät, se avsnitt 4.3.3.1). I det här fallet kan kommentarerna som sagt också ge en antydning om informanternas språkförmåga, och det kan jämföras med vad vi vet om svaren på frågorna om svenskkunskaper. Att be informanterna formulera sig med egna ord kan vidare vara ett sätt att skapa sig en klarare bild av vilka de är. Eftersom vi här har möjlighet att jämföra informanternas synpunkter på LSL före och efter uppdateringen kan vi dessutom få indikationer på om den uppdaterade LSL upplevs som en bättre ordbok.

Den inledande formuleringen i enkäten från 2011 har en annan utformning än den i 2007 års enkät:

Studie<sup>11</sup>: Här kan du skriva ned dina åsikter eller kommentarer kring *Lexins svenska lexikon*.

Studie<sup>07</sup>: Har du någon kommentar kring Lexin eller *Svenska ord*? Är det t.ex. något du saknar i ordboken? Skriv gärna det här!

Formuleringarna är emellertid inte så olika att svaren inte kan jämföras med varandra. I Studie<sup>11</sup> skriver 371 av 802 informanter en kommentar vilket motsvarar 46 %. En klar majoritet av kommentarerna i Studie<sup>11</sup>, 340 av 371, har skrivits av informanter med ett icke-svenskt modersmål. Om man ser till de informanter som tillhör LSL:s målgrupp, dvs. relativt nyligen har invandrat till Sverige och har ett annat modersmål än svenska, har 67 % av kommentarerna skrivits av dem. I Studie<sup>07</sup> väljer 110 informanter att skriva något i kommentarsfältet. Det motsvarar 31 % av de totalt 356 informanterna. I Studie<sup>07</sup> är det okänt hur många av kommentarerna som har skrivits av informanter med annat modersmål än svenska. Det vi vet är att 57 % av informanterna (182 av 318 svarande) har svenska som modersmål (se tabell 7:9). I tabell 7:11 ges uppgifter om kommentarernas antal och ordlängd i de två studierna.

TABELL 7:11. Kommentarerens antal, längd (i antal ord räknat) samt medianlängd.

Material	Totala antalet informanter	Antal kommentarer	Minst antal ord	Störst antal ord	Medianlängd
Studie <sup>11</sup>	802	371	1	168	17
Studie <sup>07</sup>	356	110	1	130	14

Av tabell 7:11 framgår att det finns såväl korta som mycket långa kommentarer men att medianlängden är 17 ord i Studie<sup>11</sup> och 14 ord i Studie<sup>07</sup>. De allra flesta kommentarer ligger inom spannet 4 till 29 ord.<sup>30</sup>

Det finns olika sätt att ta sig an den information som finns i kommentarerna. Jag har valt att först kategorisera de två uppsättningarna av kommentarer i följande typer: 1) positivt utlåtande, 2) kritik, 3) förbättringsförslag, 4) övrig anmärkning, 5) ändring av gränssnitt (förklaring ges nedan) och 6) nonsens. Därefter undersöker jag vilka informationskategorier i ordboksartiklarna som nämns explicit och huruvida just dessa kommenteras i positiva eller negativa ordalag (jfr Bjørneset 2016 som redovisar informanternas kommentarer från en undersökning om användningen av norska Lexin).

### 7.4.1 Typer av kommentarer

I tabell 7:12 redovisas absoluta och relativa frekvenser för kommentarerna i Studie<sup>11</sup> och Studie<sup>07</sup>, indelade i ovan nämnda kategorier. Kommentarer som innehåller t.ex. både kritik och positivt utlåtande har delats upp och respektive komponent har placerats i den kategori den bör ingå i. Det är förklaringen till att tabellens absoluta frekvenser är högre än det tidigare redovisade antalet kommentarer i de två studierna. För Studie<sup>11</sup> gäller att de 371 kommentarerna delats upp i 565 komponenter och för Studie<sup>07</sup> har 110 kommentarer blivit 151 komponenter.

TABELL 7:12. *Absoluta och relativa frekvenser för respektive typ av kommentar.*

Typ av kommentar	Studie <sup>11</sup>		Studie <sup>07</sup>	
	Absolut frekvens	Relativ frekvens (%)	Absolut frekvens	Relativ frekvens (%)
Positivt utlåtande	239	42	30	20
Kritik	123	22	53	35
Förbättringsförslag	115	20	48	31,5
Övrig anmärkning	69	12	17	11
Ändring av gränssnitt	15	3	–	–
Nonsens	4	1	4	2,5
<b>Total</b>	<b>565</b>	<b>100</b>	<b>151</b>	<b>100</b>

30 Müller-Spitzer (2013) redovisar en genomsnittlig kommentarslängd på 37 ord och hälften av kommentarerna har mellan 15 och 47 ord. Hennes studie riktar sig dock till en annan typ av målgrupp och frågan uppmuntrade till uttömmande och elaborerade svar (se avsnitt 4.3.3.1).

Under kategorin ”Ändring av gränssnitt” hamnar de kommentarer där det påpekas att ordboksartiklarna ser annorlunda ut än de brukar. Som tidigare redovisats (se avsnitt 5.5.1) ändrades användargränssnittets inställningar under den vecka då enkäten låg uppe på LSL:s webbsida.

I Studie<sup>11</sup> utgörs flest kommentarer av positiva utlåtanden (42 %). De uppgår till lika stor andel som de kritiska kommentarerna och förbättringsförslagen tillsammans. Anmärkningsvärt många kommentarer (69 stycken) har placerats i ”Övrig anmärkning”-kolumnen. De utgör ofta delar av en längre kommentar. Några informanter skriver t.ex. något om sig själva och sin bakgrund (”Jag har dyslexi”) och andra anger vilka andra ordböcker de använder (”använder tyda.se istället”). Ytterligare andra gör ett tillägg såsom ”exempel är mkt viktiga för invandrare” eller ”hoppas det förblir gratis att använda den”.

I Studie<sup>07</sup> utgörs 30 av de 151 kommentarerna, dvs. 20 %, av positiva utlåtanden om Lexin. De kritiska kommentarerna är nästan lika många som förbättringsförslagen. Tillsammans utgör de två typerna 66,5 % av kommentarerna. Kommentarer som hamnar under typen ”Övrig anmärkning” är antingen inte direkt relaterade till just LSL eller av mer neutral karaktär. ”Nonsens”-kommentarer innehåller ingen information av värde eller så är den inte uttolkbar. Till skillnad från vad som är fallet i Studie<sup>11</sup> bidrar de övriga kommentarerna i Studie<sup>07</sup> alltsomoftast inte med någon väsentlig information. I tabell 7:13 ges exempel från de två materialen och från respektive kategori.

TABELL 7:13. Exempel på kommentarer av varje typ i Studie<sup>07</sup> och Studie<sup>11</sup>.

Typ av kommentar	Exempel från Studie <sup>11</sup>	Exempel från Studie <sup>07</sup>
Positivt utlåtande	<i>Tack för att ni finns! Utan er blir det mycket svårt att klara mig vardagliges.</i>	<i>Jag tycker Lexin och Svenska ord är en väldigt bra sida, jag använder den väldigt ofta och det är den bästa uppslagssidan jag någonsin hittat.</i>
Kritik	<i>Jag tycker inte att ni har tillräckligt med ordförråd i lexin.</i>	<i>Ni har bara lätta ord, t.ex. hund, katt, men människor kan de här orden..jag behöver förklaring till svårare ord.</i>
Förbättringsförslag	<i>Bra om skulle tillägga mer sammansättningar.</i>	<i>det skulle vara kul om det även fanns fackord, eller ord som är knutna till vissa ämnen, ex biologi mfl.</i>
Övrig anmärkning	<i>Visste inte om bild och film.</i>	<i>Alla punkterna på fråga 8 [dvs. alla informationskategorier, min anm.] tycker jag är viktigt för att man ska ha full förståelse i ett sammanhang.</i>
Ändring av gränssnitt	<i>Det är synd att ni tagit bort ords exempel o verbs olika former.</i>	
Nonsens	<i>Huh</i>	<i>ROCK THE WORLD</i>

Den tillfälliga ändringen av gränssnittet gjordes som sagt inför Studie<sup>11</sup> (jfr figur 5:4 och 5:5 i avsnitt 5.1.1). I åtminstone 15 formulär uppmärksammas ändringen i kommentarsfältet, som t.ex. i följande kommentarer:

- (7:37) Men nu förändring, verbböjning borta. Synd!
- (7:38) Jag saknar verbböjning i Lexin.
- (7:39) De exempel plus tempus(verb) som fanns i lexikon var nyttig men tyvärr har försvunnit.
- (7:40) mycket viktig för mig som lexin tillbaka att blir med meningar och exempel, tack
- (7:41) Men sista tid finns inte information om olika form av verb, substantiv, adjektiv. Nu svårt att hitta riktig form. Tråkigt.
- (7:42) Varför har ni ändrat översättningens sätt, eftersom nu kan jag inte se t.ex. verb sina olika tempus, substantiv sina ändelser

Just dessa informanter verkar dock inte ha uppmärksammat anledningen till att ordboksartiklarna för tillfället har ett avvikande utseende. De informationskategorier som framförallt kommenteras är böjning och språkexempel och det är avsaknaden av dem som noteras.

Det som hittills redovisats indikerar att informanterna är mer nöjda med den uppdaterade versionen av LSL än med motsvarande icke uppdaterade version. De positiva kommentarerna har t.ex. ökat med 13,2 procentenheter medan den negativa kritiken har minskat med 12,4 enheter. I Hult (2014) undersöker jag också om och hur ofta informanterna explicit nämner orden för de specifika informationskategorierna i LSL samt om de omnämns i positiva eller negativa ordalag. Lemma visar sig vara den allra mest omnämnda informationskategorin i båda materialen. Alla övriga kategorier omnämns i mycket liten utsträckning, möjligen undantaget Exempel i Studie<sup>11</sup> (se vidare Hult 2014).

Önskan om fler lemman inkluderar alla övriga informationskategorier. Därför ligger det nära till hands att helt enkelt efterfråga fler ord istället för att specificera sina lexikografiska behov. Som redan påpekats kan det också finnas en osäkerhet hos informanterna med avseende på innebörden av de olika kategorierna (se avsnitt 7.2.3).

I de följande avsnitten presenteras exempel på kommentarer från Studie<sup>11</sup> och Studie<sup>07</sup>, med fokus på vad informanterna har att säga om kategorierna Lemma, Betydelse och Exempel. Dessa informationskategorier är kanske viktigast i en receptionsordbok som LSL.

## 7.4.2 Informanternas kommentarer om Lemma

De positiva kommentarerna om Lemma handlar helt enkelt om att informanterna ofta hittar de ord som de slår upp i LSL och att de är nöjda med den information de får fram. Praktiskt taget alla negativa kommentarer om Lemma i båda studierna handlar om avsaknad av uppslagsord. Många andra användarstudier vittnar också om informanternas önskemål om fler uppslagsord (se t.ex. avsnitt 4.3.1 om resultat från traditionella enkätstudier). Flera informanter är t.ex. kritiska till avsaknaden av svåra ord i LSL (jfr avsnitt 3.1.1). Några exempel på kommentarer om Lemma från 2011 och 2007 är följande:

Studie<sup>11</sup>:

- (7:43) Saknas många sammansättningsord och speciella term t.ex. fonologi, autism.
- (7:44) Svårt hitta betydelse för långa ord.
- (7:45) Men det känns som att det är de mest grundläggande orden som finns, inte andra.
- (7:46) Ibland saknas det viktiga ord.

Studie<sup>07</sup>:

- (7:47) Vissa ord man söker på finns inte med, trots att det är ett riktigt ord.
- (7:48) Det eganska många psykologiska eller politiska termer/ord som int finns.
- (7:49) Det som är dåligt att det finns inte sammansatta ord och partikelverb. Vi invandrare behöver mest hjälp med sammansatta och partikelverb som finns inte i den här lexikon.
- (7:50) Jag saknar dom svåra ord. Ord som finns är mest på nybörjar eller grundläggande nivå.

I uppdateringen av LSL utökades ordboken med cirka 3 000 uppslagsord (Malmgren 2012). Många synonymer och antonymer lades också till i ordboksartiklarna. I detta avseende har informanternas önskemål om fler lemman av olika slag uppfyllts. Informanterna uttrycker sig också oftare positivt till kategorin Lemma efter uppdateringen än före. De förblir emellertid förhållandevis kritiska till lemmaurvalet och ger mer eller mindre uttryck för samma typ av kritik 2011 som de gör 2007. Frågan är också om någon ordbok någonsin till fullo kommer att lyckas tillfredsställa användarnas behov av fler lemman. Den korta och koncisa kommentaren ”men det finns inte alla ord” är mycket talande vad gäller många användares förväntningar på LSL:s lemmaurval. Tillsammans med kommentarerna i exempel 7:47–7:50 visar denna kommentar också på att många informanter inte är helt på det klara med vilken typ av ordbok LSL är. Inte heller den kan uppfylla den spridda och lite naiva förväntningen att en ordbok ska innehålla alla ord som finns.

Enligt informanterna saknas många svåra ord i LSL. Som vi vet (se avsnitt 3.1.1) kan svårigheten hos ord ha många förklaringar och det är inte helt enkelt att förutse vilka ord som kommer att orsaka inlärningsproblem för den enskilde inläraren.

### 7.4.3 Informanternas kommentarer om Betydelse

I Studie<sup>11</sup> återfinns åtta positiva, tretton negativa och två neutrala kommentarer om informationskategorin Betydelse (som i kommentarsfältet också kan motsvaras av ord som *förklaring* eller *innebörd*). Studie<sup>07</sup> innehåller inga positiva eller övriga kommentarer utan endast åtta negativa kommentarer om samma informationskategori. Här följer några exempel:

Studie<sup>11</sup>:

- (7:51) Skulle vilja ha mer exempel på möjliga användbara prepositioner med ett ord och vilken innebörd blir om man använder viss preposition med ordet.
- (7:52) Det borde finnas möjlighet att söka betydelse/översättning på ett uttryck, inte bara enstaka ord.
- (7:53) Kan ni försöka skriva exakt ordbetydelse eller exakt synonymer.
- (7:54) Förklaringarna är bra, kort och koncist och därför enkla att förstå.

Studie<sup>07</sup>:

- (7:55) Mer exempel på ordets betydelse och hur man kan använda ordet.
- (7:56) Det saknas beskrivning av naturopat och en förståelig förklaring av rekrytera.
- (7:57) Svårt att hitta ordförklaringar.
- (7:58) Ni har bara lätta ord, t.ex. hund, katt, men människor kan de här orden.. jag behöver förklaring till svårare ord.

Inom ramen för uppdateringen ingick en grundlig revidering av uppslagsordens betydelsedifferentiering. Vidare beskrevs partikelverben mer ingående eftersom de vanligen orsakar stora bekymmer hos inlärare av svenska (se avsnitt 3.1.2.1). Som framkommer av exemplet är också användarnas förväntningar på informationskategorin Betydelse höga. Förväntningarna som uttrycks i t.ex. (7:51) och (7:55) har beaktats. I andra avseenden, som i exempel (7:56), har de inte det. En ordboksredaktör skulle kanske önska att alla användare sällade sig till informanten bakom kommentar (7:54), som tycks vara mycket nöjd.



#### 7.4.4 Informanternas kommentarer om Exempel

I Studie<sup>11</sup> återfinns nio positiva, trettio negativa och tre neutrala kommentarer om informationskategorin exempel. I Studie<sup>07</sup> omnämns kategorin Exempel endast tre gånger och samtliga i en negativ kontext. Här följer ett par citat:

Studie<sup>11</sup>:

- (7:59) För få exempel på ordens bruk. Svårt hitta synonymer, motsatser, ordspråk.
- (7:60) Ni borde ge flera uttryck och kanske några fler exempel vid varje ord eller preposition.
- (7:61) Bra exempelmeningar där man ofta hittar precis den böjning man vill ha.
- (7:62) Bättre att tillägga mer exempel, partikelverb, perfekt particip, ordspråk, slangord som man måste lära sig ock använda vardagligen.

Studie<sup>07</sup>:

- (7:63) ha gärna minst 2-3 exempel till varje ord så man förstår 100 %

Som en följd av den mer explicita betydelseuppdelningen i samband med uppdateringen av LSL placerades också ordbokens språkexempel i närmare anslutning till den specifika betydelse som illustreras. Revideringen inbegrep också att många morfologiska exempel lades till i form av genomskinliga sammansättningar och fler, och mer utbyggda, syntaktiska språkprov. I den nya versionen av LSL finns det i genomsnitt ett exempel per uppslagsord vilket är en klar förbättring i jämförelse med LSL före uppdateringen. Exempelen har dessutom byggts ut till fullständiga meningar i syfte att tydligare illustrera innebörden av ordet (se vidare avsnitt 3.2.3).

Trots revideringen är de allra flesta omdömen om språkproven i Studie<sup>11</sup> negativa. Det skulle vara mycket intressant att genomföra mera djupgående intervjuer om användares uppfattning om LSL som inlärningsordbok. Det skulle också vara spännande att följa brukare i konkreta användningssituationer där de får förklara hur de uttolkar den information som finns i ordboksartiklarna. Det förefaller t.ex. som om informanten bakom kommentaren (7:61) vänder sig till exemplen för att hitta rätt böjningsform. Är det så hen brukar gå tillväga? Exempel är kanske inte den kategori lexikografer tänker sig att användarna i första hand konsulterar för att finna information om böjning.

### 7.4.5 Sammanfattning

De uppgifter som framkommer i kommentarerna kan som sagt ge information om vilka lexikografiska behov informanterna anser sig ha och kan också erbjuda lexikografer konkreta idéer om hur ordböcker kan förbättras. Eftersom vi här har möjlighet att jämföra informanternas synpunkter på LSL före och efter uppdateringen kan vi också få indikationer på om den uppdaterade LSL upplevs som en bättre ordbok. Det finns tydliga indikationer på att så är fallet. För det första är antalet positiva kommentarer betydligt högre efter uppdateringen, dvs. i Studie<sup>11</sup>. För det andra omnämns de specifika informationskategorierna i mer positiva ordalag i det senare materialet. Därtill vet vi att majoriteten av informanterna i Studie<sup>11</sup> tillhör LSL:s primära målgrupp, vilket inte är fallet i Studie<sup>07</sup>. Det verkar alltså som att de informanter som är mer representativa med tanke på ordbokens målgrupp också är mer positiva till LSL.

Trots att informanterna verkar mer nöjda med den uppdaterade versionen av LSL innehåller de två uppsättningarna kommentarer i stort sett samma typ av kritik. Kanske är det så att man aldrig kan få för mycket av det goda eller som en informant konstaterar: ”Jag tycker att ni kan göra den bättre, men den är inte dålig”. Vidare är Lemma den informationskategori som informanterna har flest synpunkter på, medan de övriga kategorierna omnämns i mycket mindre utsträckning.

Vidare framgår av vissa kommentarer, mer eller mindre explicit, en medvetenhet om vilken typ av ordbok informanterna för närvarande använder. Det finns emellertid också kommentarer som klart och tydligt avslöjar informanternas brist på insikt om vilken typ av ordbok LSL är. Å ena sidan utgör det här kanske inget problem, eftersom LSL är gratis att använda; de som inte är nöjda med ordboken kan alltid vända sig till ett annat referensverk. Å andra sidan kanske syftet med LSL, och hela Lexin-serien för den delen, borde förtydligas på webbsidan. Åtminstone borde bilderna och filmerna lyftas fram eftersom definitioner i sig själva inte alltid är det bästa sättet att förmedla ords betydelse på, speciellt inte för andraspråksinlärare. Bilder och filmer kan däremot illustrera betydelsen av ett ord på ett bättre sätt (Lew 2010, jfr dock Franker 2013:782–784). Förvånande nog omnämns bilder och filmer inte mer än vid en handfull tillfällen bland kommentarerna. Det behöver inte betyda att dessa informationskategorier är obekanta för informanterna, men den tolkningen verkar mest rimlig eftersom bild och film också hamnar längst ner i tabellen över flest eftersökta informationskategorier (se tabell 7:6 och 7:7).

Informanternas kommentarer kan innehålla intressanta upplysningar om deras subjektiva, användningsrelaterade behov, dvs. de behov användarna själva anser sig ha. Dessa kan komma den praktiska lexikografen till godo.

Enligt funktionsläran är användarna emellertid sällan medvetna om sitt eget bästa, dvs. sina objektiva behov, och det blir lexikografens ansvar att uttolka vilka dessa är för att sedan tillfredsställa dem i ordboken (se avsnitt 2.3.3).

Slutligen kan kommentarerna avslöja något om informanternas faktiska nivå av svenskkunskaper. De flesta informanter kan ge uttryck för sina åsikter i skrift, även om det inte alltid sker på idiomatisk svenska. Det kan ju också vara så att de som har sämre kunskaper i svenska väljer att inte skriva något alls.



## 8. Delstudie III: Målgruppsanvändning

I detta tredje resultatkapitel redovisas utfallet av länkningen mellan loggfilsdata och enkätdata i 2011-materialet. Länkningen är en typ av datatriangulering och innebär i det här fallet att insamlingen av enkäter och registreringen av loggfiler har synkroniserats. Dessa har sammanlänkats i syfte att analysera enskilda informanternas aktiviteter på LSL:s webbsida (se avsnitt 5.3 och 5.6). Fokus ligger på de informanter som 1) har uppgett att de har något av Lexin-språken som sitt modersmål, 2) har anlänt till Sverige någon gång under 2000-talet och 3) uppger att de använder LSL varje dag. Sammanlagt 249 informanter uppfyller dessa kriterier och kallas i det följande för ”målgruppen” (jfr definitionen av Lexins målgrupp enligt Beijer & Fre Woldu 1997:4, 8).

Den metod som används i delstudie 3 är som redan konstaterats av mer tentativ karaktär (se avsnitt 5.3). På ett övergripande plan syftar därför gjorda iakttagelser till att ge svar på den här metodens eventuella nytta. Efter en översiktlig beskrivning av målgruppen riktas uppmärksamheten mot några enskilda informanter och deras tillhörande ordboksaktiviteter.

### 8.1 Beskrivning av målgruppen

De 249 informanterna motsvaras alltså av lika många länkningar. Dessa innehåller följaktligen dels enkätdata, dels alla de loggfilsaktiviteter som kan knytas till respektive enkät-id. I det här avsnittet ges en kortare beskrivning av de 249 informanterna med fokus på urvalskriterierna, dvs. modersmål, ankomstår och sökfrequens, medan innehållet i loggfilerna endast berörs kort.

Internetversionen av Lexin innehåller, vid sidan av LSL, 18 tvåspråkiga lexikon. Bland de 249 målgruppsanvändarnas modersmål återfinns 16 Lexin-språk. Allra flest representanter har arabiska, 64 stycken. Därefter kommer följande språk, i fallande ordning: persiska (45), spanska (31), somaliska (28), kurdiska (nord+syd) (21), ryska (16), grekiska (10), finska, serbiska och turkiska (5), kroatiska (3), albanska, azerbajdzjanska och bosniska (2) samt amhariska (1). De språk som har störst antal modersmålstalare är också de språk vars lexikon har flest uppslagningar under den månad som datainsamlingen genomförs (jfr tabell 7:10 i avsnitt 7.3.3). De två Lexin-språk som ingen av målgruppsinformerarna uppger är pashto och tigrinska. Lexikonerna från svenska till pashto och tigrinska utgör också (vid undersökningstillfället) två av de allra senaste tillskotten i Lexin.

Samtliga målgruppsinformerare har som sagt kommit till Sverige under 2000-talet. Majoriteten av dem visar sig ha angett ett ankomstår mellan 2005 och 2009 (jfr figur 7:4 i avsnitt 7.3.1).

Vidare anser sig samtliga informanter i målgruppen vara frekventa användare av LSL. En analys av loggfilsdatan visar emellertid att det verkar vara en sanning med modifikation, i alla fall när det gäller användningen av LSL och internetversionen. Nätt och jämnt en tredjedel använder LSL frekvent (fyra dagar eller fler under den aktuella veckan). Majoriteten av informanterna har istället använt LSL *färre* än fyra dagar under de sju dagar som materialet samlades in. Vid en närmare granskning av målgruppens användning framträder emellertid en något annorlunda bild. I det loggfilsunderlag som innehåller information om användarnas klickningar på länkar till uppläst uttal, bilder och filmer får man nämligen indirekt även upplysningar om informanternas (eventuella) användning av de tvåspråkiga Lexin-lexikonerna. Det visar sig då att målgruppen är frekventare användare än vad som först verkar vara fallet. Förklaringen till att informanterna har uppgett att de använder LSL dagligen kan vara att de i själva verket har svarat på frågan om hur ofta de använder Lexin i allmänhet, och inte LSL i synnerhet. En annan mycket tänkbar förklaring är förstås att LSL används från andra ip-adresser än den som är kopplad till deras enkät, t.ex. från en skola, en arbetsplats, på ett kafé eller på tåget till och från skola eller jobb. Bärbara datorer, för att inte säga smarttelefoner och surfplattor, kan ju enkelt användas i alla möjliga situationer och på många platser (jfr Holmer et al. 2015). Det är ju även fullt möjligt att informanterna har tillgång till en tryckt version av LSL (upplaga 1–3).

Loggfilerna avslöjar alltså att målgruppens användning av bilder och filmer oftare sker i samband med uppslagningar i de tvåspråkiga lexikonerna. Bilder och filmer används vidare av i stort sett alla målgruppsinformerare, oberoende av modersmål och ankomstår. De informanter som anger ett sent ankomstår använder dock länkarna till bilder och filmer i större utsträckning än de in-

formanter som har kommit till Sverige i början 2000-talet (jfr avsnitt 6.2 om navigeringsval).

Loggfilerna ger också upplysningar om huruvida informanterna ändrat i inställningarna. Denna aktivitet kallas för ”checking”. Drygt en tredjedel av målgruppen klickar på knappen ”Inställningar”. Som tidigare framgått kryssar informanterna i regel för samtliga fem informationskategorier som vid tidpunkten för datainsamlingen hade avaktiverats, dvs. böjning, exempel, konstruktioner, sammansättningar och uttryck (se vidare avsnitt 6:2).

I de följande fem avsnitten redovisas näranalyser av enskilda informanter och deras användning av LSL. Jag har medvetet valt ut informanter som har olika modersmål, är av olika kön och i varierande ålder. Därtill har jag försökt finna länkningar som kan generera belysande exempel och intressanta iakttagelser. En länkning kan innehålla allt ifrån några enstaka aktiviteter till hundratal sådana. De länkningar som slutligen valdes ut fick därför heller inte innehålla för många aktiviteter eftersom de skulle bli svårt att på ett läsvänligt sätt redovisa dem (se vidare avsnitt 5.6).

Till skillnad från i tidigare resultatkapitel redogörs här först för informanternas enkätsvar och därefter för innehållet i loggfilerna. Inledningsvis ges således en personbeskrivning utifrån informanternas enkätsvar. Därefter presenteras innehållet i loggfilerna. Först fokuserar jag på datum och tidpunkt för användningen, sedan redogörs för vilka typer av aktiviteter som utförs och slutligen ges en närmare analys av själva aktiviteterna och deras förlopp.

## 8.2 Informant 1: Mervan

### 8.2.1 Mervans enkätsvar

Informanten är en kurdisk man född 1977. Han är alltså 34 år vid tidpunkten för enkätundersökningen. Vi kan kalla honom Mervan. Mervan kom till Sverige 2004. Hans modersmål är kurdiska, vilket tillsammans med turkiska är det språk han säger sig vara bäst på. Han uppger vidare att han talar kurdiska med sin familj och kurdiska, turkiska och svenska med sina vänner. I skolan eller på jobbet använder han svenska. Han talar eller använder svenska varje dag. Mervan har genomgått sfi-utbildning, men har ingen annan utbildning i svenska språket. Han bedömer sina kunskaper i svenska som ganska goda. Mervan har en universitetsutbildning från sitt hemland och i Sverige har han läst (sfi) på Komvux. På frågan om sin huvudsakliga sysselsättning svarar han att han arbetar.

Vidare uppger Mervan att han använder LSL varje dag och han kryssar för alla de preciserade alternativen för i vilket syfte han använder ordboken: förstå ett ord i en text som han läser, få hjälp med ett ord när han skriver, få hjälp med ett ord när han ska översätta, få hjälp med ett ord han hört. Han använder alltså LSL både i situationer av reception och av produktion. På frågan om vilka informationskategorier han söker mest efter i LSL kryssar han för betydelse, motsatser och exempel. Länkarna till bilder och filmer uppger Mervan att han använder ibland. Han vet inte om han någon gång läst instruktionerna till hur man använder Lexin. På frågan om hur han löser språkliga problem om han inte använder LSL klickar han för svarsalternativen "Använder en annan ordbok på Internet", "Frågar någon" eller använder "Google eller någon annan sökmotor". Slutligen har Mervan skrivit en kommentar på 51 ord. Den lyder:

Jag tycker att det är jätte bra hemsida och det hjälper mycket till mig. Det funkar som min assistent på jobbet. Dessutom får jag hjälp att lära mig mitt modersspråk(Kurdiska) också som var förbjudet i mitt hemland med hjälp av sidan. Jag tackar jätte mycket till er. Med vänliga hälsningar.

Mervan är således en mycket nöjd användare av LSL. Av kommentaren framgår också att det svensk-kurdiska lexikonet används för att han ska lära sig modersmålet och alltså inte enbart som hjälp till att lära sig svenska språket, vilket ju är ordbokens primära syfte. Som redan konstaterats uppger Mervan att han har ganska goda svenskkunskaper och kommentaren ger ett kvitto på att hans bedömning är korrekt.

Slutligen uppger Mervan sin mejladress vilket de informanter ombads göra som kunde tänka sig att delta i fler undersökningar av Lexin.

Så långt vet vi vad Mervan uppger om sig själv och om sin användning av LSL. Vi ska nu se närmare på vilka faktiska aktiviteter han utför i ordboken.

## 8.2.2 Mervans aktiviteter i LSL

Det har redan framgått att det finns ett antal typer av aktiviteter som kan utföras på LSL:s webbsida. De aktiviteter som har registrerats under datainsamlingen är som sagt dels lexikonuppslagningar, dels navigeringsval där fokus lagts på klickningar på länkarna till "Lyssna", "Bild svenska" och "Visa film", ändringar i inställningarna samt klickningar på knappen med ett frågetecken. I kommande uppställningar redovisas även frågeformuläret som en egen aktivitetstyp.

I tabell 8:1 visas samtliga aktiviteter i LSL som genom ip-numret kan knytas till Mervans frågeformulär (jfr tabell 5:6 i avsnitt 5.6). I följande framställning antar vi för enkelhetens skull att det är Mervan som står bakom samtliga ak-



tiviteter knutna till detta nummer under den aktuella insamlingsperioden. Detsamma gäller för de andra informanterna som näranalyseras. Det är dock som redan framgått inte helt opproblematiskt att göra detta antagande (se avsnitt 5.3.2 om metodologiska överväganden).

TABELL 8:1. *Aktiviteter knutna till Mervan, enkät-id 2011:287.*

Rad	Datum	Tidpunkt	Aktivitetstyp	Aktivitet
1	3	09:29:00	Checking	Exempel
2	3	09:29:04	Checking	Konstruktioner
3	3	09:29:06	Checking	Sammansättningar
4	3	09:29:07	Checking	Uttryck
5	3	09:29:09	Checking	Böjning
6	3	09:29:46	Träff	åtaganden
7	3	09:29:55	Frågeformulär	2011:287
8	3	09:32:03	Träff	sammansättningar
9	4	11:39:08	Träff	instruktioner
10	4	12:13:37	Korr F 3	mail → mil, mal, mall
11	4	12:13:39	Träff	mall
12	7	11:06:42	Korr A	kontroversionell → kontroversiell

### 8.2.2.1 Datum och tidpunkt för användningen

I frågeformuläret uppger Mervan att han använder LSL varje dag. Av tabell 8:1 framgår emellertid att han just den här veckan utför relativt få aktiviteter över huvud taget och dessutom under endast tre av veckans sju dagar: måndag (3 oktober), tisdag (4 oktober) och fredag (7 oktober). De allra flesta aktiviteterna utförs under måndagen under ett par minuter kring halvtiotiden på morgonen. De övriga aktiviteterna utförs någon gång mellan 11.00 och 12.00 mitt på dagen tisdag och fredag. Det kan ju vara så att han de övriga dagarna använder det svensk-kurdiska lexikonet (jfr Mervans enkätkommentar). Det är förstås också fullt möjligt att han de övriga dagarna använder LSL från en annan ip-adress och kanske också från en annan typ av elektronisk enhet, t.ex. sin smarttelefon eller surfplatta. Möjligheten finns ju även att Mervan använder ett tryckt Lexin-lexikon, men den tolkningen förefaller mindre sannolik, eftersom han säger sig inte använda ordböcker i det formatet.

Att de aktiviteter som registreras under måndagen den 3 oktober kan föras till Mervan är det ingen tvekan om. De sammanfaller nämligen med ifyllande av enkäten (se rad 7). Eftersom Mervan uppger att han använder LSL varje dag är sannolikheten också stor att de övriga aktiviteterna härrör från honom.

Antalet aktiviteter är dessutom litet vilket också det talar för att få enskilda individer har använt LSL den aktuella veckan. Det som talar emot att endast Mervan står bakom aktiviteterna är möjligtvis tidpunkten de utförs, dvs. mitt på dagen och på en vardag. Han skulle lika gärna kunna ha använt en allmän dator tillgänglig för många människor. Det kan förstås också vara så att Mervan bor i ett hushåll med flera personer som använder LSL.

### 8.2.2.2 Typer av aktiviteter

Av tabell 8:1 framgår att under den aktuella insamlingsveckan har tolv aktiviteter utförts varav åtta infaller under måndagen, tre på tisdagen och endast en på fredagen. Aktiviteterna från tabell 8:1 fördelar sig på de olika typerna som framgår av tabell 8:2.

TABELL 8:2. *Typ av aktivitet och dess frekvens.*

Typ av aktivitet	Antal
Frågeformulär	1
Lexikonuppslagningar	6
- varav träff	4
- varav korr A	1
- varav korr F	1
- varav icke-träff	
Checking	5
Lyssna	
Bild svenska	
Visa film	
<b>Totalt</b>	<b>12</b>

Frågeformuläret kunde endast skickas in en gång och dess antal är alltid ett (1).

Av de sex lexikonuppslagningarna är fyra rättstavade ord som finns i LSL och som alltså ger träff i ordboken. De övriga två lexikonuppslagningarna utgörs av automatkorrigering och en korrigering med förslag.

De fem aktiviteterna av checking i tabell 8:2 innebär att Mervan har gått in i inställningarna och kryssat för fem informationskategorier.

I enkäten uppgav Mervan att han ibland tittar på bilder och filmer. Det gör han emellertid inte vid något tillfälle under den aktuella veckan.

I avsnitt 8.2.2.3 nedan beskrivs Mervans aktiviteter mer ingående.

### 8.2.2.3 Aktiviteternas förlopp

Måndagen den 3 oktober utför Mervan åtta aktiviteter under loppet av ca tre minuter, från kl. 09:29:00 till 09:32:03. De fem första aktiviteterna gör han i inställningarna där de fem avaktiverade informationskategorierna kryssas för, i följande ordning: Exempel, Konstruktioner, Sammansättningar, Uttryck och Böjning (rad 1–5). Han förefaller ha läst informationen på LSL:s förstasida om den pågående undersökningen och vad man behöver göra för att få tillbaka de fem informationskategorierna som tillfälligt avaktiverats. Därefter utför han sin första uppslagning som efterfrågar substantivet *åtaganden* (rad 6). I och med denna sökning erbjuds han att fylla i frågeformuläret, vilket är ett erbjudande han antar (rad 7). Formuläret fylls i under de följande dryga två minuterna. Sedan sker nästa uppslagning som leder till träff. Det är intressant nog ett ord som förekommer två gånger i frågeformuläret, nämligen substantivet *sammansättningar* (rad 8). Mervan har i enkäten kryssat för att sammansättningar är en av de informationskategorier som han ofta söker information om. Det kan därmed tyckas märkligt att han behöver slå upp detta ord. Kanske vill han bara bekräfta att han har förstått ordet på rätt sätt. De uppgifter om uppslagsordet *sammansättning* som han möts av i LSL torde i varje fall vara tillfredsställande, speciellt med tanke på att han har återställt artiklarna i sitt standardläge och alltså kan läsa den definition som ingår i ordboksartikeln: 'ord som består av två eller flera ord'. Motsvarande språkexempel som ges i artikeln lyder: "sammansättningar som 'trähus' och 'blåbärspaj'".

Tisdagen den 4 oktober utförs tre uppslagningar varav två leder till träff och en till korrigeringsförslag. Först slås substantivet *instruktioner* upp, vilket ger en träff (rad 9). Drygt 33 minuter senare slås substantivet *mail* upp vilket beklagligt nog inte ger träff i LSL (rad 10). Ordet finns nämligen upptaget i ordboken men bara med stavningen *mejl*. De tre alternativa förslagen på befintliga uppslagsord som ges är *mil*, *mal* och *mall*. Mervan väljer då att slå upp det sista av de tre förslagen, *mall* (rad 11). Artikeln *mall* ger dock ingen information om det ord han troligen är ute efter.

Fredagen den 7 oktober utförs en enda sökning under dagen och det är den sista för hela veckan. Då slår Mervan upp ordet *kontroversionell*, vilket automatiskt korrigeras till det rättstavade adjektivet *kontroversiell* (rad 12).

### 8.2.2.4 Kommentar

Mervans enkät är väl ifyllt men antalet aktiviteter under den undersökta veckan är litet. Det är därför svårt att göra några djupgående analyser av Mervans ordboksanvändning. Några intressanta iakttagelser kan dock göras.

Att de flesta uppslagningarna utgörs av rättstavade ord kan tyda på att de i normalfallet är ord som Mervan har stött på i en text och inte ord som han har hört och därefter försöker återge i skrift. Tre av uppslagningarna står dessutom i pluralform, vilket också kan tyda på att han stött på orden i skriftliga sammanhang.

Vidare återfinns de uppslagningar som Mervan får träff på inte bland de mest sökta lemmarna LSL (se tabell 6:5 i avsnitt 6.1.1.2). Det första ordet som påträffas är *åtagande* på placering 1 058 och av de övriga orden finner vi t.ex. *sammansättning* på plats 1 417 och *instruktion* på plats 2 673. Enligt Kelly-listan (se avsnitt 6.1.1.2) återfinns Mervans uppslagningar i spannet 3–4 000 av de vanligaste orden i svenskan.

Orden tillhör alltså inte de allra mest grundläggande svenska ordförrådet. Det är inte så överraskande med tanke på att Mervan har varit i Sverige i sju år och dessutom har genomgått sfi-utbildning. Kanske är det mer överraskande att han faktiskt fortfarande verkar tycka att Lexin tillfredsställer hans lexikografiska behov.

Det är också intressant att Mervan använder Lexins svensk-kurdiska lexikon för att lära sig kurdiska. Digitaliseringen av Lexin-ordböckerna gjorde, som tidigare beskrivits (se avsnitt 3.2.1), det möjligt att använda de tvåspråkiga ordböckerna som L1-L2-ordböcker. I Mervans fall innebär det att det svensk-kurdiska lexikonet inte bara kan användas (av kurdisktalande) för reception av svenska utan också för produktion av kurdiska (jfr också avsnitt 2.3.2).

I nästa avsnitt beskrivs en kvinnlig informant och hennes användning av LSL. Hennes aktiviteter är fler och i materialet som är knutet till henne finns också fler intressanta iakttagelser att göra.

## 8.3 Informant 2: Elira

### 8.3.1 Eliras enkätsvar

Nästa informant är en albansk kvinna född 1986. Hon är alltså 25 år vid tidpunkten för datainsamlingen. I det följande kallar jag henne Elira. Elira kom till Sverige så sent som 2010. Hennes modersmål är albanska, men hon uppger att hennes bästa språk är engelska. Albanska använder hon i kontakt med sin familj och med sina vänner, men i skolan talar hon engelska. Svenska talar eller använder hon varje dag och hon bedömer sina kunskaper i svenska som ganska goda. På frågan om vilken hennes huvudsakliga sysselsättning är svarar Elira studerande. Hon har däremot inte genomgått sfi-utbildning eller någon annan

utbildning i svenska språket. Eftersom hon kom till Sverige så sent som 2010 har hon kanske ännu inte hunnit få uppehållstillstånd och i så fall har hon heller inte kunnat delta i någon sfi-utbildning. Kanhända sker inläringen av svenska helt på egen hand.

På frågan om högsta avslutade utbildning i Sverige kryssar Elira för alternativet svensk universitetsutbildning. På frågan om högsta avslutade utbildning i hemlandet anger hon emellertid samma svar. Eftersom Elira nyss kommit till Sverige är det rimligen så att det är den utländska utbildningen som åsyftas i båda fallen. Det framgår inte exakt vilken utbildning det är hon har.

Elira kryssar för alla de preciserade alternativen för i vilket syfte hon använder ordboken: förstå ett ord i en text som hon läser, få hjälp med ett ord när hon skriver, få hjälp med ett ord när hon ska översätta en text och få hjälp med ett ord som hon hört. Hon använder således LSL för både reception och produktion. Elira uppger vidare att hon oftast är ute efter information om uttal, betydelse, exempel och konstruktioner. Elira ger emellanåt motsägelsefulla uppgifter och en bidragande orsak till det är sannolikt begränsade kunskaper i svenska språket. Vidare uppger Elira att hon ibland klickar på länkarna till ”Bild svenska”, men att hon aldrig klickar på länkarna till ”Visa film”. Hon är osäker på om hon har läst instruktionerna till Lexin. På frågan om hur hon löser språkliga problem om hon inte använder LSL svarar Elira att hon använder en annan ordbok på internet eller att hon frågar någon.

Elira har skrivit en kommentar på engelska innehållande 38 ord där hon uttrycker sin uppskattning för LSL:s omfattrika innehåll. Kommentaren lyder:

I really appreciate the richness this site offers. Not only obtaining a Swedish-Swedish explanation, but also listening part is a lot helpful. I do appreciate the examples and the structure how the word is used in the sentence.

Elira ger uttryck för att hon framför allt uppskattar definitionerna, lyssna-funktionen, språkexemplen och konstruktionsuppgifterna. Intressant nog överensstämmer detta helt och hållet med de fyra informationskategorier hon tidigare i enkäten har uppgett att hon oftast eftersöker.

En informant som Elira med, vad det verkar, mycket goda kunskaper i engelska skulle vara behjälpt av det svensk-engelska lexikonet. Förhoppningsvis har hon hittat fram till ordboken *Folkets lexikon* som i grunden utgörs av Lexins före detta svensk-engelska lexikon. Om man väljer att söka i det svensk-engelska Lexin-lexikonet (som alltså fortfarande ingår i rullistan ”Välj språk”) får man information om, och en länk till, *Folkets lexikon*. Däremot verkar Eliras självskattning av svenskkunskaper, dvs. att de är ganska goda, inte vara helt med sanningen överensstämmande. Bara det faktum att hon väljer att skriva sin kommentar på engelska tyder på det.

Elira uppger liksom Mervan sin mejladress och är alltså öppen för att bli kontaktad i fråga om framtida användarundersökningar.

### 8.3.2 Eliras aktiviteter i LSL

Vi vet nu en del om informanten vi kallar Elira. I tabell 8:3 visas de aktiviteter i LSL som genom ip-numret kan knytas till hennes frågeformulär.

TABELL 8:3. *Aktiviteter knutna till Elira, enkät-id 2011:635.*

Rad	Datum	Tidpunkt	Typ av aktivitet	Aktivitet
1	3	20:54:52	Träff	partnerskap
2	3	20:55:32	Träff	partnerskap
3	4	22:06:36	Checking	Böjning
4	4	22:06:37	Checking	Exempel
5	4	22:06:38	Checking	Konstruktioner
6	4	22:06:38	Checking	Sammansättningar
7	4	22:06:41	Checking	Uttryck
8	4	22:35:49	Träff	vart
9	4	22:36:01	Frågeformulär	2011:635
10	4	22:41:35	Träff	modersmål
11	4	22:50:49	Träff	vart
12	4	22:52:01	Träff	ställe
13	4	22:58:11	Träff	tand
14	4	22:58:24	Träff	tänder
15	4	22:58:25	Träff	tänder
16	4	22:58:45	Träff	tand
17	4	22:59:26	Träff	borsta
18	4	23:05:12	Träff	för sent
19	4	23:05:46	Träff	varandra
20	4	23:07:00	Träff	tänkaren
21	4	23:07:00	Träff	tänkaren
22	4	23:07:10	Checking	Exempel
23	4	23:07:11	Checking	Konstruktioner
24	4	23:07:14	Checking	Sammansättningar
25	4	23:07:17	Checking	Uttryck
26	4	23:07:20	Checking	Visa alltid svenskt tangentbord med å, ä, ö.

27	4	23:07:27	Träff	tänkaren
28	4	23:07:30	Checking	Böjning
29	4	23:07:30	Checking	Exempel
30	4	23:07:31	Checking	Sammansättningar
31	4	23:07:32	Checking	Konstruktioner
32	4	23:07:33	Checking	Uttryck
33	4	23:07:40	Checking	Visa alltid svenskt tangentbord med å, ä, ö.
34	4	23:07:56	Träff	tänka
35	4	23:10:19	Träff	sönder
36	4	23:11:26	Korr F 2	traig → tradig, trasig
37	4	23:11:31	Träff	trasig
38	4	23:12:03	Korr F 2	traig → tradig, trasig
39	4	23:12:05	Träff	sönder
40	4	23:12:58	Träff	knacka
41	4	23:12:58	Träff	knacka
42	4	23:14:52	Träff	stänga
43	4	23:14:54	Lyssna	stänger
44	4	23:15:42	Träff	använda
45	4	23:16:40	Träff	begagnar
46	4	23:17:23	Träff	använder
47	4	23:17:35	Träff	anspråk
48	4	23:18:19	Träff	använder
49	4	23:20:35	Träff	hämtar
50	4	23:20:35	Träff	hämtar
51	4	23:21:15	Visa film	hämtar_posten
52	4	23:21:27	Träff	hämtar
53	4	23:21:31	Visa film	hämtar_ut_medicin
54	4	23:21:40	Träff	hämtar
55	4	23:21:42	Visa film	hämtar_ut_paket
56	4	23:21:56	Träff	hämtar
57	4	23:22:07	Checking	Exempel
58	4	23:22:08	Checking	Böjning
59	4	23:22:09	Checking	Konstruktioner
60	4	23:22:11	Checking	Sammansättningar
61	4	23:22:11	Checking	Uttryck
62	4	23:22:14	Checking	Visa alltid svenskt tangentbord med å, ä, ö.
63	4	23:22:19	Träff	hämta

64	4	23:24:05	Träff	inget
65	5	23:09:31	Checking	Böjning
66	5	23:09:32	Checking	Exempel
67	5	23:09:33	Checking	Konstruktioner
68	5	23:09:33	Checking	Sammansättningar
69	5	23:09:36	Checking	Uttryck
70	9	12:13:30	Visa film	gäspar
71	9	20:47:54	Checking	Exempel
72	9	20:47:55	Checking	Böjning
73	9	20:47:56	Checking	Konstruktioner
74	9	20:47:56	Checking	Sammansättningar
75	9	20:47:57	Checking	Uttryck

Som vi ska se visar helhetsbilden av aktiviteterna i tabell 8:3 att de med mycket stor sannolikhet är kopplade till en och samma användare, dvs. Elira. I det följande ges en ingående beskrivning av aktiviteterna i tabell 8:3.

### 8.3.2.1 Datum och tidpunkt för användningen

Av tabellen kan vi utläsa att Elira i huvudsak använder LSL på kvällstid, nämligen någon gång mellan ca 21.00 och 23.30. Ett argument för att det är just Elira som utför alla aktiviteter i tabell 8:3 är den tidsmässiga regelbundenhet med vilken aktiviteterna utförs (jfr Almind 2008 som med hjälp av bland annat tidsangivelser identifierar enskilda användare). Att hon använder LSL på kvällstid talar dessutom för att hon befinner sig i sin hemmiljö. I frågeformuläret uppger hon att hon använder ordboken varje dag, men just den här veckan sker aktiviteter fyra av veckans sju dagar – måndag, tisdag, onsdag och söndag. Det stämmer å andra sidan ganska väl överens med de dagar då flest uppslagningar görs i LSL under den aktuella veckan (se tabell 6:1 i avsnitt 6.1). Som redan konstaterats kan användningen de övriga dagarna ske från en annan ip-adress i en annan miljö eller på en annan elektronisk enhet.

### 8.3.2.2 Typer av aktiviteter

I det följande antar vi som sagt att det är Elira som står bakom de totalt 75 aktiviteter som finns registrerade under den aktuella veckan. Av tabell 8:4 framgår hur aktiviteterna fördelar sig på de olika typerna.



TABELL 8:4. Typ av aktivitet och dess frekvens.

Typ av aktivitet	Antal
Frågeformulär	1
Lexikonuppslagningar	37
- varav träff	35
- varav korr A	
- varav korr F	2
- varav icke-träff	
Checking	32
Lyssna	1
Bild svenska	
Visa film	4
<b>Totalt</b>	<b>75</b>

Av de 37 lexikonuppslagningarna resulterar 35 i träffar. De övriga två är korrigering med förslag (F).

De 32 aktiviteterna av checking i tabell 8:4 fördelar sig på sex tillfällen. Vid fem av tillfällena kryssar Elira för alla de fem informationskategorier som kryssats ur, dvs. Böjning, Exempel, Sammansättningar, Konstruktioner och Uttryck. Vid tre tillfällen kryssar hon dessutom för ”Visa alltid svensk tangentbord med å, ä, ö” vilken inte är förkryssad i grundinställningarna.

I enkäten uppger Elira att hon ibland tittar på bilderna men aldrig på filmerna. Hennes aktiviteter i LSL visar emellertid något annat. Där återfinns inga klickningar på bilder men däremot har Elira vid fyra tillfällen klickat på länken till Visa film. En möjlig förklaring är att hon upptäckt filmernas existens i och med frågan i enkäten. Samtliga filmvisningar görs nämligen *efter* att hon svarat på enkäten.

### 8.3.2.3 Aktiviteternas förlopp

Måndag kväll den 3 oktober utför Elira två aktiviteter. Hon slår upp ordet *partnerskap* två gånger (se tabell 8:3, rad 1 och 2). Inga fler aktiviteter sker denna dag, åtminstone inte i LSL. Den här gången avböjer hon erbjudandet om att delta i enkätundersökningen.

Under tisdagkvällen, den 4 oktober, utför däremot Elira hela 62 av veckans totala 75 aktiviteter. Hon börjar strax efter tio på kvällen och inleder med att kryssa för de fem urkryssade informationskategorierna (rad 3–7). Det verkar alltså som hon har läst och förstått den information som delges användarna av Lexins webbsida under den aktuella veckan. Sedan gör hon ett uppehåll på

drygt 29 minuter innan hon utför nästa aktivitet och slår upp adverbet *vart* (rad 8). Det är den första uppslagningen i LSL denna kväll och erbjudande om att fylla i en enkät poppar upp. Vid detta tillfälle väljer hon att fylla i enkäten och påbörjar denna aktivitet vid klockslaget 22.36.01 (rad 9). Ifyllandet är avslutat i och med att nästa aktivitet utförs, klockan 22.41.35, då hon slår upp ordet *modersmål* (rad 10). Ordet *modersmål* används i frågeformuläret och förklaras där med att det är det första språk informanten lärt sig. Förklaringen i frågeformuläret överensstämmer med definitionen i LSL som lyder 'språk som en person är uppförd med, förstaspråk'. Av enkätsvaret att döma så förstod Elira vad som åsyftades med ordet redan då hon mötte det i enkäten. Hon har nämligen angett det språk som hon senare också uppger att hon använder när hon kommunicerar med sin familj och sina kompisar. Man kan tänka sig att Elira slår upp ordet för att hon vill försäkra sig om att hon har tolkat frågan rätt (jfr tabell 8:1 rad 7–8 och Mervans aktiviteter). Hon kan också ha haft en annan uppfattning om ordets betydelse, men att det i enkäten ändå framgick av sammanhanget vad som efterfrågades.

Efter drygt nio minuter utförs nästa uppslagning och Elira upprepar då sökningen på adverbet *vart* (rad 11). Den följande kvarten utförs tio uppslagningar som resulterar i träffar (rad 12 till 21): *ställe, tand, tänder, tänder, tand, borsta, för sent, varandra, tänkaren, tänkaren*. De två sista uppslagningarna sker alltså på samma ord och utförs också vid exakt samma klockslag, dvs. 23:07:00. Sådana här uppslagningar är sannolikt resultatet av en dubbelklickning. Användarna har således dubbelklickat på "Slå upp"-knappen (endast ett enkelklick behövs) och då har uppslagningen registrerats två gånger i loggfilerna. Man kan också slå upp vilket ord som helst inne i en ordboksartikel genom att dubbelklicka på det. I loggfilerna går det dock inte att avläsa huruvida användaren känner till och har använt denna funktion.

Därefter klickar Elira på knappen för inställningar och kryssar för informationskategorierna Exempel, Konstruktioner, Sammansättningar, Uttryck samt "Visa alltid svenskt tangentbord med å, ä, ö" (raderna 22–26). Det betyder att hon någon gång under den senaste dryga kvarten har lämnat Lexins webbsida och sedan tagit fram den igen. Hennes ändringar i inställningarna har då uppenbarligen inte sparats. Hon slår sedan åter upp ordet *tänkaren* och därefter går hon ännu en gång till inställningarna och kryssar för Böjning, Exempel, Sammansättningar, Konstruktioner, Uttryck och "Visa alltid svenskt tangentbord med å, ä, ö" (raderna 27–33). Det här är mer svårförklarligt. Det innebär att hon någon gång mellan klockslagen 23:07:20 och 23:07:30 (rad 26–28) lämnar och återkommer till webbsidan. Det är en tidsperiod på endast tio sekunder. Det är ju på intet sätt omöjligt, men framstår som underligt. Hon slår sedan upp ordet *tänka* (rad 34) vilket förmodligen är knutet till den föregående uppslagningen *tänkaren*.

Efter dryga två minuter sker sedan en uppslagning på ordet *sönder* (rad 35). Definitionen lyder 'i trasigt tillstånd'. Det förklarar nästa uppslagning, *traig* (rad 36), vilket ger utfallet korrigerig med förslag. Elira skulle här ha kunnat slå upp ordet *trasigt* som står i definitionen genom att dubbelklicka på ordet. Att hon inte gör det antyder att hon inte känner till denna funktion. De två förslagen som ges på söksträngen *traig* är *tradig* och *trasig*. Elira väljer att slå upp *trasig* (rad 37). Efter en halv minut slås återigen *traig* (rad 38) upp vilket ger ett likadant utfall som förra tillfället. Elira väljer då istället att söka på ordet *sönder* igen (rad 39). Efter en knapp minuts uppehåll sker en förmodad dubbelklick på "Slå upp"-knappen (rad 40 och 41) och verbet *knacka* slås upp.

Sedan dröjer det nästan två minuter till nästa aktivitet. Då eftersöker Elira verbet *stänga* följt av en klick på symbolen för uppläst uttal (rad 42 och 43). Efter ytterligare en knapp minut slår Elira upp verbet *använda* (rad 44). Det definieras som 'begagnar, tar i anspråk'. De tre följande uppslagningarna är just *begagnar* följt av *använder*, därefter *anspråk* och sedan *använder* igen (rad 45–48). Definitionen av *använder* utgörs av två synonyma uttryck som faktiskt är svårare än det ord som slås upp. Det visar på svårigheten att förklara redan enkla ord. Definitionen är dock till stor hjälp om Elira eftersöker just en synonym, men om hon önskar få en förklaring till verbet *använder* behövs kanske en beskrivning på enklare svenska (jfr avsnitt 3.1.2.4 om betydelsebeskrivning i ordböcker ur ett andraspråksperspektiv). I *Natur & Kulturs Svenska Ordbok* (2001:28) definieras t.ex. verbet *använda* 'ta och få hjälp av något när man gör något' (se avsnitt 3.2.2 för ett exempel på en artikel från denna ordbok). Det här är också ett exempel på hur viktiga språkproven kan vara för att illustrera ett ords betydelse och användning (jfr avsnitt 3.1.2.5 om språkprov). I LSL4 åtföljs definitionen av ordet *använder* av följande tre språkprov: *Använd alltid bilbältet!, pengarna ska användas till en studieresa och plutonium kan användas som material till en atombomb.*

Efter drygt två minuter sker nästa följd av aktiviteter som verkar höra ihop tematiskt. Först slås verbet *hämtar* upp två gånger, återigen en förmodad dubbelklickning (rad 49 och 50). Därefter klickar Elira på artikelns samtliga tre länkar för Visa film. I de fall användaren har tryckt på en länk till Bild svenska eller Visa film flyttas man till en annan webbsida där man sedan trycker på en "Spela upp"-knapp. För att återgå till "Slå upp"-sidan måste man trycka på pilen ← (dvs. 'tillbaka') längst upp till vänster i webbfönstret. När man backar från en bild- eller filmvisning (se rad 51–56) upprepas en lexikonuppslagning. Det är förklaringen till att det ser ut som att Elira slår upp ordet igen efter att hon har tittat på bilden eller filmen. De tre filmsekvenserna illustrerar verbskeendena 'hämtar posten', 'hämtar ut medicin' och 'hämtar ut paket'.

Som sista aktiviteter 4 oktober går Elira återigen till inställningarna och kryssar för exempel, böjning, konstruktioner, sammansättningar, uttryck samt

”Visa alltid svenskt tangentbord med å, ä, ö” (rad 57–62). Det innebär att hon återigen har lämnat Lexins webbsida och öppnat upp den igen sedan hon senast gjorde ändringar i inställningarna. Därefter sker två uppslagningar, först på ordet *hämta* igen och sedan på pronomenet *inget* (rad 63 och 64).

Dagen efter, onsdag den 5 oktober, är Elira endast inne på LSL:s webbsida och ändrar i inställningarna (rad 65–69). Under torsdag, fredag och lördag sker inga aktiviteter alls i LSL. Som tidigare framgått går uppslagningarna i LSL generellt ner under fredag och lördag för att återigen öka under söndagen (se tabell 6:1).

Söndag den 9 oktober använder så Elira LSL igen. Mitt på dagen, vid klockslaget 12:13:30, klickar hon på länken ”Visa film” under artikeln *gäspar* (rad 70). Hon avviker alltså från sin rutin att använda LSL på kvällstid. Sökningen föregås inte av en uppslagning i LSL, vilket innebär att hon har slagit upp ordet i något av de tvåspråkiga lexikonerna, sannolikt det svensk-albanska. På söndag kväll är Elira återigen tillbaka men de enda aktiviteter som då sker är att hon ändrar inställningarna och kryssar för de fem urkryssade informationskategorierna (rad 71–75). Det är det sista som händer under den aktuella veckan.

#### 8.3.2.4 Kommentar

Elira kom till Sverige så sent som 2010 och torde nyligen ha påbörjat sin svenskinläring. Hon förefaller dock främst använda LSL och inte det svensk-albanska lexikonet, vilket kan tyckas anmärkningsvärt. I enkätkommentaren ger hon t.ex. uttryck för sin uppskattning av de svensk-svenska förklaringarna.

Eliras sökord återfinns långt ner i den frekvenssorterade lemmalistan över lexikonuppslagningarna i LSL. Exempelvis slås orden *knacka*, *för sent*, *tänkaren* upp sammanlagt nio gånger i hela 2011-materialet, varav sex gånger av Elira. Med tanke på att hon nyligen kommit till Sverige och verkar ha mycket begränsade kunskaper i svenska språket (tvärt emot hennes självskattning) är det måhända anmärkningsvärt att hon till exempel inte slår upp några av de största verben i svenskan som till exempel *gå*, *ha* och *ta*. De verb hon slår upp är istället följande: *borsta*, *tänka*, *knacka*, *stänga*, *använda*, *begagna*, *hämta* och *gäspar*. Några av dem uppfattar jag som vanligare i tal än i skrift, t.ex. *borsta* och *gäspar*. Bara ett av orden Elira söker på påträffas i närheten av de 100 vanligaste sökningarna i 2011-materialet, nämligen verbet *använda*. Det återfinns på plats 102. I övrigt sticker verbet *begagna* ut i detta sällskap av verb. Det är högtravande och alderdomligt. Det ingår dock i definitionen av verbet *använder* och slås också upp i anslutning till sökningen på detta verb.

I en avhandlingstext finns mycket litet utrymme för spekulationer. Med detta sagt kan man föreställa sig att Elira håller på och läser en text på svenska

och under tiden som hon stöter på ord hon inte förstår slår hon upp dem i LSL. Uppslagningarna sker nämligen tidsmässigt i olika sjok. Orden som slås upp under de olika tidssjoken hör samman semantiskt. Först har vi sjoket med sökningarna *tand, tänder, tänder, tand, borsta* (rad 13–17). Sedan följer uppslagningarna *sönder, traig, trasig, traig, sönder* (rad 35–39) och därefter *använda, begagnar, använder, anspråk, använder* (rad 44–48). Sist har vi sökningarna på *hämtar* tillsammans med filmvisningarna av verbfraserna *hämtar posten, hämtar ut medicin* och *hämtar ut paket* (rad 49–56). Innan den inledande uppslagningen och efter den sista uppslagningen i respektive söksession äger ett sökuppehåll rum på en till sex minuter. De här söksessionerna utgör samttaget alltså också ett argument för att det är Elira som står bakom samtliga aktiviteter i tabell 8:3.

## 8.4 Informant 3: Luciana

### 8.4.1 Lucianas enkätsvar

Nästa informant är en spansktalande flicka som vid tidpunkten för enkätundersökningen är 15 år. Hädanefter kallas hon Luciana. Luciana kom till Sverige 2007 och var då 11 år gammal. Hennes modersmål är spanska och det är också det språk hon uppger att hon är bäst på. Hon talar spanska med sin familj, svenska och (vid behov) engelska med sina vänner. Med skolkamrater pratar hon svenska. Hon använder svenska varje dag och bedömer sina kunskaper i språket som ganska goda. Luciana har inte gått någon sfi-utbildning, eftersom hon har läst svenska inom ramen för grundskolan. Av naturliga skäl har hon ännu inte någon avslutad utbildning, varken i sitt hemland eller i Sverige och hennes huvudsakliga sysselsättning är studerande. Till frågan om utbildning i hemland har Luciana lämnat följande kommentar:

Jag började grundskolan där men utbildningen avbröts ungefär vid början årskurs fyra, då jag flyttade till ön Curaçao.

Adverbet ”där” i kommentaren torde syfta på hemlandet. Eftersom ön Curaçao ligger i Västindien kommer Luciana gissningsvis ursprungligen från ett spansktalande land i Sydamerika.

Luciana uppger att hon främst använder LSL för att förstå ett ord i en text hon läser eller för att få hjälp med ett ord när hon ska översätta en text. På frågan om vilka informationskategorier hon mest söker efter i LSL kryssar hon

för böjning, exempel och konstruktioner. De är typiskt produktionsrelaterade kategorier. I de två påföljande frågorna om just länkarna till bilder och filmer uppger hon att hon inte brukar klicka på dem. Hon har heller inte läst instruktionerna till Lexin. På frågan hur hon löser språkliga problem om hon inte använder LSL svarar hon att hon använder en annan ordbok på internet eller en tryckt ordbok, eller så frågar hon någon. Luciana har lämnat en längre kommentar som lyder på följande sätt:

Jag tycker väldigt mycket om hur detta lexikon ser ut i dag och är tacksam över att det finns för det är här jag söker de flesta sakerna jag vill veta... Det skulle såklart kunna utvidgas, och därmed förbättras, men det är ingenting som jag ser som en nackdel hos lexikonet.

Luciana är en nöjd, och som det verkar, van användare av LSL. Av kommentaren att döma kan man sluta sig till att Luciana behärskar svenska språket väl. Hon skriver mycket bra. Även Luciana har lämnat uppgift om mejladress.

## 8.4.2 Lucianas aktiviteter i LSL

Vi går nu över till de aktiviteter som via en ip-adress kan knytas till Lucianas frågeformulär. De visas i tabell 8:5.

TABELL 8:5. *Aktiviteter knutna till Luciana, enkät-id 2011:839.*

Rad	Datum	Tidpunkt	Aktivitetstyp	Aktivitet
1	3	17:38:12	Träff	biff
2	3	17:38:18	Träff	fil
3	3	20:16:42	Träff	ledsen
4	3	20:49:43	Träff	oreda
5	6	20:25:53	Träff	kulspruta
6	6	20:25:59	Frågeformulär	2011:839
7	6	20:33:41	Träff	lexikon
8	6	20:37:11	Träff	kulspruta
9	6	21:41:13	Korr F 21	återta → tårta, tenta, testa, texta, åtta, åter, åter-, -tårta, alerta, streta, starta, störta, överta, återfå, återge, återgå, återse, återstå, återtåg, återval, återgav
10	6	21:43:36	Träff	dust

Aktiviteterna under den aktuella veckan är som synes få. Det finns ändå underlag för intressanta iakttagelser. I det följande beskrivs aktiviteterna mer ingående.

#### 8.4.2.1 Datum och tidpunkt för användningen

Luciana använder precis som Elira LSL under kvällstid, vilket är en god indikation på att det är just Luciana och ingen annan som står bakom aktiviteterna knutna till hennes frågeformulär. Hon befinner sig sannolikt framför sin dator i sitt hem.

I enkäten uppger Luciana att hon använder LSL varje dag. Under den aktuella veckan registreras dock aktiviteter endast från två dagar, måndagen 3 oktober och torsdagen 6 oktober. Som redan konstaterats kan den eventuella användningen de övriga dagarna ske under andra förhållanden.

#### 8.4.2.2 Typer av aktiviteter

De olika typerna av aktiviteter och deras antal redovisas i tabell 8:6.

TABELL 8:6. *Typ av aktivitet och dess frekvens.*

Typ av aktivitet	Antal
Frågeformulär	1
Lexikonuppslagningar	9
- varav träff	8
- varav korr A	
- varav korr F	1
- varav icke-träff	
Checking	
Lyssna	
Bild svenska	
Visa film	
<b>Totalt</b>	<b>10</b>

Aktiviteterna är som sagt få och översikten kan i så måtto verka överflödig. Luciana ändrar inte i inställningarna och använder sig varken av lyssna-funktionen eller av länkarna till bild och film. De uppgifter hon lämnar i frågeformuläret stämmer överens med vad vi ser i loggfilerna vad gäller de informationskategorier hon säger sig *aldrig* eftersöka information om (jfr fråga 8 i bilaga 2). Av de nio uppslagningarna i LSL ger åtta träff. Även om uppslag-

ningarna är få så kan de kanske kopplas till Lucianas positiva omdöme om LSL som hon ger uttryck för i kommentarsfältet. Hon får nämligen träff på de uppslagningar hon gör.

#### 8.4.2.3 Aktiviteternas förlopp

Måndagen den 3 oktober gör Luciana sina två första uppslagningar. Vid klockslaget 17:38:12 slår hon upp substantivet *biff* (rad 1). I och med det kommer erbjudandet om att delta i en enkätundersökning upp på skärmen. Vid det här tillfället avstår emellertid Luciana från att delta. Istället utför hon sin nästa uppslagning på substantivet *fil* (rad 2). Därefter dröjer det två timmar och 38 minuter till nästa uppslagning som gäller adjektivet *ledsen* (rad 3) och efter ytterligare 33 minuter slår Luciana upp substantivet *oreda* (rad 4). Det är den fjärde och sista aktiviteten den dagen.

Nästa aktivitet i LSL dröjer till torsdagen den 4 oktober kl. 20:33:41 då Luciana slår upp substantivet *kulspruta*. Då dyker återigen möjligheten att svara på frågeformuläret upp och den här gången antar Luciana erbjudandet (rad 6). Ifyllandet är avslutat i och med att nästa aktivitet utförs. Då slår Luciana upp substantivet *lexikon*. Ordet används vid ett flertal tillfällen i enkäten och inte minst i själva namnet på ordboken vars användning undersöks, dvs. *Lexins svenska lexikon*. Luciana är inte ensam om att som första aktivitet efter att enkäten fyllts i slå upp ett eller flera ord som förekommer i frågeformuläret. Det har t.ex. även Mervan och Elira gjort. I de fall som undersökts har emellertid ingen informant misstolkat någon av frågorna utan svarar med den eller de uppgifter som faktiskt efterfrågas. Även i detta fall kan man alltså tänka sig att informanten är ute efter att få bekräftelse på något de i stort sett redan kan.

Följande uppslagning sker på verbet *återta* (rad 9). Det är ett ord som inte är upptaget i LSL. Luciana möts här av hela 21 förslag på ord som hon kan ha menat, t.ex. *tårta*, *tenta* och *testa*. Inget av dem slås upp. Förslagen är som tidigare beskrivits baserade på editeringsavstånd, vilket alltså innebär att de tas fram genom att bokstäver i den ursprungliga söksträngen tas bort, läggs till eller flyttas om (se avsnitt 3.2.4.2). Förutom det stora antalet förslag som ges kan man också ha synpunkter på den ordning som de presenteras i. Man skulle önska att de ord som betydelsemässigt ligger närmast sökordet skulle placeras först, till exempel *återge*. Ett annat alternativ skulle i det här fallet vara att sökordets två led slås upp, dvs. *åter* och *ta*.

Lucianas nästa, och även sista, sökning gäller substantivet *dust* (rad 10). Det är ett ord som ger träff och som definieras 'kamp, strid'. De tre söksträngarna *kulspruta*, *återta* och *dust* hör ihop tematiskt. Det troliga är att Luciana läser en text som har med militära strider att göra.



#### 8.4.2.4 Kommentrar

Luciana gör i stort sett som hon säger att hon gör enligt uppgifterna i enkäten, förutom att hon just denna vecka inte använder LSL varje dag. Det utesluter som sagt på intet sätt att hon faktiskt är en frekvent användare av LSL. Luciana söker mest på konkreta substantiv.

Samtliga uppslagningar utgörs av rättstavade svenska ord som i alla fall utom ett (*återta*) också ger träff i LSL. I den frekvenssorterade lemmalistan över loggfilssökningarna återfinns *ledsen* och *lexikon* på plats 157 och 223, medan Lucianas övriga uppslagningar är mycket mindre frekventa. Substantivet *dust* har totalt sett slagits upp vid fyra tillfällen och *kulspruta* har bara slagits upp en enda gång, och det således av Luciana. Enligt Kelly-listan är *ledsen* och *fil* vanligast av Lucianas sökord. Orden *dust* och *kulspruta* återfinns däremot inte alls.

Endast ett av sökorden, *återta*, ger inte träff i LSL, utan genererar 21 ordförslag. Uppslagningen demonstrerar problemet med att ha en rättstavningshjälp som endast tar hänsyn till editeringsavstånd. Luciana hade rimligen varit mest hjälpt av att få reda på att ordet förvisso finns i svenska språket, men däremot inte i LSL. Enligt Kelly-listan är det här verbet relativt vanligt och det skulle eventuellt kunna stå på listan över nyordsförslag till en framtida uppdatering av LSL.

## 8.5 Informant 4: Kazim

### 8.5.1 Kazims enkätsvar

Informanten bakom nästa enkät är en arabisk man som är 57 år vid tidpunkten för undersökningen. Vi kan kalla honom Kazim. Kazim kom till Sverige 2001. Hans modersmål är arabiska och det är det språk han anser sig vara bäst på. Han talar arabiska med sin familj och sina vänner. I skolan eller på jobbet talar han svenska. Kazim uppger att han talar eller använder svenska flera gånger i veckan och bedömer sina kunskaper i svenska språket som ganska bra. Han har genomgått sfi-utbildning och har också läst kursen Svenska A. På frågan om högsta avslutade utbildning i hemland anger Kazim universitet. Samma svar anger han på frågan om högsta avslutade utbildning i Sverige. Eftersom han har bott i Sverige i tio år är det fullt möjligt att han har universitetsutbildningar från två länder. Det kan också vara så som var fallet med Elira, dvs. att det är den utländska utbildningen som åsyftas. Han har inte preciserat vilken utbildning han har. Just nu är Kazim arbetsökande.

Kazim uppger vidare att han använder LSL flera gånger per dag och mest för att förstå ord i en text han läser. På frågan vilka informationskategorier han är mest intresserad av kryssar han endast för ett alternativ och det är Ordklass. Endast ett alternativ markeras också i den påföljande frågan om vilka informationskategorier som han aldrig söker uppgift om, och det är betydelse. Med tanke på att han säger sig använda LSL för reception är det svaret minst sagt förvånande. Kazim ger intryck av att ha ganska begränsade kunskaper i svenska språket, sina tio år i Sverige till trots.

Kazim uppger vidare att han aldrig brukar titta på länkarna till bild eller film och att han inte har läst instruktionerna till Lexin. På frågan om hur han löser språkliga problem om han inte använder LSL anger Kazim att han använder en tryckt ordbok.

Kazim har varken lämnat någon kommentar eller någon uppgift om mejladress.

## 8.5.2 Kazims aktiviteter i LSL

Här följer en beskrivning av de 36 aktiviteter som kan knytas till Kazims enkät. De presenteras i tabell 8:7.

TABELL 8:7. *Aktiviteter knutna till Kazim, enkät-id 2011:72.*

Rad	Datum	Tidpunkt	Aktivitetstyp	Aktivitet
1	6	20:45:34	Träff	vördnad
2	6	20:46:04	Frågeformulär	2011:72
3	6	21:07:57	Träff	amortera
4	6	21:08:09	Lyssna	amorterar
5	6	21:08:42	Lyssna	amorterar
6	6	21:09:01	Träff	amortering
7	6	21:09:29	Träff	bagatell
8	6	21:09:50	Lyssna	bagatell
9	6	21:09:51	Lyssna	bagatell
10	6	21:10:19	Träff	celeber
11	6	21:10:26	Lyssna	celeber
12	6	21:10:27	Lyssna	celeber
13	6	21:11:03	Träff	dekadent
14	6	21:12:22	Träff	fabulera

15	6	21:13:49	Icke-träff	fragmentarisk
16	6	21:14:51	Träff	fragment
17	6	21:15:46	Träff	frankera
18	6	21:15:57	Visa film	frankerar
19	6	21:16:13	Träff	frankera
20	6	21:16:35	Träff	gedigen
21	6	21:16:47	Lyssna	gedigen
22	6	21:16:49	Lyssna	gedigen
23	6	21:17:51	Träff	haverera
24	6	21:33:06	Träff	klander
25	6	21:34:45	Icke-träff	konstellera
26	6	21:35:13	Icke-träff	konstellation
27	6	23:23:30	Träff	vördnad
28	6	23:23:47	Träff	bagatell
29	6	23:24:17	Träff	celeber
30	6	23:24:46	Träff	dekadent
31	6	23:32:57	Träff	förseelse
32	6	23:41:17	Träff	förseelse
33	6	23:41:29	Träff	dekadent
34	6	23:41:32	Träff	celeber
35	6	23:50:20	Träff	dekadent
36	6	23:50:24	Träff	förseelse

### 8.5.2.1 Datum och tidpunkt för användningen

Det som förenar informanterna i detta kapitel är att de har uppgett att de använder LSL dagligen. De faktiska aktiviteterna visar, återigen, något annat. Den här veckan använder Kazim ordboken endast vid ett tillfälle, nämligen under torsdag kväll mellan klockan 20:45 och 23:50. Att aktiviteterna under denna kväll tillhör Kazim och ingen annan får sägas stå bortom allt rimligt tvivel.

### 8.5.2.2 Typer av aktiviteter

Kazims 36 aktiviteter fördelar sig på de olika typerna såsom presenteras i tabell 8:8.

TABELL 8:8. *Typ av aktivitet och dess frekvens.*

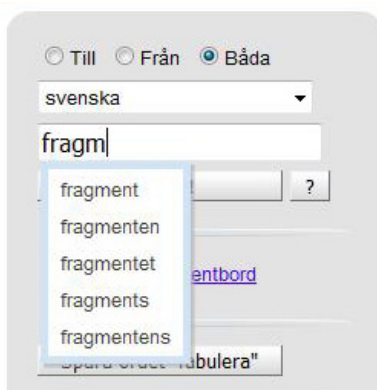
Typ av aktivitet	Antal
Frågeformulär	1
Lexikonuppslagningar	26
- varav träff	23
- varav korr A	
- varav korr F	
- varav icke-träff	3
Checking	
Lyssna	8
Bild svenska	
Visa film	1
<b>Totalt</b>	<b>36</b>

Majoriteten av Kazims aktiviteter utgörs av lexikonuppslagningar. De allra flesta av dem ger träff i LSL. Tre sökningar leder till icke-träff. De övriga nio aktiviteterna fördelar sig på åtta klickningar på knappen för uppläst uttal och en klickning på länken till Visa film. Kazim har inte genomfört några ändringar i inställningarna, vilket betyder att ordboksartiklarna han slår upp varken innehåller information om böjning, exempel, konstruktioner, sammansättningar eller uttryck. Han har inte heller klickat på några länkar till bilder.

### 8.5.2.3 Aktiviteternas förlopp

Samtliga Kazims aktiviteter utförs alltså under kvällen torsdagen den 6 oktober. Han första uppslagning, på substantivet *vördnad* (rad 1), ger träff i LSL. I samband med denna sökning kommer erbjudandet om att delta i enkätundersökningen upp. Kazim väljer då att svara på frågeformuläret (rad 2). Den följande uppslagningen är på verbet *amortera* (rad 3), ett ord han sedan lyssnar till två gånger (rad 4 och 5). Därefter modifierar han söksträngen och slår upp avledningen *amortering* (rad 6). Sedan slår Kazim upp ytterligare två ord där han lyssnar till det upplästa uttalet två gånger, först substantivet *bagatell* (rad 7–9) och sedan adjektivet *celeber* (rad 10–12). De påföljande två sökningarna

ger också träff i LSL, adjektivet *dekadent* (rad 13) och verbet *fabulera* (rad 14). Nästa uppslagning, *fragmentarisk* (rad 15), ger däremot ingen träff i LSL. Kazim följer då upp med en sökning på en del av ordet, *fragment* (rad 16). Han har återigen modifierat söksträngen, nu genom att ta bort en ändelse. Han kan ha fått hjälp av den rullista med ordformer som kommer upp när man skrivit ett par bokstäver i sökfältet (se figur 8:1).



FIGUR 8:1. Rullista med sökbara ordformer som visas allteftersom en söksträng skrivs in i sökfältet: exempel fragment/fragmentarisk. Hämtad 2016-08-10.

Det kan naturligtvis också vara så att Kazim utan att uppmärksamma denna assistans väljer att upprepa sökningen genom att söka på en del av ordet.

Följande aktivitet är en uppslagning på verbet *frankera* (rad 17). *Frankera* är ett av de ca 700 verb som har en länk till en film. Kazim väljer att klicka på länken (rad 18) och går sedan tillbaka till ordboksartikeln (rad 19). Sedan slår Kazim upp adjektivet *gedigen* och lyssnar sedan till dess uttal två gånger (rad 20–22). Därpå slår han upp verbet *haverera* (rad 23). Vid det här laget har Kazim under en dryg halvtimme i regelbunden takt slagit upp olika ord och ibland lyssnat till deras uttal. Nu upphör dock aktiviteterna under cirka en kvart. När han är tillbaka gör han under två minuter tre uppslagningar som leder till en träff, *klander* (rad 24), och två icke-träffar, *konstellera* (rad 25) och *konstellation* (rad 26). Återigen upprepar han en sökning, denna gång genom att omvandla verbet *konstellera* till ett substantiv, *konstellation*. Sedan gör Kazim ett längre uppehåll i sina aktiviteter i LSL och återkommer först nära två timmar senare. Då följer tio uppslagningar som alla ger träff i LSL. Samtliga är upprepningar av tidigare sökningar eller nya ord som upprepas gång på gång. Först upprepas de fyra sökningarna på orden *vördnad*, *bagatell*, *celeber* och *dekadent* (rad 27–30). Efter cirka sju minuter slås sedan substantivet *förseelse* upp (rad 31) och sökningen upprepas efter ungefär åtta minuter (rad 32). Därefter slås återigen *dekadent* och *celeber* upp (rad 33 och 34). Kazim två sista

aktiviteter denna kväll, och för hela veckan, är ytterligare två upprepningar. För fjärde gången slår han upp adjektivet *dekadent* (rad 35) och för tredje gången slås substantivet *förseelse* (rad 36) upp.

#### 8.5.2.4 Kommentrar

Kazim uppger att han har ganska goda kunskaper i svenska. Av enkätuppgifterna framgår dock att han inte till fullo verkar ha förstått alla frågor. Han söker emellertid på ganska svåra ord. Endast fyra av de elva unika orden ingår i Kelly-listan, och det med förhållandevis mycket låga frekvenser. I loggfilerna över de mest frekventa uppslagningarna har Kazims söksträngar också låga frekvenser. Detta kan tyda på att han behärskar svenska på en relativt avancerad nivå eller att han åtminstone försöker läsa en text på relativt avancerad nivå. Kazim har tyvärr inte lämnat någon kommentar som kan avslöja något om hans faktiska svenskkunskaper.

Kazims uppslagningar utmärker sig på flera sätt. För det första modifierar han uppslagningar som inte ger träff i LSL genom att lägga till eller ta bort en avledningsändelse, t.ex. *fragmentarisk* till *fragment* (rad 15–16) (jfr Hult 2008b om olika sökstrategier användarna tar till vid icke-träff i LSL, se även Nilzon 2016). Vid ett av tillfällena har han dessutom modifierat en söksträng som faktiskt ger träff (*amortera* till *amortering*, rad 3 och 6). För det andra utförs uppslagningarna i bokstavsordning (se t.ex. rad 6–14). Det verkar alltså som att han slår upp orden från någon slags lista. I enkäten uppger Kazim som sagt att han använder LSL för att förstå ord i en text och att ordklass är den typ av information han mest eftersöker. I så fall skulle den aktivitet han utför i stunden, dvs. slår upp ord från en alfabetiskt ordnad lista, kunna gå ut på att han tar reda på ordens ordklassstillhörighet. Men det kan då verka märkligt att han skulle behöva slå upp samma ord om och om igen.



FIGUR 8:2. *Komplett artikel för bagatell i LSL4. Hämtad 2016-08-10.*

De upprepande sökningarna är ytterligare ett utmärkande drag för Kazims uppslagningar. Det kan vara ett uttryck för frustration, då de artiklar Kazim får upp saknar de fem informationskategorierna Bøjning, Exempel, Konstruktioner,

Sammansättningar och Uttryck. I normala fall skulle han alltså få se en artikel lik den i figur 8:2. Det Kazim istället möts av är en artikel med följande utseende (se figur 8:3):



FIGUR 8:3. Manipulerad artikel för bagatell i LSL4. Hämtad 2016-08-10.

Trots att den manipulerade ordboksartikeln ger ett mycket magert intryck innehåller den fortfarande en definition av uppslagsordet, vilket ju är den avgjort viktigaste upplysningen i en receptionsordbok. Betydelse är emellertid den (enda) informationskategori som Kazim menar att han *aldrig* eftersöker.

## 8.6 Informant 5: Farah

### 8.6.1 Farahs enkätsvar

Den femte och sista informanten vars ordboksanvändning näranalyseras är en kvinna som vi kan kalla Farah. Hon kom till Sverige 2006 och var vid tidpunkten för enkätundersökningen 47 år gammal. Hennes modersmål är persiska och tillsammans med ryska är det hennes bästa språk. Hon använder persiska med sin familj. Med sina vänner talar hon svenska, ryska och persiska. I skolan eller på jobbet använder hon bara svenska. Farah uppger att hon använder svenska språket varje dag och bedömer sina kunskaper som bra. Hon har genomgått sfi-utbildning, men har ingen annan utbildning i svenska. Hennes högsta avslutade utbildning i hemlandet är på universitetsnivå. Högsta avslutade utbildning i Sverige är gymnasium. Hennes nuvarande huvudsakliga sysselsättning är studerande.

Farah använder LSL mest för att förstå ett ord i en text hon läser eller för att få hjälp med ett ord när hon skriver. På frågan om vilka informationskategorier hon mest söker efter kryssar Farah för Uttal, Sammansättningar och Exempel. På frågan om informationskategorier hon aldrig söker på kryssar hon för samtliga (!) tolv alternativ.

Farah uppger vidare att hon ibland klickar på länkarna till Bild svenska och Visa film. Hon svarar dessutom ”ja” på frågan om hon har läst instruktionerna

till Lexin. Om hon inte använder LSL för att lösa språkliga problem vänder hon sig till *Google* (eller motsvarande sökmotor på internet).

Farah har lämnat en 31 ord lång kommentar. Den lyder:

lexikon de är bra för invandrare som tolkar sitt språk på svenka,men finns mycket ord som inte i lexikon. vi behöver mer, och finns fölklaning också. tack för era hjälp.

Med formuleringen ”tolkar sitt språk på svenka” antyds att Farah (också) använder ett (eller flera) tvåspråkiga Lexin-lexikon. Hon saknar många ord och förklaringar till dem. Hon är ändå tacksam för den hjälp hon kan få (av LSL eller Lexin i allmänhet). Av kommentaren att döma har Farah begränsade kunskaper i svenska, åtminstone i skriftspråkliga sammanhang. Hon har dock bedömt sina svenskkunskaper som bra, dvs. den bästa nivån enligt svarsalternativen. Med tanke på att hon skriver en bristfällig svenska tycks självskattningen vara en aning orealistisk.

## 8.6.2 Farahs aktiviteter i LSL

I tabell 8:9 återges Farahs aktiviteter på LSL:s webbsida.

TABELL 8:9. *Aktiviteter knutna till Farah, enkät-id 2011:434.*

Rad	Datum	Tidpunkt	Aktivitetstyp	Aktivitet
1	9	07:08:34	Checking	Exempel
2	9	07:08:41	Checking	Sammansättningar
3	9	07:08:43	Checking	Böjning
4	9	07:09:00	Visa film	joggar
5	9	08:03:08	Checking	Visa alltid svenskt tangentbord med å, ä, ö
6	9	13:17:01	Träff	burit
7	9	13:17:37	Frågeformulär	2011:434
8	9	13:49:41	Träff	burit
9	9	13:51:55	Träff	burit
10	9	13:52:36	Korr F 10	buret → uret, -uret, boret, bäret, budet, buset, burit, buren, -buret, burget
11	9	13:52:45	Träff	-buret
12	9	14:03:20	Bild svenska	badrum trio sv
13	9	14:03:24	Bild svenska	dynamic_text / badrum.swf trio sv
14	9	14:03:24	Bild svenska	menu_text / badrum.swf trio sv
15	9	14:04:07	Lyssna	w256.mp3 / badrum.swf trio sv
16	9	14:06:42	Bild svenska	hygien / badrum.swf trio sv



### 8.6.2.1 Datum och tidpunkt för användningen

Samtliga Farahs aktiviteter utförs under en och samma dag, nämligen söndagen den 9 oktober. Det rimmar illa med uppgiften om att hon använder LSL varje dag. Hon påbörjar sina ordboksaktiviteter strax efter sju på morgonen, närmare bestämt 07:08:34. Efter ett uppehåll på några timmar återkommer hon till LSL omkring kvart över ett på dagen. Därefter gör hon ytterligare ett uppehåll för att återkomma en sista gång ca tio i två.

### 8.6.2.2 Typer av aktiviteter

I tabell 8:10 visas Farahs 16 aktiviteter uppdelade på de olika typerna.

TABELL 8:10. *Typ av aktivitet och dess frekvens.*

Typ av aktivitet	Antal
Frågeformulär	1
Lexikonuppslagningar	5
- varav träff	4
- varav icke-träff	
- varav korr A	
- varav korr F	1
Checking	4
Lyssna	1
Bild svenska	4
Visa film	1
<b>Totalt</b>	<b>16</b>

Aktiviteterna är få under perioden för datainsamlingen. Samtliga utförs under en och samma dag. De fördelar sig i nästan lika delar mellan lexikonuppslagningar, checking och klickningar på länkar till ljud, bild eller film. I enkäten uppger Farah att hon ibland tittar på bilder och filmer, vilket hon enligt loggfilererna bevisligen också gör.

### 8.6.2.3 Aktiviteternas förlopp

Som redan framgått använder Farah LSL endast under söndagen den 9 oktober. Aktiviteter utförs vid tre olika tillfällen denna dag.

Tidigt på morgonen kryssar Farah för tre av de avaktiverade informationskategorierna: Exempel (rad 1), Sammansättningar (rad 2) och Böjning (rad 3). Det kan betyda att hon har tagit till sig informationen att det pågår en undersökning av LSL, att inställningarna har manipulerats och hur man ska göra om man vill ta tillbaka de informationskategorier som avaktiverats. Farah kryssar dock för endast tre av de fem avaktiverade kategorierna. Två av kategorierna, Exempel och Sammansättningar, återfinns också bland de kategorier som Farah enligt uppgifterna i enkäten ofta eftersöker information om.

Därefter klickar Farah på länken till Visa film för verbet *joggar*. Hon kan emellertid inte göra det om hon inte först slår upp ordboksartikeln *joggar*. Men någon sådan aktivitet ser vi inte i tabell 8:9. Klickningen på länken till filmen för *joggar* föregås alltså av en uppslagning i något av de tvåspråkiga lexikonerna, antagligen det svensk-persiska. Hittills har vi alltså inte några belägg för att Farah använder LSL över huvud taget.

Följande aktivitet genomförs efter uppemot en timme. Då aktiverar Farah rutan för ”Visa alltid svenskt tangentbord med å, ä, ö” (rad 5). Den här informationen är inte aktiverad i grundinställningarna. Därefter gör Farah ett långt uppehåll i sina aktiviteter.

Efter drygt fem timmar utförs nästa aktivitet. Då slår Farah upp supinumformen av verbet *bära*, dvs. *burit* (rad 6). Det går inte att veta om Farah här aktivt har valt att använda LSL. Om man inte väljer att slå upp i något av de tvåspråkiga lexikonerna slås sökordet nämligen automatiskt upp i LSL. Sökordet *burit* ger två träffar i LSL, dels på hänvisningslemmat *burit* som hänvisar till verbet *bär*, dels på själva uppslagsordet *bär*.

Lexikonuppslagningen *burit* innebär att hon nu erbjuds att delta i enkätundersökningen och hon väljer att göra det (rad 7). Efter ungefär en halvtimme sker så nästa aktivitet. Då slår Farah återigen upp ordet *burit* (rad 8). Sökningen upprepas ytterligare en gång efter drygt två minuter (rad 9). Farah verkar inte tillfredsställd med de träffar hon får. Inför nästa uppslagning har sökordet ändrats från *burit* till *buret* (rad 10). Den sökningen ger dessvärre ingen träff i LSL utan leder till korrigeringsförslag, den här gången tio till antalet: *uret*, *-uret*, *boret*, *bäret*, *budet*, *buset*, *burit*, *buren*, *-buret*, *burget*. Farah väljer att gå vidare med förslaget *-buret* (rad 11). Denna uppslagning ger två träffar i LSL, dels artikeln *bil* där sammansättningen *bilburen* återfinns, dels artikeln *rullstol* under vilken det sammansatta ordet *rullstolsburen* förekommer.

Farahs sista fem aktiviteter är klickningar på länkar till bilder och ljud utan några föregående lexikonuppslagningar i LSL. Om en länkaktivitet inte föregås av en uppslagning på det aktuella ordet i LSL innebär det som sagt att en sökning i ett tvåspråkigt lexikon har skett. Uppslagningarna i ett tvåspråkigt lexikon, högst sannolikt det svensk-persiska, har i sin tur lett till att bildtemat för Kök och badrum klickats fram. Väl där klickar Farah bl.a. på texten *en dusch*

alternativt på duschkranen på bilden och får uttalet uppläst (rad 15). Efter cirka två och en halv minut utför Farah sin sista aktivitet för dagen. Under huvudtemat Kök och badrum finns tre rubriker: Kök, Badrum och Hygien. Farah har nu varit inne på illustrationerna under rubriken Badrum och från den bilden har hon klickat sig vidare till den nomenklatoriska illustrationen som ges under rubriken Hygien (rad 16), se figur 8:3 (se avsnitt 3.1.2.8 för ett annat exempel på en sådan illustration).



FIGUR 8:4. Nomenklatorisk illustration till rubriken Hygien.

#### 8.6.2.4 Kommentar

De uppgifter som Farah lämnar i enkäten ger uppfattningen att hon har begränsade kunskaper i svenska. Loggfilsaktiviteterna avslöjar också att det inte är LSL som Farah i första hand verkar använda utan ett tvåspråkigt Lexin-lexikon, troligen det svensk-persiska. Det vet vi eftersom hon har klickat på länkar till bilder utan att det föregås av en uppslagning i LSL. Vid det tillfälle som enkäten fylls i har hon inte aktivt valt ett tvåspråkigt lexikon utan sökordet har slagits upp i LSL och därför visas länken till enkäten.

Man kan med rätta fråga sig varför de informanter som har ett modersmål som representeras av ett Lexin-lexikon inte använder detta tvåspråkiga lexikon, utan slår upp i LSL. Farah är ett exempel på en informant som faktiskt *inte* verka använda LSL, utan framför allt ett tvåspråkigt lexikon. Hon uppvisar också ganska begränsade kunskaper i svenska och det naturligaste för henne torde vara att i första hand använda tvåspråkiga ordböcker. Mervan och Luciana däremot behärskar svenska bättre och att de väljer att slå i LSL är därför inte så svårt att förstå.

## 8.7 Sammanfattning

Det som är gemensamt för de fem enskilda informanterna är att de har ett annat modersmål än svenska, att de har kommit till Sverige under 2000-talet och att de uppfattar sig som frekventa (dagliga) användare av LSL. Av loggfilsaktiviteter att döma har dock ingen av dem använt LSL varje dag utan som mest fyra dagar (Elira). Mervan, Elira, Luciana, Kazim och Farah uppvisar sinsemellan både likheter och olikheter i sina sökbeteenden. Mervan använder t.ex. bara LSL på dagtid, medan Elira och Kazim endast söker i LSL på kvällstid. Tre av informanterna har aktiverat de avaktiverade informationskategorierna. Länkarna till uppläst uttal, bilder och filmer utnyttjas mycket lite av de fem informanterna, men desto mer av målgruppen i stort (som utgörs av 249 informanter). Endast Farah har klickat på samtliga tre typer av länkar. Kazim är den som i synnerhet har använt sig av lyssna-funktionen.

Målgruppsinformanterna verkar också befinna sig i olika typer av användningssituationer. Mervans tre inledande söksträngar står i plural vilket antyder att han har stött på orden i den just den böjningsformen i någon typ av text. Eliras uppslagningar sker tidsmässigt i olika sjok och inom de olika tidssjoken hör sökorden ihop semantiskt. Det är troligt att hon sitter och läser en text på svenska och slår upp de ord hon inte förstår. I Lucianas uppslagningar anas en liknande användarsituation. Kazim däremot förefaller befinna sig i en helt annan brukssituation. Han ser ut att hämta sina uppslagsord från en alfabetiskt ordnad ordlista. Uppslagningarna följer nämligen en alfabetisk ordning (*am-ortera, bagatell, celeber* etc.). Efter en första runda upprepar han några av söksträngarna, i liknande ordningsföljd. Tolkningen förstärks av att samtliga uppslagna ord står i grundform.

Målgruppsinformanternas typer av uppslagningar skiljer sig som synes åt sinsemellan, men också i förhållande till de mest frekventa uppslagningarna i hela 2011-materialet. Där återfinns t.ex. många verb där söksträngen sammanfaller med uppslagsformen, dvs. presens (jfr tabell 6:3 i avsnitt 6.1.1.1). Så

är inte fallet för de fem målgruppsinformerarnas uppslagningar. Faktum är att dessa har fler likheter med de mest frekventa uppslagningarna i 2008-materialet, än med dem i 2011-materialet. De flesta uppslagningar utgörs av substantiv och söksträngarna sammanfaller dessutom oftast med ordens grundform. Homografer är ovanliga. Skillnaderna mellan de två materialen verkar av den anledningen inte bara kunna beskrivas i termer av före och efter uppdateringen av LSL.



## 9. Sammanfattning och framåtblick

I denna avhandling undersöks bruket av internetversionen av *Lexins svenska lexikon*, LSL. LSL är en enspråkig ordbok, avsedd för inlärare av svenska på relativt avancerad nybörjarnivå. Den ingår i ett större, statligt finansierat projekt, Lexin, som också har till uppgift att ta fram tvåspråkiga ordböcker för nyanlända. LSL reviderades 2008–2011. Ett övergripande syfte med studien är att ta reda på om och hur LSL används av målgruppen. Föreliggande undersökning är den första omfattande studien om just denna ordboks användning.

Syftet med undersökningen är tredelat, närmare bestämt att 1) undersöka användningen av LSL, 2) granska vilka det är som använder LSL samt 3) närstudera enskilda informanternas sökbeteende. De tre delsyftena bildar utgångspunkt för tre delstudier. Studien genomförs med hjälp av en kombination av metoder – loggfilsanalys och webbaserat frågeformulär. Materialet för undersökningen utgörs av loggfilsdata från 2008 och 2011 och enkätdata från 2007 och 2011, dvs. före och efter den omfattande revideringen av LSL. Materialet från 2011 är inte bara kvantitativt större, utan innehåller också fler typer av information. Därför står 2011-materialet i fokus genom hela avhandlingen.

### 9.1 Användningen av LSL

I den första delstudien redogörs för undersökningen av *användningen* av LSL3 (dvs. icke reviderade LSL) och LSL4 (dvs. reviderade LSL). Detta görs med hjälp av loggfilsanalys. Två typer av loggfiler registreras på webbsidan för LSL,

dels lexikonuppslagningar (typ a), dels navigeringsval (typ b). En stor fördel med metoden loggfilsanalys är att det finns en likhet mellan undersökningssituationen och en verklig brukssituation, representerad av faktiska användare och deras aktiviteter på LSL:s webbsida. Materialet består av en veckas uppslagningar i början av oktober 2008 respektive 2011. I analysen fokuserar jag på följande frågor:

- Vilka lexikonuppslagningar är mest frekventa när det gäller söksträngar som ger träff i LSL? Vad kännetecknar dem?
- Vilka söksträngar ger inte träff i LSL? Vad är utmärkande för dem?
- Vilka likheter och skillnader uppvisar de två loggfilsmaterialen beträffande de två första frågorna? Verkar revideringen av LSL ha haft någon effekt på användningen av LSL?
- Vilka navigeringsval utnyttjas av användarna?

Det finns uppenbara likheter mellan de två materialen. Exempelvis sker lexikonuppslagningarna efter samma tidsmässiga mönster. Söktrycket är högst måndag till torsdag och går sedan ner under helgen. Vad gäller tidpunkt på dygnet utförs det markant fler uppslagningar mellan cirka 09.00 och 15.00 samt mellan cirka 19.00 och 22.00. Vidare är den vanligaste söksträngen i båda materialen substantiv i obestämd form singular och därefter kommer verb i infinitiv och adjektiv i grundform. Trots att många böjningsformer är sökbara i Lexin slår användarna ändå främst på ordens grundform. Jag gör tolkningen att detta kan bero på att användarna antingen inte känner till att det går att söka på böjningsformer eller att de är så pass vana ordboksanvändare att de vet att uppslagsord normalt står i grundform.

På andra punkter är däremot de två materialen påfallande olika. Det gäller både de mest frekventa söksträngarna som resulterar i en direkt träff i ordboken och de vanligaste uppslagningarna som leder till någon typ av korrigering eller till en icke-träff. En typisk träff i 2008-materialet är ett abstrakt substantiv i grundform, t.ex. *perspektiv*. En typisk träff i 2011-materialet kännetecknas istället av att den utgörs av en böjningsform, kanske ett verb i presens och inte sällan en homograf ordform, såsom i *far*. I 2011-materialet återfinns också, jämfört med i 2008-materialet, mer basala och vanliga ord i betydelsen ord som man lär sig tidigt i sin svenskinläring och som är mycket frekventa i svenskan. Det rör sig t.ex. om mycket vanliga verb med flera grundbetydelse och vanligen många överförda användningar, inte sällan förekommande i olika typer av fasta fraser. Söksträngarna i sig utgörs dock mycket sällan av en fras. I 2011-materialet är andelen flerordssökningar omkring 4 %. En förklaring till att andelen flerordsuppslagningar inte är större kan vara att användarna gene-



rellt sett är högst medvetna om hur (tryckta) ordböcker är strukturerade på makronivå, dvs. att uppslagsord i regel består av ett enda ord.

De mest frekventa uppslagningar som generar en korrigerings eller resultat i rena icke-träffar uppvisar som sagt också olikheter. Kännetecknande för 2008-orden är att de är rättstavade svenska ord i grundform. Det rör sig t.ex. om *estetik*, *integration* och *pedant*. Även 2011-orden är många gånger rättstavade svenska ord och fraser i grundform, såsom *involvera*, *ont om*, *pennfodral* och *snirkliga utsmyckningar*. Flertalet av uppslagningarna utgörs dock av felstavade ord, inte sällan böjningsformer. Några exempel på sådana är *\*argsnit*, *\*defenition*, *\*hårdigt*, *\*tragidin* och *\*tyvär*. Här återfinns dessutom utländska ord och fraser t.ex. *encumber*, *tikanis* och *what r u doing*. När en sökning inte ger en direkt träff i LSL, men heller inte en ren icke-träff, aktiveras rättstavningshjälpen. Den är baserad på editeringsavstånd och inte på uttal eller utifrån några semantiska kriterier. Att basera rättstavningshjälpen endast på editeringsavstånd visar sig fungera vid mindre felstavningar, förutsatt att motsvarande rättstavade ord återfinns i ordboken. Men i de fall det sökta ordet inte ingår i LSL blir utfallet av en korrigerings många gånger missvisande. Användarna av LSL skulle vara mycket behjälpta av en rättstavningshjälp som kunde hantera uttalsrelaterade felstavningar. I loggfils materialet återfinns t.ex. söksträngarna *\*ovasät*, *\*livnadsstanndar* och *\*kanäkabrod* bland de uppslagningar som inte ger en träff i LSL. I lemmauppsättningen för LSL ingår dock de tre lemmarna *oavsett*, *levnadsstandard* och *knäckebröd*.

Jag drar slutsatsen att olikheterna i de två materialen sannolikt beror på användargruppens sammansättning. Användarna av LSL3 förefaller till stor del vara modersmålstalande skolungdomar. Ett avgörande faktum i sammanhanget är att ett svensk-engelskt lexikon fortfarande ingick i Lexins tvåspråkiga utbud 2008, men togs bort 2010 i samband med uppdateringen av LSL. Användarna av LSL4 verkar däremot i högre grad utgöras av inlärare på en relativt avancerad nybörjarnivå, dvs. den primära målgrupp LSL riktar sig till.

Med navigeringsval avses för det första klickningar på länkar till uppläst uttal, bilder och filmer. Dessa länkar verkar utnyttjas av användarna i begränsad omfattning. Med navigeringsval avses också de ändringar som användarna gör under fliken ”Inställningar” och klickningar på frågetecknet, ”?”. Under den vecka datainsamlingen genomfördes manipulerades inställningarna och följande informationskategorier klickades bort från ordboksartiklarna: Böjning, Exempel, Konstruktioner, Sammansättningar och Uttryck. På Lexins startsida informerades användarna om den pågående undersökningen och hur de skulle återställa inställningarna i standardläge, om så önskades. Manipuleringen av inställningarna gjordes i ett försök att få mer detaljerad kunskap om dels användarnas sökbeteende, dels vilka specifika informationskategorier användarna är intresserade av. Det visar sig att användarna helt enkelt väljer att aktivera

samtliga fem informationskategorier som vid tidpunkten för datainsamlingen hade avaktiverats, dvs. Böjning, Exempel, Konstruktioner, Sammansättningar och Uttryck. Intrycket är att detta görs utan någon större reflektion. Användarna tycks bara vilja ta tillbaka det som saknas. Någon klarare bild informanternas lexikografiska behov på kategorinivå kunde därför inte uppnås.

## 9.2 Användarna av LSL

I delstudie 2 undersöks *vilka det är som använder* LSL. Metoden som används är webbaserat frågeformulär. Jag önskar få upplysningar om användarna själva och hur deras användningssituation ser ut. Omständigheterna kring min undersökning borgar för ett gott gensvar på frågeformuläret och även för sanningsenliga svar. Uppskattningsvis torde det t.ex. inte ta lång tid att fylla i enkäten. Informanterna är dessutom anonyma, jag som intervjuare är inte synlig och ämnet för undersökningen är inte känsligt, men av tillräcklig stor vikt för att engagera. Det som efterfrågas är också ganska okomplicerad information och deltagarna kan förväntas vara i stånd att läsa och förstå frågorna, i synnerhet då de har anpassats till ordbokens målgruppsanvändare, dvs. personer som har grundläggande, men inte avancerade, kunskaper i svenska språket. Av informanternas svar att döma verkar vissa frågor ändå ha varit svåra att förstå. En framtida enkät bör förmodligen även vara tillgänglig på engelska, möjligen också på de större invandrarspråken.

I fokus för delstudien står en enkätundersökning från 2011, Studie<sup>11</sup>. Den innehöll 20 frågor. För ungefär hälften av frågorna finns också ett jämförelsematerial i form av en enkätundersökning från 2007 av LSL3, här Studie<sup>07</sup>. Den innehöll 10 frågor. I delstudie 2 önskar jag besvara frågor som:

- Används LSL av den åsyftade målgruppen?
- Vilka lexikografiska behov ger informanterna uttryck för att de har?
- Vilka likheter och skillnader uppvisar de två uppsättningarna av enkätmaterial? Finns det olikheter som kan knytas till revideringen av LSL?

Det står för det första klart att både LSL3 och LSL4 används av den åsyftade målgruppen, men av informanternas uppgifter att döma i olika stor utsträckning. De typiska användarna av LSL3 är skolungdomar eller studenter i 15–29-årsåldern som har svenska som modersmål. De använder LSL frekvent och ger uttryck för att ha lexikografiskt relevanta behov av både receptiv och produktiv art. Betydelse är den informationskategori som är mest eftersökt,

men också Stavning och Exempel. Inte sällan har informanterna lämnat en kommentar. Där framgår det i så fall att de är ganska nöjda med ordboken men att det saknas många ord, kanske i synnerhet ord med fackspråklig karaktär.

De typiska användarna av LSL4 är i 30-årsåldern och har ett annat modersmål än svenska, t.ex. arabiska, persiska eller spanska. De har en universitetsutbildning från sitt hemland och har kommit till Sverige under tidsperioden 2005–2011. De har i regel hunnit genomgå sfi-utbildning. Det är inte helt osannolikt att deras utländska utbildning har validerats gentemot det svenska utbildningsväsendet eller den svenska arbetsmarknaden. De studier som nu bedrivs skulle i så fall kunna vara av kompletterande karaktär. Vidare använder de LSL frekvent och har lexikografiska behov som rör både reception och produktion. Betydelse är den informationskategori som är mest eftersökt, därefter Exempel och Böjning. Inte sällan har de lämnat en kommentar där det framgår att de är mycket tacksamma och nöjda användare av LSL, mer så än 2007-användarna, men att ordboken skulle kunna innehålla förklaringar av fler (svåra) ord.

Jag drar slutsatsen att LSL4 i högre utsträckning än LSL3 används av dem som tillhör den åsyftade målgruppen, dvs. inlärare på relativt avancerad nybörjarnivå. Förändringar i informantgruppens sammansättning relateras till stor del till borttagningen av det svensk-engelska lexikonet.

### 9.3 Målgruppsanvändning på individnivå

I delstudie 3 närstuderas *enskilda informanters sökbeteende* i syfte att få djupare kunskap om användarnas ordboksanvändning på individnivå. Här ingår endast material från 2011. I första steget urskiljs en målgrupp av informanter utifrån modersmål, ankomstår till Sverige och sökfrequens. De uppgår till 249 informanter. Från målgruppen väljs sedan fem informanter ut vars länkningsnäranalyseras, här med de fiktiva namnen Mervan, Elira, Luciana, Kazim och Farah. De utvalda länkningarna representeras av informanter med olika modersmål, av olika kön och i varierande ålder.

Inga generella slutsatser kan naturligtvis dras på ett så litet informanturval. Det är heller inte meningen. Avsikten är att betrakta ordboksanvändning från ytterligare ett perspektiv. Loggfilsanalysen och enkätstudien genererade två olika typer av kvantitativa data, medan länkningen istället resulterar i ett kvalitativt underlag. Bruket av flera metoder är tänkt att ge en mer fullständig och komplett bild av det som undersöks.

I delstudie 3 behandlas följande frågeställningar:

- Vilka likheter och skillnader uppvisar de enskilda informanternas sökbeteende?
- Överensstämmer informanternas lexikografiska självbild med hur de faktiskt använder LSL?
- Kan den aktuella metoden generera relevanta och tillämpliga resultat i anknytning till praktiskt lexikografiskt arbete?

Det som är gemensamt för de fem enskilda informanterna är således att de har ett annat modersmål än svenska, att de har kommit till Sverige under 2000-talet och att de uppfattar sig som frekventa (dagliga) användare av LSL. Av loggfilsaktiviteterna att döma har dock ingen av dem använt LSL varje dag utan som mest fyra dagar (Elira). Vidare använder Mervan t.ex. bara LSL på dagtid medan Elira och Kazim endast söker i LSL på kvällstid.

De specifika ord som målgruppsinformanterna har slagit upp skiljer sig inte bara åt sinsemellan, utan också i förhållande till de mest frekventa uppslagningarna i hela 2011-materialet. Där återfinns t.ex. många verb där söksträngen sammanfaller med uppslagsformen, dvs. presens. Så är inte fallet för målgruppsinformanternas uppslagningar. Faktum är att dessa har fler likheter med de mest frekventa uppslagningarna i 2008-materialet, än med dem i 2011-materialet. Skillnaderna mellan de två materialen verkar av den anledningen inte bara kunna beskrivas i termer av före och efter uppdateringen av LSL, eller kopplas till att modersmålstalare gör på ett sätt och inlärare på ett annat. Användningen av LSL uppvisar till synes en något mer komplex bild än så.

Länkarna till uppläst uttal, bilder och filmer utnyttjas i mycket liten omfattning av de fem målgruppsinformanterna. När det gäller de fem urkryssade informationskategorierna har tre av fem målgruppsinformanter återaktiverat dessa. Målgruppsinformanterna förefaller vidare befinna sig i olika typer av användningssituationer. Eliras uppslagningar sker t.ex. tidsmässigt och tematiskt i olika sjuk. Kazim däremot förefaller befinna sig i en helt annan brukssituation. Han ser ut att hämta sina uppslagsord från en alfabetiskt ordnad lista.

Målgruppsinformanternas lexikografiska självbild visar sig överensstämma på vissa punkter med hur de faktiskt använder LSL, och på andra punkter inte alls. Det är dock inte oväntat att svaren på frågor om självskattning (t.ex. svenskkunskaper) och användarbeteende (t.ex. sökfrequens) som inte helt överensstämmer med hur det faktiskt verkar förhålla sig.

I delstudie 3 ingick också att utvärdera metoden att länka loggfilsdata och enkätdata i ett försök att komma närmare enskilda informanter och deras användning av LSL. Frågan är om den aktuella metoden kan generera relevanta och tillämpliga resultat i anknytning till praktiskt lexikografiskt arbete. Svaret är ja, i alla fall indirekt. Metodkombination, inklusive datatrianguleringen i

delstudie 3, har i hög grad bidragit till att ge en mer komplett och nyanserad bild av användningen av LSL, och kanske också av Lexin i allmänhet. Vi har fått tre olika perspektiv på ordboksanvändning. Loggfilsanalysen gav en relevant överblick över användarnas vanligaste uppslagningar. Enkätundersökningen resulterade i en viktig, om än generaliserad, profilering av användarna. Till sist möjliggjorde länkningen att jag på nära håll (tentativt) kunde följa enskilda användare och deras ordboksaktiviteter.

## 9.4 Framtidens LSL och vidare forskning

I den allmänna teoretiska bakgrunden presenterades två grundläggande perspektiv på lexikografisk teori, en lingvistikorienterad inriktning och en autonom inriktning, som ser lexikografen som självständig i förhållande till andra vetenskapsgrenar. Inom båda dessa orienteringar har användarna kommit att spela en central roll. Genom användarprofilering försöker man karakterisera den typiska användaren och hans syfte med användningen samt vilka lingvistiska och lexikografiska förkunskaper hen besitter. Därjämte bedrivs användarforskning (jfr fotnot 2) som inbegriper alla de metoder som tas i anspråk för att ta reda på vad människor gör när de konsulterar sina ordböcker.

I översikten över tidigare forskning framkom att majoriteten av användarstudierna utgår från icke-sannolikhetsurval. Så har även varit fallet i föreliggande undersökning. Jag har däremot försökt hitta fram till den mest ändamålsenliga urvalsstrategin utifrån de praktiska förutsättningar som rådde. Beträffande enkätundersökningen verkar informanterna i Studie<sup>11</sup> utgöra ett representativt urval åtminstone i ett avseende: de språk som har störst antal modersmålstalare bland informanterna sammanfaller med de Lexin-språk vars lexikon har flest uppslagningar under den månad som datainsamlingen genomförs. Informanturvalet förefaller således åtminstone vara representativt vad gäller fördelningen på informanternas modersmål med tanke på hur den verkliga användningen ser ut. Studiens resultat kan likväl inte generaliseras till bruket av andra ordböcker och i strikt mening kanske inte heller till bruket av LSL. Med detta sagt kan studier baserade på ett icke-representativt urval trots allt bidra med relevant information om möjliga tendenser som kan gälla för större grupper och utgöra ett underlag för vidare utveckling av tidigare studier och teorier.

Min förhoppning är att föreliggande undersökning utgör ett bidrag till den mer teoretiska användarforskningen och att studiens resultat lämnar kunskaper av värde för den praktiska lexikografen. Studien visar bl.a. att förbättringarna i samband med uppdateringen uppskattas av informanterna. Om re-

resurser finns kan man gå vidare i samma riktning. Exempelvis finns uppslag till nyord bland såväl de korrigerade uppslagningarna som de rena icke-träffarna. En god idé vore också att göra alla böjningsformer sökbara. Det har också med tydlighet framgått att rättstavningshjälpen kan förbättras. Vidare uttrycker många informanter starka önskemål om fler uppslagsord i LSL. Lexinprojektet har emellertid begränsade resurser och ordböckerna lär inte utvidgas med fler lemman inom överskådlig tid. En klar förbättring av Lexin skulle dock kunna uppnås genom att uppslagningar som inte ger direkt träff i ordboken togs bättre om hand. Användarna skulle exempelvis, i de fall det är aktuellt, kunna få information om att ordet som eftersöks finns i svenska språket, om än inte i lemmalistan för Lexin. I bästa fall kan användaren också upplysas om att ordet förekommer i en annan ordbok, och även få namnet på eller länken till denna. Vidare innebar revideringen av LSL3 en uppgradering av idiomerna till sublemman. En naturlig utveckling av sökfunktionen i Lexin kunde vara att göra dessa existerande flerordsuttryck sökbara. Slutligen skulle bildmaterialet i Lexin må bra av en översyn och uppdatering.

Framtida forskning kan bedrivas i olika riktningar. Det skulle t.ex. vara intressant att upprepa studien och fokusera på bruket hos de användare som *inte* har ett Lexin-språk som modersmål och följaktligen *måste* använda LSL. Det talas ca 200 språk i Sverige idag (Parkvall 2016). Det ligger långt utanför Lexinprojektets uppdrag att utarbeta tvåspråkiga lexikon till alla minoritetsspråk i Sverige, även om flera av dem har ett mycket stort antal talare. LSL fyller alltså här en oerhört viktig funktion. En sådan replikerande studie skulle kunna göras på befintligt material. Det befintliga materialet skulle vidare kunna utvidgas med nya loggfiler med jämna mellanrum (förslagsvis vart tredje år) och så småningom utgöra ett omfattande underlag för jämförelser över tid. Därutöver innehåller det befintliga materialet ca 9 000 flerordsuppslagningar som det inte fanns utrymme att undersöka närmare. Det skulle vara spännande att analysera dem utifrån dels ett andraspråksperspektiv, dels ett allmänt lingvistiskt perspektiv.

Ett annat sätt att gräva djupare i användningen av LSL är att ta till andra metoder. Mer djupgående och bakomliggande aspekter på ordboksanvändning skulle t.ex. kunna uppnås med intervjuer. Det befintliga materialet innehåller exempelvis mejladresser till flera hundra informanter som uppgav att de kunde tänka sig att delta i fler användarundersökningar.

Ett tredje alternativ är att vidga perspektivet ännu mer. Min undersökning tangerar flera närliggande forskningsområden, inte minst svenska som andraspråk men också tillämpad informationsteknologi och utbildningsvetenskap. Ett potentiellt framtida forskningsämne som gränsar till samtliga områden rör användningen av Lexin i anslutning till sfi-undervisningen. I läroplanen för sfi kan man t.ex. läsa följande: ”Inom utbildningen i svenska för invandrare ska

eleven utveckla sin kompetens att använda olika digitala verktyg och hjälpmedel för information, kommunikation och lärande” (Skolverket 2012:8). I det här sammanhanget fyller förstås elektroniska ordböcker i allmänhet och Lexin i synnerhet en avgörande funktion. Det skulle i det här sammanhanget t.ex. vara givande att få direktkontakt med användarna ute ”på fältet”, t.ex. genom att följa några sfi-elever under en vanlig studievecka.

De forskningsidéer som lagts fram här är inriktade på fler empiriska studier av just Lexin. I framtida forskning bör naturligtvis siktet vidgas ifråga om både empiri och teori. Det skulle förslagsvis vara intressant att närmare undersöka de skilda lexikografiteoretiska perspektiven (företrädda å ena sidan av Sue Atkins och Michael Rundell, å andra sidan av Herbert Ernst Wiegand och Sven Tarp) på användaranpassning och vilka reella konsekvenser de får för den innehållsmässiga utformningen av den slutgiltiga produkten, dvs. ordboken.





## Efterord

Så. Doktorerandet är till ända. Man kan fråga sig vad det är jag har sysslat med egentligen. Svaret är enkelt. Jag har ägnat mig åt de riktigt stora frågorna: Är jag en anhängare av funktionsläran? Ska jag lyxa till det med finkaffe idag eller nöja mig med institutionskaffet? Vad ska jag göra med mitt liv? Ska jag redovisa de här resultaten i en tabell eller ett diagram? Kan jag verkligen med att skicka den här texten till handledarna? Och vilka ska jag tacka i efterordet? Ja, det finns förstås några personer jag vill tacka särskilt.

Professor Sven-Göran Malmgren var huvudhandledare fram till sin pensionering. Därefter övertog Emma Sköldberg huvudhandledarskapet. Emma och Sven-Göran! Utan er – ingen avhandling. Min tacksamhet över att ha fått ta del av er enastående klokskap och gedigna erfarenhet är oändlig. Jag är såklart obeskrivligt glad över att vara klar, men jag kommer verkligen att sakna de lärorika och underhållande handledningsstunderna.

Ett innerligt tack går förstås också till Ingegerd Enström, min slutseminariegranskare. Din noggranna läsning och inte minst din sakkunskap inom andraspråksforskning bidrog till en avsevärd förbättring av det slutgiltiga manuset. Tack, Ingegerd!

Redaktör för avhandlingen har Barbro Wallgren Hemlin varit. Stort tack Barbro, för din mycket uppmärksamma redaktörläsning! Sven Lindström har ansvarat för sättning av manus. Sven, tack för en förstklassig insats! Hallfríður Helgadóttir har snabbt och säkert språkgranskat avhandlingens engelska summary och abstract. Kærar þakkir!

Just den här undersökningen hade över huvud taget inte varit möjlig att genomföra om inte Christian Mattsson, f.d. projektledare för Lexin, gett mig

sitt uppmuntrande medgivande. Professor Viggo Kann, tillika webbansvarig för Lexin, har inte bara varit behjälplig vid insamlingen av material till undersökningen utan har även tålmodigt svarat på mina många frågor av mer eller mindre genomtänkt slag. En ovärderlig insats har också utförts av Christian Sjögren, f.d. systemutvecklare vid Lexikaliska institutet, i arbetet med att bearbeta materialet till ett analyserbart skick. Inte heller får jag glömma att tacka de användare av *Lexins svenska lexikon* som valde att bli mina informanter. Jag är er alla ett stort tack skyldig!

Tack vare donations- och resestipendier har jag kunnat delta i kurser och konferenser i Sverige och utomlands. Tack Adlerbertska Stipendiestiftelsen och Stiftelsen Paul och Marie Berghaus donationsfond för mycket generösa bidrag i samband med dessa resor. Slutligen erhöll jag ett doktorandstipendium från Kungliga och Hvitfeldtska stiftelsen för att färdigställa denna avhandling. Varmt tack!

Under åren på institutionen har jag haft min arbetsplats både ”uppe på” LT-gatan och ”nere på” Humanisten och jag har trivts fantastiskt bra på båda ställena. Hjärtligt tack alla kolleger för det!

Det hela började trots allt på Lexikaliska institutet. Till samtliga gamla och nya kolleger vill jag rikta ett innerligt tack för de år som hittills varit. Ett speciellt tack till Sture Berg för många trevliga arbetsstunder i morfologins tjänst.

Och så det kära doktorandkollegiet. Doktorander kommer och går, men doktorandgemenskapen består. Tack särskilt till er som bistått med hejarop, såväl mental som fysisk stöttning och kaffe, kaffe, kaffe! Ni är verkligen guld värda. En särskild hälsning vill jag dessutom framföra till de doktorander som jag under årens lopp har delat rum med, Inga-Lill Grahn och Louise Holmer: tack för skratt, språkande och – inte minst – rumssämja! Slutligen vill jag nämna Elisabet Engdahl som var studierektor för forskarutbildningen under mina första år som doktorand. Tack Elisabet för stöd och omtanke i en period av utmattnings. Det ordnade sig!

I livet utanför institutionen finns en familj som jag håller mycket av. En mamma och tre bröder, Niklas och våra två små pojkar. Niklas: tack för skjutsen! (Inte minst i symbolisk mening.) Tom och Fred, ögontröst och hjärtansfröjd – love you. Och så käre morbror Lars, från en doktor till en annan ;-): du förblir min förebild i livet.

Göteborg den 20 november 2016

Anki

# Summary

## Introduction

This thesis presents a study on the use of the online learner's dictionary *Lexin Swedish dictionary*, hereafter LSL. LSL is a monolingual dictionary and a part of Lexin, a project financed by the Swedish government. Lexin consists of a series of dictionaries including LSL and about twenty bilingual dictionaries. The largest bilingual dictionaries, including LSL, contain about 30,000 lemmas. The Lexin dictionaries have been available online since the middle of the 1990s. The project was initiated at the end of the 1970s with the object of producing print dictionaries particularly suited for the teaching of Swedish to immigrants. The printed dictionaries as well as the online version of Lexin include about 1,800 pictures divided into 31 themes, such as *Cooking and meal time*, *City and traffic* and *Nature and landscape*. The internet version also includes audible pronunciation. In addition, more than 700 verbs are illustrated by means of short film sequences (animations). The target group for the Lexin dictionaries are immigrants who have recently come to Sweden and are about to learn Swedish but the dictionaries are also meant to satisfy the need for improved knowledge of Swedish held by immigrants who have lived in Sweden for a longer period of time (Beijer & Fre Woldu 1997:8). Whereas the bilingual dictionaries are clearly intended for learners at absolute beginners' level, the monolingual LSL rather aims at users that are relatively advanced beginners. The bilingual dictionaries cover larger minority languages in Sweden where the market supply of bilingual dictionaries is limited.

In 2008–2011, the LSL underwent a comprehensive revision. There were three general principles governing the update (Malmgren 2012, Hult et al. 2010): 1) to further emphasise the distinctive features of a learner's dictionary, 2) to avoid grammatical terminology – even more and 3) in case of conflict between the reception and production aspects, the reception aspect would have preference. Approximately 3,000 new lemmas were added in the revision. Special attention was also given to the description of particle verbs. They generally cause difficulties for second language learners of Swedish.

In short, LSL is a suitable object of inquiry for two reasons. The dictionary has a well-defined target group and an explicit purpose. Thus, the intended use is very clear. What is more, the dictionary is frequently used.

The thesis has a threefold aim, namely to 1) investigate the *use* of LSL, 2) find out who the *users* of LSL are, and 3) take a closer look at *individual informant's search behaviour*. This in turn, forms the basis for three sub studies.

## Method and data

Those who conduct user research in lexicography, i.e. examine dictionary users and dictionary use, have been greatly inspired by the methodology used in social science research (cf. Lew 2002, Tarp 2008a). In this study I make use of two methods namely, web-based questionnaires and observation (by registration of log files), two common methods for survey research. Questionnaires are well suited for factual questions and issues concerning people's opinions. Behavioural issues are better examined by direct observation. Thus, registration of look-ups in a dictionary is the most suitable way of getting access to authentic data about dictionary use on the internet. A third method that was used is data triangulation. In this case it means that the data generated from the former methods are combined and generate a third material. Basically, the purpose of the triangulation is to get a more complete and nuanced picture of the use of LSL. This third method also has a tentative purpose and the procedure of linking log file data and questionnaire data will be evaluated.

As already stated, LSL was thoroughly revised in 2008–2011. The body of the information consists of log file data from 2008 and 2011 and questionnaire data from 2007 and 2011, that is, from before and after the updating project. The 2011 data is not only quantitatively larger, but includes more types of information. For that reason, the 2011 data is the centre of attention throughout this thesis.

## Theoretical perspectives on lexicography

There are two basic perspectives on lexicography, one linguistically oriented approach and one autonomous approach that consider lexicography to be independent in relation to other branches of science. According to the autonomous lexicographical theories, the purpose of a dictionary is defined on the basis of a profile of potential types of users and types of user situations. One might argue that this actually is what seriously practicing lexicographers do, regardless of which theoretical approach they support. Lorentzen & Trap Jensen (2006:200) simply call this way of working “common sense”. In recent decades, the dictionary function and the user oriented perspective have become a priority in both practical and theoretical lexicography.

In my opinion, there is no doubt that the starting point for all (language) lexicography is linguistics. The linguistically oriented approach is therefore of great relevance in the present study. However, the ideas behind the more autonomous lexicographical theories are also of great importance. That includes, for example, specific lexicographical knowledge about dictionary typology and a uniform lexicographical terminology, as well as theoretical perspectives which presuppose the user and the use of the dictionary.

## Learner’s dictionaries and vocabulary learning

One important distinction in second language research is the one between a *second language* and a *foreign language*. A second language is learned in the environment where it is used and where it is the main language of communication, whereas a foreign language is learned in an environment where the language is *not* used as the main language of communication (Abrahamsson 2009:15). LSL is primarily a dictionary for second language users. The long term goal for second language learners is to achieve very good language skills including a large and varied vocabulary, but these learners also have a short term, acute need to use the language to communicate with their environment (Enström 1996:11).

Indisputably, dictionaries can be eminent tools in the process of vocabulary learning. When learning vocabulary in a non-native language, dictionaries are especially valuable tools. Second language learners and native speakers do not differ greatly in terms of what words they are confronted with every day. However, when it comes to dictionaries, second language learners are in need of a more explicit and comprehensive description of the language in question, than the native speaker. This applies above all to culture-specific words. Even though beginners benefit the most from bilingual dictionaries, a monolingual

dictionary might at relatively early stages be an important complement to the bilingual dictionary, provided that the style of defining is adapted to the learners' limited language skills. The purpose of LSL is to be such a dictionary.

## Previous research in dictionary use

Traditionally, user studies have focused on users in an educational context – not least because dictionaries are frequently used in this environment – and especially students who have English as their foreign language (e.g. Béjoint 1981). Furthermore, most user studies about learner's dictionaries have concentrated on rather advanced learners and students (e.g. Wingate 2002). However, there is a lack of studies of dictionary use among less qualified users. The dictionary being investigated here, i.e. LSL, is aimed at learners of Swedish at a relatively advanced beginner's level and they are primarily learning Swedish as a second language and not as a foreign language.

Web questionnaire surveys are relatively few in number considering how popular this method has been traditionally (some examples are Lorentzen & Theilgaard 2012, Müller-Spitzer 2013 and Holmer & Sköldberg 2016). In these studies the majority of informants are in their 30s and well-educated. They primarily look for information about meaning, spelling, inflection and examples. Rambøll (2012) reports on a user survey of the Lexin dictionaries in Norway. According to this study, the average Lexin user has lived in Norway for two to five years. The bilingual dictionaries most used (beside Norwegian–English) are Norwegian to Arabic, Polish, Somali, Tigrinya and Russian. The dictionaries are most often used in educational settings. Similar to most user studies, there is a wish for an extended lemma selection. On the whole, users are very content with the Lexin dictionaries.

An early log file study is one done by de Schryver & Joffe (2004). This study mainly focuses on users' look-ups and emphasises their practical applicability for dictionary makers. For instance, lemma lacunae can be discovered among the unsuccessful look-ups. More recent log file studies, for example Almind (2008), have concentrated on analysing the search behaviour of individual users (cf. de Schryver et al. 2006).

## Sub study 1

Sub study 1 examines the *use* of LSL by means of log files documenting dictionary look-ups (log file type a) and choices of navigation (log file type b) on the LSL website. As already indicated, a great advantage of this method is that

there is a similarity between the situation of inquiry and an authentic user situation, represented by actual users and their activities on LSL's website. The data consists of seven days of web site activities at the beginning of October 2008 (representing LSL3, i.e. the non-revised LSL) and in 2011 (LSL4, i.e. the revised LSL) respectively. In total, the two sets of data contain nearly 360,000 look-ups and about 100,000 choices of navigation. Sub study 1 focuses on the following questions:

- What searches are most frequent when it comes to successful look-ups in LSL? What are their characteristics?
- What searches are unsuccessful in LSL? What are their characteristics?
- What similarities and differences are shown in the two sets of log file data regarding the aforementioned questions? Has the revision of LSL had any effect on the use of LSL?
- Do users make use of the choices of navigation? Do they, for example, click on the links to pictures and animations?

There are indeed similarities in the two sets of data. For example, searches are performed following the same temporal pattern. Users are most energetic on Monday to Thursday, and then less active during the weekend. As for time of day, significantly more look-ups are made between 09.00 and 15.00 and between 19.00 and 22.00. Furthermore, the most common hit in both sets of data is nouns in the indefinite singular, followed by verbs and adjectives in basic forms. Consequently, in spite of the fact that inflectional forms are searchable, the search string usually consists of the basic form. This may be because users either do not know that it is possible to make inflectional searches or that they are such accustomed dictionary users that they know that the entry form normally corresponds to the basic form.

In other respects the two sets of data are strikingly different. This applies both to the most frequent search strings resulting in a direct hit in the dictionary and the most frequent unsuccessful look-ups. A typical hit in the 2008 data is an abstract noun in basic form, for example *perspektiv* ('perspective'). However, a typical hit in the 2011 data consists of an inflectional form, perhaps a verb in the present tense, not infrequently a homograph like for example *far* ('is/are going' or 'father'). Moreover, the 2011 data contains more central and common words than the 2008 data, for instance very frequent polysemic verbs like *do*, *go* and *have*, which are also often used in a figurative sense and they generally appear in different kinds of fixed phrases. The search string itself, however, rarely consists of a phrase. The proportion of multi word searches is about 4% of the total amount of look-ups in the 2011 data. This could imply that users generally are very aware of how (printed) dictionaries are structured at macro level, i.e. that the entry form usually consists of one single word.

Searches not resulting in direct hits in LSL3 or LSL4 can 1) be automatically corrected and directed to an existing lemma, 2) generate a list of two or more suggestions of orthographically similar lemmas, or 3) lead to a pure non-hit. The proportion of unsuccessful searches in the 2008 data is 38% and in the 2011 data 27%. The improved success rate is largely due to the fact that more inflectional forms have been made searchable in the updated version of LSL.

The most frequent unsuccessful searches in the 2008 data are typically correctly spelled Swedish words in basic form, for example *estetik* ('aesthetics') and *pedant* ('pedant, perfectionist'). The corresponding 2011 data also include correctly spelled Swedish words, such as *involvera* ('involve') and *pennfodral* ('pen case'). But, this data also consists of many inflected and/or misspelled search strings, for example *\*argsnit* (corresponding to a spelling like *\*'ill-temreped'*) and *\*defenition* (*\*'defenition'*). Moreover, among the most usual unsuccessful searches we also find non-Swedish words and phrases like *encumber*, *tikanis* and *what r u doing*.

The module for handling misspellings is based only on transposition of letters and not on pronunciation. This seems to work perfectly in case of minor misspellings, provided that the corresponding correctly spelled word is found in the dictionary. In addition, when the searched word is not included in the lemma selection of the LSL, a user can be faced with a long list of often quite irrelevant suggestions of existing lemmas (as they are based on orthographic resemblance alone). The users of LSL, and – perhaps even more – the users of the bilingual Lexin dictionaries, would benefit greatly from a fuzzy search tool that could handle misspellings related to pronunciation.

I conclude that the differences in the two sets of log file data are likely to depend on the composition of the two groups of informants. The users of LSL3 seem to consist largely of native-speaking (young) students. A decisive factor in this context is that a Swedish-English Lexin dictionary was still included in the Lexin website in 2008, but was removed in 2010 in connection with the update of the LSL. Users of LSL4 appear to belong, to a greater extent, to the target group of LSL: learners at a relatively advanced beginner's level.

The choices of navigation, log file type b, include clickings on links to audible pronunciation, pictures and film sequences. These links seem to be used very little. Choices of navigation also include changes in the settings where specific information categories can be activated or inactivated. During the data collection the settings were deliberately manipulated. Five information categories were removed: inflection, examples, compounds, valency and fixed phrases. Users were instructed how to change the settings back to default if they wanted to. This experiment was made in an attempt to obtain more detailed knowledge of both the users' search behaviour, and which specific information categories users find most important. As it turns out, many users



did nothing at all, some complained about the missing information (in most cases inflection), and some actually changed the settings. However, after doing this they generally, within a few seconds, activated all of the removed information categories. It seems the users just wanted to have back what was lost. Consequently, no deeper knowledge is revealed about the users' lexicographical needs of specific information categories.

## Sub study 2

In sub study 2 *users* of LSL are surveyed by means of a web questionnaire. I wish to find out more about the users' background and their situations of use. The given conditions have the prospect of ensuring a good response to the survey. For example, the respondents are anonymous and the matter of investigation is not a sensitive one, but hopefully sufficiently important for users to be interested. Moreover, I ask for quite uncomplicated information and the respondents can be expected to be able to read and understand the questions, especially as the wordings have been adapted to the target user group of the dictionary, i.e. people who have a basic command of Swedish. Judging from the informants' responses, it seems certain wordings were still difficult to understand. A future questionnaire should perhaps be available in English and, possibly, also in some of the major immigrant Lexin languages.

Sub study 2 involves two sets of questionnaire data. Firstly, 802 questionnaires from 2011 (below Studie<sup>11</sup>, representing LSL4), and secondly, for comparison, 356 questionnaires from 2007 (below Studie<sup>07</sup>, representing LSL3). This study addresses issues like:

- Is LSL used by the intended target group?
- What lexicographical needs are expressed by the users?
- What similarities and differences are shown in the two sets of questionnaire data? Are the differences due to the revision of LSL?

Both LSL3 and LSL4 are used by the target group, but, to judge from the questionnaire data, the updated version of LSL, i.e. LSL4, is used slightly more. The typical users of LSL3 are, as was presumed from the log file analysis, (young) students 15 to 29 years old who have Swedish as their native language. They frequently use LSL and have lexicographical needs of both receptive and productive nature. Meaning is the information category most often searched for, but also spelling and examples. They wrote comments where they made clear that they are quite satisfied with the dictionary, but also indicate that many words are missing, especially words of specialist nature.

The typical LSL4 users are in their early 30s and have a mother tongue other than Swedish, possibly Arabic, Persian or Spanish. They have a university degree from their native country. They came to Sweden sometime between 2005 and 2009 and have taken the course "Swedish for beginners", which means they have a residence permit in Sweden. At the moment they are studying, perhaps with the purpose of adapting their degree to the educational system and labour market in Sweden. Moreover, the typical informants in Study<sup>11</sup>, as in Study<sup>07</sup>, are frequent users of LSL4 and have lexicographical needs of both receptive and productive character. Similarly, meaning is the most wanted information category, thereafter examples and inflection. They wrote comments saying that they are very satisfied with LSL4, but that the dictionary is in need of more (difficult) words.

Therefore, analogous to the differences in the two sets of log file data, I conclude that LSL4, to a larger extent than LSL3, is used by those who belong to the intended target group. Even here, the changes in the two groups of informants seem mainly to be related to the removal of the Swedish-English dictionary.

### Sub study 3

In sub study 3 the *individual search behaviour* of five target group users is examined more closely with the purpose of gaining a deeper knowledge about dictionary use on an individual level. This study involves only data from the year 2011. Firstly, a target group of informants is selected on the basis of their native language, year of arrival and search frequency. They amount to 249 informants. From this target group five informants are chosen whose linkings are analysed in more detail, here with fictitious names: Mervan, Elira, Luciana, Kazim and Farah. The chosen linkings are represented by informants with different native tongues, of different gender and of varying ages. Of course, no general conclusions can be drawn from such a small sample of informants, but that is not the purpose of this study. It examines dictionary use and users from yet another angle and the results from sub studies 1 and 2 can be further nuanced. Sub study 3 focuses on the following issues:

- What similarities and differences are shown in the search behaviour of these five users?
- Is there coherence between the informants' lexicographical self-image and their actual use of LSL?
- Can this method generate relevant and applicable results in connection with practical lexicography?

Thus, what the five target group users have in common is that they have a different native tongue than Swedish, they have come to Sweden at the beginning of the 21st century and they regard themselves as daily users of LSL<sup>4</sup>. However, judging from their log file activities, none of them has actually used the dictionary on a daily basis, rather four days at the most (Elira). Consequently, on this point the lexicographical self-image of the target group informants is not consistent with how often they really use LSL. What is more, they use the dictionary at different times of the day. Mervan, for example, only uses LSL<sup>4</sup> during the daytime, whereas Elira and Kazim exclusively perform look-ups in the evening.

Moreover, the target group informants' look-ups differ from each other as well as compared to the most frequent look-ups in the entire 2011 data. Here we find, for example, many verbs where the search string is the present tense, which, in fact, corresponds to the entry form in LSL. This is not the case for the target group informants' look-ups. The fact is that these look-ups have more in common with the most frequent look-ups in the 2008 data, than with the 2011 data. For instance, most look-ups consist of a noun, the search string corresponds to the words' basic forms, and homographs are rarely found. As a result, the differences between the two sets of data from 2008 and 2011 respectively, cannot be solely explained in terms of before and after the revision of LSL. The use of LSL, as appears evident from the data, uncovers a slightly more complex picture than that.

In addition, the five informants find themselves in different types of user situations. For example, Elira's look-ups occur in time-based and thematic chunks. Kazim, on the other hand, appears to find himself in a totally different user situation. He seems to get his entry words from an alphabetically organised list, and then he searches for these words in that same order. Furthermore, three of the informants have activated the removed information categories. Finally, links to audible pronunciation, pictures and film sequences are used very rarely by the target group informants.

One important aim of the sub study 3 was to evaluate the method of linking log file data with questionnaire data in an attempt to get closer to individual informants and their use of LSL. The question is whether this method can generate relevant and applicable results to practical lexicography. I find the answer to be, yes. Combining methods and triangulating data have indeed resulted in a more complete and balanced picture of the use of LSL, and perhaps even of Lexin in general. We have looked at dictionary use from three different perspectives. Firstly, the log file analysis produced a general view of the users' most frequent look-ups. Secondly, the questionnaire survey resulted in an important, though generalised, user profile. Lastly, the data triangulation (i.e. the linking of the two types of data) made it possible to follow individual users and their dictionary activities at close quarters.

## Future LSL and further research

The majority of user studies are based on non-probability samples. This study is no exception. I have tried to arrive at the most appropriate selection strategy possible based on given prerequisites. As far as the questionnaire is concerned, the informants of Study<sup>11</sup> seem to be a representative sample of LSL users, at least in one respect: the (non-Swedish) languages with the greatest number of native speakers among the informants coincide with the Lexin-languages whose dictionary has the highest number of look-ups (in the data collection period). The results of the study may still not be generalised to describe the use of other dictionaries, and strictly speaking, not even the use of LSL. Nevertheless, studies based on non-probability samples can contribute with relevant information about possible tendencies that may apply to larger groups and provide a basis for further development of earlier studies and theories.

Hopefully, this thesis is a contribution to the more theoretical user research, and the results will provide valuable knowledge for practical lexicography. The study shows, among other things, that the improvements in connection with the update are appreciated by the informants. If possible, one can go further in the same direction. Lists of unsuccessful searches can generate suggestions for new lemmas. A good idea would also be to make *all* inflections searchable. In addition, the fuzzy search tool should not only be based on edit distance, but also on pronunciation related misspellings. Finally, the picture theme would benefit from a review and update.

Future research can be conducted into various directions. It would, for example, be interesting to replicate the study and focus on the users who do not have a Lexin-language as their mother tongue and, hence, *must* use the LSL. There are about 200 minority languages in Sweden today. Understandably, the Lexin project cannot possibly prepare bilingual dictionaries in all these languages, although several of them have a very large number of speakers. Evidently, LSL fulfils an extremely important function. Another way to dig deeper into the use of LSL is to resort to other methods. More profound and underlying aspects of dictionary use could for example be achieved with interviews. In addition, the data at hand could also, periodically, be further extended with new log files (preferably every three years) which would eventually, provide a comprehensive basis for comparisons over time.

Another approach would be to broaden the research perspective even more. My investigation touches several nearby areas of research, not least Swedish as a second language but also applied information technology and educational science. A potential future research topic adjacent to all areas is that of language learning for recently arrived immigrants. The syllabus for the "Swedish for beginners" course states for example that students are expected to develop

their skills in using various digital tools related to learning. In this context electronic dictionaries in general, and LSL or Lexin in particular, serve a most important function.

The research ideas presented so far have focused on more empirical studies of LSL. Naturally, the research should be broadened empirically as well as theoretically. As a suggestion, it would be interesting to examine more closely the various metalexicographical theories on user customization, and the actual consequences on a microstructural level in the design of the final product, i.e. the dictionary.



## Referenser

- Abel, Andrea, Chiara Vettori & Natascia Ralli (eds) 2014. *Proceedings of the XVI EURALEX International Congress: The User in Focus. 15-19 July 2014, Bolzano/Bozen*. Bolzano/Bozen: EURAC research.
- Abrahamsson, Niclas 2009. *Andraspråksinläring*. Lund: Studentlitteratur.
- Allén, Sture 1972. *Tiotusen i topp. Ordfrekvenser i tidningstext*. (Data linguistica 6.) Stockholm: Almqvist & Wiksell.
- Allén, Sture 1977. *Våra viktiga ord. Basordlista med utbytesord*. Stockholm: Esselte studium.
- Almind, Richard 2008. Søgemønstre i logfiler. *LexicoNordica* 15, s. 33–55.
- Atkins, B. T. Sue (ed.) 1998. *Using Dictionaries. Studies of Dictionary Use by Language Learners and Translators*. (Lexicographica Series Maior 88.) Tübingen: Max Niemeyer Verlag.
- Atkins, B.T. Sue & Krista Varantola 1998a. Language Learners Using Dictionaries: The Final Report on the EURALEX/AILA Research Project on Dictionary Use. In: B.T. Sue Atkins (ed.), s. 21–81.
- Atkins, B.T. Sue & Krista Varantola 1998b. Monitoring Dictionary Use. In: Atkins, B.T. Sue (ed.), s. 83–122.
- Atkins, B. T. Sue & Michael Rundell 2008. *The Oxford Guide to Practical Lexicography*. Oxford: Oxford University Press.
- Barnhart, Clarence L. 1962. Problems in Editing Commercial Monolingual Dictionaries. In: Householder, Fred W. & Sol Saporta (eds), *Problems in Lexicography. Report of the Conference on Lexicography Held at Indiana University, November 11–12, 1960*. Bloomington: Indiana University, s. 161–181.

- Beijer, Mai & Kiros Fre Woldu 1997. *Detta är LEXIN. Lexikon för invandrare. Lägesbeskrivning våren 1997*. Stockholm: Skolverket.
- Béjoint, Henri 1981. The Foreign Student's Use of Monolingual English Dictionaries: A Study of Language Needs and Reference Skills. *Applied Linguistics* 2(3), s. 207–222.
- Béjoint, Henri 2000. *Modern Lexicography. An introduction*. Oxford: Oxford University Press.
- Béjoint, Henri 2010. *The Lexicography of English: from origins to present*. Oxford: Oxford University Press.
- Benbow, Timothy 1990. Report on the New Oxford English Dictionary User Survey. *International Journal of Lexicography* 3(3), s. 155–203.
- Berg, Sture 1978. *Olika lika ord: svenskt homograflexikon*. (Data Linguistica 12.) Stockholm: Almqvist & Wiksell.
- Berg, Sture & Yvonne Cederholm 2001. Att hålla på formerna. Om framväxten av Svensk morfologisk databas. I: Allén, Sture, Sture Berg, Sven-Göran Malmgren, Kerstin Norén & Bo Ralph (red.), *Gäller stam, suffix och ord. Festskrift till Martin Gellerstam den 15 oktober 2001*. (Meijerbergs arkiv för svensk ordforskning 29.) Göteborg, s. 58–69.
- Berg, Sture, Louise Holmer & Ann-Kristin Hult 2008a. SAOL Plus – a new Swedish electronic dictionary. In: Bernal, Elisenda & Janet De Cesaris (eds), s. 291–296.
- Berg, Sture, Louise Holmer, Ann-Kristin Hult, Susanne Lindstrand & Christian Sjögreen 2008b. Sextiofemårigare? Kring adjektivformer i SAOL Plus. I: Jóhannesson, Kristinn et al. (red.), s. 42–49.
- Bergenholtz, Henning 1994. Grundordbog til genbrug. *LexicoNordica* 1, s. 243–247.
- Bergenholtz, Henning 2007. Kvaliteten af de leksikografiske definitioner i ordbøger, leksika og encyklopædier. *LexicoNordica* 14, s. 11–34.
- Bergenholtz, Henning & Sven-Göran Malmgren 2005. Förord. *LexicoNordica* 12, s. 1–5.
- Bergenholtz, Henning & Sven Tarp 2003. Two opposing theories: On H. E. Wiegand's recent discovery of lexicographic functions. *Hermes – Journal of Linguistics* 31, s. 171–196.
- Bergenholtz, Henning & Mia Johnsen 2005. Log Files as a Tool for Improving Internet Dictionaries. *Hermes – Journal of Linguistics* 34, s. 117–141.
- Bergenholtz, Henning & Mia Johnsen 2007. Log Files Can and Should Be Prepared for a Functionalistic Approach. *Lexikos* 17, s. 1–20.
- Berkov, Valeriy 1996. Passive vs. Active Dictionary. A revision. In: Gellerstam, Martin et al. (eds), s. 547–550.
- Bernal, Elisenda & Janet De Cesaris (eds) 2008. *Proceedings of the XIII EURALEX International Congress, (Barcelona 15–19 July 2008)*. (Sèrie



- activitats 20.) Barcelona: Universitat Pompeu Fabra, Institut universitari de lingüística aplicada.
- Bjørneset, Tove 2016. LEXIN i Norge – Hva sier brukerne? I: Gudiksen, Asgerd & Henrik Hovmark (red.), s. 89–99.
- Bjørneset, Tove & Ásta Svavarsdóttir 2005. Nord-Lexin. I: Fjeld, Ruth Vatvedt & Dagfinn Worren (red.), *Nordiske studiar i leksikografi 7. Rapport frå Konferanse om leksikografi i Norden, Volda 22.–24. mai 2003*. Oslo, s. 52–66.
- Bogaards, Paul 1990. Où cherche-t-on dans le dictionnaire? *International Journal of Lexicography* 3(2), s. 79–102.
- Bogaards, Paul 1991. Word frequency in the search strategies of French dictionary users. *Lexicographica* 7, s. 202–212.
- Bogaards, Paul 1998. What type of words do language learners look up? In: Atkins, B. T. Sue (ed.), *Using Dictionaries. Studies of Dictionary Use by Languages Learners and Translators*. (Lexicographica Series Maior 88.) Tübingen: Max Niemeyer Verlag, s. 151–157.
- Bogaards, Paul 2003. [Recension av] Ursula Wingate: The Effectiveness of Different Learner Dictionaries. An Investigation into the Use of Dictionaries for Reading Comprehension by Intermediate Learners of German. *International Journal of Lexicography* 16(2), s. 200–201.
- Bogaards, Paul 2010. Lexicography: Science without Theory? In: de Schryver, Gilles-Maurice (ed.), *A Way with Words: Recent Advances in Lexical Theory and Analysis – A Festschrift for Patrick Hanks*. (Menha Linguistics Series.) Kampala: Menha Publishers, s. 313–332.
- Bogaards, Paul & Willem A. van der Kloot 2001. The use of grammatical information in learners' dictionaries. *International Journal of Lexicography* 14(2), s. 97–121.
- Borin, Lars, Markus Forsberg & Johan Roxendal 2012. Korp – the corpus infrastructure of Språkbanken. In: Calzolari, Nicoletta et al. (eds), s. 474–478.
- Boyd, Sally & Kari Fraurud 2010. Challenging the homogeneity assumption in language variation analysis. Findings from a study of multilingual urban spaces. In: Auer, Peter & Jürgen Erich Schmidt (eds), *Language and Space – an International Handbook of Language Variation*. Berlin: De Gruyter Mouton, s. 686–706.
- Byrman, Gunilla, Hans Lindquist & Magnus Levin (red.) 2000. *Korpusar i forskning och undervisning. Rapport från ASLA:s höstsymposium, Växjö, 11–12 november 1999*. Uppsala: ASLA.
- Calzolari, Nicoletta, Khalid Choukri, Thierry Declerck, Mehmet Uğur Doğan, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odiijk & Stelios Piperidis (eds) 2012. *Proceedings of the Eighth International Confer-*

- ence on Language Resources and Evaluation, LREC 2012, Istanbul, Turkey, 23–25 May. Istanbul: European language resources distribution agency.
- Collins Cobuild English Dictionary 1987. Sinclair, John M. (ed.). London: Harper Collins.
- Cruse, Alan D. 1986. *Lexical Semantics*. Cambridge: Cambridge University Press.
- Denscombe, Martyn 2016. *Forskningshandboken – för småskaliga forskningsprojekt inom samhällsvetenskaperna*. 3 uppl. Lund: Studentlitteratur. Översatt av Per Larsson.
- Domeij, Richard & Ola Karlsson 2013. Delaktig i det digitala samhället – ett flerspråkigt perspektiv. I: Hyltenstam, Kenneth & Inger Lindberg (red.), s. 519–543.
- Dolezal, Fredric Thomas & Don R. McCreary 1999. *Pedagogical Lexicography Today. A Critical Bibliography on Learners' Dictionaries with Special Emphasis on Language Learners and Dictionary Users*. (Lexicographica Series Maior 96.) Tübingen: Max Niemeyer Verlag.
- Doughty, Catherine J. & Michael H. Long (eds) 2003. *The Handbook of Second Language Acquisition*. Oxford: Blackwell Publishing.
- Dziemianko, Anna 2010. Paper or Electronic? The Role of Dictionary Form in Language Reception, Production and the Retention of Meaning and Collocations. *International Journal of Lexicography* 23(3), s. 257–273.
- Dziemianko, Anna 2012. On the use(fulness) of paper and electronic dictionaries. In: Granger, Sylviane & Magali Paquot (eds), s. 319–341.
- Dykstra, Anne & Tanneke Schoonheim (eds) 2010. *Proceedings of the XIV Euralex International Congress, Leeuwarden, 6–10 July 2010*. Leeuwarden/Ljouwert: Fryske Akademy.
- Enström, Ingegerd 1996. *Klara verba. Andraspråkslärares verbanvändning i svenskan*. (Nordistica Gothoburgensia 18.) Göteborg: Acta Universitatis Gothoburgensis.
- Enström, Ingegerd 2013a. *Att undervisa om ord. En studie av lärares behov, förutsättningar och kunskapsutveckling*. (GU-ISS Forskningsrapporter från Institutionen för svenska språket.) Göteborg: Göteborgs universitet, Institutionen för svenska språket.
- Enström Ingegerd 2013b. Ordförråd och ordinläring – med särskilt fokus på avancerade inlärare. I: Hyltenstam, Kenneth & Inger Lindberg (red.), s. 169–195.
- Enström, Ingegerd 2016. *Ordens värld. Svenska ord – struktur och inläring*. 2 uppl. Uppsala: Hallgren & Fallgren.
- Farø, Ken & Henrik Gottlieb 2007. Socioleksikografi – eller: ”Det sekundære og afledte er ordbøgerne”. Ny dansk disputats. *LexicoNordica* 14, s. 189–199.

- Farø, Ken & Henrik Lorentzen 2009. De oversete og mishandlede ordforbindelser – hvilke, hvor og hvorfor? *LexicoNordica* 16, s. 75–102.
- Fjeld, Ruth Vatvedt & Lars S. Vikør 2008. *Ord og ordbøker. Ei innføring i leksikologi og leksikografi*. Kristiansand: Høyskoleforlaget.
- Fjeld, Ruth Vatvedt & Julie Mathilde Torjusen (eds) 2012. *Proceedings of the 15th EURALEX International Congress 7–11 August 2012*. Oslo: Department of Linguistics and Scandinavian Studies, University of Oslo.
- Folkets lexikon*. <<http://folkets-lexikon.csc.kth.se/folkets>>. Hämtat 19 november 2016.
- Franker, Qarin 2013. Att utveckla litteracitet i vuxen ålder – alfabetisering i en flerspråkig kontext. I: Hyltenstam, Kenneth & Inger Lindberg (red.), s. 771–816.
- Frick, Nils & Sten Malmström 1983 [1976]. *Språkklyftan. Hur 700 ord förstås och missförstås*. Stockholm: Tiden.
- Garlén, Claes 1985. Lexikon för invandrare. *Språkvård* 1985(2), s. 21–25.
- Gass, Susan M. & Alison Mackey (eds) 2011. *The Routledge Handbook of Second Language Acquisition*. New York: Routledge.
- Gellerstam, Martin 1988. Verb syntax in a dictionary for second-language learning. I: Gellerstam, Martin (ed.), *Studies in Computer-Aided Lexicology*. (Data Linguistica 18.) Stockholm: Almqvist & Wiksell, s. 103–123.
- Gellerstam, Martin 1995. Kommentar till ordboken. I: *Svenska ord med uttal och förklaringar*. 3 uppl., s. 762–776.
- Gellerstam, Martin 1999. LEXIN – lexikon för invandrare. *LexicoNordica* 6, s. 3–18.
- Gellerstam, Martin, Jerker Järborg, Sven-Göran Malmgren, Kerstin Norén, Lena Rogström & Catarina Røjder Papmehl (eds) 1996. *Euralex '96 Proceedings I–II. Papers submitted to the Seventh EURALEX International Congress on Lexicography in Göteborg, Sweden*. Göteborg: Göteborg University, Department of Swedish.
- Golden, Anne 2006. Minoritets elever og ordforrådet i lærebøker. I: Lindberg, Inger & Karin Sandwall (red.), *Språket och kunskapen – att lära på sitt andraspråk i skola och högskola*. (Rapporter om svenska som andraspråk (ROSA) 7.) Göteborg: Institutet för svenska som andraspråk, s. 115–127.
- Golden, Anne 2011. *Ordforråd, ordbruk og ordlæring*. 2 uppl. Oslo: Gyldendal.
- Golden, Anne & Anne Hvenekilde 1983. *Rapport fra Prosjektet Lærebokspråk*. Oslo: Senteret for språkpedagogikk, Universitetet i Oslo.
- Gouws, Rufus H. 2005. Meilensteine auf dem historischen Weg der Metalexikographie. *Lexicographica* 21, s. 158–178.

- Gouws, Rufus H. 2012. Theoretical Lexicography and the *International Journal of Lexicography*. *International Journal of Lexicography* 25(4), s. 450–463.
- Granger, Sylviane & Magali Paquot 2010. *e-Lexicography in the 21<sup>st</sup> Century: New Challenges, New Applications. Proceedings of the eLex 2009, Louvain-la-Neuve, 22–24 October 2009*. Louvain: Université catholique de Louvain, Centre for English Corpus Linguistics.
- Granger, Sylviane, & Magali Paquot (eds) 2012. *Electronic Lexicography*. Oxford: Oxford University Press.
- Grinstead, Wren Jones 1915. An experiment in the learning of foreign words. *Journal of Educational Psychology* 6(4), s. 242–245.
- Gudiksen, Asgerd & Henrik Hovmark (red.) 2016. *Nordiske Studier i Leksikografi 13. Rapport fra 13. Konference om Leksikografi i Norden, København 19.–22. maj 2015*. København: Københavns Universitet.
- Gull, Maria & Britt Klintonberg 2004. +46. *Kurs C. +lärare*. Stockholm: Almqvist & Wiksell.
- Hammarberg, Björn 2013. Teoretiska ramar för andraspråksforskning. I: Hytlenstam, Kenneth & Inger Lindberg (red.), s. 27–84.
- Hannesdóttir, Anna Helga 1998. *Lexikografihistorisk spegel. Den enspråkiga svenska lexikografins utveckling ur den tvåspråkiga*. (Meijerbergs arkiv för svensk ordforskning 23.) Göteborg: Göteborgs universitet: Meijerbergs institut för svensk etymologisk forskning.
- Hatherall, Glyn 1984. Studying dictionary use: some findings and proposals. In: Hartmann, R.R.K. (ed.), s. 183–189.
- Hartmann, Reinhard R.K. 1983. The bilingual learner's dictionary and its uses. *Multilingua* 2(4), s. 195–201.
- Hartmann, Reinhard R.K. (ed.) 1984a. *LEXeter '83 Proceedings: Papers from International Conference on Lexicography at Exeter, 9–12 Sept. 1983*. (Lexicographica Series Maior 1.) Tübingen: Niemeyer.
- Hartmann, Reinhard R.K. 1984b. Preface. In: Hartmann, R.R.K. (ed.).
- Hartmann, Reinhard R.K. 1990. A quarter of a century's lexicographical conferences. In: Magay, Tamás & Judit Zigány (eds), *BudaLEX'88. Papers from the Third International EURALEX Congress*, Budapest, 4–9 September 1988, Budapest: Akademiai Kiado, s. 569–575.
- Hausmann, Franz Josef 1977. *Einführung in die Benutzung der neufranzösischen Wörterbücher*. (Romanistische Arbeitshäfte.) Tübingen: Max Niemeyer Verlag.
- Hausmann, Franz Josef, Oskar Reichmann, Herbert Ernst Wiegand & Ladislav Zgusta (eds) 1989–1991. *Wörterbücher / Dictionaries / Dictionnaires. Ein internationales Handbuch zur Lexikographie / An International Encyclopedia of Lexicography / Encyclopédie internationale de lexicogra-*

- phie I–III.* (Handbücher zur Sprach- und Kommunikationswissenschaft 5.1–5.3.) Berlin/New York: Walter de Gruyter.
- Henriksen, Birgit 1995. Hvad vil det sige at kunne et ord? Ordförståelse og ordbeherskelse. *Sprogforum. Tidsskrift for sprog- og kulturpedagogik* 1(3), s. 12–18.
- Holmboe, Henrik 1999. Det danske bidrag till NORDLEXIN. *LexicoNordica* 6, s. 3–17.
- Holme, Idar Magne & Bernt Krohn Solvang 1997. *Forskningsmetodik. Om kvalitativa och kvantitativa metoder.* 2 uppl. Lund: Studentlitteratur. Översatt av Björn Nilson.
- Holmer, Louise & Emma Sköldberg 2014. Appifiering till allas lycka? Om danska ordboksappar med särskilt fokus på DDO. *LexicoNordica* 21, s. 235–252.
- Holmer, Louise, Ann-Kristin Hult & Emma Sköldberg 2015. Spell-checking on the fly? On the use of a Swedish dictionary app. In: Kosem, Iztok, Miloš Jakubiček, Jelena Kallas & Simon Krek (eds.), s. 356–371.
- Holmer, Louise & Emma Sköldberg 2016. Synonymer.se i fokus – om användningen av en svensk ordbokssajt. I: Gustafsson, Anna W., Lisa Holm, Katarina Lundin, Henrik Rahm & Mechtild Tronnier (red.), *Svenskans beskrivning 34. Förhandlingar vid trettiofjärde sammankomsten, Lund den 22–24 oktober 2014.* Lund, s. 215–228.
- Holtti, David 2003. *Sambälle Opinion Massmedia Västra Götaland 2003.* Göteborg: SOM-institutet, s. 503–517.
- Hornby, Albert Sidney 1965. Some problems of lexicography. *English Language Teaching* XIX, s. 104–110.
- Householder, Fred W. 1962. Summary report. In: Householder, Fred W. & Sol Saporta (eds), *Problems in Lexicography. Report of the Conference on Lexicography Held at Indiana University, November 11–12, 1960.* Bloomington, s. 279–282.
- Hovdenak, Marit 2008. Dei norske LEXIN-ordbøkene. *LexicoNordica* 15, 219–234.
- Hulstijn, Jan H. & B. T. Sue Atkins 1998. Empirical Research on Dictionary Use in Foreign-Language Learning: Survey and Discussion. In: Atkins, B. T. Sue (ed.), s. 7–19.
- Hult, Ann-Kristin 2008a. Användarna bakom loggfilerna. Redovisning av en webbenkät i Lexin online Svenska ord. *LexicoNordica* 15, s. 73–91.
- Hult, Ann-Kristin 2008b. Från ord till handling. En studie i ordboksanvändning på nätet. I: Svavarsdóttir, Ásta, Guðrun Kvaran, Gunnlaugur Ingólfsson & Jón Hilmar Jónsson (red.), *Nordiske studier i leksikografi 9. Rapport fra konferanse om leksikografi i Norden, Akureyri 22.–26 maj 2007.* Reykjavík, s. 243–257.

- Hult, Ann-Kristin 2012. Old and New Study Methods Combined. Linking Web Questionnaire with Log Files from the *Swedish Lexin Dictionary*. In: Fjeld, Ruth Vatvedt & J. M. Torjusen (eds.), s. 922–928.
- Hult, Ann-Kristin 2014. The Authentic Voices of Dictionary Users – Viewing Comments on an Online Learner’s Dictionary Before and After Revision. In: Abel, Andrea, Chiara Vettori & Natascia Ralli (eds), s. 237–247.
- Hult, Ann-Kristin, Sven-Göran Malmgren & Emma Sköldberg 2010. Lexin – a report from a recycling lexicographic project in the North. In: Dykstra, Anne & Tanneke Schoonheim (eds), s. 800–809.
- Hyltenstam, Kenneth & Inger Lindberg (red.) 2013. *Svenska som andraspråk – i forskning, undervisning och samhälle*. 2 uppl. Lund: Studentlitteratur.
- Hyltenstam, Kenneth & Inger Lindberg 2013. Förord. I: Hyltenstam, Kenneth & Inger Lindberg (red.). Lund: Studentlitteratur, s. 7–19. *International Journal of Lexicography*. Oxford: Oxford University Press (1988–).
- Japac, Lilli, Antti Ahtiainen, Jan Hörngren, Håkan Lindén, Lars Lyberg & Per Nilsson 1997. *Minska bortfallet*. Örebro: Statistiska Centralbyrån.
- Jóhannesson, Kristinn, Hans Landqvist, Aina Lundqvist, Lena Rogström, Emma Sköldberg & Barbro Wallgren-Hemlin (red.) 2008. *Nog ordat. Festskrift till Sven-Göran Malmgren den 25 april 2008*. (Meijerbergs arkiv för svensk ordforskning 34.) Göteborg: Göteborgs universitet, Meijerbergs institut för svensk etymologisk forskning.
- Johnson, Samuel 1747. *The Plan of an English Dictionary. Preface*. Lynch, Jack (ed.). <<https://andromeda.rutgers.edu/~jlynch/Texts/plan.html>>. Hämtat 29 februari 2016.
- Josephson, Olle 1982. *Svåra ord. En undersökning av förståelsen av 153 ord från ekonomiska, sociala och politiska sammanhang*. (Meddelanden från Institutionen för nordiska språk, MINS 11.) Stockholm: Stockholms universitet.
- Järborg, Jerker 2007. Om ord och ordkunskap. I: Lindberg, Inger & Sofie Johansson Kokkinakkis (red.), s. 61–100.
- Kann, Viggo 2016. *Algorithms, Data Structures and Complexity: Labb 2*. <<https://www.kth.se/social/course/DD1352/subgroup/adk13/page/lab-2-5>>. Hämtat 9 november 2016.
- Karlsson, Fred 1992. SWETWOL: A comprehensive Morphological Analyser for Swedish. *Nordic Journal of Linguistics* 15, 1–45.
- Kilgariff, Adam, Milos Husak, Katy McAdam, Michael Rundell & Pavel Rychly 2008. GDEX: Automatically Finding Good Dictionary Examples in a Corpus. In: Elisenda Bernal & Janet DeCesaris (eds), s. 1421–1432.
- Kipfer, Barbara Ann 1987. Dictionaries and the intermediate student: Communicative needs and the development of user reference skills. In: Cowie,

- Anthony (ed.), *The dictionary and the language learner: Papers from EUR-ALEX seminar at the University of Leeds, 1–3 April 1985*. (Lexicographica Series Maior 17.) Tübingen: Max Niemeyer Verlag, s. 44–54.
- Kobayashi, Chiho 2008. The use of pocket electronic and printed dictionaries: A mixed-method study. In: Bradford Watts, K., T. Muller & M. Swanson (eds.), *JALT 2006 Conference Proceedings*. Tokyo: JALT, s. 769–783.
- Korp 2016. <[https://spraakbanken.gu.se/korp/#?stats\\_reduce=word&cqp=&corpus=gp2008,gp2011](https://spraakbanken.gu.se/korp/#?stats_reduce=word&cqp=&corpus=gp2008,gp2011)>. Hämtat 30 juni 2016.
- Koyama, Toshiko & Osamu Takeuchi 2004. Comparing electronic and printed dictionaries: How the difference effected EFL learning. *JACET Bulletin* 38, s. 33–46.
- Kosem, Iztok, Jelena Kallas, Polona Gantar, Simon Krek, Margit Langemets & Maria Tuulik (eds) 2013. *Electronic lexicography in the 21<sup>st</sup> century: thinking outside the paper. Proceedings of the eLex 2013 conference, 17–19 October, Tallinn, Estonia*. Ljubljana/Tallinn: Trojina, Institute for Applied Slovene Studies/Eesti Keele Instituut.
- Kromann, Hans-Peder, Theis Riiber & Poul Rosbach. 1984a. 'Active' and 'passive' bilingual dictionaries: the Ščerba concept reconsidered. In: Hartmann, Reinhard R.K. (ed.), s. 207–215.
- Kromann, Hans-Peder, Theis Riiber & Poul Rosbach 1984b. Überlegungen zu Grundfragen der zweisprachigen Lexikographie. In: Wiegand, Herbert Ernst (ed.), *Studien zur neuhochdeutschen Lexikographie V*. Hildesheim/New York: Olms, s. 159–238.
- Kwary, Deny Arnos 2012. Adaptive Hypermedia and User-Oriented Data for Online Dictionaries: A Case Study on an English Dictionary of Finance for Indonesian Students. *International Journal of Lexicography* 25(1), s. 30–49.
- Källström, Roger 2012. *Svenska i kontrast. Tvärspråkliga perspektiv på svensk grammatik*. Lund: Studentlitteratur.
- Køhler Simonsen, Henrik 2011. User Consultation Behaviour in Internet Dictionaries: An Eye-Tracking Study. *Hermes – Journal of Language and Communication Studies* 46, s. 75–101.
- Køhler Simonsen, Henrik 2014. Brugerne er allerede mobile! I: Fjeld, Ruth Vatvedt & Marit Hovdenak (red.), *Nordiske studier i leksikografi 12. Rapport fra Konferanse om leksikografi i Norden. Oslo 13.–16. august 2013*. Oslo: Novus, s. 416–429.
- Køhler Simonsen, Henrik 2015. Den mobile ordbogsbruger. I: Gudiksen, Asgerd & Henrik Hovmark (red.), s. 227–238.
- Laufer, Batia 1997a. The lexical plight in second language reading: Words you don't know, words you think you know, and words you can't guess.

- In: Coady, James & Thomas N. Huckin (eds), *Second Language Vocabulary Acquisition*. Cambridge: Cambridge University Press, s. 20–34.
- Laufer, Batia 1997b. What's in a word that makes it hard or easy: some intralexical factors that affect the learning of words. In: Schmitt, Norbert & Michael McCarthy (eds), s. 140–155.
- Labov, William 1972. *Sociolinguistic patterns*. (University of Pennsylvania publications in conduct and communication 4.) Philadelphia: University of Pennsylvania Press.
- Lakoff, George & Mark Johnson 2003 [1980]. *Metaphors we live by*. Chicago: University of Chicago Press.
- Lernia 2016. *Svenska för invandrare*. <https://www.lernia.se/utbildning/utbildningstyper/Sfi---svenska-for-invandrare>>. Hämtat 20 november 2016.
- Lew, Robert 2002. Questionnaires in dictionary use research: A reexamination. In: Braasch, Anna och Claus Povlsen (eds), *Proceedings of the Tenth EURALEX International Congress. EURALEX 2002, Copenhagen, Denmark, August 13-17*. Vol. 1. Copenhagen: Center for Sprogteknologi, Copenhagen University, s. 267–271.
- Lew, Robert 2004. *Which Dictionary for Whom? Receptive Use of Bilingual, Monolingual and Semi-Bilingual Dictionaries by Polish Learners of English*. Poznań: Motivex.
- Lew, Robert 2008. Lexicographic Functions and Pedagogical Lexicography: Some Critical Notes on Sven Tarp's "Lexicography in the Borderland between Knowledge and Non-Knowledge". In: Iwan, Krystyna & Iwona Korpaczewska (eds), *Przegląd Humanistyczny. Pedagogika. Politologia. Filologia*. Szczecin: Szczecińska Szkoła Wyższa Collegium Balticum, s. 114–123.
- Lew, Robert 2010. Multimodal Lexicography: The Representation of Meaning in Electronic Dictionaries. *Lexikos* 20, s. 290–306.
- Lew, Robert 2013. Online dictionary skills. In: Kosem, Iztok et al. (eds.), s. 16–31.
- Lew, Robert under utg. Opportunities and limitations of user studies. In: Tiberius, Carole & Carolin Müller-Spitzer (eds), *Research into dictionary use / Wörterbuchbenutzungsforschung. 5. Arbeitsbericht des wissenschaftlichen Netzwerks „Internetlexikografie“*. Mannheim: Institut für Deutsche Sprache, s. 5–15.
- Lew, Robert 2015. Research into the use of online dictionaries. *International Journal of Lexicography* 28(2), s. 232–253.
- Lew, Robert and Roger Mitton. 2011. 'Not the Word I Wanted? How Online English Learners' Dictionaries Deal with Misspelled Words' in Kosem, Iztok and Karmen Kosem (eds.), *Electronic Lexicography in the 21st Century: New Applications for New Users. Proceedings of eLex 2011, Bled, 10-12 November 2011*. Ljubljana: Trojina, Institute for Applied Slovene Studies, s. 165-174.



- Lew, Robert & Gilles-Maurice de Schryver 2014. Dictionary users in the digital revolution. *International Journal of Lexicography* 27(4), s. 341–359. *Lexicographica = Lexicographica. International Annual for Lexicography*. Tübingen: Niemeyer (1984–).
- LexicoNordica = LexicoNordica. Tidskrift om leksikografi i Norden*. Utgitt av Nordisk forening for leksikografi i samarbeid med Nordisk språksekretariat (1994–).
- Lexin Frågor och svar 2016. *Varför har engelska Lexin försvunnit och ersatts av Folkets lexikon?* <<http://lexin.nada.kth.se/lexin/#about=1;main=3>>. Hämtat 25 juli 2016.
- Lexinstatistik 2016. *Statistik för användning av Lexin* <<http://lexin.nada.kth.se/lexin/statistik.html>>. Hämtat 2 februari 2016.
- Lexins svenska lexikon* [Svenska ord] 2011. 4 uppl. <<http://lexin.nada.kth.se/lexin/>>. Hämtat 9 november 2016.
- Lieberman, Anatoly 1990. Valerij P. Berkov. ”Russko-norvežskij slovar’/Russisk-norsk ordbok”. Review. *Scandinavian Studies* 62(2), s. 230–231.
- Lindberg, Inger & Sofie Johansson Kokkinakkis 2007 (red.). *OrdiL. En korpusbaserad kartläggning av ordförrådet i läromedel för grundskolans senare år*. (Rapporter om svenska som andraspråk (ROSA) 8.) Göteborg: Göteborgs universitet, Institutionen för svenska språket, Institutet för svenska som andraspråk.
- Lindberg, Inger 2007. Forskning om läromedelsspråk och ordförrådsutveckling. I: Lindberg, Inger & Sofie Johansson Kokkinakkis (red.), s. 13–60.
- Ljunggren, Magnus & Sven-Göran Malmgren 2007. En ny svensk-rysk och rysk-svensk ordbok. *LexicoNordica* 14, s. 225–238.
- Ljungman, Sofie 2012. *Varför är svåra ord svåra? En undersökning av sexton ord från högskoleprovets ORD-del*. Magisteruppsats. Göteborg: Institutionen för svenska språket. <<http://hdl.handle.net/2077/31245>>. Hämtat 9 november 2016.
- Loenheim, Lisa 2012. [Recension av] Lexins svenska lexikon. *LexicoNordica* 19, s. 295–316.
- Lorentzen, Henrik & Lars Trap-Jensen 2006. Den Danske Ordbog: en ordbog for eksperter, godtfolk og almindeligt interesserede. *LexicoNordica* 13, s. 197–208.
- Lorentzen, Henrik & Liisa Theilgaard 2012. Online dictionaries – how do users find them and what do they do once they have? I: Fjeld, Ruth Vatvedt & Julie Mathilde Torjusen (red.), s. 654–660.
- LSL3 = [Internetversionen av] *Svenska ord: med uttal och förklaringar* 1995. 3 uppl. Stockholm: Myndigheten för skolutveckling/Liber.
- LSL4 = *Lexins svenska lexikon* [Svenska ord] 2011. 4 uppl. <<http://lexin.nada.kth.se/lexin/>>. Hämtat 9 november 2016.

- McMillan dictionary* 2016 [2002]. <<http://www.macmillandictionary.com/learn/clear-definitions.html>>. Hämtat 9 november 2016.
- Malmgren, Sven-Göran 1994. Språkprovets form och funktion i svenska betydelseordböcker från Östergrens Nusvensk ordbok till Svensk ordbok. *LexicoNordica* 1, s. 107–117.
- Malmgren, Sven-Göran 1995. Svar till Knud Troels Thomsen. *LexicoNordica* 2, s. 300–302.
- Malmgren, Sven-Göran 1999. De svenska LEXIN-ordböckerna ur ett (ställföreträdande) användarperspektiv. *LexicoNordica* 6, s. 79–90.
- Malmgren, Sven-Göran 2000. Korpusar i ordboksarbete och grammatikforskning. I: Byrman, Gunilla m.fl. (red.), s. 192–207.
- Malmgren, Sven-Göran 2003. Begå eller ta självmod? Om svenska kollokationer och deras förändringsbenägenhet. *Språk och Stil* 13(1), s. 123–168.
- Malmgren, Sven-Göran 2009. On production-oriented information in Swedish monolingual defining dictionaries. In: Nielsen, Sandro & Sven Tarp (eds), *Lexicography in the 21st century. In Honour of Henning Bergenholtz*. Amsterdam/Philadelphia: John Benjamins, s. 93–102.
- Malmgren, Sven-Göran 2012. Från Svenska ord 3 (Lexin) till Svenska ord 4. I: Eaker, Birgit, Lennart Larsson & Anki Mattisson (red.), *Nordiska studier i lexikografi 11. Rapport från Konferensen om lexikografi i Norden, Lund 24–27 maj 2011*. Lund, s. 454–465.
- Malmgren, Sven-Göran & Maria Toporowska Gronostaj 2009. Valensbeskrivning i svenska ordböcker – och några andra. *LexicoNordica* 16, s. 181–196.
- von Martens, Monica & Kristian Blenselius 2016. Morfologi för lärarstudenter. *LexicoNordica* 23, s. 105–118.
- Martin, Willy 2007. Government policy and the planning and production of bilingual dictionaries: The ‘Dutch’ approach as a case in point. *International Journal of Lexicography* 20(3), s. 221–237.
- Martin, Willy & Anne Tamm 1996. OMBI: An Editor for Constructing Reversible Lexical Databases. In: Gellerstam, Martin et al. (eds), s. 675–687.
- Martola, Nina 2002. Två nya lättillgängliga ordböcker. *LexicoNordica* 9, s. 253–270.
- Měchura, Michal Boleslav 2008. Giving them what they want: search strategies for electronic dictionaries. In: Bernal, Elisenda & Janet DeCesaris (eds), s. 1295–1300.
- Melka, Francine 1997. Receptive vs. productive aspects of vocabulary. In: Schmitt, Norbert & Michael MacCarthy (eds), s. 84–102.
- Moon, Rosamund 2007. Sinclair, lexicography, and the Cobuild project: The application of theory. *International Journal of Corpus Linguistics* 12(2), 159–181.

- Müller-Spitzer, Carolin 2013. Contexts of dictionary use. In: Kosem, Iztok et al. (eds), s. 1–15.
- Nagy, William 1997. On the role of context in first- and second-language vocabulary learning. In: Schmitt, Norbert & Michael McCarthy (eds), s. 64–83.
- Nation, Paul 2013. *Learning Vocabulary in Another Language*. 2 uppl. Cambridge: Cambridge University Press.
- Nationalencyklopedin* 2016. <<http://www.ne.se/uppslagsverk/encyklopedi>>. Hämtat 13 november 2016.
- Natur och Kulturs Svenska Ordbok* 2001. Köhler, Per-Olof, Birgitta Hene & Ulla Messelius (red.). Stockholm: Natur och Kultur.
- Navarro, Gonzalo 2001. A guided tour to approximate string matching. *ACM computing surveys* (CSUR) 33:1, 31–88.
- Nesi, Hilary 1996. The role of illustrative examples in productive dictionary use. *Dictionaries* 17, s. 198–206.
- Nesi, Hilary 1999. A user's guide to electronic dictionaries for language learners. *International Journal of Lexicography* 12(1), s. 55–66.
- Nesi, Hilary 2000a. On screen or in print? Students' use of a learner's dictionary on CD-ROM and in book form. In: Howarth, P. & R. Herington (eds), *EAP Learning Technologies*. Leeds: Leeds University Press, s. 106–114.
- Nesi, Hilary 2000b. *The Use and Abuse of EFL Dictionaries. How learners of English as a foreign language read and interpret dictionary articles*. (Lexicographica Series Maior 98.) Tübingen: Max Niemeyer Verlag.
- Nesi, Hilary 2012. Alternative dictionaries: Uncovering dark practices. In: Granger, Sylviane & Magali Paquot (eds), s. 363–378.
- Nesi, Hilary & Paul Meara 1994. The effect of language background and culture on productive dictionary use. In: Martin, Willy, Willem Meijs, Margreet Moerland, Elsemiek ten Pas, Piet van Sterkenburg & Piek Vossen (eds), *Euralex 1994. Proceedings. Papers submitted to the 6th EUR-ALEX International Congress in Amsterdam, The Netherlands, 30 August – 3 September, 1994*. Amsterdam: Vrije Universiteit, s. 577–585.
- Nesi, Hilary & Richard Hail 2002. A study of dictionary use by international students at a British university. *International Journal of Lexicography* 15(4), s. 277–305.
- Nikula, Henrik 1995. Exemplens funktioner i ordböcker. I: Svavarsdóttir, Ásta, Guðrún Kvaran, Gunnlaugur Ingólfsson & Jón Hilmar Jónsson (red.) *Nordiske studier i leksikografi 3. Rapport fra Konferanse om leksikografi i Norden, Reykjavík 7–10 juni 1995*. Reykjavík, s. 311–319.
- Nikula, Kristina 2008. Kvinnligt och manligt. Illustrationerna i *Bonniers svenska ordbok*. I: Jóhannesson, Kristinn et al. (red.), s. 304–311.

- Nikula, Kristina 2012. Samspelet mellan text och bild i enspråkigt svenska ordböcker. *LexicoNordica* 19, s. 103–122.
- Nilzon, Björn 2016. *Söka, slå och förstå... Språkintrouktionselevens bruk av elektroniska ordböcker och översättningsverktyg*. Kandidatuppsats. Göteborg: Institutionen för svenska språket.
- NLO = Bergenholtz, Henning, Ilse Cantell, Ruth Vatvedt Fjeld, Dag Gundersen, Jón Hilmar Jónsson & Bo Svensén 1997. *Nordisk leksikografisk ordbok*. Oslo: Universitetsforlaget.
- NoKSO = *Natur och Kulturs Svenska Ordbok* 2001. Köhler, Per-Olof, Birgitta Hene & Ulla Messelius (red.). Stockholm: Natur och Kultur.
- Nygaard, Lars & Ruth Vatvedt Fjeld 2008. Analyse av søkelogger for bedre søkumuligheter i elektroniske ordbøker. *LexicoNordica* 15, s. 57–72.
- Pálfi, Lorand-Levente & Sven Tarp 2009. Lernerlexikographie in Skandinavien – Entwicklung, Kritik und Vorschläge. *Lexicographica* 25, s. 135–154.
- Parkvall, Mikael 2016. *Sveriges språk i siffror: vilka språk talas och av hur många?* (Språkrådets skrifter 20.) Stockholm: Morfem och Språkrådet.
- Piotrowski, Tadeusz 2009. [Review of] Sven Tarp. Lexicography in the Borderland between Knowledge and Non-Knowledge. General Lexicographical Theory with Particular Focus on Learner's Lexicography. *International Journal of Lexicography* 22(4), s. 480–486.
- Prentice, Julia 2010. *På rak sak. Om ordförbindelser och konventionaliserade uttryck bland unga språkbrukare i flerspråkiga miljöer*. (Göteborgsstudier i nordisk språkvetenskap 13.) Göteborg: Institutionen för svenska språket, Göteborgs universitet.
- Prentice, Julia & Emma Sköldberg 2013. Flerordsenheter – ur ett andraspråksperspektiv. I: Hyltenstam, Kenneth & Inger Lindberg (red.), s. 197–220.
- Quirk, Randolph 1973. The social impact of dictionaries in the UK. *Annals of the New York Academy of Sciences* 211, s. 76–88.
- Rambøll 2012 (på oppdrag for Utdanningsdirektoratet). *Brukerundersøkelse av LEXIN – ordbøker for minoritetsspråklige*. <[www.udir.no/globalassets/upload/5/lexin.pdf](http://www.udir.no/globalassets/upload/5/lexin.pdf)>. Hämtat 12 november 2016.
- Ritchie, William C. & Tej K. Bhatia (eds) 2009. *The New Handbook of Second Language Acquisition*. 2 uppl. Bingley, UK: Emerald.
- Rogström, Lena 2000. Användarstrategier i OSA-databasen. I: Byrman, Gunilla m.fl. (red.), s. 255–265.
- Rosch, Eleanor 1977. Human Categorization. In: Warren, Neil (ed.), *Advances in Cross-Cultural Psychology* 1. London: Academic Press, s. 1–72.

- Rundell, Michael 2012. 'It works in practice but will it work in theory?' The uneasy relationship between lexicography and matters theoretical. In: Fjeld, Ruth Vatvedt & Julie Mathilde Torjusen (eds), s. 47–92.
- SAOL 13 = *Svenska Akademiens ordlista* 2006. 13 uppl. Stockholm: Norstedts i distr.
- SAOL 14 = *Svenska Akademiens ordlista* 2015. 14 uppl. Stockholm: Norstedts i distr.
- SAOL Plus [cd-rom] 2007. Stockholm: Norstedts Akademiska Förlag i distr.
- SCB Statistik 2015. *Statistikskolan: Mer bortfall i statistiken*. <[http://www.scb.se/sv\\_/Hitta-statistik/Artiklar/Statistikskolan-Mer-bortfall-i-statistiken](http://www.scb.se/sv_/Hitta-statistik/Artiklar/Statistikskolan-Mer-bortfall-i-statistiken)>. Hämtat 8 augusti 2016.
- Ščerba, Lev V. 1995 [1940]. Towards a General Theory of Lexicography. *International Journal of Lexicography* 8(4), s. 314–350.
- Schemann, Hans 1989. Das phraseologische Wörterbuch. In: Hausmann, Franz Josef et al. (eds), s. 1019–1032.
- Schmitt, Norbert & Michael McCarthy (eds) 1997. *Vocabulary: Description, Acquisition and Pedagogy*. Cambridge: Cambridge University Press.
- de Schryver, Gilles-Maurice & David Joffe 2004. On How Electronic Dictionaries are Really Used. In: Williams, Geoffrey & Sandra Vessier (eds), *Proceedings of the eleventh EURALEX international congress, EURALEX 2004, Lorient, France, July 6-10*. Lorient: UBS, s. 187–196.
- de Schryver, Gilles-Maurice, David Joffe, Pitta Joffe & Sarah Hillewaert 2006. Do Dictionary Users Really Look Up Frequent Words? – On the overestimation of the Value of Corpus-based Lexicography. *Lexikos* 16, s. 67–83.
- Searle, John R. 1969. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.
- Sinclair, John M. (ed.) 1987. *Looking Up. An account of the COBUILD Project in lexical computing and the development of the Collins COBUILD English language dictionary*. London: Harper Collins.
- Sinclair, John M. (ed.) 2004. *How to use corpora in language teaching*. (Studies in Corpus Linguistics 12.) Amsterdam: Benjamins.
- Skolverket 2012. *Kommunal vuxenutbildning i svenska för invandrare. Kursplan och kommentarer*. <<http://www.skolverket.se/publikationer?id=2939>>. Hämtat 6 december 2016.
- Skolverket 2016. *Statistik om utbildning i svenska för invandrare*. <<http://www.skolverket.se/statistik-och-utvardering/statistik-i-tabeller/utbildning-i-svenska-for-invandrare>>. Hämtat 9 november 2016.
- Sköldberg, Emma 2004. *Korten på bordet. Innehålls- och uttrycksmässig variation hos svenska idiom*. (Meijerbergs arkiv för svensk ordforskning 31.) Göteborg: Göteborgs universitet, Meijerbergs institut för svensk etymologisk forskning.

- Smit, Maria 2001. Herbert Wiegand with a “Metalexicographical Panga” in the Jungle: An Unlocking of Wörterbuchforschung. *Lexikos* 11, s. 296–310.
- SND = *Svensk Nationell Datatjänst* 2016. <<https://snd.gu.se/sv/>>. Hämtat 20 november 2016.
- Stein, Gabriele 1991. Illustrations in dictionaries. *International Journal of Lexicography* 4(2), s. 99–127.
- Stroud, Christopher 1979. Kontrastiv lexikologi. I: Hyltenstam, Kenneth (red.) *Svenska i invandrarperspektiv*. Lund: Gleerups, s. 174–218.
- Summers, Della 1988. The role of dictionaries in language learning. In: Carter, Ronald & Michael McCarthy (eds), *Vocabulary and Language Teaching*. London: Longman, s. 111–125.
- Svenska ord: med uttal och förklaringar* 1984. 1 uppl. Stockholm: Skolöverstyrelsen & Statens inst. för läromedelsinformation/Solna: Esselte studium.
- Svenska ord: med uttal och förklaringar* 1992. 2 uppl. Stockholm: Statens skolverk och Norstedts.
- Svenska ord: med uttal och förklaringar* 1995. 3 uppl. Stockholm: Myndigheten för skolutveckling/Liber.
- Svanlund, Jan 2009. *Lexikal etablering: en korpusundersökning av hur nya sammansättningar konventionaliseras och får sin betydelse*. (Stockholm studies in Scandinavian Philology, N.S., 52.) Stockholm: Acta universitatis Stockholmiensis.
- Sveningsson, Malin, Mia Lövheim & Magnus Bergqvist 2003. *Att fånga Nätet. Kvalitativa metoder för Internetforskning*. Lund: Studentlitteratur.
- Svensén, Bo 2001. Markeringsangivelser i ordböcker: några strukturella aspekter. I: Gellerstam, Martin, Kristinn Jóhannesson, Bo Ralph & Lena Rogström (red.), *Nordiska Studier i Lexikografi 5. Rapport från Konferens om lexikografi i Norden, Göteborg 27.–29. maj 1999*. Göteborg, s. 357–375.
- Svensén, Bo 2004 [1987]. *Handbok i lexikografi. Ordböcker och ordboksarbete i teori och praktik*. 2 uppl. Stockholm: Norstedts Akademiska Förlag.
- Tarp, Sven 1999. Lørnerordbøger for indvandrere og andet godtfolk. *LexicoNordica* 6, s. 107–132.
- Tarp, Sven 2006. *Leksikografien i grænselandet mellem viden og ikke-viden: Generel leksikografisk teori med særlig henblik på lørnerleksikografi*. Århus: Handelshøjskolen i Århus, Center for Leksikografi.
- Tarp, Sven 2008a. Kan brugerundersøgelser overhoved afdække brugernes leksikografiske behov? *LexicoNordica* 15, s. 5–32.
- Tarp, Sven 2008b. *Lexicography in the Borderland between Knowledge and Non-Knowledge. General Lexicographical Theory with Particular Focus on*

- Learner's Lexicography*. (Lexicographica Series Maior 34.) Tübingen: Max Niemeyer Verlag.
- Tarp, Sven 2010. Reflections on the Academic Status of Lexicography. *Lexikos* 20, s. 450–465.
- Tomaszczyk, Jerzy 1979. Dictionaries: users and uses. *Glottodidactica* 12, s. 103–119.
- Tono, Yukio 2011. Application of Eye-tracking in EFL Learners' Dictionary Look-up Process Research. *International Journal of Lexicography* 24(1), s. 124–153.
- Toporowska Gronostaj, Maria 2005. Elektroniska ordböcker i Sverige: nutid och framtid. *LexicoNordica* 12, s. 87–107.
- Trap-Jensen, Lars, Henrik Lorentzen & Nicolai H. Sørensen 2014. An odd couple – Corpus frequency and look-up frequency: what relationship? In: Kosem, Iztok & Michael Rundell (eds.), *Slovensčina 2.0*. Vol. 2(2). Trojina, Institute for Applied Slovene, Slovenia, s. 94–113.
- Userneeds 2016. *Kvinnor är mer benägna att svara på undersökningar via mobiltelefon*. <<http://userneeds.se/webstatus/nyheter/women-are-likely-to-answer-online-surveys>>. Hämtat 14 november 2016.
- Varantola, Krista 1998. Translators and their Use of Dictionaries. In: B.T. Sue Atkins (ed.), 179–192.
- Verlinde, Serge & Jean Binon 2010. Monitoring Dictionary Use in the Electronic Age. In: Dykstra, Anne & Tanneke Schoonheim (eds), s. 1144–1151.
- Viberg, Åke 2004. Lexikal utveckling i ett andraspråk. I: Hyltenstam, Kenneth & Inger Lindberg (red.), *Svenska som andraspråk – i forskning, undervisning och samhälle*. 1 uppl. Lund: Studentlitteratur, s. 197–220.
- Volodina, Elena & Sofie Johansson Kokkinakkis 2012. Introducing the Swedish Kelly-list, a new lexical e-resource for Swedish. In: Calzolari, Nicoletta et al. (eds), s. 1040–1046.
- Weschler, Robert & Chris Pitts 2000. An experiment using electronic dictionaries with EFL students. *The internet TESL Journal* 6. <[www.iteslj.org/Articles/Weschler-ElectroDict.html](http://www.iteslj.org/Articles/Weschler-ElectroDict.html)>. Hämtat 19 november 2016.
- Welker, Herbert Andreas 2006. *O uso de dicionários. Panorama geral das pesquisas empíricas*. Brasília: Thesaurus.
- West, Michael P. 1935. *The new method English dictionary; explaining the meaning of 24,000 items within a vocabulary of 1,490 words*. London: Longmans/Green.
- Wiegand, Herbert Ernst 1977. Nachdenken über Wörterbücher – Aktuelle Probleme. In: Wiegand, Herbert Ernst, Günther Drosdowski & Helmut Henne (eds), *Nachdenken über Wörterbücher*. Mannheim: Bibliographisches Institut, s. 51–102.

- Wiegand, Herbert Ernst 1998. *Wörterbuchforschung. Untersuchungen zur Wörterbuchbenutzung, zur Theorie, Geschichte, Kritik und Automatisierung der Lexikographie*. Berlin: Walter de Gruyter.
- Wiegand, Herbert Ernst 2001. Was eigentlich sind Wörterbuchfunktionen? Kritische Anmerkungen zur neueren und neuesten Wörterbuchforschung. *Lexicographica* 17, s. 217–248.
- Wikipedia 2016. *Spearmans rangkorrelation*. <[https://sv.wikipedia.org/wiki/Spearmans\\_rangkorrelation](https://sv.wikipedia.org/wiki/Spearmans_rangkorrelation)>. Hämtat 19 november 2106.
- Wikström, Inger 1991. *Att läsa för skolan och för livet. Skönlitteratur och ämnesundervisning i samhällsorienterade ämnen – ett integrerat arbetssätt i andraspråksundervisningen*. Stockholm: Skriptor Förlag.
- Wingate, Ursula 2002. *The Effectiveness of Different Learner Dictionaries. An Investigation into the Use of Dictionaries*. (Lexicographica. Series Maior 112.) Tübingen: Max Niemeyer.
- Wolfer, Sascha, Alexander Koplenig, Peter Meyer & Carolin Müller-Spitzer 2014. Dictionary Users do Look up Frequent and Socially Relevant Words. Two Log File Analyses. In: Abel, Andrea, Chiara Vettori & Natascia Ralli (eds), s. 281–290.
- Zgusta, Ladislav 1971. *Manual of Lexicography*. (Janua linguarum. Series Maior 39.) Prague: Mout.



# Bilaga 1. Frågeformuläret från 2007

1. Ålder:

- upp till 14.  15-19.  20-29.  30-39.  40-49.  50-59.  60 eller mer.
- 

2. Kön:

- Kvinnligt.  Manligt.
- 

3. Sysselsättning:

- Studerande.  Yrkesarbetande.  Arbetsökande.  Annat.

Om du svarat Annat, vilken sysselsättning har du?

4. Modersmål:

5. Om du har annat modersmål än svenska, hur bedömer du dina kunskaper i svenska språket?

- Dåliga.  Ganska bra.  Bra.  Mycket bra.
- 

6. Din högsta utbildning i Sverige:

- Grundskola (eller motsvarande).  Gymnasium (eller motsvarande).  Folkhögskola (eller motsvarande).  Universitet (eller motsvarande).

Din högsta utbildning i ditt hemland (om annat än Sverige):

- Grundskola (eller motsvarande).  Gymnasium (eller motsvarande).  Folkhögskola (eller motsvarande).  Universitet (eller motsvarande).
- 

7. Hur ofta brukar du söka i Lexins svenska ordbok Svenska ord?

- Detta är första gången.  Högst en gång i månaden.  Flera gånger i månaden.  Flera gånger i veckan.
-

7. Hur ofta brukar du söka i Lexins svenska ordbok Svenska ord?

- Detta är första gången.  Högst en gång i månaden.  Flera gånger i månaden.  Flera gånger i veckan.

---

**Resten av frågorna gäller dig som har sökt i Svenska ord förut.**

8. Vilken typ av information söker du oftast efter? Du kan kryssa i flera alternativ.

- Stavning.  Uttal.  Uttal via ljudfunktion.  Böjning.  Ordklass.  Betydelse.  Exempel.  
 Stående uttryck.  Sammansättningar/avledningar.  Bildtema.  Annat.

Om du svarat Annat, vad tänker du då på?

---

9. I vilka situationer brukar du använda Svenska ord?

- För att förstå ett ord i en text som jag läser.  För att få hjälp med ett ord när jag skriver en text.  För att få veta hur ett ord uttalas.  Annan situation.

Om du svarat Annan situation, vad tänker du då på?

---

10. Brukar du få tag på den information du söker efter i Svenska ord?

- Ja.  Nej.  Delvis.

---

Har du någon kommentar kring Lexin och Svenska ord? Är det till exempel något du saknar i ordboken? Skriv gärna det här!

# Bilaga 2. Frågeformuläret från 2011

## Enkät om användning av Lexin

Just nu pågår en undersökning av användningen av *Lexins svenska lexikon*, som nyligen har uppdaterats. Svara på några frågor kring dig och din användning av lexikonet så är du med och gör det ännu bättre!

Vid flervalsfrågorna klickar du i önskad ruta (till vänster om det valda alternativet). Om du vill skriva i en kommentarruta måste du först flytta markören in i rutan.

Skicka iväg din enkät genom att klicka på knappen märkt "Skicka enkät" längst ner.

*Tack för hjälpen!*

---

1. Födelseår:

2. Kön:

- Kvinna.
  - Man.
- 

3. Högsta avslutade utbildning i Sverige:

- Ingen avslutad utbildning.
- Grundskola.
- Gymnasium.
- Folkhögskola.
- Komvux (eller motsvarande).
- Universitet/högskola.
- Annan (ange nedan).

Annan utbildning:

4. Vad är din huvudsakliga sysselsättning?

- Jag studerar.
- Jag arbetar.
- Jag söker arbete.
- Annan (ange nedan).

Annan huvudsaklig sysselsättning:

5. Hur ofta brukar du söka i *Lexins svenska lexikon*?

- Detta är första gången.
- Högst en gång i månaden.
- Flera gånger i månaden.
- Flera gånger i veckan.
- Flera gånger per dag.

6. Använder du mest *Lexins svenska lexikon* för att (ange max två alternativ):

- förstå ett ord i en text som du läser?
- få hjälp med ett ord när du skriver en text?
- få hjälp med ett ord när du ska översätta en text?
- få hjälp med ett ord som du hört?
- något annat (ange nedan).

Ev. annan anledning:

7. Vilka typer av information söker du mest efter i *Lexins svenska lexikon* (ange max tre alternativ):

- Uttal.
- Ordklass.
- Böjning.
- Förkortningar.
- Betydelse.
- Motsatser.
- Sammansättningar.
- Exempel.
- Konstruktioner.
- Uttryck.
- Bild.
- Film.

8. Finns det information som du aldrig söker efter i *Lexins svenska lexikon* (kryssa för så många du behöver)?

- Uttal.
- Ordklass.
- Böjning.
- Förkortningar.
- Betydelse.
- Motsatser.
- Sammansättningar.
- Exempel.
- Konstruktioner.
- Uttryck.
- Bild.
- Film.

---

9. Brukar du titta på bilderna som finns till vissa ord (dvs klickar du på länken *Bild svenska*)?

- Ofta.
- Ibland.
- Aldrig.

---

Brukar du titta på filmerna som finns till vissa ord (dvs klickar du på länken *Visa film*)?

- Ofta.
- Ibland.
- Aldrig.

10. Har du någon gång läst instruktionerna till *Lexin*, där det står om hur man ska tolka och förstå den information som finns i Lexinordböckerna?

- Ja.
- Nej.
- Vet inte.

---

11. Hur löser du språkliga problem om du inte använder *Lexins svenska lexikon*? (Du får kryssa för flera.)

- Använder en annan ordbok på *Internet*.
- Använder en tryckt ordbok.
- Frågar någon.
- Använder Google eller någon annan sökmotor på *Internet*.
- Annat sätt (ange nedan).

Ev. annat sätt:

---

**Fråga 12 till 19 riktar sig till dig som har annat modersmål (=det första språk du lärde dig) än svenska. Övriga kan gå till fråga 20.**

---

12. Vilket år kom du till Sverige? (Galler dig som har flyttat till Sverige från ett annat land.)

13. Har du genomgått SFI-utbildning? (SFI=svenska för invandrare)

- Ja.
- Nej.

---

Har du gått någon annan kurs eller utbildning i svenska språket?

- Ja (ange vilken nedan).
- Nej.

Ev. annan kurs eller utbildning i svenska:

14. Vilket är ditt modersmål?

15. Vilket språk är du bäst på?

16. Hur ofta talar eller använder du svenska?

- Nästan hela tiden.
- Varje dag.
- Flera gånger i veckan.
- Mer sällan.

17. Vilket/vilka språk talar du främst...

med din familj?

med dina vänner?

med kompisarna i skolan eller på jobbet?

18. Hur bra bedömer du att dina kunskaper i svenska är?

- Mycket bra.
- Bra.
- Ganska bra.
- Ganska dåliga.
- Dåliga.

19. Högsta avslutade utbildning i ditt hemland:

- Ingen avslutad utbildning.
- Grundskola (eller motsvarande).
- Gymnasium (eller motsvarande).
- Folkhögskola (eller motsvarande).
- Universitet/högskola (eller motsvarande).
- Annan (ange nedan).

Annan utbildning:

Ev. kommentar:

---

20. Här kan du skriva ned dina åsikter eller kommentarer kring *Lexins svenska lexikon*:

---

Lämna gärna ditt namn och din e-postadress om du kan tänka dig att vara med i fler undersökningar av Lexin:

---

**Stort tack för hjälpen!**

*OBS! Det just ifyllda formuläret kan bara skickas iväg en gång!*

