

MEIJERBERGS ARKIV

FÖR SVENSK ORDFORSKNING

utgivet av
styrelsen för Meijerbergs institut
vid Göteborgs universitet
genom
Anna Helga Hannesdóttir

43

GÖTEBORG 2019



Att tolka det sammansatta

Befästning och mönster i första- och andraspråkstalaras
tolkning av sammansättningar

Lisa Loenheim

Redaktör
Emma Sköldberg

Ingår som del 43 i Meijerbergs arkiv för svensk ordforskning

© Lisa Loenheim och
Meijerbergs institut för svensk etymologisk forskning
Göteborgs universitet
Box 200
SE-405 30 Göteborg

ISSN 0348-7741
ISBN 978-91-87850-02-8

Sättning: Sven Lindström
Omslagsbild: Anna Niklasson
Porträttfoto: Magnus Loenheim
Tryck: Repro Lorensberg, Göteborg 2019

Förord

Ända sedan jag var liten har jag drömt om att någon gång få ge ut en bok. Det känns stort att jag nu får göra det, och det finns många personer som jag vill tacka för hjälpen på vägen.

Varmt tack till mina handledare Anna Hannesdóttir och Sofia Tingsell. Anna, du har ställt de stora vetenskapsteoretiska frågorna, samtidigt som du har ett imponerande sinne för detaljer och språklig precision. Tack för alla noggranna genomläsningar och hjälpen med att vässa texten. Jag uppskattar särskilt din stadiga vägledning och stora tillgänglighet i slutfasen av arbetet. Sofia, du är en mästare på att se helheter och de stora dragen i både texter och teorier: tack för alla smarta idéer som väglett arbetet teoretiskt såväl som metodologiskt, liksom stilsäkra textkommentarer som bidragit till att förtydliga den röda tråden i avhandlingen. Tack också för all pepp under den här tiden.

En annan person som bidragit stort till denna avhandling är statistik-konsult Nils-Gunnar Pehrsson som gett experthjälp när det gäller den statistiska analysen och presentationen av den. Tack för ditt engagemang i mitt avhandlingsprojekt!

A heartfelt thank you goes to Professor Alexander Onysko, who contributed to this study as an informal “remote consultant” at an early stage of this project. Thanks for valuable – and enjoyable – discussions concerning the outline of the study, especially the questionnaire.

Jag vill också lyfta fram den hjälp jag fått av Judy Ribeck Nyström, som så generöst tillgängliggjort och utfört sökningar för min räkning i en lärobokskorpus. Tack för det Judy, och tack för de trevliga åren som rumskamrater!

Många tack till Eli Anne Eiesland som var granskare på mitt slutseminarium. Du bidrog med värdefulla synpunkter som gav vägledning framåt. Dessutom gjorde du seminariet till en trevlig tillställning som gav energi inför återstoden av arbetet.

Tack till alla elever som deltagit i studien och som gjort denna avhandling möjlig, och tack till deras lärare som välkomnat mig till sina klasser!

Jag vill också tacka ytterligare några kolleger och tidigare kolleger som bidragit till denna avhandling: Emma Sköldberg för en mycket ambitiös redaktörläsning, Sven Lindström för en proffsig insats och gott samarbete i samband med sättningen av manus, Hallfríður Helgadóttir för snabb

och stilsäker språkgranskning av summaryn, Anki Hult för referensgranskning (utöver stöttning och vänskap), Anna Winlund för läsning av inledningen, Anja Allwood för språklig respons på abstractet och Linda Sjöström för praktisk hjälp med allt runt doktorandanställningen.

Ytterligare ett tack går till Meijerbergs institut för svensk etymologisk forskning för att min avhandling får ingå i serien *Meijerbergs arkiv för svensk ordforskning*.

Tack vare stipendier från Stiftelsen Paul och Marie Berghaus donationsfond och Stiftelsen Henrik Ahrenbergs studiefond har jag kunnat åka på konferenser. Från Kungliga och Hvitfeldtska stiftelsen erhöll jag ett doktorandstipendium avsett för att färdigställa avhandlingen. Varmt tack!

Under större delen av doktorandtiden har jag delat lokaler med ISA-/FlerSva-gänget. Jag har uppskattat alla fikor, goda samtal och glada skratt tillsammans med er! Särskilt tack till Carina Carlund som var ett stort stöd i en tung period: du har en social fingertoppskänsla utöver det vanliga.

Jag vill också rikta ett innerligt tack till Marcel Bernhardsson för hjälpen under denna period med att lyfta blicken och få perspektiv på tillvaron.

Det som jag uppskattat allra mest under de här åren är gemenskapen i doktorandgänget. Det har varit fantastiskt att få vara en del av detta gäng! Jag vill särskilt nämna "årskollegan" Joel Olofsson: jag är glad över allt vi kunnat dela. I den intensiva slutfasen har doktorandfikorna och tillfället att prata om stort och smått gett ett välkommet avbrott. Särskilt tack till Henrietta Adamsson Eryd och Malin Sandberg för gemensamma arbetshelger med trevligt utbyte på sluttampen.

Sammanfattningsvis är jag väldigt nöjd med min tid på forskarutbildningen; tack Anders-Börje Andersson och Elisabet Engdahl för att ni uppmuntrade mig att söka!

Även familj och släkt har bidragit direkt och indirekt till denna avhandling. Magnus Loenheim har hjälpt mig i arbetet med tabeller och figurer: tack för det! Anna Niklasson har gjort omslagsbilden. Tack, kära kusin, för att du gjorde boken som jag drömt om till en väldigt fin bok!

Tack mamma och pappa för allt gott ni gett och ger mig i livet. Särskilt tack till mamma för praktisk hjälp i slutskedet av avhandlingstiden. Till mina systrar Helena och Klara: jag är glad att ni finns där.

Till sist och framför allt: kärlek till mina närmsta! Magnus, ditt osvikliga stöd på alla sätt och alla plan har varit ovärderligt, I owe you! Vilma och Ludvig: vilken tur jag har som har er!

Göteborg den 30 november
Lisa Loenheim

Title: Att tolka det sammansatta. Befästning och mönster i första- och andraspråkstälarens tolkning av sammansättningar
English title: Interpreting composite structures. Entrenchment and patterns in the interpretation of Swedish compounds by L1 and L2 speakers
Language: Swedish, with an English summary
Author: Lisa Loenheim

Abstract

This thesis presents a cognitive questionnaire-based study on the interpretation of compounds. The analysis is based on the interpretations from 190 Swedish-speaking students in upper secondary schools. The overall aim is to test the assumption of usage-based approaches that linguistic phenomena become entrenched in the mind of the speaker as a result of repetition. To be able to analyse such entrenchment effects both established compounds and novel compounds with many meaning potentials are included, and interpretations from speakers with various exposure time to Swedish (L1 and L2 speakers) are analysed.

The result of this study indicates that the meaning potentials of compounds are evaluated only when necessary. If the compound is entrenched as a unit, no analysis is needed. If the compound is unfamiliar to the language user, he/she makes use of an entrenched linguistic template, i.e. a specific (similar) compound or a low-level or higher-level schema, as an analogy base in the interpretation.

This study confirms that the frequency of the compound and the user's exposure time to Swedish are factors of great impact: frequent compounds result in more concordant interpretations than less frequent ones, and L1 speakers are more concordant in their interpretations than L2 speakers. The discrepancy between the speaker groups (L1 and L2) concerning established compounds is a result of entrenchment differences of these specific compounds (due to various exposure time to Swedish), whereas the discrepancy concerning the interpretation of novel compounds reflects various degrees of entrenchment of linguistic templates that can serve as analogy bases.

This study further indicates that the right-headedness pattern is not as established for Swedish compounds as previously assumed. This is especially true for adjectival compounds.

KEY WORDS: Swedish compounds, meaning potential, cognitive linguistics, L1/L2, questionnaire, compositionality, transparency, schema, analogy, linguistic template

Meijerbergs institut för svensk etymologisk forskning
University of Gothenburg

ISSN 0348-7741

ISBN 978-91-87850-02-8



Innehåll

1. Inledning	1
1.1. Studiens upplägg och syfte	2
1.2. Sammansättning som fenomen	3
1.3. Termer och begrepp	5
1.4. Avhandlingens disposition	7
2. Forskningsöversikt: sammansättningar	9
2.1. Semantiska relationer mellan leden	10
2.2. Tolkning av sammansättningar	12
2.3. Egenskapsprojicering och relationslänkning	16
2.4. Sammansättningens semantiska huvudled	19
2.4.1. Förhållningssätt till schemat för huvudledets placering ..	20
2.4.2. Transfer från L1 till L2 vid identifieringen av huvudledet	23
2.5. Sammansättningar ur ett L2-perspektiv	25
3. Teoretiskt ramverk	29
3.1. Bruksbaserade kognitiva perspektiv	29
3.1.1. Språkförmågan – integrerad med andra kognitiva förmågor	30
3.1.2. Betydelse, kompositionalitet och analyserbarhet	32
3.1.3. Schema och konstruktion	34
3.1.4. Frekvens, minnesrepresentation och befästning	36
3.1.5. Schema eller analogi?	39
3.1.6. Saliens och befästning	43
3.1.7. Konventionalisering i relation till befästning	44
3.1.8. Kognitiva perspektiv på flerspråkighet	44
3.1.9. Befästning – inte bara en fråga om frekvens	48
3.2. Lexikala perspektiv	48
3.2.1. Kompositionalitet, meningspotential och genomskinlighet	49
3.2.2. Lexikalisering och konventionalisering	53
3.2.3. Konventionalisering och meningspotential ur ett flerspråkighetsperspektiv	57

4. Metod och material	61
4.1. Enkätstudiens utformning	62
4.2. Hypoteser	62
4.3. Undersökta sammansättningstyper	64
4.3.1. Nominala sammansättningar	67
4.3.2. Adjektiviska sammansättningar	73
4.4. Referenskorpus	77
4.5. Enkätens orddel	78
4.6. Enkätens språkliga bakgrundsfrågor	80
4.7. Etiska överväganden: information, samtycke och anonymitet ..	82
4.8. Instruktion till och genomförande av enkäten	82
4.9. Skolor och gymnasieprogram	84
4.10. Informanterna och deras språkliga bakgrund	87
4.10.1. Informantgrupper och kategoriseringsprinciper	87
4.10.2. Flerspråkighetsbakgrund och L2	90
4.10.3. Svenska som L1 (SVE) trots andra modersmål hos föräldrarna	92
4.10.4. Flerspråkig (SVE+) trots ett angivet modersmål	93
4.10.5. SVE trots svenska och ytterligare modersmål i familjen	94
4.10.6. Tvetydiga uppgifter om modersmål och startålder för svenskan	95
4.11. Kvantitativ analys av resultaten i enkätens flervalsdel	96
4.11.1. Statistisk metod	96
4.11.2. Bortfall	97
4.12. Analys av fritexttolkningarna i relation till svaren i enkätens flervalsdel	98
5. Tolkningar i enkätens flervalsdel	99
5.1. Nominala sammansättningar	100
5.1.1. Huvudled i nominala sammansättningar – typen <i>spelkort</i> och <i>vagnbarn</i>	101
5.1.2. Huvudled i treledade sammansättningar – typen <i>barnbokshylla</i> och <i>sommardagdröm</i>	110
5.1.3. Semantisk relation i sammansättningar som betecknar behållare – typen <i>fruktorg</i> och <i>blomlåda</i>	113
5.1.4. Jämförande egenskapsprojicering eller tematisk relationslänkning – typen <i>corpörn</i> och <i>citronvatten</i> ..	117
5.1.5. Identitet/jämförande egenskapstolkning eller tematisk relationslänkning – typen <i>kvinnopräst</i> och <i>herrfrisör</i> ..	121

5.1.6. Agens vid verbavledning – typen <i>polismisshandel</i> och <i>studentförakt</i>	126
5.1.7. Betydelsesammansmälta AN-sammansättningar – typen <i>grönyta</i> och <i>småkakor</i>	129
5.2. Adjektiviska sammansättningar	133
5.2.1. Huvudled i AA-sammansättningar – typen <i>gulvit</i> och <i>svartvit</i>	133
5.2.2. Jämförelse eller orsak i NA-sammansättningar – typen <i>iskall</i> och <i>solvarm</i>	139
5.2.3. Sammansättningar med <i>-rädd</i>	144
5.2.4. Sammansättningar med <i>-fri</i>	145
5.2.5. Sammansättningar med <i>-färdig</i>	147
5.3. Sammanfattning	152
6. Tolkningar i enkätens fritextdel	155
6.1. Sammansättningar som förekommer i enkätens båda delar	156
6.2. Val av semantisk relation mellan leden	159
6.2.1. Material, innehåll eller ändamål	160
6.2.2. Jämförande egenskapsprojicering eller tematisk relationslänkning – djur och artefakter	160
6.2.3. Jämförande egenskapstolkning, identitet eller tematisk relationslänkning – personbetecknande samman- sättningar	161
6.2.4. Agens vid verbavledning	163
6.2.5. Betydelsesammansmälta AN-sammansättningar	164
6.2.6. Jämförelse	164
6.3. Sammansättningens huvudled i informanternas fritexttolkningar	166
6.4. Meningspotentialen hos en sammansättning – exemplet <i>vagnbarn</i>	168
6.5. Skillnader mellan L1- och L2-talarnas tolkningar i fritext- och flervaldsdelen	169
6.6. Sammanfattning	170
7. Analys	173
7.1. Frekvens och befästning av schematiska mönster och specifika instanser	174
7.1.1. Identifiering av huvudled i nominala sammansättningar	174
7.1.2. Segmentering av treledade sammansättningar	178

7.1.3. Semantisk relation i sammansättningar som betecknar behållare	179
7.1.4. Jämförande egenskapsprojicering eller tematisk relationslänkning	181
7.1.5. Identitet/jämförande egenskapstolkning eller tematisk relationslänkning i personbetecknande samman- sättningar	183
7.1.6. Agens vid verbavledning	184
7.1.7. Betydelsesammansmälta AN-sammansättningar	186
7.1.8. Identifiering av huvudledet i AA-sammansättningar ..	187
7.1.9. Jämförelse eller orsak vid NA-sammansättningar	189
7.1.10. Sammansättningar med <i>-räd</i>	191
7.1.11. Sammansättningar med <i>-fri</i>	191
7.1.12. Sammansättningar med <i>-färdig</i>	192
7.1.13. Sammanfattning	193
7.2. Jämförelse	196
7.3. Saliens	197
7.4. Blockering	200
7.5. Omvärldskunskap	201
7.6. Konklusion	203
8. Sammanfattande diskussion	209
8.1. Befästning, analogi och mönster vid tolkning av sammansättningar	210
8.2. Konventionalitet – en fråga om tid och input	211
8.3. Att identifiera det semantiska huvudledet	213
8.4. Mångfald av meningspotentialer eller konventionella mönster?	216
8.5. Genomskinlighet, kompositionalitet och befästning	217
9. Slutsatser	219
Summary	223
Referenser	235
Bilaga 1	245
Bilaga 2	261
Bilaga 3	265

Tabeller

TABELL 2:1.	Sammanställning över terminologin avseende egenskaps- och relationstolkning i tidigare studier	17
TABELL 4:1.	Undersökta sammansättningar i kategorin <i>huvudled i nominala sammansättningar</i>	68
TABELL 4:2.	Undersökta sammansättningar i kategorin <i>huvudled i treledade nominala sammansättningar</i>	69
TABELL 4:3.	Undersökta sammansättningar i kategorin <i>semantisk relation i sammansättningar som betecknar behållare</i> ...	70
TABELL 4:4.	Undersökta sammansättningar i kategorin <i>jämförande egenskapsprojicering eller tematisk relationslänkning</i> ...	70
TABELL 4:5.	Undersökta sammansättningar i kategorin <i>identitet/ jämförande egenskapstolkning eller tematisk relationslänkning i personbetecknande sammansättningar</i>	71
TABELL 4:6.	Undersökta sammansättningar i kategorin <i>agens vid verbaledning</i>	72
TABELL 4:7.	Undersökta sammansättningar i kategorin <i>betydelse-sammansmälta AN-sammansättningar</i>	73
TABELL 4:8.	Undersökta sammansättningar i kategorin <i>huvudled i AA-sammansättningar</i>	74
TABELL 4:9.	Undersökta sammansättningar i kategorin <i>jämförelse eller orsak i NA-sammansättningar</i>	75
TABELL 4:10.	Undersökta sammansättningar i kategorin <i>NA-sammansättningar med -rädd</i>	75
TABELL 4:11.	Undersökta sammansättningar i kategorin <i>sammansättningar med -fri</i>	76
TABELL 4:12.	Undersökta sammansättningar i kategorin <i>sammansättningar med -färdig</i>	76
TABELL 4:13.	Fördelning mellan studerade korpusar	78

TABELL 4:14. Beskrivning av ingående skolor och klasser i enkätstudien	86
TABELL 4:15. Antal informanter, fördelade efter informantgrupp	88
TABELL 5:1. Informanternas svar på nominala sammansättningar som testar identifiering av huvudled	102
TABELL 5:2. Informanternas svar vid treledade sammansättningar med och utan fog	111
TABELL 5:3. Informanternas val av semantisk relation vid sammansättningar som betecknar behållare	114
TABELL 5:4. Informanternas svar på sammansättningar som testar valet mellan egenskapsprojicering och relationslänkning.....	118
TABELL 5:5. Informanternas svar på sammansättningar som testar valet mellan identitets-/egenskapstolkning och relationslänkning	122
TABELL 5:6. Samband mellan identitetstolkningar av <i>kvinnopräst</i> och <i>manspräst</i> i informantgruppen som helhet	124
TABELL 5:7. Samband mellan identitetstolkningar av <i>kvinnopräst</i> och <i>manspräst</i> , redovisat per informantgrupp	125
TABELL 5:8. Informanternas svar på verbavledda NN-sammansättningar som testar tolkningen av agens	127
TABELL 5:9. Informanternas svar vid betydelsesammansmälta AN-sammansättningar	130
TABELL 5:10. Informanternas svar vid AA-sammansättningar	135
TABELL 5:11. Informanternas val av jämförelse- eller orsaksrelation vid NA-sammansättningar	140
TABELL 5:12. Informanternas svar vid sammansättningar med <i>-rädd</i>	145
TABELL 5:13. Informanternas svar vid sammansättningar med <i>-fri</i> ...	146
TABELL 5:14. Informanternas svar vid sammansättningar med <i>-färdig</i>	148
TABELL 5:15. Samband mellan tolkningar av <i>-färdig</i> i informantgruppen som helhet	149
TABELL 5:16. Samband mellan tolkningar av <i>-färdig</i> , redovisat per informantgrupp	151

Figurer

FIGUR 5:1.	Fördelning av tolkningar med huvudledet till höger, redovisat per informantgrupp	107
FIGUR 5:2.	Fördelning av tolkningar med huvudledet till höger, redovisat efter startålder	109
FIGUR 5:3.	Fördelning av tolkningar i enlighet med fogeregeln, redovisat per informantgrupp	112
FIGUR 5:4.	Fördelning av konventionella tolkningar av betydelse-sammansmälta AN-sammansättningar	133



1

Inledning

Ständigt och med lätthet bildar vi nya sammansättningar. Till sin natur är dessa ofta mångtydiga. Ett ord fogas till ett annat – och ett tredje uppstår. Man kan vanligen härleda helhetens betydelse ur de ingående delarna, men det är däremot inte *förutsägbart* vilken betydelse som uppstår när två ord fogas ihop i en sammansättning. Oförutsägbarheten följer av uttryckets komplexitet. För det första kan sammansättningens ingående ord rymma flera betydelser. För det andra kan de ingående ordens betydelser modifieras av samförekomsten med det andra ordet, så att det ena eller båda orden får en något annorlunda betydelse som sammansättningsled än vad de hade var för sig. För det tredje kan det finnas flera teoretiskt tänkbara relationer mellan de ingående leden. För att exemplifiera med två sammansättningar är det iskalla vattnet 'kallt som is', medan den solvarma klippan är 'varm av solen'. Två till synes jämförbara sammansättningar skiljer sig således åt med avseende på den semantiska relationen mellan leden. Medan *iskall* förklaras med en jämförande parafras är relationen mellan leden i *solvarm* en orsaksrelation.

När en talare spontant bildar en helt ny sammansättning är den motiverad av kontexten, dvs. den konkreta språksituationen, och har sin betydelse motiverad av denna. Också sådana sammansättningar kan genom

frekvens och spridning få en mer kontextberoende, *konventionaliserad* betydelse, dvs. en betydelse som språkbrukarna i en språkgemenskap är någorlunda överens om. I fallet med *iskall* och *solvarm* har sammansättningarna konventionaliserats med olika relationer mellan leden.

Sammansättningar är ett gränshenomen mellan lexikon och grammatik. I egenskap av sådana gränshenomen har de under de senaste decennierna tilldragit sig ett ökande intresse. Sammansättningsforskning har växt fram som ett eget, tvärdisciplinärt fält, där lingvister, psykolingvister och psykologer ägnat sig åt forskning som syftat till att dels kartlägga och kategorisera sammansättningar, dels (via experimentella studier) få en inblick i den mentala representationen av språket. Som Plag (2003:132) formulerar det rymmer forskningsfältet många svårlösta frågor: ”compounding is a field of study where intricate problems abound, numerous issues remain unsolved, and convincing solutions are generally not so easy to find”.

Min studie är kognitivt inriktad och fokus ligger på hur sammansättningar – etablerade och mer tillfälliga – tolkas. I vilken mån analyseras en sammansättning kompositionellt (dvs. betydelsen pusslas fram utifrån de ingående leden) och i vilken mån hanteras den som en helhet?

1.1. Studiens upplägg och syfte

För att undersöka hur sammansättningar tolkas har jag genomfört en empirisk studie, närmare bestämt en enkätstudie, med talare av svenska. Studien omfattar både talare som har svenska som sitt förstaspråk (L1) och talare som har svenska som sitt andraspråk (L2). Skälen till detta är två. Det första handlar om representativitet och min intention att belysa hur svenska sammansättningar uppfattas av olika användare i det flerspråkiga Sverige av i dag. Det andra skälet är att ett kontrasterande L1–L2-perspektiv ger analysmässiga fördelar, eftersom man kan jämföra tolkningar från talare som har varierande exponeringstid för svenskan. Analysen baseras på 190 enkäter där gymnasieungdomar med svenskspråkig respektive flerspråkig bakgrund tolkar 85 sammansatta ord, både etablerade sammansättningar och tillfälliga bildningar. (Enkäten i sin helhet återfinns som bilaga 1.)

Syftet med studien är att undersöka vilka strukturella, semantiska och lexikala mönster som aktualiseras i informanternas tolkning av de undersökta sammansättningarna. Syftet är vidare att studera vilka kognitiva principer som vägleder informanterna i deras tolkningar. När det gäller de etablerade sammansättningarna undersöker jag vilken betydelse deras frekvens har för graden av samstämmighet i informantgruppen,

vilket korresponderar med hur internaliserade, eller med den kognitiva grammatikens terminologi, befästa, de är hos individen. När det gäller de tillfälliga (påhittade) sammansättningarna analyserar jag vilken hjälp informanterna har av lexikala, semantiska och strukturella mönster på olika abstraktionsnivåer (från specifika sammansättningar till abstrakta mönster eller scheman) som den obekanta sammansättningen kan relateras till när den tolkas.

Den enkät som informanterna besvarade innehåller två delar: en del med frågor med angivna tolkningsalternativ (hädanefter *flervalsdelen*) och en del med öppna frågor (hädanefter *fritextdelen*). I enkätens flervalsdel undersöker jag styrkeförhållanden mellan utvalda alternativ ur sammansättningarnas meningspotentialer (dvs. deras potentiella betydelser; se vidare avsnitt 3.2.1). Där studeras generella drag och regelbundenheter i informanternas tolkningar. I enkätens fritextdel formulerar informanterna själva förklaringar till ett urval sammansättningar. Dessa förklaringar illustrerar mångfalden av potentiella betydelser hos sammansättningarna och kompletterar därmed resultaten från flervalsdelen.

Min undersökning är avsedd som ett bidrag till sammansättningsforskningen genom att jag tillför ett kognitivt perspektiv och även beaktar flerspråkighetsaspekter. Utgångspunkten för studien är att jag vill testa antagandet inom kognitiv lingvistik om att frekvent upprepning av språkliga fenomen leder till att de blir befästa i språkbrukarens mentala lexikon. Genom att jämföra tolkningarna hos en grupp som har svenska språket med sig från början (både en- och flerspråkiga L1-talare) och en grupp som har svenskan som andraspråk, kan jag undersöka vilken betydelse exponering och exponeringstid har för graden av befästning av enskilda sammansättningar och sammansättningsmönster. De hypoteser som denna studie utgår från beskrivs i metodkapitlet, avsnitt 4.2. Hypoteserna är indelade i tre grupper: 1) hypoteser som rör sammansättningarna, 2) hypoteser som rör tolkningarna i informantgruppen som helhet och 3) hypoteser avseende skillnader i tolkning mellan informantgrupperna.

I ett större perspektiv vill jag – utifrån min jämförande studie med talare av olika språklig bakgrund (L1/L2) – bidra med empiri till teoretiska antaganden inom kognitiv lingvistik, lexikal semantik och flerspråkighetsforskning.

1.2. Sammansättning som fenomen

Sammansättningen som fenomen uppmärksammades redan under 300-talet f.Kr. av indiska grammatiker, och sammansättning är en produktiv ordbildningsprincip i moderna germanska språk. En sammansättning

kan definieras som ett ”ord som kan delas upp i minst två (ordliknande) huvuddelar som vardera innehåller minst ett rotmorfem” (Malmgren 1994:32). Såväl förledet som efterledet kan utgöras av ett rotmorfem, en avledning eller en sammansättning (Teleman, Hellberg & Andersson 1999, hädanefter SAG, del 1:221). Sammansättningar utmärks vidare av ett språkspecifikt prosodiskt mönster som särskiljer dem från fraser. I svenskan har sammansättningarna en sammansättningsaccent som ger en uttalsmässig skillnad mellan ”en sjuk sköterska” och ”en sjuksköterska” (Liljestrand 1993:35). För svenskans del tillkommer dessutom att sammansättning ibland markeras genom en fog mellan för- och efterled, t.ex. ett inskjutet *-s* som i *havsbad* eller en tillagd fovevokal som i *nattetid*. Vissa rotmorfem tappar ändelsevokalen vid sammansättning, jfr *flicknamn*. Om rotmorfemet slutar på *a* kan slutvokalen ersättas med en fovevokal (*e*, *o* eller *u*), såsom i *sagobok* (Malmgren 1994:37).

Sammansättningarna i svenskan kan delas in i olika strukturella undergrupper. De flesta räknas som determinativa, vilket innebär att efterledet betraktas som sammansättningens grammatiska och semantiska huvudled, medan förledet utgör underordnad modifierare, t.ex. *gungstol* ’stol som man kan gunga i’ (SAG 2: 42, 186). Utöver dessa identifierar Teleman (1970) tre grupper av icke-determinativa sammansättningar. *Kopulativa* kallar han de adjektiviska sammansättningar som har två grammatiska och semantiska huvudled, jämför *svensk-rysk* om en ordbok mellan svenska och ryska eller *sötsur* om en sås som är till lika delar söt och sur. I SAG går dessa under beteckningen *additiva* (SAG 2:188). Vidare särskiljer Teleman possessiva sammansättningar (*bahuvrihi*), där det semantiska huvudet finns utanför sammansättningen och sammansättningen anger vad det semantiska huvudet har, t.ex. *blåklocka* om en växt med en blå klocka och *rödstrumpa* om en kvinna med röda strumpor och, utvidgat, med röda åsikter (Teleman 1970:37). I en alternativ analys skulle de possessiva sammansättningarna kunna ses som determinativa sammansättningar med det enda särdraget att de uppbar en metonymisk semantik, där delen (den blåa klockan respektive den röda strumpan) får stå för helheten (växten respektive kvinnan), vilken utgör det semantiska huvudet (se Mellenius 1997:21–22). Teleman urskiljer ytterligare en mindre grupp av icke-determinativa sammansättningar, nämligen imperativiska som *krypin*, *tittut* och *hålligång* (1970:37). Kännetecknande för de svenska sammansättningarna är att de i normalfallet skrivs som ett ord utan bindestreck. Undantaget är kopulativa och imperativiska sammansättningar, där ett bindestreck mellan för- och efterled är möjligt, som i *svensk-rysk* respektive *kryp-in*.

Indelningen i determinativa, kopulativa och possessiva sammansättningar ger delvis en förenklad bild, eftersom en sammansättnings seman-

tiska struktur inte existerar oberoende av de språkbrukare som tolkar den. Hur en sammansättnings struktur uppfattas i ett språksamhälle är, som Josefsson (2005) framhåller, ytterst en fråga om förhandling och konventionalisering:

Det är viktigt att understryka att sammansättningar inte i sig själva är determinativa, kopulativa eller possessiva, utan att ett visst tolkningsalternativ är det naturliga eller har blivit konventionaliserat, d.v.s. blivit det allmänt accepterade inom en viss språkgemenskap. (Josefsson 2005:87)

Utifrån informanternas tolkningar av ett urval sammansättningar analyserar jag hur de förefaller uppfatta sammansättningarnas semantiska och strukturella mönster, t.ex. om en sammansättning bestående av två adjektiviska led uppfattas som determinativ eller kopulativ. Sammansättningar är, som framgått, ett komplext fenomen i gränslandet mellan lexikon och grammatik. Det rör sig om en morfologisk struktur bestående av minst två ord (eller rotmorfem), där de ingående leden bär en uppsättning semantiska drag. Det är i den kognitiva dimensionen och genom individens konceptualisering eller begreppsbildning som den sammansatta betydelsen framträder. Ur sammansättningens meningspotential – och mot bakgrund av strukturella och semantiska mönster – framstår en viss tolkning av leden och relationen mellan dem som den mest rimliga.

1.3. Termer och begrepp

I avhandlingen studerar jag sammansättningar av olika struktur. När jag omtalar *nominala sammansättningar* refererar jag till sammansättningar där efterledet består av ett nominalt, dvs. substantiviskt, led. Med *adjektiviska sammansättningar* avses sammansättningar där efterledet utgörs av ett adjektiviskt led.

Om den nominala sammansättningen har ett nominalt förled kallar jag den för en NN-sammansättning (t.ex. *husbåt*). En nominal sammansättning som har ett adjektiviskt förled kallas AN-sammansättning (t.ex. *finkläder*). VN-sammansättning avser en nominal sammansättning med ett verbalt förled (t.ex. *springflod*).

Följaktligen får en adjektivisk sammansättning med ett nominalt förled (t.ex. *snövit*) beteckningen NA. En sammansättning som består av ett adjektiviskt förled och ett adjektiviskt efterled (t.ex. *svartvit*) kallar jag för en AA-sammansättning. Om den adjektiviska sammansättningen innehåller ett verbalt förled kallas den för VA-sammansättning (t.ex. *gråtfärdig*).

I avhandlingen använder jag omväxlande termerna *huvud* och *huvudled* för det led i sammansättningen som är det grammatiskt och semantiskt överordnade. Tolkningar som följer strukturen där ledet till höger i sammansättningen utgör sammansättningens semantiska huvud kallar jag för *högertolkningar*. Följaktligen benämner jag de tolkningar där det vänstra ledet utgör det semantiska huvudledet *vänstertolkningar*.

Från den kognitiva lingvistikens använder jag *schema* om en språklig mall som abstraherats ur och är gemensam för flera språkliga instanser, t.ex. [N-rädd] för *hundrädd* och *fågelrädd*. I detta fall markerar [N] att luckan kan instantieras av olika nomen. Jag använder termen *schema* synonymt med *mönster* (se vidare avsnitt 3.1.3).

Token används i denna avhandling med två något olika betydelser. När jag redogör för antalet token i den referenskorpus som används som utgångspunkt för frekvensbedömningar avses en enhet som i text omgärdas av spatium. I kognitiv lingvistik används tokenfrekvens om antalet instantieringar av ett specifikt ord eller uttryck i ett språkbrukssammanhang, t.ex. antalet förekomster av tokenet *moonless* eller *moonless night*. Detta kontrasteras mot typfrekvens som är ett mått på antalet olika, unika typer som kan fylla en lucka i ett schema eller en konstruktion, t.ex. *moon-*, *fruit-* och *hope-* i [N]-luckan i konstruktionen [N-less].

När det gäller informanternas tolkning av de undersökta sammansättningarna i studien kan jag endast indirekt säga något om den, eftersom tolkning är en kognitiv process. Min analys utgår snarare från en realisering av denna process i form av ett konkret språkligt yttrande. I fritextdelen formulerar informanterna själva förklaringen till ett urval sammansättningar, och dessa förklaringar speglar, åtminstone delvis, deras tolkning. I flervaldsdelen av enkäten presenteras informanterna däremot inför färdiga svarsalternativ, och därmed styrs deras tolkning i vissa riktningar. Syftet med denna testdesign är att analysera specifika aspekter (t.ex. om informanterna väljer en parafra som speglar en struktur där ledet till höger eller ledet till vänster utgör sammansättningens huvudled).

I analysen av informanternas tolkningar tillämpar jag distinktionen L1/L2. Med *L1* avses det språk som en individ tillägnat sig först, medan *L2* avser ett språk som tillägnats efter etableringen av ett förstaspråk. *L1*, *förstaspråk* och *modersmål* används synonymt i avhandlingen, liksom *L2* och *andraspråk* (se vidare diskussion om dessa begrepp i 3.1.8). När jag använder termerna *L1* och *L2* i denna undersökning utgår jag från *svenskan*, dvs. L1- och L2-talare syftar på talare som har svenskan som första- respektive andraspråk. Distinktionen mellan *tillägnande* och *inläring* är inte central i den här studien. Studien handlar om

informerarnas språkliga intuition eller kunskap, oberoende av om denna inhämtats via tillägnande eller inläring.

1.4. Avhandlingens disposition

Avhandlingen omfattar nio kapitel. I detta inledande kapitel har studiens övergripande syfte presenterats och sammansättningen som fenomen introducerats. Jag har också redogjort för min användning av några för avhandlingen centrala begrepp och termer. I kapitel 2 ges en forskningsöversikt över sådan sammansättningsforskning som är relevant för denna studie. Där presenteras dels studier som rör semantisk klassificering, dels studier avseende tolkningen av sammansättningar. De redovisade tolkningsstudierna avser talare med olika språklig bakgrund (L1- och L2-talare) och varierande grad av språklig erfarenhet (vuxna och barn). Kapitel 3 utgör avhandlingens teorikapitel. I detta kapitel redogör jag för de bruksbaserade, kognitiva perspektiv som föreliggande studie utgår från. Jag redogör även för centrala begrepp från lexikal semantik, som jag använder mig av i analysen. I båda dessa delar (kognitiv teori och lexikal semantik) behandlas också flerspråkighetsaspekter. I kapitel 4 beskrivs undersökningens metod och material. Kapitel 5 är ett kvantitativt resultatkapitel, där distributionen av svaren i enkätens flervalsdel redovisas. I kapitel 6, som är ett mer syntetiserande resultatkapitel, relateras svaren i den styrda flervalsdelen till svaren i fritextdelen av enkäten, där informanterna fått formulera egna förklaringar. Därefter följer ett analyskapitel, kapitel 7, där resultaten i enkäten diskuteras i relation till studiens teoretiska ramverk. Kapitel 8 innehåller en sammanfattande diskussion om studiens bidrag till kognitiva, lexikala och flerspråkiga perspektiv. I det avslutande kapitel 9 presenteras slutsatserna från studien. Därefter följer en sammanfattning på engelska.



2

Forskningsöversikt: sammansättningar

I många språk i världen är sammansättning en av de mest produktiva processerna för att bilda nya ord (se t.ex. Booij 2010:93, Scalise & Vogel 2010:1). Sammansättning är emellertid inte ett universellt fenomen (Bauer 2009:344). Trots att produkten av denna ordbildningsprocess – dvs. sammansättningarna – är så grundläggande i många språk, har intresset för dem varit förhållandevis svagt bland språkforskare fram till de senaste decennierna, då sammansättningar tilldragit sig desto mer uppmärksamhet.

I det följande presenteras ett urval olika perspektiv som anlagts på sammansättningar. Generellt ger Lieber & Štekauers (2009) antologi en god översikt över sammansättningsforskning under början av 2000-talet. I antologin presenteras en mångfald av teoretiska och deskriptiva perspektiv på sammansatta ord. Antologin innehåller också en fyllig del med typologiska beskrivningar av sammansättningar i ett antal språk från olika delar av världen. Även Scalise & Vogel (2010) har i en tvärdisciplinär antologi samlat bidrag med olika infallsvinklar på sammansättningar som fenomen.

Inledningsvis i detta kapitel presenteras forskning som syftat till att klassificera sammansättningar efter den semantiska relation som föreligger mellan leden (avsnitt 2.1). Därefter redogörs i avsnitt 2.2 för några stu-

dier om tolkning av sammansatta ord som har betydelse för föreliggande avhandling. I avsnitt 2.3 fokuseras två övergripande och konkurrerande tolkningsstrategier, som identifierats hos språkbrukare vid nominala sammansättningar, nämligen jämförande egenskapsprojicering och tematisk relationslänkning. I det därpå följande avsnittet redovisas forskningsresultat från studier avseende identifiering av det grammatiska och semantiska huvudledet i en sammansättning (avsnitt 2.4). Detta avsnitt är vidare uppdelat i en del som rör tillägnande och förhållningssätt till huvudledets placering i modersmålet (avsnitt 2.4.1) och en del i vilken kontrastiva aspekter på identifieringen av huvudledet behandlas (avsnitt 2.4.2). Avsnitt 2.5 behandlar svenskans sammansättningar ur ett L2-perspektiv.

2.1. Semantiska relationer mellan leden

Under 1960- och 1970-talet handlade sammansättningsforskningen i hög grad om relationstypologier, dvs. semantisk klassificering av sammansättningar. Genom klassificering ville man belysa vilka semantiska relationer mellan leden som aktualiserades i sammansättningar. Utgångspunkten var att alla sammansättningar kunde kategoriseras efter ett begränsat antal strikt definierade semantiska relationer mellan leden (t.ex. DEL–HELHET, SYFTE och PLATS). Eiesland (2016:25) beskriver denna forskningsinriktning som ”the list approach”, eftersom det handlar om en listning av de relationer som förekommer i sammansättningar. Det är värt att notera att dessa försök att klassificera sammansättningar efter semantisk relation var avgränsade till att omfatta sammansättningar som belagts i språkbruket, snarare än att illustrera alla relationer som är tänkbara när en sammansättning bildas. Dessutom omfattade de i de flesta fall endast nominala sammansättningar där såväl förled som efterled utgörs av ett substantiv. På svenskt område presenterade dock Teleman (1970) en klassificering som även omfattar adjektiviska och verbala sammansättningar.

Den forskning som låg bakom relationstypologierna genomfördes ofta mot en generativ bakgrund och utifrån antagandet om att sammansättningar uppstår genom transformation från underliggande relativ- och prepositionsfraser (se t.ex. Lees 1960, Levi 1978). Sammansättningar som inte genererades av grammatiken betraktades som avvikelser som fick hanteras av lexikonet. Den kategorisering som Levi (1978) presenterar har fått stort genomslag och använts i senare psykolingvistisk forskning (se Eiesland 2016:27).

I andra typologier räknade man med ett begränsat antal relationer mellan konstituenterna utan att utgå från förekomsten av transformationella

regler. Eiesland (2016:26) och Ryder (1994:19) lyfter fram Warrens (1978) typologi som ett intressant sådant exempel, eftersom hennes klassificering baseras på frekvenser i ett korpusmaterial. Warrens typologi är också mer finindeldad än de typologier som Lees (1960) och Levi (1978) presenterar, vilka är synnerligen övergripande och därmed vaga. Li (1971) upprättar en typologi som inkluderar kontrastiva aspekter. Han presenterar 24 relationer som förekommer i både engelska och kinesiska sammansättningar (se även Eiesland 2016:29–30).

Semantiska klassificeringar av sammansättningar har senare genomförts också inom ramen för ett kognitivt, bruksbaserat paradig. Ryder (1994) utgår från informantstudier och använder experimentella metoder. Benczes (2006) gör en systematisering av metaforiska och metonymiska sammansättningar. Materialet består av sammansättningar som excerperats ur ordböcker och ur en internetbaserad samling neologismer från amerikansk press. Eiesland (2016) presenterar en klassificering av norska nominala sammansättningar där båda led utgörs av ett substantiv (dvs. NN-sammansättningar), grundad på en analys av 2 000 sammansättningar från korpusen NoWaC (Norwegian Web as Corpus). Av de 14 relationer som identifieras är PLATS ELLER TID den mest frekvent förekommande (15,8 %), följt av UTMÄRKANDE ELLER KARAKTERISERANDE DEL (11,85 %), SYFTE (10,65 %), IDENTITET (9,75 %) och JÄMFÖRELSE (8,2 %) (Eiesland 2016:109).

I klassificeringarna har, som nämnts, nominala sammansättningar fokuserats. Det är symptomatiskt att de typologier som gjort anspråk på att ge en uttömmande lista över de semantiska relationerna i sammansättningar enbart upptar konkreta, nominala sammansättningar – men även i senare sammansättningsforskning har man i förvånansvärt liten grad behandlat exempelvis adjektiviska sammansättningar (se dock t.ex. Teleman 1970, Forza, Guevara & Scalise 2009, Loenheim 2016, Niday 2017).

I en mindre pilotstudie baserad på korpusmaterial (Loenheim 2016) har jag för NA-sammansättningar analyserat styrkeförhållanden mellan olika semantiska (och grammatiska) relationer, dvs. jag har undersökt med vilken frekvens olika semantiska relationer förekommer i NA-sammansättningar. Materialet består av fem korpusar ur Språkbanken (Språkbanken/Korpus; se även Borin et al. 2012; se tabell 1 i bilaga 2). Genomgången av 293 NA-sammansättningar med minst en träff per miljon token visar att den förstärkande relationen (som i *jättebra*) är den mest typ- och teckenfrekventa relationen bland studiens 24 sammansättningskategorier. Den semantiskt närliggande jämförande relationen är den fjärde mest typfrekventa och den femte mest teckenfrekventa. Tillsammans utgör förstärkande och jämförande sammansättningar 26 % av de 293 sammansättningarna. NA-sammansättningar med en orsaksrelation, som

också är relevanta för föreliggande avhandling, har en blygsam representation (se tabell 2 i bilaga 2 för en sammanställning). Genomgången av NA-sammansättningar med efterledet *-kall* respektive *-varm* i samma korpusar visar att den förstärkande relationen är den starkast representerade med avseende på typfrekvens, följt av den jämförande relationen och den orsaksrelaterade (se tabell 3 och 4 i bilaga 2). Även i Nidays (2017) studie om adjektiviska sammansättningar i norskan är jämförelse och förstärkande relation de mest frekvent förekommande relationerna.

Sammanfattningsvis har den klassificerande sammansättningsforskningen blivit alltmer empirisk och korpusbaserad, inriktad på att beskriva det brukliga snarare än att göra anspråk på att ge en heltäckande beskrivning av alla relationer som aktualiseras i sammansättningar. Det transformationella perspektiv som låg bakom sådana relationstypologier som de hos Lees (1960) och Levi (1978) betraktas numera som utdaterat av generativa lingvister (se Eiesland 2016:27). Antagandet om att sammansättningar bildas genom transformationer från underliggande satser har falsifierats i studier av barns språkutveckling, där det framgår att sammansättningar kommer före relativsatser i barnens språk (se Søgaard 2005:320). Dessutom tar många forskare avstånd från antagandet om ett begränsat antal möjliga relationer mellan för- och efterled i sammansättningar (se t.ex. Downing 1977:840–841, Murphy 1988:555–557, Kay & Zimmer 1990:239–240, Costello & Keane 2000:325, Plag 2003:148). I stället betonas sammansättningsrikas tolkningsmöjligheter.

2.2. Tolkning av sammansättningar

Sammansättningsforskningen i stort har som sagt rört sig i riktning mot mer empiriska undersökningar av faktiskt språkbruk. I linje med detta har även språkbrukares tolkningar av sammansättningar kommit att tilldra sig större intresse.

Downing (1977) sätter tydligt språkbrukarna i centrum. Hennes studie är pragmatiskt inriktad med fokus på tolkning och produktion av sammansättningar. Receptionsdelen innehåller sammansättningar som är etablerade i olika grad (1977:817–818). I studien framträder en mer komplex bild av sammansättningsrikas semantiska relationer än den bild som tidigare presenterats i semantiska klassificeringar. Downing kommer fram till att många av informanternas tolkningar är ”at best REDUCIBLE to underlying relationships as suggested by Li and others, but only with the loss of much of the semantic material considered by the subjects to be relevant or essential to the definitions” (Downing 1977:826; hennes emfas). Informanternas tolkningar ger med andra ord mer detaljerad semantisk

information om sammansättningarna än den information som följer av klassificeringarna. Dessutom förekommer tolkningar som svårigen låter sig inordnas i de föreslagna klassificeringarna, jämför *cow-tree* som ges fritextförklaringen 'a tree that cows like to rub up against' (Downing 1977:826–828, citatet från s. 827).

Ett annat intressant resultat som Downing noterar är att informanterna i hög grad gör tolkningar med en jämförelserelation mellan leden, när de i ett kontextfritt tolkningstest ges i uppgift att ge förklaringar till ett antal nya, oetablerade sammansättningar. Detta avviker från resultatet i en annan del av studien, där hon gör en bedömning av den semantiska relationen mellan leden hos belagda men icke etablerade sammansättningar som eliciterats ur ett insamlat material med texter av skilda slag. Bland dessa förekommer inte jämförelse som semantisk relation mellan leden. Downing drar slutsatsen att informanterna använder jämförelserelationen 'an x like y' som en "last-resort relationship" när ingen annan relation mellan leden är uppenbar. Den oetablerade sammansättningen *autumn pebble* ges exempelvis förklaringen 'many-colored pebble' (Downing 1977:830).

Även Ryder (1994) har språkbrukarna i fokus i sin avhandling. Studien är banbrytande på så sätt att Ryder tillämpar schemateori i analysen av sammansättningsbruk. Ur ett kognitivt perspektiv söker hon identifiera språkbrukarnas strategier när de skapar och tolkar nya sammansättningar. Produktions- och tolkningsprocesserna beskrivs i termer av ett "ordered chaos", i vilken språkliga mallar eller förebilder (eng. *linguistic templates*) ger vägledning. Den specifika mall som används för att tolka eller producera en ny sammansättning kallar Ryder för analogibas (eng. *analogy base*). En sådan analogibas kan utgöras av en enskild sammansättning (t.ex. *doghouse*) eller en mall baserad på en grupp av sammansättningar med ett gemensamt kärnord (t.ex. *sea* i sammansättningar som *sea lion*, *seaman* och *seaweed* eller *house* i *boathouse*, *warehouse*, *tree house*). Även omvärldskunskapen är betydelsefull i tolkningsprocessen, framhåller Ryder (1994:79–80, 89–90). Ryders modell kommer att presenteras mer ingående i avsnitt 3.1.5.

Hur språkbrukare tolkar nya sammansättningar har också undersökts i psykolingvistiska studier. Utifrån experiment med nybildade sammansättningar har Gagné & Shoben (1997) presenterat CARIN-modellen (Competition Among Relations In Nominals; se även Gagné 2002, Gagné & Spalding 2007, 2010). Modellen utgår från att vi lagrar relationspreferenser, dvs. information om vilka relationer specifika element brukar ha till varandra, t.ex. att *choklad* som modifierare brukar ange vad huvudledets referent är gjord av, medan *berg* som modifierare brukar markera plats. I CARIN-modellen antas språkbrukarna lagra relationspreferenser för modifierarna i högre grad än för huvudleden. Detta innebär att det blir

lättare att tolka en ny sammansättning om kombinationen överensstämmer med relationspreferenserna för modifieraren. Modellen har fått visst stöd i reaktionsstudier som visar att informanterna snabbare kunde avgöra om en ordkombination var en tolkningsbar engelsk sammansättning i de fall där kombinationen överensstämde med relationspreferenserna för modifieraren, medan relationspreferenserna för huvudledet hade ingen, eller mycket begränsad, inverkan på reaktionstiden (Gagné & Shoben 1997:73, 80–81, 83–85).

CARIN-modellens antagande om att mönstren för modifierarna är viktigare än mönstren för efterleden, när språkbrukare tolkar nya sammansättningar, har dock ifrågasatts i andra studier. Maguire et al. (2010a, 2010b) bekräftar genom experiment CARIN-modellens antaganden om en kortare reaktionstid i de fall där den nya kombinationen överensstämmer med relationspreferenser. Men forskarna visar att relationspreferenserna verkar på en mer abstrakt nivå. De gäller semantiska kategorier, snarare än enskilda ordled. En ny sammansättning som *stone squirrel* tolkas gärna som en instantiering av det produktiva mönstret [substance–object], vilket ger relationen <made of>. I denna tolkningsprocess utgår tolkningen från de båda semantiska kategorierna 'substans' och 'objekt'. Preferensen för detta mönster är så stark att det slår igenom även när det resulterar i en pragmatiskt mindre rimlig tolkning. I ett experiment tolkas exempelvis *leather needle* i hög grad som en 'nål av läder'. Forskarnas analys är att en sådan tolkning baseras på semantiska kategorier i en tolkningsprocess som involverar både modifieraren och huvudledet. De argumenterar för en interaktionell statistisk effekt där båda konstituenterna spelar en viktig roll (Maguire et al. 2010a:290, 294, Maguire et al. 2010b:64–66, 68–69).

I en annan experimentell studie har Gagné & Spalding (2010) visat att det vid tolkningen av sammansättningar finns ett samband mellan antalet möjliga ledrelationer och processningstiden. Ju fler potentiella, konkurrerande ledrelationer en sammansättning har, desto längre processningstid (2010:287, 299).

Sammansättningar har också uppmärksammats i barnspråksstudier. Mellenius (1997) redovisar resultaten från en sådan studie som rör barns produktion och tolkning av sammansättningar. Hon konstaterar att barn tidigt både förstår och producerar sådana. NN-sammansättningar är den struktur som kommer först hos barnen, redan före tre års ålder, följt av sammansättningar av annan struktur (1997:75–76).

Receptionsdelen av Mellenius (1997) studie omfattar två delar. Dels omfattar den ett språkligt experiment baserat på bildidentifiering, där 60 barn ges i uppgift att identifiera det semantiska huvudledet i ett antal sammansättningar. Dels omfattar den en longitudinell studie med 71 barn, där

Mellenius undersöker vilka semantiska relationer som är prefererade hos barnen i tolkningen av ett antal sammansättningar.

I den longitudinella delstudien som behandlar semantiska relationer följer Mellenius (1997) en årskull med 71 barn, som hon träffar vid fyra tillfällen, en gång om året, från det att de är 6–7 år gamla till det att de är 9–10 år gamla. Barnen får förklara samma sammansättningar vid varje undersökningstillfälle. Det rör sig om tillfälliga och mer etablerade sammansättningar (t.ex. *ormgräs/gräsorm*, *trädfågell/fågelträd* och *skoläder/lädersko*). Svaren analyseras sedan utifrån relationerna PLATS, GENITIV, LIKHET, PREFERENS, MATERIAL, SYFTE, TVÅNG, ÄGANDE, SAMEXISTENS, KÄLLA och SMAK. Av dessa visar sig PLATS och MATERIAL vara de relationer som används mest frekvent i barnens tolkningar (1997:110, 115–117, 155). Ett intressant resultat är att de nybildade sammansättningarna *gräsorm* och *marsipanapelsin* får alltfler materialtolkningar i takt med att barnen blir äldre (1997:122, 147). Det är också tydligt att samstämmigheten, som förväntat, är betydligt större vid de etablerade sammansättningarna (t.ex. *pälsdjur*, *lädersko* och *fågelträd*) än vid de nya sammansättningarna (t.ex. *trädfågel*, *gräsorm* och *lushår*). Samstämmigheten ökar också något (men måttligt) från det första undersökningstillfället till det sista, särskilt vid de etablerade sammansättningarna, som får alltfler konventionella tolkningar. Den tydligaste utvecklingen i barnens svar är emellertid att andelen svar som speglar strukturen med sammansättningens huvud till höger ökar i takt med att barnen blir äldre (Mellenius 1997:111–112, 124–126, 155; se vidare avsnitt 2.4.1).

Svanlund (2009) analyserar i en kombinerad korpus- och enkätstudie etableringsprocessen för nio nya sammansättningar i svenskan i början av 2000-talet, bl.a. *curlingförälder* och *stafettläkare*. Samstämmigheten i informanternas förklaringar av de relativt nya orden är förhållandevis stor. Svanlund drar av detta slutsatsen att etableringsprocessen påverkar ett ords användnings- och tolkningspotential på så sätt att några tolkningar ur potentialen successivt framträder som mer sannolika. Etableringsprocessen påverkar därigenom sannolikheten för att en viss sammansättning ska användas på ett visst sätt och i en viss betydelse. Samtidigt kan kontextuella faktorer bidra till att modulera etablerade betydelser. Till skillnad från vad som antas inom CARIN-modellen tyder resultaten på att mönstren för sammansättningarnas efterled påverkar tolkningarna mer än mönstren för modifierarna (Svanlund 2009:219, 221, 241–242).

Sammanfattningsvis illustrerar receptionsstudierna om sammansättningar att det mesta är möjligt. Möjligt är dock något annat än brukligt. Om en sammansättning konventionaliserats i en viss betydelse konkurrerar denna ut andra potentiella betydelser (se vidare 3.2.2). Vid nybildade sammansättningar som inte genomgått konventionalisering är flera

alternativa betydelse tillgängliga, men samtidigt visar flertalet studier att mönster på olika nivåer förefaller påverka tolkningarna och styra dem i vissa riktningar. Enligt Ryder (1994) kan både modifierare och huvudled från bekanta sammansättningar eller sammansättningsfamiljer fungera som analogibaser vid tolkningen av nya sammansättningar, liksom mer abstrakta mönster. Somliga forskare betonar relationspreferenserna för modifierarna (CARIN-modellen, Gagné & Shoben 1997), medan andra noterar att mönstren för huvudledet förefaller påverka tolkningarna mest (se Svanlund 2009:219–220). Maguire et al. (2010a, 2010b) argumenterar för en interaktionell modell där semantiska kategorier från sammansättningens båda domäner samverkar. Det som förenar dessa perspektiv är att man framhåller hur tidigare språkliga erfarenheter, och lagrade mönster från dessa, påverkar våra tolkningar när vi möter nya sammansättningar. Med hjälp av sådana lagrade mönster omvandlas potentiell betydelse till aktiverad betydelse i en viss kontext.

För att återknytta till relationstypologierna framstår det som meningslöst att på förhand försöka fastställa ett begränsat antal semantiska relationer mellan leden i sammansättningar, eftersom sammansättningar kan tolkas på så många olika sätt (se t.ex. Downing 1977:840–841). Empiriskt grundade typologier, såsom Eieslands (2016) klassificering, kan emellertid bidra med att ur det möjliga skilja ut det brukliga. Därmed kan de illustrera konventionaliserade mönster i språket.

2.3. Egenskapsprojicering och relationslänkning

Två övergripande och konkurrerande tolkningsstrategier som identifierats för nominala sammansättningar är å ena sidan strategier där modifieraren antas ange egenskapsdimensioner, å andra sidan strategier där ledens referenter antas fylla olika funktionella roller i en direkt, tematisk relation till varandra (se t.ex. Gagné & Shoben 1997, Wisniewski & Love 1998, Costello & Keane 2000). Svanlund (2009) benämner strategierna *egenskapsprojicering* respektive *relationslänkning*. Egenskapsprojicering innebär att en egenskap som är typisk för modifierarens referent projiceras på referenten för huvudledet, dvs. huvudledets referent jämförs i något avseende med modifierarens referent. Med egenskapsprojicering blir betydelsen hos *tigerhaj* sålunda 'haj som är randig som en tiger'. Egenskapsprojicerande sammansättningar fungerar därmed jämförande. Relationslänkning innebär att det i stället förekommer en tematisk relation mellan ledens referenter, där de intar olika funktionella roller,

genom vilka de förhåller sig till varandra. *Tigerhaj* i betydelsen 'haj som jagar tigrar' är ett exempel på sådan relationslänkning, där hajen antar en agentiv subjektroll medan tigern utgör verbhandlingens objekt (Svanlund 2009:25).

Terminologin på området framstår som något förvirrande. För det första är Svanlunds (2009) term *egenskapsprojicering* en modifiering av "property interpretation" hos Wisniewski & Love (1998). Svanlunds term fokuserar den process genom vilken en eller flera egenskaper från modifierarens referent överförs till huvudledets referent. För det andra är motsatsförhållandet inte uppenbart mellan å ena sidan egenskapstolkning/egenskapsprojicering, där leden inte antas stå i direkt relation till varandra, och å andra sidan relationslänkning, där en sådan relation antas föreligga. I kontrast till det antagna motsatsförhållandet skulle även egenskapstolkning kunna betraktas som baserad på en relation mellan ledens referenter, nämligen en jämförande relation. En sådan användning noteras hos Ryder (1994), som gör bruk av termen *likhet*, och hos Eiesland (2016) som använder termen *jämförelse* motsvarande *egenskapstolkning* hos Wisniewski & Love (1998) och *egenskapsprojicering* hos Svanlund (2009). Ryder (1994) och Eiesland (2016) ger ingen paraplyterm motsvarande *tematisk relation* eller *relationslänkning* för de relationer som inte hör hemma under beteckningen *likhet* respektive *jämförelse*. En översikt över terminologin ges i tabell 2:1.

TABELL 2:1. *Sammanställning över terminologin avseende egenskaps- och relationstolkning i tidigare studier.*

Wisniewski & Love (1998)	Svanlund (2009)	Downing (1977)	Ryder (1994)	Eiesland (2016)
Egenskapstolkning <i>robin hawk</i> 'hök med rött bröst som en rödhake'	Egenskapsprojicering <i>hökuggla</i> 'uggla som till utseendet liknar en hök'	Jämförelse <i>pumpkin bus</i> 'buss som ser ut som en pumpa'	Likhet <i>tiger-lion</i> 'lejon som är randigt som en tiger'	Jämförelse <i>leopardskilpadde</i> 'sköldpadda som är mönstrad som en leopard'
Tematisk relation <i>robin hawk</i> 'hök som jagar rödhakar'	Relationslänkning <i>dwhök</i> 'hök som (antas) jaga duvor'

Med viss modifiering av terminologin men i anslutning till tidigare forskning använder jag i det följande termen *jämförande egenskapsprojicering* för sådana tolkningar där huvudledets referent jämförs med en utmärkande

egenskap hos modifierarens referent. Jag kommer omväxlande att använda termen *jämförande egenskapstolkning* för samma semantiska relation, särskilt vid personbetecknande sammansättningar, där det sällan går att urskilja *en* specifik egenskap hos förledets referent som projiceras på huvudledets referent. När personbetecknande sammansättningar fungerar jämförande antas huvudledets referent snarare uppbära flera av de egenskaper som normalt tillskrivs förledets referent (jfr *pojkflicka*). Mot detta ställer jag termen (*tematisk*) *relationslänkning* om relationer som bygger på att ledens referenter intar olika funktionella roller i förhållande till varandra (dvs. om relationer som *inte* bygger på jämförelse).

För svenskans del diskuterar Svanlund (2009:25) hur beteckningarna för olika fågelarter fördelar sig mellan ett egenskapsprojicerande schema (*hökuggla*, *sparvuggla*) och ett relationslänkande jaktschema (*duvhök*, *sparvhök*). Svanlund (2013) ställer frågan enligt vilket mönster en påhitad art som *korporn* skulle tolkas.

I valet mellan jämförande egenskapsprojicering och tematisk relationslänkning får omvärldskunskapen ofta avgöra vilket av tolkningsalternativen som är det mest rimliga. Många gånger är det förstås mindre självklart än i fallet med *tigerhaj* vilken tolkning som är rimligast ur ett semantiskt eller pragmatiskt perspektiv, och olika forskare har gjort olika bedömningar av vilken strategi som är den mest grundläggande och tillgängliga för språkbrukarna. Enligt Downing (1977) och Gagné & Shoben (1997) är tematisk relationslänkning språkbrukarens förstahandsval. Jämförande egenskapstolkning är något språkbrukaren tar till endast som en sista utväg, om ingen tematisk relation förefaller rimlig, såsom en "last-resort or default relationship" (Downing 1977:830). Wisniewski och Love (1998) tillbakavisar denna hypotes som de kallar för "the last resort hypothesis". Genom ett flertal experiment visar de att egenskapstolkning är en grundläggande tolkningsstrategi som är jämbördig med tematisk relationslänkning.

I ett experiment där Ryder (1994) undersöker betydelsen av språkliga förebilder eller mallar i tolkningen av nya sammansättningar framkommer det att både det etablerade jaktschemat (relationslänkning) och det etablerade likhetsschemat (egenskapsprojicering) har starkt inflytande på tolkningarna när informanterna fritt formulerar sina tolkningar av ett antal djur-djur-sammansättningar. Vid sammansättningarna *rabbit-dog*, *cow-wolf* och *squirrel-hound* fördelar sig tolkningarna förhållandevis lika mellan relationslänkning och egenskapsprojicering. Ryder diskuterar inflytande från etablerade sammansättningar med *-hund*, vilka ofta följer ett relationslänkande mönster. Det finns dock en tendens i Ryders material att sammansättningar där förledet betecknar ett djur med ett karakteristiskt drag (t.ex. *elephant-fish*, *tiger-lion*, *kangaroo-horse*) ges egenskapsproji-

cerande, dvs. jämförande, tolkningar, i betydligt högre grad än relationslänkande tolkningar (Ryder 1994:84, 117–118, 355–357, 359–362).

Eiesland (2016:130–131) konstaterar i sin omfattande korpusstudie av norska sammansättningar att jämförande egenskapsprojicering är särskilt vanlig vid sammansättningar som betecknar djur och växter, dvs. vid sammansättningar där huvudledets referent är ett djur eller en växt. I sådana sammansättningar jämförs ofta djuret/växten med artefakter (t.ex. *gaffelmose*), andra växter/djur (*leopardskilpadde*) eller personer/roller (*kongemakrell*). Liknande resultat presenteras av Wisniewski & Love (1998). I en studie av 1 400 sammansättningar, fördelade på lika delar sammansättningar som betecknar djur, växter och artefakter, identifierar forskarna en särskilt stor andel egenskapsprojicering (42,5 %) vid sammansättningar som betecknar djur och växter (Wisniewski & Love 1998:195).

2.4. Sammansättningsens semantiska huvudled

Regeln (eller schemat) avseende det grammatiska och semantiska huvudledets placering i sammansättningar, har uppmärksammats i språkvetenskapliga sammanhang, ofta utifrån ett generativt eller psykolingvistiskt perspektiv. För svenska språkets del har man antagit att svenskan följer en universell preferens för en sammansättningsstruktur med huvudledet till höger och en underordnad modifierare till vänster (SAG 2:42, 186, 526). Detta har närmast betraktats som ett axiom. Sålunda har frågan inte undersökts empiriskt i större utsträckning, särskilt inte vad gäller sammansättningar av annan struktur än nominala, och särskilt inte när det gäller vuxnas språk. Williams (1981:248) beskriver den högerställda strukturen i termer av Right Hand Head Rule. I generativa modeller har denna regel betraktats som en absolut regel som ingen modersmålstalare avviker från (se även Selkirk 1982).

Även om den empiriska forskningen varit begränsad, förekommer det ändå studier som uppmärksammar frågan om hur modersmålstalare tillägnar sig och förhåller sig till regeln eller schemat för huvudledets placering. Resultat från några studier av detta slag redovisas i det följande (avsnitt 2.4.1).

I kontrastiva studier har man undersökt eventuella transfereffekter mellan språk med olika placering av sammansättningsens grammatiska och semantiska huvudled. Resultat från kontrastiv forskning av detta slag redovisas i det följande (avsnitt 2.4.2).

2.4.1. Förhållningssätt till schemat för huvudledets placering

Mellenius (1997) behandlar barns produktion och reception av nominala sammansättningar. En aspekt som studeras är hur barnen förhåller sig till regeln om det semantiska huvudledets placering i sammansatta ord. Mellenius visar att barn tidigt kan identifiera huvudledet i en sammansättning. Hon redovisar resultat från ett experiment med bildidentifiering (influerat av Clark, Gelman & Lane 1985), där 60 barn i åldrarna 2–5 år fick i uppgift att peka ut den bild som överensstämde med en specifik sammansättning. Experimentet omfattade 24 tillfälliga NN-sammansättningar med kända ordled. Samtliga barn hade svenska som sitt enda modersmål. För varje testord (t.ex. *ormglas*) fick barnen välja mellan fyra bilder som illustrerade referenten för antingen sammansättningens huvudled (*glas*), sammansättningens modifierare (*orm*), ett objekt som påminde om det objekt som huvudet betecknade (*mugg*) eller ett objekt som påminde om det objekt som modifieraren betecknade (*fisk*) (Mellenius 1997:94–97). I åldersgruppen 3 år och 4 månader till 3 år och 11 månader (dvs. 3;4–3;11) hade barnen i genomsnitt 80 % korrekt identifiering av huvudledet, jämfört med 69 % i åldersgruppen 2;8–3;3 och 56 % i åldersgruppen 2;0–2;7. Mellenius konstaterar att barnens förmåga att utläsa sammansättningens huvudled ökar i takt med att de blir äldre, och att den är förhållandevis etablerad i åldersgruppen 3;4–3;11, där medelåldern är 3;8 år (Mellenius 1997:98–99).

Det är värt att notera att en del av tolkningarna i Mellenius (1997) experiment är semantiskt, snarare än strukturellt, avvikande. Den vanligaste feltypen i experimentet är tolkningar där förledet pekas ut som huvudled, men det förekommer också tolkningar, särskilt hos de yngsta barnen, där ett objekt som liknar huvudledets referent pekas ut som huvudled (t.ex. bilden av en mugg vid sammansättningen *ormglas*). Även dessa tolkningar räknas som felaktiga, men feltypen säger egentligen ingenting om informantens behärskning av strukturen modifierare-huvud (1997:100–101).

Eftersom Mellenius studie följer samma testdesign som Clark, Gelman & Lane's (1985) studie är också resultaten jämförbara. I deras experiment identifieras huvudledet på korrekt sätt av barnen i 82 % av fallen i åldersgruppen 3;4, att jämföra med 48 % i åldersgruppen 2;4 (Clark, Gelman & Lane 1985:86). Resultaten från den engelska och den svenska studien följs således åt och tyder på att barn tidigt tillägnar sig strukturen för huvudledets placering i modersmålets sammansättningar.

Andra studier visar emellertid att det tar flera år innan barn till fullo förstår hur modifierare och huvud hänger ihop och tillsammans bildar en konceptuell enhet. Nicoladis (2003) presenterar ett experiment där 35 barn i 3–4-årsåldern fick i uppgift att välja den bild som motsvarade

referenten för en ny sammansättning, t.ex. *dragon box*. De kunde välja mellan fyra bilder: en där modifierarens referent var avbildad, en där huvudledets referent var avbildad, en där både modifierarens och huvudledets referenter var avbildade men vid sidan av varandra och slutligen en bild som motsvarade sammansättningens betydelse, där modifierarens och huvudledets referenter var integrerade. Treåringarna valde i betydligt större utsträckning än fyraåringarna en bild som motsvarade den ena av referenterna (t.ex. bilden av en drake eller en låda vid *dragon box*). Både tre- och fyraåringarna valde i cirka en tredjedel av fallen en bild där båda referenterna förekom, men vid sidan av varandra, t.ex. en bild som visade en drake intill en låda, i stället för en låda dekorerad med drakar. Hälften av treåringarna och två tredjedelar av fyraåringarna valde den bild som speglade sammansättningens betydelse, dvs. där ledens referenter var integrerade med varandra. Sammanfattningsvis visar detta att barnens uppfattning om en sammansättning som en konceptuell enhet av två integrerade delar ännu är under utveckling i 3–4-årsåldern (Nicoladis 2003:45, 47). Krott (2009:131) refererar en studie av Parault, Schwanenflugel & Haverback (2005). I ett experiment med nya NN-sammansättningar visar forskarna att barn ännu i sexårsåldern i hög grad ger förklaringar där de ingående komponenterna förekommer sida vid sida snarare än integrerade med varandra, t.ex. när *book magazine* ges förklaringen ”a big magazine and a little book”.

I sin longitudinella delstudie om semantiska relationer i sammansättningstolkningar (se avsnitt 2.2), följde Mellenius 71 barn från det att de var 6–7 år gamla till det att de var 9–10 år gamla. Hon analyserar hur de förklarar de semantiska relationerna mellan delbegreppen i ett antal etablerade respektive nya sammansättningar. I en kvantitativ analys av samma material fokuserar Mellenius utvecklingen när det gäller identifieringen av sammansättningens huvudled. Medelvärden avseende tolkningar med huvudledet till vänster sjunker signifikant från mätningen vid sju års ålder (1,44 av 8 vänstertolkningar) till mätningen vid åtta års ålder (0,78 av 8 vänstertolkningar). Analysen omfattar de 54 barn som deltagit vid samtliga fyra testtillfällen (Mellenius 1997:111–112).

Inte bara i barnstudier, utan också i informantstudier med vuxna (Ryder 1994, Onysko 2016) förekommer i språk med en högerstruktur tolkningar där förledet betraktas som sammansättningens huvudled, i opposition mot förutsägelse hos Williams (1981:248) och Selkirk (1982); se även Costello & Keane (2000:325).

I ett av Ryders (1994) experiment uppgår vänstertolkningarna till 3–5 % av det totala antalet tolkningar. Utifrån detta resultat ifrågasätter hon föreställningen om att vänstertolkningar skulle kunna avfärdas som performansfel (1994:138).

I en annan informantstudie, där 116 informanter tolkar tolv nybildade engelska sammansättningar visar Onysko (2016) att NN-sammansättningarna som har ett agentivt huvud – *bucket philosopher* och *board poet* – tolkas med huvudledet till höger i betydligt högre grad än de övriga. Detta ser Onysko som en indikation på att sammansättningar med agentivt huvud är mer resistent än andra mot att tolkas enligt en omvänd vänsterstruktur (Onysko 2016:491). Med andra ord förefaller det kognitivt tilltalande för informanter att tolka sammansättningens agentiva element som sammansättningens huvud. Detta visar – inte helt oväntat (men i strid med Williams 1981:248 och Selkirk 1982) – att inte bara strukturella utan också semantiska aspekter påverkar språkbrukarna när de identifierar sammansättningens huvudled.

Något som diskuterats för svenskans del är i vilken grad fogen mellan konstituenterna i längre sammansättningar ger vägledning om det semantiska huvudledets räckvidd. Det är utmärkande för svenskan (och flera andra germanska språk) att man kan bilda sammansättningar med fler än två konstituent. Om förledet i en sammansättning i sig är en sammansättning används vanligen en fog, t.ex. *s* framför efterledet, så att *barnbok + hylla* blir *barnbokshylla*. Detta förhållande går under beteckningen Tamms fogeregeln efter språkvetaren Fredrik Tamm. Noreen (1906:418–420) ger en detaljerad översikt över undantagen från denna regel. Huruvida fogeregeln tillämpas eller inte har att göra med formen på ordslutet i det sammansatta förledet, konstaterar han (se även Josefsson 1997:59–61). I modern tid förekommer regeln som en skrivrekommendation (se Karlsson 2017:148). Utifrån regeln om foga-*s* framför led tre har man i tidigare forskning noterat den strukturella skillnaden mellan orden i sammansättningspar som *barnbokhylla*–*barnbokshylla*. Utifrån fogeregeln är strukturen hos *barnbokshylla* {*barnbok + s + hylla*}, med *hylla* som sammansättningens huvudled, vilket ger betydelsen 'en hylla för barnböcker'. En *barnbokhylla* å andra sidan, utan foga-*s*, borde enligt samma regel analyseras {*barn + bokhylla*}, vilket ger betydelsen 'en bokhylla för barn' (se Bergman 1962:43). Det förefaller emellertid svårt att finna sådana sammansättningspar där fogen fungerar semantiskt särskiljande mellan de två potentiella strukturerna.

Sammanfattningsvis visar flertalet studier (avseende svenska och engelska) att barn tidigt lär sig att identifiera huvudledet i nominala sammansättningar. Hur det ser ut i adjektiviska sammansättningar är oklart, eftersom sådana sammansättningar uppmärksammats i mycket liten grad (se dock Gleitman & Gleitman 1970). Vidare har det diskuterats om förekomsten respektive frånvaron av ett fogaelement i treledade sammansättningar skulle kunna markera det semantiska huvudledets räckvidd, men frågan har inte undersökts empiriskt.

I tidigare sammansättningsforskning har det också framkommit att det förefaller lättare att identifiera huvudledet i nybildade sammansättningar om huvudledets referent är agentiv (se Onysko 2016), vilket talar för att inte bara grammatiska utan också semantiska eller kognitiva aspekter har betydelse för hur sammansättningen tolkas.

2.4.2. Transfer från L1 till L2 vid identifiering av huvudledet

Frågan om hur man identifierar sammansättningens huvudled är intressant ur ett kontrastivt flerspråkighetsperspektiv. I flera kontrastiva studier har man undersökt i vilken mån språkbrukares tolkningar i deras L2 påverkas av sammansättningsstrukturen i deras L1. Resultaten från sådana studier har betydelse för denna studie där en stor del av informantgruppen är flerspråkiga. Tidigare har man antagit att transfereffekter förekommer mellan språken hos en flerspråkig individ endast när det finns en strukturell överlappning mellan språken, på så sätt att språken delar en grammatisk konstruktion. Transfer av en grammatisk konstruktion från språk A till språk B har ansetts kunna ske endast när konstruktionen också förekommer i språk B vid sidan av en konkurrerande konstruktion, dvs. när språk B har två alternativa uttryckssätt (Müller 1998; Döpke 1998). I kontrast till detta antagande påvisar flera nyare studier om tolkning och produktion av sammansättningar förekomst av transfereffekter även i fall där språken har olika sammansättningsstruktur, dvs. även i fall där det inte föreligger någon strukturell överlappning mellan språken när det gäller huvudledets placering.

Ett exempel på sådana transfereffekter framkommer hos Nicoladis (2002). Nicoladis redovisar resultaten av ett produktionstest med informanter i åldern 3–4 år, där 25 barn tvåspråkiga i franska och engelska matchades parvis (efter ålder) med informanter i en kontrollgrupp av 25 enspråkiga, engelsktalande barn. Hon konstaterar tydliga transfereffekter hos de tvåspråkiga barnen, trots att ingen strukturell överlappning föreligger mellan språken. I produktionen av engelska sammansättningar, som eliciterades utifrån bilder, producerade de tvåspråkiga barnen nästan dubbelt så många sammansättningar med huvudledet till vänster som kontrollgruppen, vilket Nicoladis förklarar med inflytande från franskan, där huvudet står till vänster och bestämningen till höger i NN-sammansättningar (med och utan preposition) och nominalfraser (Nicoladis 2002:843–844, 858, 861–862).

Foroodi-Nejad & Paradis (2009) genomförde ett liknande produktionstest med fokus på huvudledets placering i sammansättningar. Barn som var tvåspråkiga i persiska och engelska jämfördes med enspråkiga kontrollgrupper av persisk- respektive engelsktalande barn. Studien om-

fattade 52 barn i fyraårsåldern. De undersökta språken i Foroodi-Nejad & Paradis (2009) studie skiljer sig på ett avgörande sätt från språken hos Nicoladis (2002). Persiska sammansättningar har vanligen huvudledet till vänster, men även högerställning förekommer, vilket gör att det finns en strukturell överlappning mellan persiskan och engelskan. I enlighet med hypotesen om en strukturell överlappning kunde man därmed förvänta sig transfereffekter från engelskan (endast huvudet-till-höger-sammansättningar) till persiskan (både huvudet-till-höger och huvudet-till-vänster), men inga transfereffekter i motsatt riktning. I kontrast till detta antagande påvisar Foroodi-Nejad & Paradis (2009) transfereffekter inte bara från engelska till persiska, utan också från persiska till engelska, trots att det persiska mönstret med huvudet till vänster inte har någon motsvarighet i engelskan. Sålunda förefaller de tvåspråkiga barnen påverkas av sina två tillägnade mönster för sammansättningsstruktur i produktionen av både engelska och persiska sammansättningar, oavsett om det föreligger en strukturell överlappning eller inte (Foroodi-Nejad & Paradis 2009:416, 419–420, 423).

Onysko (2016) lyfter svårigheterna med att undersöka identifiering av huvudledet i sammansättningar utifrån bildbaserade produktionstest som hos Nicoladis (2002) och Foroodi-Nejad & Paradis (2009). Han menar att bilden kan frammana olika konceptualiseringar avseende fokuserat föremål (eng. *figure*) och bakgrund (eng. *ground*) hos olika individer. I de nämnda studierna fick barnen först se en bild av ett föremål, t.ex. en bil, följt av en bild av ett annat föremål, t.ex. ballonger. Därefter visades en bild av de två föremålen integrerade, t.ex. en bil med ballonger avbildade utanpå. Bilden förväntades frammana sammansättningen *balloon car* hos barnen (utifrån ett förberedande och kontrollerande test med vuxna). Ett metodologiskt problem är att barnen skulle kunna uppfatta ballongen (snarare än bilen) som det viktiga föremålet (*figure*) som råkade befinna sig på en bil (*ground*). Utifrån en sådan konceptualisering skulle *car balloon* spegla en struktur med huvudet till höger, jfr '(drawings of) a balloon found on a car' (Onysko 2016:474; se även Langacker 2008:57, 72). Nicoladis (2002:859) tar själv upp frågan om att vissa av sammansättningarna i hennes test skulle kunna konceptualiseras på olika sätt, och möjligtvis olika av barn och vuxna.

Onysko undersöker vidare förekomsten av tvärspråkligt inflytande, med avseende på sammansättningsstruktur, utifrån en alternativ test-design, där 116 vuxna informanter från två typologiskt skilda språk – maori och engelska – tolkar ett antal nybildade, engelska sammansättningar i ett receptionstest. Maori har en sammansättningsstruktur där huvudledet är placerat till vänster i sammansättningen i motsats till sammansättningsstrukturen i engelska. I studien jämfördes en grupp av två-

språkiga informanter med engelska och maori som språk med enspråkiga kontrollgrupper i maori respektive engelska samt en tvåspråkig kontrollgrupp i engelska och ytterligare ett språk annat än maori. Testet innehöll sammansättningar där kända led från bekanta semantiska domäner kombinerades (t.ex. väder + kroppsdel, djur + byggnad, instrument + profession) och bildade nya sammansättningar (t.ex. *fog ear*, *spider cafeteria*, *board poet*). Studien visar att de informanter som är tvåspråkiga i maori och engelska har signifikant fler tolkningar med huvudet till vänster än å ena sidan de enspråkiga kontrollgrupperna, å andra sidan den tvåspråkiga kontrollgruppen i engelska och ett annat språk än maori. Onysko drar slutsatsen att resultatet speglar ett tvärspråkligt inflytande på engelskan från ett, med avseende på sammansättningsstruktur, typologiskt kontrasterande språk. Studien bekräftar att transfereffekter kan förekomma hos flerspråkiga personer även om ingen strukturell överlappning föreligger mellan strukturerna i de respektive språken (Onysko 2016:479–481, 487–488, 490–492). Med hänvisning till Grosjean (2001) ser Onysko detta som ett tecken på att båda språken hela tiden är aktiverade hos en tvåspråkig individ:

On a broader scale, the results suggest that the diverse languages of bilingual and multilingual speakers are always active and can shine through on different levels of language use. (Onysko 2016:492)

Sammanfattningsvis har man i ett antal kontrastiva studier identifierat transfereffekter avseende tolkningen av sammansättningens huvudled. Språkbrukarnas L1 har visat sig påverka deras identifiering av huvudledet i L2. Transfereffekter har konstaterats även i de fall där det inte föreligger någon strukturell överlappning mellan de aktuella språken.

2.5. Sammansättningar ur ett L2-perspektiv

I avsnitt 2.4.2 strax ovan redovisades internationell forskning som anlagt ett kontrastivt flerspråkighetsperspektiv på identifieringen av huvudledets placering i sammansättningar. I det följande redovisas forskning som anlagt ett flerspråkighets- eller L2-perspektiv på svenskans sammansättningar.

Det är dokumenterat sedan tidigare att svenskans sammansättningar kan vara problematiska ur ett L2-perspektiv. De är svårtolkade av flera skäl: de är ofta långa och kan därmed vara svåra att segmentera, och såväl de ingående leden som relationen mellan dem kan rymma polysemi. De semantiska relationerna mellan för- och efterled är särskilt oförutsägbara

för L2-talaren som i lägre grad än L1-talaren är bekant med konventionerna i det aktuella språket.

I ett ofta citerat exempel reflekterar en tolvåring, som har svenska som sitt L2, över betydelsen hos uttrycket ”ovanstående figur” i en lärobok – är det måne en figur som är ’ovan vid att stå?’ (Hene 1993:8). I det här fallet förefaller polysemin hos *figur* leda informanten i personifierande riktning. Därmed blir det fritt fram för en tolkning där det homografa förledet i sammansättningen tolkas som ett adjektiv som beskriver en egenhet hos figuren (personen?) – den är ovan att stå – snarare än ett adverb som tillsammans med efterledet skapar ett rumsligt uttryck som rör en abstrakt figur.

Inom ramen för projektet Ord i L (Ord i läromedel, Lindberg & Johansson Kokkinakis 2007; se vidare avsnitt 4.4) undersöker Holmegaard (2007) hur första- och andraspråkselever på högstadiet och gymnasiet, samt en mindre grupp av flerspråkiga studenter på universitetsnivå, segmenterar och tolkar ett antal sammansatta ord från lärobokstexter i samhällsorienterande ämnen. I ett segmenteringstest, där uppgiften är att identifiera sammansättningarnas två huvuddelar, lyckas eleverna förhållandevis väl, men spridningen i elevgruppen är stor. Den sammansättning som visar sig vara svårast att analysera, *motreformation*, segmenteras avvikande (*motre-formation*) i närmare två tredjedelar av fallen. Avvikande segmentering av *sammanslagning* (*sammans-lagning*) och *valsensation* (*vals-ensation*) tolkar Holmegaard som ett försök hos informanterna att identifiera och dela av sammansättningarna vid ett förväntat fuge-s (2007:142, 145–146). Men som jag ser det skulle tolkningarna likaväl kunna vara ett resultat av att *vals* och *lagning* är mer identifierbara ord för informanterna än *val* och *slagning*.

I ett matchningstest där informanterna ges i uppgift att matcha nio sammansättningar med nio ordförklaringar bereder ord som *förhandlingsvilja*, *ställföreträdare*, *fackföreningsrörelse* och *oppositionspartier* svårigheter för informanterna (cirka 50 % avvikande svar), och i ett flervalstest med fyra alternativa förklaringar till testorden visar sig *tillverkningsindustri*, *justitiedepartement* och *samhällsinflytande* vara relativt svårtolkade ord (Holmegaard 2007:148–151). Holmegaard kopplar resultatet till en låg förekomst av orden i lärobokstexterna. Hon lyfter dessutom fram polysemin i några av sammansättningarna som en komplicerande faktor. Skillnaderna mellan de elever som har svenska som förstaspråk och de elever som har svenska som andraspråk är försumbara i segmenteringstestet men värda att notera i de övriga testen. En sammanslagning av resultatet för matchningstestet och flervalstestet (totalt 15 sammansättningar) visar att 64 % av eleverna med svenska som modersmål har 10–15 konventionella tolkningar av sammansättningarna, medan motsvarande siffra är 44 %

för eleverna med annat modersmål än svenska (2007:150, 152–154). Holmegaards konklusion är att ”långa ord med flera stavelser i SO-ämnena är svåra att förstå för många elever”. Hon kommer fram till att detta gäller både eleverna med svenska som förstaspråk och eleverna med svenska som andraspråk, men hon lyfter särskilt andraspråkselevernars svårigheter (2007:158).

Holmegaard (2007:151–152) konstaterar vidare att de ord som förefaller svårast för eleverna med svenska som andraspråk är allmänna och abstrakta, såsom *inflytande* (i *samhällsinflytande*) och *rörelse* (i *fackföreningsrörelse*). L2-talares svårigheter att förstå sådana abstrakta men icke-fackspråkliga ord har tidigare uppmärksammats i en norsk studie av ordförrådet i läromedel (Golden & Hvenekilde 1983). Problemet med den här typen av ord (t.ex. *gnida* i fysik och *hamn* i geografi) är att lärarna ur sitt förstaspråksperspektiv inte uppfattar dem som problematiska och därmed inte förklarar dem, till skillnad från utpräglade fackspråkliga ord som pochar på en förklaring (Golden 2006:117). En annan svårighet som Holmegaard (2007:150) noterar är förekomsten av polysemi. Särskilt komplicerat blir det för L2-talare när ett polysemt ord i en sammansättning används i en mindre vanlig betydelse (t.ex. *rörelse* i *fackföreningsrörelse*).

Relationspolysemin hos sammansättningar är en stötesten som blir särskilt påtaglig ur ett flerspråkighetsperspektiv, eftersom förståelsen av den sammansatta betydelsen förutsätter kännedom om vilken betydelse-relation, av flera teoretiskt tänkbara, som konventionaliserats för den specifika instansen. Enligt Josefsson (2005:84) kännetecknas konventionaliserade sammansättningar av att betydelse-relationen mellan förled och efterled har ”fixerats inom en språkgemenskap, d.v.s. en större grupp av människor är någorlunda överens om vad ordet betyder”. Men sådana konventionaliseringsprocesser är alltså språkspecifika (de gäller inom en språkgemenskap), och kan inte översättas från ett språk till ett annat. Eftersom det är språkspecifikt och kulturberoende vilken betydelse-relation som har konventionaliserats för en viss sammansättning måste denna läras in för varje ord separat (se Enström 2013:187). Språket innehåller många fall av sammansättningar som påminner om varandra men har icke-parallella relationer, såsom i ordparet *snöskor* – *läderskor*. För ogenomskinliga sammansättningar, såsom *nyckelpiga* och *jordgubbe*, tillkommer ytterligare tolkningssvårigheter, eftersom betydelsen inte är härledbar ur delarnas betydelse (Enström 2013:187; se vidare 3.2.1 om genomskinlighetsbegreppet). Fördelen vid ogenomskinliga sammansättningar är emellertid att de i hög grad behandlas i ordböckerna. Även Enström uppmärksammar att segmenteringen av sammansättningar kan vålla problem för inläraren, eftersom det kan vara svårt att identifiera gränsen mellan för- och efterled i sammansättningar som *tesil* och *talkör*.

Dessutom kan sammansatta ord få en annan betydelse än när leden används som två separata ord. Så är fallet med sammansättningen *underarmen* som skiljer sig semantiskt från prepositionsfrasen *under armen* (Enström 2013:187–188).

Det är känt att ord som betecknar konkreta referenter är mer lättillgängliga än abstrakta ord, eftersom de har en mer entydig relation mellan uttryck och referent (se Enström 2013:185, Pavlenko 2009:144). Sammansättningar är till sin natur polysema, och det är tänkbart att inlärare i större utsträckning känner till konkreta, bokstavliga betydelser än mer abstrakta betydelser hos dem. Möjligen är det en sådan effekt som syns i L2-tolkningen av *ovanstående figur* hos Hene. I ett annat exempel i samma studie tolkar en annan tolvårig andraspråkstalare *naturtillgångar* på ett bokstavligt sätt som 'små vägar till naturen' (Hene 1993:7–8).

Sammanfattningsvis är sammansättningar en utmaning ur ett andraspråks- eller flerspråkighetsperspektiv. Det rör sig om långa ord som kan vara svåra att segmentera och det är dessutom oförutsägbart – och språk-specifikt – med vilken relation mellan leden som de konventionaliserats.

3

Teoretiskt ramverk

I det här kapitlet presenteras det teoretiska ramverket för min studie. Kapitlet innefattar en beskrivning av de bruksbaserade, kognitiva perspektiv på språk som denna studie utgår från (avsnitt 3.1). Här redogörs också för lexikala perspektiv på kompositionalitet, meningspotential och konventionalisering (avsnitt 3.2). Teorikapitlets båda delar innehåller vardera ett avsnitt som beskriver den kognitiva modellen respektive de lexikala/semantiska begreppen ur ett flerspråkighetsperspektiv.

3.1. Bruksbaserade kognitiva perspektiv

Den här studien tar sin utgångspunkt i ett kognitivt paradig. I analysen av mitt material använder jag mig av verktyg från kognitiv grammatik och konstruktionsgrammatik, som är systerteorier inom detta paradig. I kognitiva teorier betraktas språkförmågan som integrerad med andra kognitiva förmågor hos människan. Kognitiv forskning som är teoretiskt bruksbaserad (eng. *usage-based*) utgår dessutom från att språkliga erfarenheter skapar och påverkar de kognitiva representationerna av språket. Det som är frekvent i språket, men också det som är *salient* (dvs. det som har stort uppmärksamhetsvärde; se avsnitt 3.1.6), antas bli mer befäst i det

mentala lexikonet (se Kemmer & Barlow 2000). Processer av befästning (eng. *entrenchment*) omfattar dels enskilda ord och uttryck, dels mönster eller scheman (Langacker 2017:47–48, Schmid 2017:22–23).

I inledningen till detta kapitel redogör jag för hur sambandet mellan den språkliga förmågan och andra kognitiva förmågor beskrivs inom kognitiv grammatik (avsnitt 3.1.1). Avsnitt 3.1.2 handlar om betydelse, kompositionalitet och analyserbarhet i kognitiv grammatik. Därefter följer ett kortare avsnitt om begreppen schema respektive konstruktion (avsnitt 3.1.3). Betydelsen av token- och typfrekvenser för befästningen av enskilda ord, uttryck och scheman beskrivs i avsnitt 3.1.4. Därefter följer ett avsnitt om förhållandet mellan schemainstantiering och analogi (avsnitt 3.1.5). Sambandet mellan saliens och befästning diskuteras i det därpå följande avsnittet (avsnitt 3.1.6). I avsnitt 3.1.7 ställer jag konventionalisering i relation till befästning. Därefter följer ett avsnitt som handlar om flerspråkighet ur ett kognitivt perspektiv (avsnitt 3.1.8). Beskrivningen av det kognitiva paradigmet avrundas med ett sammanfattande avsnitt där jag konstaterar att befästning inom detta paradigm inte bara är en fråga om frekvens (avsnitt 3.1.9).

3.1.1. Språkförmågan – integrerad med andra kognitiva förmågor

I kognitiv grammatik utgår man från att språket omfattar ”a structured inventory of conventional linguistic units” (Langacker 1987:57). Språket är närmare bestämt uppbyggt av symboliska strukturer, dvs. strukturer som har en betydelsesida (”semantic pole”) och en uttryckssida (”phonological pole”). Detta gäller orden i språket, men också såväl mindre enheter (morfem) som större enheter (satsstrukturer). Grammatiken handlar om hur sådana symboliska enheter kombineras för att bilda komplexa uttryck (Langacker 2008:5, 15). De symboliska enheterna i den kognitiva grammatiken motsvaras av konstruktioner inom konstruktionsgrammatiken (se vidare avsnitt 3.1.3).

Eftersom grammatik och lexikon är uppbyggda på samma sätt med en form- och en betydelsesida, råder det inte någon artskillnad, utan bara en gradskillnad, mellan grammatik och lexikon, och många språkliga uttryck, såsom idiom och sammansättningar, kan inte enkelt hänföras till antingen grammatiken eller lexikonet. Följaktligen avvisas den generativa grammatikens antagande om en separat språkmodul som hanterar syntax åtskilt från lexikon och semantik. Samma utgångspunkt har konstruktionsgrammatiken (se avsnitt 3.1.3). Syntax och semantik betraktas som nära förbundna, och samma kognitiva processer antas hantera såväl

ord som komplexa uttryck (Langacker 2008:18–19). Jag delar i grunden detta synsätt, och förhåller mig till informanternas tolkningar som instantieringar av mönster på olika schematicitetsnivåer, även om jag i vissa fall skiljer mellan mer strukturella (grammatiska) och mer semantiska (lexikala) mönster i analysen.¹

Kognitiv teori utmärks av att språkförmågan ses som en integrerad del i människans kognitiva förmågor. Den språkliga förmågan antas bygga på psykologiska/kognitiva förmågor såsom perception, minne och kategorisering. Några centrala kognitiva processer som vi använder för att förstå och bilda språkliga uttryck är *association*, *automatisering*, *schematisering* och *kategorisering* (Langacker 2008:8, 16–17).

Med associationen som mentalt redskap kan man göra en koppling mellan den fonologiska och den semantiska sidan hos ett ord, varvid man får tillgång till ordet som en symbolisk enhet (Langacker 2008:16).

Automatisering inträffar när man genom repetition lärt sig något så väl, att man inte behöver tänka medvetet på delmomenten i en process, t.ex. att knyta skorna. På språklig nivå motsvaras detta av sådan repetition av ett språkligt uttryck som leder till att uttrycket blir befäst som en *enhet* i språkbrukarens mentala lexikon (Langacker 2008:16, 2017:40; se vidare avsnitt 3.1.4).

Förekomsten av enheter i det mentala lexikonet får dock inte tolkas statistiskt. Enheterna ska inte förstås som byggstenar, utan definieras i den kognitiva grammatiken som etablerade mönster av processningsaktivitet, och sådana enheter uppkommer genom gradvisa befästningsprocesser. Språkliga element som blir befästa som enheter står för stabilitet genom potentialen till rekursion (Langacker 2017:40–41; se även Schmid 2015:7–8). När ett uttryck (t.ex. *moonless night*) blivit befäst som enhet, behöver det inte analyseras i sina delar, varvid saliensen (dvs. uppmärksamhetsvärdet) för de ingående delarna (*moon*, *-less* och *night*) minskar (Langacker 2008:16–17; se vidare avsnitt 3.1.6 om saliensen).

Schematisering innebär att språkbrukaren ur ett antal bruksinstanser finner gemensamma drag och bildar ett mer abstrakt schema eller begrepp. Ur erfarenheten av romerska ringar i gymnastiksalen, ringen på fingret och ringen som man kan sitta i bildas genom schematisering det abstrakta begreppet ”circular entity”. Från instanserna *moonless*, *hopeless* och *fruitless* kan schemat [N-less] abstraheras. När ett sådant abstrakt schema blivit befäst som en enhet kan det i motsatt riktning fungera som mall för nya instantieringar (Langacker 1999:93, 2008:17, 24; 2017:44; se vidare

1 Ett liknande förhållningssätt, när det gäller konstruktionsgrammatisk analys, framkommer hos Lyngfelt (2018:8–9), som argumenterar för att man mycket väl kan göra en distinktion mellan lexikala konstruktioner och fraskonstruktioner, utan att för den skull betrakta dem som fundamentalt olika enheter som hanteras av separata kognitiva moduler.

avsnitt 3.1.3 om scheman). Schematiseringen är intimt förknippad med en annan kognitiv process, nämligen kategorisering.

Kategorisering innebär att företeelser jämförs och sorteras efter likhet. Utifrån kriteriet tillräcklig likhet bildas kategorier som används för att hantera nya företeelser eller språkliga instanser. Följaktligen är jämförelse en grundläggande, kognitiv princip (se Langacker 1999:94, 2008:17, Tuggy 2007:86–87, Schmid 2015:11).

Sammanfattningsvis betonas sambandet mellan språklig förmåga och andra kognitiva förmågor, och individens aktiva roll i språklig såväl som i annan kognitiv aktivitet lyfts fram: ”[c]reating a novel expression is not necessarily different in fundamental character from problem-solving activity in general” (Langacker 1987:65).

3.1.2. Betydelse, kompositionalitet och analyserbarhet

I den kognitiva grammatiken betonar man att betydelser inte finns i orden eller i grammatiken. Betydelser finns i huvudet hos språkbrukarna och uppstår genom konceptualisering eller begreppsbildning (Langacker 2008:4, 27). Eftersom individens kognition sätts i centrum är betydelsen hos ett uttryck ytterst dess betydelse för en enskild språkbrukare. Men individens begreppsbildning uppstår inte i ett vakuum. Konceptualiseringar sker i hjärnan, men de sträcker sig utanför den, framhåller Langacker, eftersom de är konceptualiseringar av världen. När vi talar konceptualiserar vi inte bara det specifika, språkliga yttrandet utan också den vida kontexten för yttrandet, t.ex. samtalspartnerns förväntade kunskap och intentioner. Vid begreppsbildning tar vi med andra ord in världen och vår samlade kunskap om den. Konceptualiseringen sträcker sig också utanför språkbrukarens hjärna eftersom betydelser uppstår i diskurser genom social interaktion och förhandling med andra språkbrukare (Langacker 2008:28–30). Denna förhandling om betydelse sker inte förutsättningslöst, utan utgår från individernas kunskap om vad ordet normalt betyder: ”While everything may be negotiable, something has to be learned and conventionalized as a basis for negotiation” (Langacker 2008:30). En instans i en viss betydelse blir en etablerad del av språket (dvs. en enhet) först när den är befäst i språkbrukares mentala lexikon och konventionaliserad för medlemmarna i en språkgemenskap (Langacker 2008:38).

Kognitiva perspektiv utgår från en encyklopedisk semantik. Språklig betydelse ses som nära sammanflätad med kunskap om och erfarenhet av världen (se Langacker 2008:39–41, Bybee 2013:52–53). När vi tolkar företeelser och fenomen (inklusive språkliga) gör vi det utifrån den upplevelse av världen som vi erfar med alla våra sinnen (Langacker 2008:32–33). Langacker lyfter fram att våra konceptualiseringar sker utifrån kognitiva

domäner. Domäner som rymd, tid, temperatur, lukt och smak kategoriserar Langacker som basala. Som icke-basala domäner räknar han sensoriska, emotiva och kinestetiska erfarenheter. Till de icke-basala domänerna hör även abstrakta begrepp som nås via intellektet (t.ex. begreppet 'rättvisa'). De icke-basala domänerna betraktar Langacker som jämförbara med begreppet *ram* (eng. *frame*) hos Fillmore (1982) (Langacker 2008:44–47).

Vid sammansatta ord och uttryck involveras alla de domäner som är aktuella för de ingående simplexen. Men också nya domäner kan väckas eller gå från perifera till salienta i det sammansatta uttrycket. För sammansättningen *fågelbur* blir associationen att fåglar kan hållas i bur mer salient än för de ingående delarna *fågel* respektive *bur*. Likaledes väcker sammansättningen *läppstift* kulturella föreställningar som inte är salienta i de ingående delarna *läpp* och *stift* (Langacker 2008:53).

Langacker framhåller i många sammanhang problemen med antagandet om en strikt kompositionalitetsprincip, där den sammansatta betydelsen betraktas som förutsägbar utifrån de ingående delarna. Ett sådant antagande fungerar möjligen i en modulär språkmodell, där lexikala enheter ses som tydligt avgränsade byggstenar med fasta betydelser, men inte i en kognitiv språkmodell, där betydelser är emergenta och antas uppstå genom konceptualisering i en ständigt pågående växelverkan mellan semantisk och pragmatisk kunskap (se Langacker 2008:39–40). Han diskuterar också språkets kompositionalitetsprincip i relation till sammansatta ord. Han konstaterar inledningsvis att helhetens betydelse inte är lika med *summan* av de ingående delarna, så som det ibland föreslås i litteraturen om sammansättningar, eftersom olika ordning mellan leden ger olika betydelse (Langacker 2008:168).

I den mån en sammansättning är kompositionell, i betydelsen förutsägbar, är den det för att den följer ett konventionaliserat, överordnat (eller konstruktionellt) schema (se vidare avsnitt 3.1.3 om schemabegreppet). I premodifierande språk följer sammansättningar exempelvis ett schema som stadgar att det sista ledet i sammansättningen i normalfallet är det profilerade ledet, dvs. det led som bär upp huvudbetydelsen (Langacker 2008:168–169). Med ett exempel från svenskan är en *husbåt* 'en typ av båt', medan ett *båthus* är 'en typ av hus', dvs. i det första fallet profileras 'båt', i det andra 'hus'. Begreppet profilerat led i kognitiv grammatik betecknas med mer traditionella termer *huvud* eller *huvudled*, som är de termer som används i denna avhandling.

Enligt det överordnade schemat för sammansättningar är det profilerade elementet, dvs. huvudet, på något sätt knutet till förledets modifierande komponent, men det är inte förutsägbart vilken semantisk relation som föreligger mellan ledens referenter. En sammansättning som *jar lid* 'burklock' väcker inte nödvändigtvis den mest uppenbara associationen – 'ett lock

som används för att täcka en burk’ – hos alla talare. *Jar lid* skulle t.ex. kunna tolkas som ett lock som är dekorerat med en bild av en burk (Langacker 2008:169). För att ta ett annat exempel betecknar sammansättningen *blackbird* inte vilken svart fågel som helst, såsom i frasen *black bird*, utan betecknar en specifik fågelart. Därför innebär kompositionell betydelse i Langackers definition att helhetens betydelse är motiverad, dvs. härledbar ur de ingående leden – på ett sådant sätt som specificeras i ett överordnat schema – men den är däremot inte förutsägbar. Det råder partiell snarare än fullständig kompositionalitet (Langacker 1987:448–450, 2008:169):

The issue of compositionality must therefore be formulated in terms of whether the composite structure derives from component structures in the manner specified by a constructional schema. (Langacker 2008:169)

Langacker framhåller vidare att det kan uppstå en diskrepans mellan ett uttrycks förväntade betydelse och dess realiserade betydelse, eftersom de ingående komponenterna och vad han kallar det konstruktionella schemat inte är de enda tillgängliga resurserna för att förstå det. Andra resurser i tolkningsprocessen är omvärldskunskap och imaginära strategier, såsom när man använder sig av metaforik och metonymi i tolkningen (se vidare Langacker 2008:170). Med konstruktionellt schema avser han ett överordnat, abstrakt schema (Langacker 1991:3).

Kompositionalitet definierar Langacker som en objektiv företeelse: helhetens betydelse som en funktion av två betydelseelement. *Analyserbarhet* å andra sidan definieras som språkbrukarens upplevelse av kompositionalitet, och syftar närmare bestämt på ”the extent to which speakers are cognizant (at some level of processing) of the contribution that individual component structures make to the composite whole” (Langacker 1987:457). Analyserbarheten kan försvinna när helheten i en kompositionell struktur överskuggar delarna, eller åtminstone kan saliensen för de ingående delarna minska. Detta inträffar när en sammansättning blir befast som en enhet. Men omvänt kan språkbrukare också vara medvetna om de ingående delarna i en struktur som bara uppvisar delvis kompositionalitet, dvs. en struktur kan vara analyserbar trots att den inte är till fullo kompositionell (Langacker 1987:457, 2008:17, 2017:42–43). Begreppet analyserbarhet sammanfaller delvis med begreppet genomskinlighet i den lexikala semantiken, vilket diskuteras i avsnitt 3.2.1.

3.1.3. Schema och konstruktion

Schemabegreppet är centralt inom kognitiv grammatik, och schema kan definieras som ”a superordinate concept, one which specifies the basic

outline common to several, or many, more specific concepts” (Tuggy 2007:83). Mönster och schanan tar form genom att språkbrukare kategoriserar likartade språkliga instanser i en process som kan benämnas schematisering: ”[p]atterns and regularities of any sort, at any level of specificity, reside in schematic units abstracted from occurring expressions” (Langacker 2008:219). I motsatt riktning instantieras schanat när det får utgöra mall för produktionen eller tolkningen av ett specifikt uttryck. När det gäller den formella sidan kan ett sådant schanat innehålla fasta led och luckor (eng. *slots*) som ska fyllas, och på innehållssidan antas helheten få en viss betydelse.² I en sammansättningskontext skulle man kunna exemplifiera detta med [N-låda], där *låda* utgör ett fast led och [N] markerar en lucka som skulle kunna fyllas med olika substantiv som betecknar material, t.ex. *glas* eller *plast*, varvid helheten får betydelsen ’en låda av N’. I ett alternativt schanat för [N-låda] instantieras [N] av ett substantiv som anger innehåll, t.ex. *frukt*, varvid helheten får betydelsen ’låda med frukt’. När ett sådant schanat blivit befäst som en enhet fungerar det som en mall för att hantera nya uttryck av samma slag (se Langacker 2008:24). Langacker betonar att det är svårt att göra en gränsdragning mellan schanat och deras instantieringar:

They [schemas] are better seen as being **immanent** in their instantiations (i.e. as “lying within” them). What I mean by this is that schemas reside in certain aspects of the processing activity in which their instantiations reside. (Langacker 2008:174; hans emfas)

Langacker antar att schanat med låg abstraktionsgrad såväl som specifika instantieringar samexisterar med mer övergripande schanat i den mentala grammatiken och framhåller vår förmåga att konceptualisera situationer på varierande nivåer av schemacitet (Langacker 1987:134–135).

Schanabegreppet har stora likheter med begreppet konstruktion inom konstruktionsgrammatik, som används om konventionaliserade par av form (syntax, morfologi, prosodi) och betydelse (semantik, pragmatik) (Fillmore et al. 1988, Croft 2001, Goldberg 2006, 2013). Liksom i den kognitiva grammatiken tar man avstånd från en strikt åtskillnad mellan lexikon och syntax. I konstruktionsgrammatiken ser man i stället språket som uppbyggt av konstruktioner på olika schemacitetsnivåer, där konstruktionens luckor är lexikalt fyllda i olika grad. Konstruktioner

² I den kognitiva grammatiken antar man även förekomsten av schanat utan luckor, t.ex. schanat för de olika ordklasserna (Langacker 2008:219–220). Relationen mellan schanat och instans antas inte bara gälla symboliska enheter utan också fonologiska och semantiska. Den semantiska enheten ’animal’ är till exempel schematisk i relation till den underordnade semantiska enheten ’dog’ (Heyvaert 2009:239).

definieras som form–betydelse-par ”at varying levels of complexity and abstraction” (Goldberg 2013:17).

Konstruktioner kan vara helt eller delvis lexikalt fyllda (*banana* respektive *V-ing*). På en mer abstrakt nivå förekommer konstruktioner i form av fyllda (*kick the bucket*) eller delvis fyllda idiom och uttryck (*the Xer the Yer*). De mest schematiska konstruktionerna är konstruktioner för hela satser, där ingen av konstruktionens luckor är lexikalt fyllda (t.ex. den ditransitiva konstruktionen S V Obj1 Obj2) (Goldberg 2013:17). Lyngfelt et al. (2014:271) presenterar en motsvarande beskrivning för svenska konstruktioner på olika schematicitetsnivåer. Det går således utmärkt att på en lägre schematicitetsnivå beskriva sammansättningar och sammansättningsmönster i termer av konstruktioner, där för- och efterleden kan fyllas av olika morfologiska eller lexikala enheter (se t.ex. Booij 2009, 2010, 2013). Jag har emellertid valt att inte använda begreppet konstruktion, utan att tala om sammansättningar som instantierar vissa strukturella och semantiska mönster eller scheman, i enlighet med tidigare kognitiv sammansättningsforskning (t.ex. Ryder 1994).

3.1.4. Frekvens, minnesrepresentation och befästning

I kognitiv lingvistik antas frekvent upprepning av ett mönster eller en instans leda till en automatisering av de associationer som mönstret eller ordet väcker hos individen, t.ex. associationen mellan form och betydelse. Detta innebär att såväl formen som betydelsen blir mer tillgänglig för språkbrukaren (se Kemmer & Barlow 2000). Med Langackers terminologi leder repetitionen av en språklig struktur (t.ex. ett ord eller ett uttryck) till en successivt ökande grad av befästning som kan resultera i att strukturen blir en enhet. Såväl enskilda ord och uttryck som scheman kan etableras som enheter (Langacker 1987:57–59, 2017:40–41, 44, 53; se även Schmid 2017:23). När en språklig struktur blivit befäst som en enhet innebär det att automatisering uppnåtts:

A **unit** is a structure that a speaker has mastered quite thoroughly, to the extent that he can employ it in largely automatic fashion, without having to focus his attention specifically on its individual parts or their arrangement. Despite its internal complexity, a unit constitutes for the speaker a “pre-packaged” assembly; because he has no need to reflect on how to put it together, he can manipulate it with ease as a unitary entity. It is effectively simple, since it does not demand the **constructive effort** required for the creation of novel structures. Psychologists would speak of a “habit”, or say that “automatization” has occurred. (Langacker 1987:57; hans emfas)

En enhet hos Langacker är med andra ord etablerad, ”färdigförpackad” och automatiserad (se även Schmid 2007:118; jfr också definitionen av befästning hos Schmid 2015:10). Langacker betonar emellertid att det inte finns någon skarp gräns mellan etablerade och oetablerade lingvistiska enheter, utan betonar att det rör sig om gradskillnader, eftersom befästning och konventionalisering är graduella fenomen (Langacker 2008:38). Han framhåller vidare att befästningen av ett uttryck leder till minskad saliens för de ingående delarna: helheten går då före delarna (Langacker 2008:17, 2017:42–43). Detta förhållande är direkt applicerbart på förståelsen av sammansatta ord: om man känner till ordet som helhet finns det ingen anledning att pussla sig fram till betydelsen.

Forskare inom bruksbaserade teorier utgår från att våra språkliga minnesrepresentationer, dvs. befästningen av enskilda ord, uttryck och scheman, påverkas av token- och typfrekvenser. Tokenfrekvens rör frekvensen för en specifik sträng och leder till befästning av enskilda ord och uttryck, till exempel verbet *drive* eller uttrycket *drive me crazy*. Befästningen innebär att det sker en automatisering av associationen mellan form och betydelse. Typfrekvens rör antalet olika, unika instanser, eller typer, i en schematisk lucka, t.ex. antalet adjektiv eller adjektivfraser som kan instantiera [AP]-luckan i konstruktionen [DRIVE NP/PRON AP]. Hög typfrekvens korrelerar med schemabefästning och produktivitet, dvs. uppkomsten av nya instanser (Bybee 2013:59–62, Langacker 2017:47–48, Schmid 2017:14).

När det gäller schemabefästning kan upprepade instanser som *give me the doll*, *give me the book*, *give me the cup* (dvs. hög typfrekvens) leda inläraren att abstrahera fram schemat [give me X] (Schmid 2017:14).

När det gäller sambandet mellan typfrekvens och produktivitet är det två faktorer som spelar in. För det första ger hög typfrekvens en bredare bas för analogi. För det andra innebär hög typfrekvens att talarna måste processa instanser med låg tokenfrekvens, vilket leder till aktivering och därmed förstärkning av schemat eller konstruktionen (Bybee 2013:59–62).

Översatt till ordbildning leder hög tokenfrekvens för *moonless* till befästning av denna specifika avledning. Hög typfrekvens genom repetitionen av olika unika instanser som slutar på *-less*, t.ex. *moonless*, *hopeless*, *fruitless*, *treeless* och *cordless* leder däremot till befästning av ett överordnat schema [N-less] (Langacker 2008:24).

Kognitiv grammatik och konstruktionsgrammatik förenas i antagandet om att konkreta språkliga erfarenheter påverkar den kognitiva representationen av språket. Exakt hur abstraktionen från det konkreta till det abstrakta går till beskrivs på olika sätt i olika modeller. Inom ramen för konstruktionsgrammatik presenterar Bybee (2013) en exemplarmodell (eng.

exemplar representation model). I denna modell har de konkreta språkliga token som vi möter, tillsammans med de kontextuella omständigheter som omgärdade instansen, en direkt inverkan på den mentala representationen. Sålunda ges pragmatiken stort utrymme i Bybees modell, som är en modell för minnesrepresentation i vid bemärkelse (Bybee 2013:52–53).

Språkinstanter (dvs. token) som bedöms lika lagras som exemplar. Exemplaren är abstraktioner över enskilda token och förekommer på olika schematicitetsnivåer. De kan t.ex. utgöras av en vokal, ett ord eller ett uttryck. Dessa exemplar är olika starka beroende på hur många token som bygger upp dem. Därmed har frekvensen en direkt inverkan på den språkliga representationen. I nästa led grupperas specifika exemplar i exemplarmoln (eng. *exemplar clouds*) baserat på likhet och frekvens. De schematiska luckorna i en konstruktion utgörs av sådana exemplarmoln. Starka exemplar som är uppbyggda av många token utgör kärnan i exemplarmolnet. Exemplarmolnen kan formeras på fonetiska, semantiska, pragmatiska eller kontextuella grunder. För varje ord eller konstruktion är dessa exemplarmoln från olika domäner länkade, så att språkbrukaren kan registrera samförekomster av ett ord eller en fras i dessa olika domäner (Bybee 2013:53–54, 57).

Exemplarmolnet för ett specifikt ord skulle inkludera alla betydelser och kontexter i vilken ordet erfarits, vilket kan jämföras med Allwoods (2003) definition av *meningspotential* (se avsnitt 3.2.1). Men information som inte blir förstärkt genom repetition kan gå förlorad (Bybee 2013:54).

Langacker diskuterar exemplarmodeller i relation till sin teori. Enligt honom råder det ingen principiell skillnad mellan exemplarteori och kognitiva modeller som utgår från scheman som strukturerande princip, eftersom båda perspektiven härleder språkförmågan ur processningsaktivitet och förekomsten av etablerade enheter med varierande grad av befästning. Exemplarmolnen i exemplarteorin skulle enligt Langacker lika gärna kunna betraktas som lågnivåschema (Langacker 2009:637, 2017:54; se vidare nästa avsnitt 3.1.5).

Sammanfattningsvis, och översatt till min sammansättningsforskning, har tokenfrekvenser betydelse för befästningen av en specifik sammansättning (som en enhet eller ett exemplar). Genom befästning automatiseras associationen mellan form och betydelse. (Hög) typfrekvens (eller förekomsten av flera exemplar) för en lucka i ett sammansättningsschema innebär att detta kan instantieras av nya sammansättningar, eller att en ny sammansättning tolkas enligt ett befäst mönster, såsom när *skrikfärdig* tolkas med samma relation mellan leden som *gråtfärdig*. Detta kan inträffa via schemainstantiering eller analogi.

3.1.5. Schema eller analogi?

I klassiska språkvetenskapliga teorier finns en tradition av att skilja mellan abstrakta, produktiva mönster och mer konkreta, analogiska mönster, i synnerhet vad gäller språklig produktion. Som produktiva räknas bildningar som genereras via regelbundna, abstrakta scheman eller regler, medan analogiska bildningar anses vara oregelbundna och formerade efter en konkret förebild. I Langackers kognitiva grammatik är det emellertid ingen substantiell skillnad mellan schemabaserade och analogibaserade beskrivningar. Langacker framhåller att även analogisk bildning förutsätter tillgång till ett schema som får ligga till grund för den nya formen. Den enda skillnad som kvarstår mellan schemabaserad och analogisk formering är huruvida schemat finns lagrat som en enhet (schemabasering) eller behöver extraheras ur konkreta exempel (analogi) (Langacker 1987:446–447; se även Tuggy 2007:100–102).

Langacker antar vidare att scheman med låg abstraktionsgrad såväl som specifika instantieringar samexisterar med mer övergripande scheman i den mentala grammatiken. Han framhåller emellertid betydelsen av scheman på låg abstraktionsnivå (eng. *lower-level schemas*) och betonar att dessa lågnivåskeman ”may on balance be more essential to language structure than higher-level schemas representing the broadest generalizations” (Langacker 2000:29; se även Heyvaert 2009:235). Liknande resonemang förs av Ryder (1994:89, 193, 197–198) och Dabrowska (2010) som finner empiriska belägg för ytnära analogi.

I linje med sin exemplarmodell framhåller Bybee att nya instantieringar av en konstruktion ofta uppstår genom analogibildning utifrån specifika exemplar, snarare än genom produktiv bildning utifrån abstrakta representationer (Bybee 2013:57–59). Analogi definierar hon som ”the process by which a speaker comes to use a novel item in a construction” (Bybee 2010:57). Utvidgningen av en konstruktion kommer att baseras på de prototypiska exemplaren i exemplarmolnet, dvs. de exemplar som är uppbyggda av många token, t.ex. *crazy* i AP-luckan i [DRIVE NP/PRON AP]. Konstruktioner med hög tokenfrekvens kan emellertid komma att övergå i automatiska helfraser. I händelse av en sådan automatisering aktiveras inte exemplarmolnet, vilket leder till förlorad produktivitet och i förlängningen till förlorad analyserbarhet (Bybee 2013:57–59, 61).

Booij (2010:94–95) konstaterar att det ibland är svårt att avgöra om en ny sammansättning är bildad i analogi med en specifik sammansättning eller utifrån ett schema. Han exemplifierar med Watergate-skandalen där sammansättningen *Watergate* gett upphov till andra bildningar med *-gate* som betecknar specifika politiska skandaler. När flera sådana bildningar

etablerats är det oklart om en ny instans är bildad i analogi med en specifik sammansättning, eller om den utgår från ett ytnära schema [X-gate].

Vilken schematicitetsnivå språkbrukare förlitar sig på är en teoretiskt intressant fråga, inte minst när det gäller sammansättningsforskning, eftersom sammansättningar är ett gränsfenomen mellan strukturella och lexikala mönster. Ryder (1994) tar sig an denna fråga i sin avhandling om tolkning av sammansatta ord, där informanter med engelska som modersmål tolkar tillfälliga bildningar på engelska.

Ryder utgår från att språkbrukare använder sig både av språkliga mallar (eng. *linguistic templates*) och semantiska scheman (eng. *semantic information schemas*), dvs. omvärldskunskap, i sina tolkningar. Språkliga mallar förekommer på olika schematicitetsnivåer. Den specifika språkliga mall som används för att tolka eller producera en ny sammansättning kallar hon för tolkningens analogibas. Analogibasen kan exempelvis utgöras av en enskild instans (t.ex. *doghouse*) eller en ordfamilj med ett gemensamt kärnord som för- eller efterled (t.ex. *sea lion, seaman* och *seaweed; boathouse, warehouse, tree house*). Om tolkningen sker utifrån en ordfamilj kan den baseras på en individuell instans i en sådan ordfamilj eller utgå från en abstraktion över de enskilda instanserna, t.ex. schemat [sea-X] eller [X-house]. Vilken hjälp lyssnaren får av kärnordet i en sammansättning beror på dess signalreliabilitet (eng. *cue reliability*). Signalreliabilitet handlar om grad av förutsägbarhet. Hög signalreliabilitet innebär att kärnordet i en sammansättning alltid, eller i hög grad, påverkar sammansättningens betydelse på samma sätt. På en mer abstrakt nivå än ordfamiljens kan språkbrukare göra tolkningar utifrån scheman som [contained–container] eller [part–whole]. På den mest abstrakta, schematiska nivån är den symboliska strukturens fonologiska pol ospecificerad och den semantiska polen ger en maximalt schematisk specificering: 'an X that has some relation to Y'. Detta abstrakta schema innebär att efterledet X tenderar att utgöra det profilerade ledet, dvs. det semantiska huvudledet (Ryder 1994:79–83, 89–90).

I Ryders modell utgår all tolkning från analogibaser. Sålunda upphävs distinktionen mellan analogi och schemabaserad tolkning. Hon framhåller emellertid att förebilderna på en ytnära nivå av pragmatiska skäl framstår som viktigast, eftersom de kan ge detaljerad information om betydelsen hos den nya sammansättningen:

The lowest levels of templates seem most likely as analogy bases. In cases where the context provides very little information, the more detailed templates allow the listener to arrive at a very rich detailed meaning for the new compound. (Ryder 1994:89)

Samtidigt kan det vara svårt för språkbrukaren att hitta en förebild på ytnära nivå, eftersom sådana förebilder är baserade på enstaka instanser. Om språkbrukaren inte har tillgång till någon ytnära analogibas kan hon behöva aktualisera ett mer abstrakt schema (Ryder 1994:87–89).

Ryder exemplifierar bruket av olika abstraktionsnivåer med ett resonemang om hur en språkbrukare skulle tolka en tillfällig bildning som *water sock*. Sammansättningen är skapad i analogi med *wind sock*, som betecknar en sockliknande artefakt som indikerar vindens riktning och styrka. Om språkbrukaren som möter *water sock* är bekant med *wind sock* är det troligt att denna sammansättning används som analogibas så att *water sock* tolkas som ett sockliknande redskap för mätning av vattenströmmarnas riktning och styrka. Om språkbrukaren inte har tillgång till denna analogibas är det tänkbart att tolkningen i stället baseras på en annan konventionaliserad sammansättning som hon associerar till, t.ex. *vattenballong*, varvid en helt annan betydelse framträder: 'ett sockliknande objekt fyllt med vatten som kan användas i lek och kastas på någon, vanligtvis använt av barn på barnkalas'. Om språkbrukaren inte kan associera till någon konventionaliserad sammansättning är hon hänvisad till en mer generell mall, t.ex. [naturelement–klädesplagg] 'klädesplagg som skyddar mot naturelement' (som instantieras av exempelvis *regnjacka* och *solhatt*). Utifrån ett sådant schema får *water sock* betydelsen 'socka som skyddar fötterna mot väta'. Om språkbrukaren inte har tillgång till detta schema kan tolkningen göras utifrån ett ännu mer generellt schema [substance that can be contained–container] 'behållare som innehåller en substans' (som i *ciderflaska*), varvid den nya sammansättningen får betydelsen 'socka eller sockliknande objekt som innehåller vatten'. Denna betydelse påminner om den betydelse som framträder när *vattenballong* används som analogibas, men den är betydligt mer generell, eftersom analogibasen på denna schematiska nivå ger mycket mindre detaljerad information. Därför kan ytnära analogibaser, som ger kontextuell information, vara mer användbara, såsom framhölls i citatet ovan (Ryder 1994:87–89).

Enligt Ryder är grupper av etablerade instanser som delar ett kärnord tacksamma analogibaser när man tolkar en ny sammansättning. De är språkliga mallar på låg nivå och ger därmed detaljerad semantisk information, men de är ändå tillräckligt frekventa för att finnas lagrade. Särskilt användbara är kärnord som har hög signalreliabilitet, eftersom analogibasen i så fall utgörs av en semantiskt homogen grupp av sammansättningar. Vid låg signalreliabilitet kan den analogiska tolkningen däremot avvika från den avsedda betydelsen, eftersom kärnordet inte pekar ut en specifik betydelse, såsom exemplifierades i fallet med *water sock* (Ryder 1994:89).

Vilken generalitetsnivå som språkbrukaren använder sig av i tolkningen av en ny sammansättning är svårt att avgöra, och språkliga och semantiska scheman samverkar:

[...] when a person is confronted with a new *X + box* compound and decides it means 'a box containing X', there is no way to know whether he made this decision based on a linguistic template or on a semantic information schema for boxes containing things. And even if we were sure he used a linguistic template, there is no way to determine from his response alone whether the template was a low-level one, such as an individual compound like *cigar box*, or *X + box*, or the more general pattern *Contained + Container*, which can be based not only on compounds ending in *box*, but also on such terms as *potato sack*, *silver drawer* [...]. In fact, it is likely that all these patterns, both semantic information schemas and linguistic templates, play a role in his interpretation. (Ryder 1994:145)

Även om det således är svårt att avgöra vilken språklig mall språkbrukaren stöder sig på, finner Ryder vissa empiriska belägg för att språkbrukare föredrar ytnära analogibaser framför mer abstrakta (Ryder 1994:193, 197).

Oavsett schematicitetsnivå hos den språkliga mall som språkbrukaren använder som analogibas behöver den betydelse som analogibasen bär med sig checkas mot omvärldskunskap, dvs. det som Ryder kallar för semantiska informationsscheman. Dessa gäller den symboliska enhetens semantiska pol och kan delas in i händelsescheman (eng. *event schemas*), entitetsscheman (eng. *entity schemas*) och särdragsscheman (eng. *feature schemas*). Ordet *läkare* kan exempelvis väcka associationer till händelsescheman som [PATIENT BESÖKER LÄKARE] eller [LÄKARE UTFÖR OPERATION]. Entitetsscheman är mer generella och rör gemensamma drag som abstraheras från de individuella händelseschemana, t.ex. att läkaren är [KLÄDD I VIT ROCK] och [VÄLUTBILDAD]. Ur händelse- och entitetsscheman kan särdragsscheman för karakteristiska och särskiljande drag abstraheras, t.ex. [INTELLIGENT] eller [VÄLBESTÄLLD] när det gäller läkaren. Särdragsscheman framstår tydligare i samband med konkreta företeelser, som har mer entydiga särdrag, såsom färg, form och storlek. Entitetsschemana [BOLL] och [APELSIN] framkallar exempelvis särdragsschemat [RUND] (Ryder 1994:68–72, 89–90).

När språkbrukaren tolkar en ny sammansättning söker hon ett schema som är gemensamt för båda leden och som kan användas för att fylla i den betydelse som den språkliga mallen bidrar med. Om inget schema som är gemensamt för båda leden kan appliceras kan språkbrukaren skapa ett nytt eller ackommodera ett etablerat schema, t.ex. använda sig av det perifera händelseschemat [HUND DRICKER VIN] för att kunna tolka sammansättningen *wine-dog*. Som en sista utväg kan språkbrukaren använda

sig av ett semantiskt skelett, dvs. göra en mycket schematisk tolkning, t.ex. 'en X som finns i Y' eller – i fallet med hunden – 'en hund som har med vin att göra' (Ryder 1994:89–93).

Ryder framhåller att särdragsscheman är lättillgängliga och har stor potential att användas, vilket leder till en likhetsrelation mellan leden (se tabell 2:1 i avsnitt 2.3). En *äppelkudde* skulle utifrån ett särdragsschema kunna tolkas som en 'rund och röd kudde'. I Ryders experiment med nya sammansättningar är likhetstolkningar representerade i alla undersökta kategorier (1994:90–91, 125). Likhetstolkningar är en stark defaulttolkning om inget centralt schema är tillgängligt för språkbrukaren, eftersom "computing similarities between different objects is a fairly common human cognitive activity" (Ryder 1994:124).

Sammanfattningsvis är det inte helt enkelt att avgöra om produktion och tolkning av sammansättningar sker via scheman eller analogi. Inom det kognitiva paradigmet är det heller ingen avgörande skillnad mellan schemabasering och analogi. Forskningsläget är sådant att många kognitivt inriktade forskare lyfter fram betydelsen av ytnära analogibaser (låg-nivåsscheman eller specifika enheter/exemplar) i tolkningen och produktionen av nya uttryck, inte minst när det gäller specifika instanser, såsom sammansättningar. Också i barnspråksstudier har betydelsen av analogi lyfts fram, både vad gäller produktion och tolkning av sammansättningar (se t.ex. Mellenius 1997:63, Krott 2009:132–134, Krott et al. 2009:108, Tomasello 2003).

3.1.6. Saliens och befästning

Som framgått är frekvens en viktig faktor i befästningsprocesser, men det är ingen nödvändig förutsättning. En språkinstans kan bli befäst redan första gången man möter den om den är salient, dvs. om den har högt uppmärksamhetsvärde.

Ontologisk saliens avser utmärkande, framträdande och särskiljande drag hos fenomenen i världen, t.ex. vingarna hos fågeln eller ränderna hos zebran, dvs. iögonfallande drag. Salienta fenomen i världen tenderar att också bli kognitivt salienta, eftersom de lättare än mindre ontologiskt salienta fenomen framkallar motsvarande mentala representationer (Schmid 2007:120).

Ontologiskt salienta egenskaper hos en referent avspeglas i etablerade konnotationer för motsvarande ord, vilka manifesteras i produktiv ordbildning. Bildningar med *-haj* som efterled – t.ex. *fnanshaj*, *börshaj* och *korthaj* – refererar sannolikt till hajens rovdjursinstinkt (jfr sammansättningarna *loan shark*, *land shark* och *card shark* i engelska). Etablerade bildningar som *zebrafink*, *zebrafisk* och *zebramussla* får i vardagsspråket

sällskap av bildningar som *zebrabyxor* och *zebraväska*, där den gemensamma nämnaren är referensen till zebrans randighet.

Langacker presenterar vad man skulle kunna kalla för en saliens-hierarki, där han graderar saliens- eller uppmärksamhetsvärdet för olika fenomen i världen: ”a whole is more salient than its parts; a physical object is more salient than an abstract entity; and a person has maximal cognitive salience” (Langacker 1991:171).

3.1.7. Konventionalisering i relation till befästning

Ur ett kognitivt perspektiv är befästning av ett uttryck en psykologisk process, medan konventionalisering är ett socialt fenomen. Konventionalisering av specifika uttryck sker i en språkgemenskap när de sprids och språkbrukarna blir alltmer överens om deras betydelse. Konventionaliseringen innebär att det uppstår en socialt etablerad koppling mellan uttrycket (formen) och dess betydelse (se vidare avsnitt 3.2.2 för ett lexikalt perspektiv på konventionalisering).

Schmid (2015) argumenterar för att konventionalisering inte är det samma som distribuerad befästning, eftersom dessa processer verkar i olika domäner. För det första, framhåller han, sker befästning i individers hjärnor, medan konventionalisering sker i språkgemenskaper. För det andra rör processer av befästning individers associationsmönster, och den mentala representationen av befästa enheter kan skilja sig åt mellan individer, med avseende på t.ex. betydelse och funktion. Sociala konventionaliseringsprocesser, å andra sidan, verkar på yttrandetyper (dvs. ytnära abstraktioner över enskilda yttranden), och konventionalisering är ”blind” för detaljerna i mentala representationer. En talare som använder ett ord bidrar till ordets spridning, men det är oklart hur ordet kommer att representeras i olika individers hjärnor. För det tredje skiljer sig befästning och konventionalisering åt med avseende på de krafter som ligger bakom processerna. Psykologiska krafter som benägenheten att jämföra och kategorisera sammanhänger med befästning. Huruvida en nybildning får fäste och blir konventionaliserad i en språkgemenskap sammanhänger med hur sociala nätverk är organiserade, eftersom det påverkar spridningen och normeringen av uttryck (Schmid 2015:11, 16).

3.1.8. Kognitiva perspektiv på flerspråkighet

I min studie gör jag bruk av distinktionen mellan förstaspråk (L1) och andraspråk (L2). Jag utgår från Abrahamssons (2009:13) definition där inlärningsordningen mellan språken är det särskiljande kriteriet, och där

andraspråket definieras som ”vilket språk som helst som tillägnats *efter* det att förstaspråket etablerats eller börjat etableras” (hans notation).

Distinktionen mellan L1 och L2 har emellertid problematiserats, både utifrån teoretiska och mer praktiska perspektiv. Forskare har uppmärksammat svårigheterna med att dra en entydig gräns mellan ett första- och ett andraspråk. Komplexiteten i att kategorisera olika typer av flerspråkighet har framhållits, eftersom flerspråkighetsbakgrunder kan se mycket olika ut, t.ex. vad gäller startålder (eng. *age of onset*) för det nya språket, platsen för den tidiga inläringen (hemma eller i förskola/skola) och modersmål hos föräldrarna (se t.ex. Fraurud & Boyd 2011). Dessutom har distinktionen mellan L1 och L2 kommit att ifrågasättas alltmer inom sociolingvistiska och pedagogiska ramverk – delvis på ideologiska grunder. I distinktionen mellan L1 och L2 har man identifierat en enspråkighetsnorm och ett bristperspektiv på flerspråkighet. Som alternativ har forskare och praktiker (t.ex. verksamma lärare) lyft fram hur flerspråkiga individer utnyttjar alla sina språkliga resurser i en process av transspråkande (eng. *translanguaging*), där gränserna mellan språken delvis upphävs (se t.ex. Svensson 2017, García & Wei 2018, Yoxsimer Paulsrud et al. 2018).

Att L1/L2-distinktionen är svårhanterlig framgår i föreliggande studie, där informanter med till synes liknande språklig bakgrund skiljer sig åt i rapporteringen av sina modersmål (svenska och ett ytterligare modersmål respektive annat modersmål än svenska; se vidare 4.10.2). I den mån L1/L2-distinktionen förstärker en enspråkighetsnorm eller speglar ett bristperspektiv på flerspråkighet är detta problematiskt. Det resursperspektiv på flerspråkighet som utmärker forskning om transspråkande är ideologiskt tilltalande. Utifrån ett kognitivt perspektiv är det emellertid relevant för mig att använda distinktionen L1–L2 som ett analytiskt verktyg – utan att jag därmed anammar en enspråkighetsnorm – eftersom det har betydelse för den mentala representationen om språket är ett första- eller ett andraspråk.

Ellis (2006) presenterar det kognitiva ramverket CREED, där akronymen står för att L2-tillägnande är Construction-based, Rational, Exemplar-driven, Emergent och Dialectic. Sålunda följer L1- och L2-tillägnandet samma principer. Att tillägna sig ett språk handlar om att tillägna sig *konstruktioner*, som mappar form och betydelse. Språkprocessningen är *rationell* i så måtto att inlärarens tolkning omedvetet påverkas av tidigare språkliga erfarenheter som inverkat på den språkliga representationen. Således bedömer inläraren sannolikheten för en viss tolkning utifrån erfarna frekvensmönster och kontextuella mönster för konstruktionen. Att inläringen är *exemplardriven* innebär att abstrakta representationer (scheman/konstruktioner/kategorier) emanerar ur exemplar av olika prototypiskhet. Prototypiska exemplar är mer lättillgängliga än andra i inlär-

ningsituationen. Att språket är *emergent* innebär att det är ett komplext och dynamiskt system som konstrueras i en ständigt pågående, interaktiv process. När det gäller en emergent språklig representation skiljer sig de kognitiva omständigheter som gäller vid andraspråksinläring från situationen vid förstaspråksinläring. Till skillnad från förstaspråksinläringens ”tabula rasa”-situation beskriver Ellis andraspråkstillägnande som en situation av ”tabula repleta”, där erfarenheter och associationer från förstaspråket färgar andraspråket:

In the first few years of life, the human learning mechanism optimizes its representations of first language from the cumulative sample of first language input. One result of this process is that the initial state for SLA is a tabula repleta; it is no longer a plastic system, it is one that is already tuned and committed to the L1. Our later experience is shaded by prior associations; it is perceived through the memories of what has gone before. Since the optimal solution for L2 is not that for L1, SLA suffers from various types of L1 interference. Transfer phenomena pervade SLA. (Ellis 2006:109; hans emfas)

Ett språk tillägnas via de associationer som väcks utifrån input. Detta gäller andraspråket likaväl som förstaspråket. Men L2-inläraren är påverkad av konkurrerande associationer från sitt L1. När det gäller referens till tid, exempelvis, tenderar L2-inlärare att snabbt tillägna sig temporala adverb, eftersom sådana dels är bekanta från deras L1, dels är salienta och tydliga. Tempusböjningen av verb är däremot en mindre pregnant tempusmarkör och därmed en mer svårtillgänglig struktur. L2-inläraren förlitar sig inledningsvis enbart på adverbial, medan redundanta tempusmarkörer överskuggas. Detta exemplifierar inlärningsfenomenet blockering (eng. *blocking*) och anses vara ett tydligt exempel på hur erfarenheter från L1 färgar inläringen av L2:

[...] associative learning induces such phenomena of learned selective attention as overshadowing and blocking, latent inhibition, perceptual learning, interference, and other effects of salience, transfer and inhibition. All of these mechanisms filter and color the perception of the second language. Thus the shortcomings of the L2 endstate are rational when seen through the lenses of the L1, although again, they are the result of many interacting parts and processes. (Ellis 2006:110)

Den sista bokstaven i akronymen CREED står för *dialectic*. L2-inläringen är dialektisk i så måtto att språkinstanter från den dittills uppnådda nivån i inlärar- eller interimsspråket bryts mot den evidens som pragmatisk, lingvistisk och metalingvistisk respons bidrar med (Ellis 2006:112).

När det gäller det mentala lexikonet är skillnaden mellan första- och andraspråksinläring en fråga om skillnader avseende begreppsbild-

ning. Medan förstaspråksinläraren samtidigt lär sig om världen och de språkliga uttrycken för företeelser och förhållanden i världen, har andraspråksinläraren redan (i varierande grad, beroende på ålder) tillgång till begrepp och deras språkliga motsvarigheter på förstaspråket. Inlärningen av ett andraspråk handlar emellertid inte bara om att lära sig nya etiketter för kända begrepp och referenter, eftersom det kan råda tvärspråkliga skillnader mellan begrepp i olika språk, vilket aktualiserar frågan om konceptuell representation. Pavlenko redogör för forskning som fokuserar lingvistiska kategorier (eng. *linguistic categories*), dvs. konceptuella mentala representationer som är sammanlänkade med ord. Forskningen gäller således mappningen mellan referenter i världen och ord (Pavlenko 2009:125–126).

Experiment visar att inlärningen av ett ord i L2 är oproblematisk så länge det råder fullständig eller nästan fullständig konceptuell ekvivalens mellan första- och andraspråket. Uppgiften blir då att länka L2-ordet till ett redan etablerat begrepp. Positiv L1-transfer underlättar denna process. Inlärningen av en L2-form är svårare vid partiell ekvivalens eller konceptuell icke-ekvivalens. Vid partiell ekvivalens kan skillnaden bestå i att kategorier är avgränsade på olika sätt i olika språk. Ett exempel på detta är engelskans *jar* som motsvaras av spanska *frasco* som utöver 'jar' också inkluderar 'bottle' och 'container'. Partiell ekvivalens råder också när ett begrepp på ett språk motsvaras av två i ett annat, t.ex. såsom i fallet med det engelska verbet *know* som motsvaras av *veta* och *känna* på svenska. Studier visar att det är svårast i de fall språkbrukaren ska gå från en vid kategori i sitt förstaspråk till en snävare på andraspråket (t.ex. från *frasco* till *jar*), eller när en kategori i förstaspråket motsvaras av två i andraspråket (t.ex. när *know* ska uttryckas med antingen *veta* eller *känna*) (Pavlenko 2009:134–138). Konceptuell icke-ekvivalens mellan språk råder när ett begrepp i ett språk saknar en konceptuell motsvarighet i ett annat. Pavlenko beskriver ett experiment som indikerar att 'privacy' i engelskan inte har någon motsvarighet på ryska. Icke-ekvivalens kan leda till konceptuell utveckling, dvs. en utveckling av en ny mental representation som möjliggör en mappning ord–referent. Alternativt kan icke-ekvivalens föranleda en konceptuell omstrukturering, varvid inläraren justerar gränserna för en kategori (Pavlenko 2009:139–141).

Pavlenko (2009) presenterar en modell för den konceptuella representationen av två språk hos den tvåspråkiga individen. Modellen gäller mappningen mellan referenter i världen och ord. Pavlenko utgår från att det finns L1-specifika och L2-specifika lingvistiska kategorier. Med avseende på dessa råder det konceptuell icke-ekvivalens mellan språken. Därutöver förekommer gemensamma lingvistiska kategorier. För dessa råder full eller partiell ekvivalens. Konceptuella representationer för språken är med

andra ord helt och hållet delade, partiellt överlappande eller fullständigt språkspecifika. Förekomsten av språkspecifika konceptuella representationer medför konceptuell icke-ekvivalens mellan språk (2009:146–148). Detta kan ha betydelse för hur de flerspråkiga informanterna i min studie tolkar enkätens sammansättningar.

3.1.9. Befästning – inte bara en fråga om frekvens

Som framgått utgår forskare inom det kognitiva paradigmet från att såväl språkliga som encyklopediska erfarenheter påverkar den mentala representationen av språket. Repetition är en viktig faktor som påverkar huruvida språkliga token blir befästa som enheter. Frekvensen för ord och uttryck är betydelsefull, eftersom det statistiskt sett är större chans att språkbrukaren mött ett ord upprepade gånger om det är frekvent. (Att mäta frekvens är dock behäftat med svårigheter, och det är omöjligt att mäta hur ofta en enskild individ mött ett specifikt ord; se vidare avsnitt 4.4.) Nära sammanhängande med faktorn frekvens är faktorn tid. Det är upprepade möten med ett ord över tid (dvs. ackumulerad frekvens) som kan leda till befästning av ett ord eller uttryck. Det är därför L1/L2-distinktionen är viktig i min studie, eftersom den speglar skillnader i exponeringstid för det svenska språket. L1/L2-distinktionen är också betydelsefull ur ett kontrastivt perspektiv. L2-talarens språk påverkas av befästa språkliga och konceptuella mönster från förstaspråket, dvs. L2-talarna har att beakta andra, konkurrerande mönster och andra delar i meningspotentialerna, vilket påverkar tillägnet av andraspråket. Utöver dessa faktorer har saliensen hos ord och uttryck inverkan på graden av befästning hos talare. Ord som speglar pregnanta egenskaper blir lättare befästa, liksom ord som väcker känslor (t.ex. svordomar).

3.2. Lexikala perspektiv

Från den lexikala semantiken har diskussionen om kompositionalitet, meningspotential och genomskinlighet betydelse för min studie. Förutom dessa begrepp diskuteras i det följande begreppen lexikalisering och konventionalisering.

3.2.1. Kompositionalitet, meningspotential och genomskinlighet

Kompositionalitetsprincipen i språket innebär att helhetens betydelse följer av de ingående delarna. Kompositionalitet beskrivs ofta som resultatet av en bottom up-process utifrån generella ordbildningsregler genom vilken ord kombineras och bildar en fras med en specifik betydelse, som i sin tur kombineras till en sats med en specifik betydelse. Lexikon och grammatik samverkar så, att betydelsen uppstår genom de ingående delarna och hur de kombineras. I en sammansättningskontext är det som redan nämnts viktigt att notera att helhetens betydelse inte är lika med *summan* av de ingående delarna, såsom ibland föreslås, eftersom olika ordning mellan leden ger olika betydelse (se Langacker 2008:168).

Det är knappast kontroversiellt att påstå att ett uttrycks betydelse kan härledas ur de ingående delarna, men frågan är hur strikt man tolkar kompositionalitetsprincipen, dvs. hur *förutsägbar* man antar att betydelsen är utifrån de ingående delarna.

Costello & Keane (2000) refererar en traditionell syn på kompositionalitetsprincipen. Enligt denna råder strikt förutsägbarhet, dvs. helhetens betydelse betraktas som given utifrån de ingående delarna:

According to the principle of compositionality, anyone who knows the meaning of each word in a complex expression should need no further information to grasp the meaning of the expression. (Costello & Keane 2000:336)

Denna definition av kompositionalitetsprincipen är stark, menar de, och det är lätt att hitta exempel där den framstår som alltför stark. Costello & Keane argumenterar i stället för ett graduellt kompositionalitetsbegrepp (2000:336–340).

Kompositionalitetsprincipens begränsningar blir uppenbara i samband med sammansättningar, eftersom dessa till sin natur nästan alltid är flertydiga. Därmed föreligger ett stort antal potentiella, kompositionella betydelser, vilket innebär att kompositionalitetsprincipen är otillräcklig som förklaring till vår förmåga att förstå nya komplexa ord (se Svanlund 2009:34).

Ett begrepp som sammanhänger med kompositionalitet och analyserbarhet och som brukar användas i ordbildningslitteraturen är *genomskinlighet*. Traditionellt sett brukar en sammansättning räknas som genomskinlig ”om den klart kan delas upp i de ingående ordlederna, och förtrogenhet med ordlederna är tillräckligt för att tolka betydelsen hos sammansättningen” (Svanlund 2009:35). Enligt Malmgren (1994:24) är genomskinliga sammansättningar ”lätt begripliga för vem som helst som

kan de båda ord som ingår i dem” (Malmgren 1994:24). Utifrån sådana definitioner bedöms *jordhög* vara en genomskinlig sammansättning, medan *jordgubbe* beskrivs som en ogenomskinlig sammansättning (se Enström 2016:103). Kompositionalitet är alltså en förutsättning för genomskinlighet. Vid kompositionella, determinativa sammansättningar är helheten hyponym till efterledet: en *jordhög* är ’en hög av jord’. Så är inte fallet vid ogenomskinliga sammansättningar, såsom *jordgubbe* (*jordgubbe* är inte en ’gubbe i eller av jord’). Många ogenomskinliga sammansättningar är ogenomskinliga p.g.a. metaforik, t.ex. när en person liknas vid solen och beskrivs som en *solstråle*. Kompositionella sammansättningar brukar vara produktiva, vilket innebär att de ingående leden är utbytbara (jämför *jordhög/stenhög* med varierande förled, respektive *stenhögl/stengårdsgård* med varierande efterled).

Genomskinlighet har en praktisk betydelse inom lexikografen. Eftersom allt inte får plats i en ordbok blir genomskinlighet ett starkt kriterium för vilka sammansättningar som ska behandlas i lexikografiska verk. I urvalet prioriteras ogenomskinliga sammansättningar framför genomskinliga. I praktiken upptas emellertid även många genomskinliga sammansättningar i lexikografiska verk (se Svensén 2004:81).

När man skiljer mellan å ena sidan ogenomskinliga sammansättningar och å andra sidan regelbundna, genomskinliga sammansättningar är risken stor att man överdriver sammansättnings genomskinlighet utifrån den kunskap man redan har om de konventionaliserade betydelserna. Som Svanlund (2002) framhåller kan man ifrågasätta genomskinligheten hos många av de sammansättningar som definieras som genomskinliga i ordbildningsläror. Han nämner exemplen *daghem* (hämtat ur Thorell 1981) och *legevisitt* (hämtat ur Bakken 1998), där all information om de inblandade parterna och relationen mellan dem knappast är synlig i själva sammansättningen (Svanlund 2002:18–19). Ett annat exempel som jag identifierat är *solvarm* som Malmgren (1994:24) betecknar som en genomskinlig sammansättning, trots att den kan uttolkas med olika relationer mellan leden. Genomskinlighetsbegreppet är förrädisk, menar Svanlund (2002:23–24), eftersom det speglar den förenklade föreställningen om att sammansättnings betydelser enkelt kan räknas ut enligt en matematisk princip. Men polysemi kan förekomma både hos ordleden och i relationen mellan dem. Därför är det enligt Svanlund mer fruktbart att använda sig av den kognitiva semantikens motiveringsbegrepp: sammansättnings betydelser är vanligen motiverade och kan härledas ur de ingående leden, men de är inte förutsägbara (2002:34). Som framgått tidigare (avsnitt 3.1.2) innebär motiveringsbegreppet hos Langacker att det råder partiell snarare än fullständig kompositionalitet (Langacker 1987:448–450). Min uppfattning är att denna syn på kom-

positionalitet fungerar bättre på sammansättningar än en strikt tolkad kompositionalitetsprincip, där kompositionalitet sammankopplas med förutsägbarhet.

Upplevelsen av sammansättnings genomskinlighet (vilken delvis sammanfaller med Langackers analyserbarhetsbegrepp) har att göra med hur lätt vi aktiverar relevanta kunskapsramar och språkliga mönster, och hur lätt det är att kombinera dessa, framhåller Svanlund (2009). Ibland räcker den minikontext som leden ger varandra; ibland avgör olika typer av kontextuella faktorer, snarare än ordleden i sig. Som genomskinliga räknar Svanlund sammansättningar där vi lätt genererar en föredragen default-betydelse, där bara en av flera möjliga betydelser aktiveras (2009:35–36). Ur detta perspektiv är genomskinlighet en fråga om subjektiv upplevelse, snarare än en objektiv egenskap som kan tillskrivas sammansättningen i sig.

Kontexten har som framgått stor betydelse vid tolkningen av sammansättningar. Betydelsen hos enskilda sammansättningar kan sägas uppstå i gränsen mellan språklig kunskap och omvärldskunskap. Eftersom det ofta finns flera potentiella betydelsereationer mellan för- och efterled får sammanhanget eller omvärldskunskapen konsulteras för att fastställa att *hästskor* är 'ett slags skor avsedda för häst' medan *krokodilskor* är 'skor tillverkade av krokodil(skin)' – för att ge ett exempel på potentiell relationspolysemi.

Forskare som velat ge teoretisk tyngd åt den inverkan som kontexten har på ett uttrycks betydelse har etablerat eller använt sig av begrepp som *betydelse-* och *meningspotential*.

Hanks (2000) för en teoretisk diskussion om kontextens inverkan på ordens betydelse. Han framhåller att polysemi och betydelse- eller meningspotentialer (snarare än fasta betydelser) är grundläggande i språket. I artikeln "Do Word Meanings Exist?" lyder konklusionen enligt följande:

Yes, word meanings do exist, but traditional descriptions are misleading. Outside the context of a meaning event, in which there is participation of utterer and audience, words have meaning potentials, rather than just meaning. The meaning potential of each word is made up of a number of components, which may be activated cognitively by other words in the context in which it is used. These cognitive components are linked in a network which provides the whole semantic base of the language, with enormous dynamic potential for saying new thing and relating the unknown to the known. (Hanks 2000:214)

Snarare än betydelser bär orden alltså potentiella betydelser som först i en specifik språksituation kan övergå i realiserad betydelse.

Ett liknande resonemang förs av Norén & Linell (2006) som för svenskans del etablerar termen *meningspotential*. Forskarna ifrågasätter

antagandet om en strikt kompositionalitetsprincip, där betydelsen hos ett uttryck enkelt kan härledas ur de ingående orden. Situationen är mer komplex, menar de, eftersom orden påverkar varandras potentialitet:

[...] att olika uttryck, t.ex. *lexem*, i en utsaga inte bara bidrar till tolkningen av hela utsagan utan också förstärker eller försvagar [...] utvalda delar av varandras *potentialer*. Det innebär att utsagans innebörd inte kan härledas från summan av elementära betydelser hos orden plus kombinationsstrukturer enligt något slags kompositionalitetsprincip, utan att den emergerar ur samspelet mellan orden; genom att selektera och förstärka delar av varandras potentialitet blir innebörden av hela utsagan mer än vad som blir resultatet av att addera delarna var för sig till varandra. (Norén & Linell 2006:15–16; deras emfas)

Allwood (2003) går så långt att han tillskriver omvärldskunskapen samma betydelse som den språkliga kunskapen i sin beskrivning av ordens ”meaning potential”, dvs. vad orden potentiellt kan betyda. I hans definition av meningspotential ingår all information som ett ord har använts för att uttrycka – på ett individuellt och ett socialt plan:

The meaning potential, then, does not result from trying to find a generally valid type meaning for a word; rather, it is the union of individually or collectively remembered uses. [...] A consequence of this approach is that no attempt is made to distinguish between lexical and encyclopedic information in terms of the kind of information that is contained in the meaning potential. Meaning potentials contain both kinds of information – information deriving from use of language and information deriving from other experience with the world. (Allwood 2003:43)

Norén & Linell (2006:9) invänder mot Allwoods (2003) antagande om att meningspotentialen är lika med summan av individens och språkgemenskapens användningar av ordet. Enligt dem är meningspotentialer strukturerade: det finns kärnaspekter i betydelsen och mer perifera aspekter.

Ordens meningspotentialer, såsom de beskrivs av Allwood (2003), är i det närmaste allomfattande. Men i en specifik språksituation aktiveras bara delar av denna potentialitet och blir till aktualiserad betydelse:

[...] activated actual meaning only exists in context, and outside of context, there is only potential meaning. [...] When used, a linguistic expression activates its meaning potential through cognitive operations whose function is to achieve compatibility between the meaning potential of a particular expression, the meaning potentials of other expressions, and the extralinguistic content. Contextually determined meanings thus result from partial activation of the meaning potentials of the expressions guided by cognitive operations. (Allwood 2003:52–53)

Distinktionen mellan potentiell betydelse (hos själva uttrycket) och aktiverad betydelse (i den specifika språksituationen) känns igen från Hanks (2000) resonemang.

Genom begreppet meningspotential problematiseras begrepp som kompositionalitet och genomskinlighet. En sammansättning kan knappast ha en kompositionell, i betydelsen förutsägbar, betydelse, eftersom det rör sig om ett komplext uttryck där såväl delar som helhet har meningspotentialer som innefattar både språklig och encyklopedisk kunskap (se Allwood 2003:43). I vilken mån vi uppfattar en sammansättning som genomskinlig har dels att göra med hur bekanta vi är med den, dels har det – vid tillfälliga sammansättningar – att göra med vilka ledtrådar vi har tillgång till i form av språklig och extralingvistisk kontext, samt i vilken mån vi kan analysera den nya sammansättningen enligt kända språkliga mönster. Med andra ord tar vi hjälp av tidigare erfarenheter som gjort avtryck i uttryckets meningspotential. Allwoods definition av meningspotentialen som ett resultat av individens och kollektivets samlade språkliga erfarenheter har kritiserats av Norén & Linell (2006) för att vara alltför vidlyftig. Jag instämmer inte i denna kritik. Som jag tolkar Allwood (2003) är meningspotentialen visserligen mycket vid. Den innefattar alla individuella och kollektiva erfarenheter av det språkliga uttrycket, vilket kan jämföras med Bybees diskussion om mental representation av enskilda instanser i form av exemplar och exemplarmoln. Men i ett visst sammanhang aktiveras bara delar av meningspotentialen och blir till realiserad betydelse genom individens kognition. Beskrivningen av hur potentiell betydelse övergår i realiserad överensstämmer med den kognitiva grammatikens antagande om betydelse som en följd av individens konceptualisering (se avsnitt 3.1.2).

Hur språkbruk och meningspotential (eller *betydelsepotential*) påverkar varandra i en ständig växelverkan illustreras i följande citat från Rudebeck (2006): "[...] varje användning av ett ord är beroende av ordets tidigare användning i språkkollektivets historia och i individens liv, samtidigt som den förändrar ordets fortsatta betydelsepotential" (Rudebeck 2006:128).

3.2.2. Lexikalisering och konventionalisering

Lexikalisering är en term som kommit att användas i flera skilda betydelser. I det följande redogör jag för olika perspektiv på lexikalisering och klargör hur jag kommer att använda mig av termen.

När det gäller sammansättningar har termen *lexikaliserad* ofta använts om sådana sammansättningar vars betydelse inte tydligt kan härledas ur de ingående leden, dvs. sammansättningar som framstår som semantiskt genomskinliga eller idiosynkratiska. Denna användning av termen före-

kommer exempelvis hos Bauer (1983) och Bakken (1998) som beskriver lexikalisering i termer av en diakron process från nybildning, via konventionalisering till idiosynkrasi och semantisk ogenomskinlighet. Även i Rogströms (1998) lexikografhistoriska avhandling lyfts idiosynkrasi fram som lexikaliseringsprocessens resultat. Lexikalisering definieras där som ”[d]en process som resulterar i att ett ord slutligen etableras i en viss form med en viss betydelse som inte är given eller kan förstås enbart utifrån ordets formsida” (1998:205–206).

I en annan definition innebär lexikalisering att ett visst begrepp eller fenomen i världen får ett specifikt uttryck, ett *lexem*. Det klassiska exemplet är att *dygn* är ett lexikaliserat uttryck i svenskan, till skillnad från i engelskan där samma semantiska innehåll måste uttryckas med en fras *day and night* (Simonsen et al. 1988:308; se också Svanlund 2002:8).

Ytterligare ett perspektiv på lexikalisering tar fasta på graden av etablering, dvs. graden av konventionaliserad fasthet i förbindelsen mellan uttryck och innehåll. Lexikalisering i denna bemärkelse motsvaras av konventionalisering hos Bakken (1998) och institutionalisering hos Bauer (1983) (se vidare Svanlund 2002:8).

Motsägelsefullt nog används således lexikalisering å ena sidan för att beteckna uppkomsten av en etablerad koppling mellan innehåll och uttryck, å andra sidan för att beteckna idiosynkrasi och semantisk ogenomskinlighet. Svanlund (2002) argumenterar övertygande för uppfattningen att termen *lexikalisering* borde reserveras för ordens etablerings- eller konventionaliseringsfas, snarare än att beteckna uppkomsten av idiosynkrasier, eller det han kallar för *anomalier*. Innan denna diskussion analyseras närmare kommer Bakkens (1998) lexikaliseringsmodell att presenteras, eftersom den utgör en viktig bakgrund till Svanlunds argumentation.

Bakken definierar lexikalisering som ”en gradert prosess der en sammensetnings betydning går over fra å være komposisjonell og motivert, til å bli ikke-komposisjonell og demotivert” (1998:61). Allteftersom processen fortskrider framstår kopplingen mellan uttryck och betydelse som mer och mer arbiträr. Den interna morfologiska strukturen blir alltmer irrelevant för språkbrukarna – och sammansättningen kan till sist komma att uppfattas eller hanteras som ett simplex (Bakken 1998:61, 71–72). Lexikaliseringsprocessen, menar hon, inleds med uppkomsten av en ”spontanlaging”, dvs. en ny bruksinstans av en tidigare icke känd sammansättning. Sammansättningen har referens men inledningsvis ingen konventionaliserad betydelse, utan förståelsen uppstår genom kontexten (Bakken 1998:71–72, 76–77).

I lexikaliseringsprocessens första steg inleds konventionaliseringen. En spontanbildning blir konventionaliserad när det har uppstått en social enighet om en given förbindelse mellan uttryck och ett innehåll, och när

denna förbindelse kan etableras oavhängigt av kontext. Sammansättningar kan vara svagt eller starkt konventionaliserade, och graden av konventionalisering är kopplad till antalet bruksinstanser. Enligt Bakken är de konventionaliserade sammansättningarna dessutom svagt lexikaliserade: de hör till ett socialt sanktionerat språkbruk (inte bara ett individuellt), och de kan utgöra analogibaser för nybildning. I princip är de konventionaliserade sammansättningarna att betrakta som enheter. Enligt Bakken kan det dock vara svårt att avgöra om svagt konventionaliserade sammansättningar som *damesko* och *vinterkläder* hämtas fram som fasta enheter eller om de bildas kompositionellt i ögonblicket. Det rör sig om motiverade och produktiva sammansättningar som i princip skulle kunna nybildas för varje bruksinstans. Med hänvisning till Langacker framhåller Bakken att den systemdefinierade kompositionaliteten inte behöver försvinna vid konventionalisering; ledens bidrag till helhetens betydelse kan utgöra en aspekt av det som blir konventionaliserat. Däremot är det troligt att den individuella analyserbarheten påverkas negativt av konventionaliseringen, eftersom en konventionaliserad koppling mellan uttryck och innehåll ger automatisering och denna automatisering fungerar som en motkraft till analyserande verksamhet. Konventionaliserade sammansättningar kan alltså mycket väl ha en kompositionell betydelse – men samtidigt är konventionaliseringen en nödvändig förutsättning för nästa steg i lexikaliseringsprocessen, nämligen betydelsesammansmältning (Bakken 1998:71–74, 78).

Konventionaliserade sammansättningar kan i lexikaliseringsprocessens andra fas drabbas av betydelsesammansmältning och icke-kompositionalitet. Det innebär att helhetens betydelse inte längre är en direkt funktion av de ingående leden. Även om betydelseaspekterna hos leden är fortsatt relevanta för helhetens betydelse, så har helheten en mer specificerad betydelse än vad de ingående leden antyder. Konventionaliserade och kompositionella sammansättningar kan ges cirkeldefinitioner av typen ”en *damesko* är en sko för damer” men för betydelsesammansmälta, icke-kompositionella sammansättningar fångar sådana cirkeldefinitioner inte hela betydelsen, jämför ”?ett *rödvin* är ett vin som är rött”. Så länge ett led i sammansättningen kan bytas ut mot en synonym är sammansättningen kompositionell. När detta inte längre är möjligt är det ett tecken på att sammansättningen är icke-kompositionell och att den genomgått betydelsesammansmältning, jämför *rödvin*?*purpurvin*. Som en följd av betydelsesammansmältningen blir den interna strukturen irrelevant (Bakken 1998:80–83).

Betydelsesammansmältning kan – men måste inte – leda till fas tre i Bakkens lexikaliseringsmodell, nämligen demotivering. En demotiverad sammansättning definieras av Bakken som en icke-kompositionell sam-

mansättning där ett eller båda leden inte står i något motiveringsförhållande till helhetens betydelse. Sammansättningar är starkt demotiverade när båda leden förefaller märkliga, såsom i *marreritt* ('mardröm'). I svagt demotiverade sammansättningar har bara ett av leden oklar betydelse. Så är fallet med sammansättningarna *brudgom* och *sjusover*. Demotivering kan ske av språkliga orsaker: andraledet i *brudgom* finns inte längre kvar som simplex. Demotivering kan dessutom vara resultatet av förändrade uppfattningar om världen. Sammansättningen *marreritt* 'mardröm' var motiverad så länge man trodde på ett övernaturligt kvinnoväsen – en *mara* – som satte sig på den sovandes bröst. Slutligen kan demotivering inträffa när helhetens betydelse ändrats så mycket att ledens betydelse och helhetens inte längre hänger ihop, jämför *stafkarl* 'man med stav' som övergått till *stakkar* 'person värd medömkan'. Konventionalisering och betydelsesammansmältning, dvs. icke-kompositionalitet, är en förutsättning för demotivering; det är svårt att tänka sig demotivering *före* icke-kompositionell betydelse. Däremot måste inte icke-kompositionella och betydelsesammansmälta sammansättningar genomgå demotivering. Sammansättningarna *brandman* och *rödvin* är betydelsesammansmälta men motiverade (Bakken 1998:84–88).

Svanlund (2002) invänder mot den processorienterade, linjära beskrivningen av lexikalisering hos Bakken (1998). Han menar att modellen förutsätter att lexikalisering i betydelsen 'uppkomst av anomalier' är något som sker successivt och som drabbar redan etablerade, regelbundna sammansättningar. Därmed, menar Svanlund, bortser man från möjligheten att sammansättningar bär på anomalier redan från början, t.ex. att *blåbär* från början uppstått som beteckning för något mer specifikt än ett blått bär vilket som helst. Dessutom finns exempel på att nya, anomaliserade betydelser hos ord uppstår plötsligt, snarare än gradvis, exempelvis på grund av genomslag via ett enskilt teveprogram (Svanlund 2002:25–26).

Svanlunds (2002) huvudkritik mot tidigare användningar av begreppet rör dock det tänkta sambandet mellan lexikalisering och ogenomskinlighet. *Lexikaliserad* har använts om sådana sammansättningar vars betydelse inte tydligt kan härledas ur de ingående leden, dvs. semantiskt ogenomskinliga sammansättningar som bär på anomalier. Bakom en sådan syn, framhåller Svanlund, ligger antagandet om att bara det som är oregelbundet behöver listas i det mentala lexikonet, eftersom grammatiken behandlar resten. Denna uppfattning vilar på en strikt dikotomi mellan grammatiken (som behandlar regelbundenheter) och lexikonet (som hanterar oregelbundenheter). I kontrast till detta framhåller Svanlund att många forskare i dag tar avstånd från en sådan strikt dikotomi mellan grammatik och lexikon, och i stället antar att det mentala lexikonet också rymmer regelbundna, etablerade sammansättningar (Svanlund 2002:10–11). Som

framgått gäller detta i hög grad det kognitiva paradigmet (se avsnitten 3.1.1 och 3.1.3). Svanlund uppmärksammar ytterligare ett problem med att definiera lexikalisering som anomalisering. Om man skiljer mellan å ena sidan lexikaliserade, ogenomskinliga sammansättningar och å andra sidan regelbundna, genomskinliga sammansättningar är risken stor att man överdriver sammansättningsgenomskinlighet utifrån den kunskap man redan har om de etablerade sammansättningarna (2002:18–19; se vidare avsnitt 3.2.1).

Svanlunds konklusion är att termen *lexikalisering* inte bör användas om semantiska avvikelser utan reserveras för etableringsfasen, närmare bestämt de ”processer som förvandlar vissa uttryck till etablerade lexikonheter” (2002:41). Lexikalisering enligt Svanlunds syn motsvarar därmed konventionalisering hos Bakken. I sina empiriska undersökningar använder han själv ofta termen *etablering* i stället för *lexikalisering*, för att markera att det är den aspekten av lexikaliseringsbegreppet som avses (se t.ex. Svanlund 2009:17–18, 38).

Lexikaliseringsgraden menar Svanlund (2002:8) kan mätas antingen utifrån ett specifikt uttryck (hur fast bestämd eller etablerad uttryckets betydelse är) eller utifrån ett specifikt innehåll (hur självklart det är vilket ord vi använder för att beteckna företeelsen). I min undersökning är det förra perspektivet viktigast, dvs. lexikalisering som ett mått på etablerad betydelse. Jag kommer däremot inte att använda *lexikaliserad* som beteckning för semantiskt ogenomskinliga sammansättningar, eftersom det rimmar illa med det kognitiva paradigmet, där man antar att det mentala lexikonet också rymmer sammansättningar som skulle kunna pusslas ihop utifrån produktiva ordbildningsregler. Termen *lexikaliserad* reserveras i min studie för de fall där det uppstått en social enighet om en given förbindelse mellan uttryck och innehåll, dvs. där ett specifikt innehåll fått en fast beteckning som är socialt sanktionerad. Jag kommer också att använda termerna *konventionaliserad* och *etablerad* i samma betydelse.

3.2.3. Konventionalisering och meningspotential ur ett flerspråkighetsperspektiv

I ett flertal svenska studier har man påvisat en lägre grad av konventionalitet i andraspråkstales svenska jämfört med modersmålstales, även i de fall där inlärarna uppnått en avancerad nivå.

Ekberg (2013) framhåller att en lägre grad av konventionalitet och idiomacitet i andraspråkstales språk än i förstaspråkstales inte nödvändigtvis innebär förekomst av många grammatiska eller lexikala avvikelser. Det kan röra sig om mer subtila skillnader, eftersom idioma-

ticitet handlar om att ”*använda* språket på det sätt som en infödd talare gör” (2013:260). Ekberg redovisar resultat från projektet Svenskan hos barn födda i Sverige av invandrade föräldrar, vilket omfattar tvåspråkiga barn i årskurs 5–9. Hon konstaterar att de tvåspråkiga barnen använder ett mindre idiomatiskt språk än den enspråkiga kontrollgruppen. På den lexikala nivån noterar hon omskrivningar som *mammas mamma* i stället för *mormor*, och *port* i stället för *gräns*. Men på en mer övergripande nivå noterar hon att de tvåspråkiga barnen uppvisar större variation och kreativitet när de beskriver bilder och återberättar bildserier än de enspråkiga barnen. De enspråkiga barnen använder sig av fler konventionaliserade och färdigbildade uttryck, dvs. de är mer stereotypa eller förutsägbara i sin språk användning. Ekberg drar slutsatsen att idiomatiskt språkbruk utöver grammatikalitet också handlar om användningsfrekvenser. De tvåspråkiga barnens språk framstår som mindre idiomatiskt för att de underanvänder vissa konventionaliserade uttryck och konstruktioner. Hon konstaterar att stereotypi i vissa fall faktiskt är det som utmärker ett idiomatiskt språkbruk (Ekberg 2013).

Att flerspråkiga personer är mindre konventionella i sitt språkbruk sammanhänger med att de har tillgång till andra mönster och vidare meningspotentialer än personer som bara har ett modersmål. Eftersom tillägandet av ett andraspråk är en situation av *tabula repleta* (Ellis 2006) är det förväntat att mönster från förstaspråket inverkar på andraspråket. Med Onyskos (2016:492) ord kan man förvänta sig att ”the diverse languages of bilingual and multilingual speakers are always active and can shine through on different levels of language use”.

När det gäller tolkning av sammansättningar ställs frågan om konventionalisering på sin spets. Till min studie har jag valt ut sammansättningar som består av simplex som är lättillgängliga och frekventa, men relationspolysemin innebär att betydelsen inte framgår enbart av de ingående leden. När det gäller de etablerade sammansättningarna kräver förståelsen av dem ofta att man känner till den konventionaliserade kopplingen mellan leden. Hur sammansättningar konventionaliserats är oförutsägbart, och konventionaliseringsprocesser är språkspecifika och kan inte översättas från ett språk till ett annat. Därför måste varje sammansättning läras in separat (se Enström 2013:187).

Det finns indikationer på att L2-talares språkbruk skiljer sig från L1-talares med avseende på graden av kompositionell analys. Enligt Wray (2008:19) analyserar L2-talare språklig input i högre grad än L1-talare. Det resulterar i att de lagrar mindre enheter i sitt mentala lexikon än L1-talare, som utöver ord också lagrar färdiga fraser och uttryck. Eftersom L2-talarna utgår mer från enskilda ord blir de mer beroende av regler som styr hur dessa kan kopplas samman (se även Wray 2002:206–209).

Översatt till min sammansättningsforskning skulle detta kunna innebära att L2-talarna i högre grad än L1-talarna förlitar sig på kompositionalitet och försöker pussla ihop sammansättnings betydelse utifrån de ingående leden, vilket – som vi sett – är en mindre framgångsrik strategi, eftersom sammansättningar sällan är kompositionella i betydelsen förutsägbara utifrån de ingående leden. Polysemin hos sammansättningar innebär att även skenbart genomskinliga sammansättningar har flera potentiella betydelser.

Den tillgång till vidare meningspotentialer som utmärker L2-talarens språk medför L1–L2-skillnader med avseende på semantiskt djup i ordförståelsen. Ordförståelsen är i en tillägnandefas partiell, snarare än precis (se Henriksen 1999:311). Ord för konkreta företeelser tenderar att vara lättare att tillägna sig än mer abstrakta, eftersom de aktualiserar färre potentiella betydelser och i högre grad än abstrakta är konceptuellt ekvivalenta mellan språk (Pavlenko 2009:144; se även Enström 2013:185).



4

Metod och material

I detta kapitel redovisas metoden och materialet för den enkätstudie om sammansatta ord som föreliggande avhandling baseras på. I kapitlet presenteras principerna bakom enkätstudiens utformning (avsnitt 4.1) tillsammans med hypoteserna för undersökningen (avsnitt 4.2). Därefter följer en genomgång av de strukturella och semantiska mönster eller aspekter som jag undersöker hos de aktuella sammansättningarna och som legat till grund för indelningen i tolv kategorier av sammansättningar (avsnitt 4.3). Den korpus som utgjort referensmaterial vid bedömningen av de enskilda sammansättningarnas frekvens beskrivs i avsnitt 4.4.

I avsnitten 4.5–4.6 redogörs för enkätens orddel och den del som innehåller frågor om informanternas språkliga bakgrund. Etiska överväganden och det praktiska genomförandet av datainsamlingen beskrivs i avsnitten 4.7–4.8. Därefter följer en beskrivning av de skolor som ingår i studien (avsnitt 4.9). Avsnitt 4.10 handlar om studiens informanter. Jag redogör för indelningen av dem i olika informantgrupper, samt diskuterar de metodologiska svårigheter och gränsdragningsproblem som jag mött i samband med denna kategorisering. I avsnitt 4.11 redogörs för den kvantitativa analysen av resultaten från enkätens flervaldsdel. Där beskrivs de statistiska test som tillämpas på materialet. I avsnitt 4.12 beskriver jag principerna för analysen av tolkningarna i enkätens fritextdel.

4.1. Enkätstudiens utformning

I föreliggande studie undersöks sammansättningar ur ett receptionsperspektiv. Jag analyserar närmare bestämt hur en grupp talare av svenska tolkar ett urval sammansättningar, med syftet att dels belysa strukturella och semantiska mönster hos sammansättningarna, dels analysera talarnas tolkningsstrategier. Informantgruppen består av 202 ungdomar som studerar på ett studieförberedande program på gymnasiet. Datainsamlingen genomfördes på sex gymnasieskolor i 15 klasser och undervisningsgrupper under 2015.

Tidigare forskning om sammansatta ord har i hög grad behandlat nominala sammansättningar, och ofta konkreta sådana (se Svanlund 2009:28). Här studeras utöver nominala sammansättningar också adjektiviska. Enkäten inkluderar nominala och adjektiviska sammansättningar där förledet tillhör någon av de öppna ordklasserna substantiv (N), adjektiv (A) och verb (V), dvs. följande strukturer: NN, VN, AN, NA, VA och AA. Verbala sammansättningar är också intressanta, men de inkluderas inte i denna studie.

Mitt undersökningsobjekt är kompositionellt uppbyggda sammansättningar. Till den enkät som använts för att samla in material har jag valt sammansättningar som uppfyller följande kriterier:

- de kan förstås utan metaforisk tolkning
- helhetens betydelse är potentiellt kompositionell
- de är inte självklart genomskinliga eller analyserbara, på grund av relationspolysemi och (i vissa fall) på grund av polysemi hos de ingående leden

Enkäten består dels av en fritextdel, där informanterna själva förklarar vad var och en av 20 sammansättningar betyder, dels av en flervalsdel där informanterna väljer mellan två eller tre färdiga svarsalternativ med kontextmeningar till 69 sammansättningar. Urvalet av ord och valet av definitionsformat är delvis baserat på erfarenheter från en mindre pilotstudie med 34 informanter som genomfördes under hösten 2013 (se vidare avsnitt 4.5).

4.2. Hypoteser

Genom enkäten undersöks strukturella och semantiska mönster hos sammansättningarna. Undersökningen utgår från ett antal hypoteser, dels hypoteser med fokus på sammansättningarna, dels hypoteser som

rör tolkningarna. Under hypotes 1 presenteras antaganden om svenska sammansättningar, deras struktur och betydelser (1a–c). Hypotes 2 rör tolkningarna i informantgruppen som helhet (2a–c), och hypotes 3 rör skillnader i tolkning mellan L1- och L2-talarna av svenska (3a–e).

1. Hypoteser avseende sammansättningarna, deras struktur och betydelser:
 - a) Det grammatiska och semantiska huvudledets placering sist i nominala sammansättningar är en starkt etablerad struktur.
 - b) I adjektiviska AA-sammansättningar är det grammatiska och semantiska huvudledet och dess placering inte lika etablerat som i nominala sammansättningar.
 - c) Frekventa sammansättningar har en mer konventionaliserad betydelse än mindre frekventa sammansättningar och tillfälliga bildningar, där flera alternativ i meningspotentialen är öppna.
2. Hypoteser avseende tolkningarna i informantgruppen som helhet:
 - a) Huvudledets placering sist i nominala sammansättningar får stort genomslag i informantgruppens tolkningar.
 - b) Vid AA-sammansättningar är spridningen i informantgruppens tolkningar större än vid nominala sammansättningar, med avseende på identifieringen av sammansättningens grammatiska och semantiska huvudled.
 - c) Frekventa sammansättningar ges mer samstämmiga tolkningar i informantgruppen än mindre frekventa.
3. Hypoteser avseende skillnader i tolkning mellan de informanter som har svenska som förstaspråk (L1) och de informanter som har svenska som andraspråk (L2):
 - a) L1-talarnas tolkningar av nominala sammansättningar speglar i högre grad än L2-talarnas den etablerade strukturen där efterledet utgör sammansättningens semantiska och grammatiska huvudled.
 - b) L1-talarna följer i högre grad än L2-talarna gängse principer för huvudledets placering i adjektiviska sammansättningar, dvs. de undviker i högre grad tolkningar som speglar en struktur med huvudledet till vänster.
 - c) Vid etablerade och frekventa sammansättningar är samstämmigheten större i tolkningarna från L1-talarna än i tolkningarna från L2-talarna.
 - d) Vid etablerade men mindre frekventa sammansättningar är skillnaderna mellan L1- och L2-talarnas tolkningar större än vid mer frekventa sammansättningar.
 - e) Tillfälliga bildningar tolkas i analogi med betydelsen hos etablerade sammansättningar i högre grad av L1-talarna än av L2-talarna.

4.3. Undersökta sammansättningstyper

I följande avsnitt beskrivs de undersökta sammansättningarna. Enkäten består av 85 sammansättningar som jag delat in i tolv kategorier beroende på den semantiska eller strukturella relationen mellan leden. Sju av kategorierna innehåller nominala sammansättningar, och de resterande fem innehåller adjektiviska sammansättningar. I detta avsnitt ges en övergripande beskrivning av de sammansättningstyper som enkäten innehåller. Därefter följer en noggrannare beskrivning för varje grupp av sammansättningar i avsnitten 4.3.1.1–4.3.2.5.

En strukturell aspekt som analyseras är hur etablerat det schema är som styr det grammatiska och semantiska huvudledets placering. Det undersöker jag genom att testa vilket led som informanterna tolkar som sammansättningens huvudled i tolv nominala sammansättningar (t.ex. *båthus* och *spelkort*). Tre av dessa tolv möter informanterna i fritextdelen, medan resterande nio förekommer i flervalsdelen av enkäten. Informanternas samlade svar ger information om vilken status schemat [efterled = huvudled] har vid nominala sammansättningar i informantgruppen som helhet, som hypotes 2a är avsedd att testa, vilket i förlängningen ger information om schemats etablering i språkgemenskapen (se hypotes 1a). I nästa steg undersöks om det finns skillnader mellan informantgruppernas tolkningar: i vilken grad speglar de olika gruppernas tolkningar gängse principer för huvudledets placering? Därigenom testas hypotes 3a, dvs. antagandet om att informanterna med svenska som förstaspråk i större utsträckning än informanterna med svenska som andraspråk följer schemat [efterled = huvudled] i sina tolkningar. Ur ett kognitivt perspektiv är sådana skillnader förväntade som ett resultat av skillnader i befästning av dels det övergripande schemat för huvudledets placering, dels de enskilda instanserna.

Hur informanterna identifierar det semantiska huvudledet i en sammansättning analyseras för ytterligare en typ av nominala sammansättningar. Jag undersöker tolkningen av treledade sammansättningar (t.ex. *barnbokhylla–barnbokshylla*) för att se om närvaron eller frånvaron av fog påverkar det semantiska huvudledets räckvidd.

I de övriga grupperna av nominala sammansättningar är mer ytnära, semantiska aspekter och lågnivåskematan i fokus, snarare än abstrakta strukturer och övergripande scheman. I fokus här står de semantiska relationerna mellan leden. Jag analyserar vad tolkningarna från informantgruppen som helhet säger om meningspotentialen hos sammansättningarna. Därigenom testas hypotesen om att hög frekvens korrelerar med samstämmighet i informantgruppen som helhet (hypotes 2c), vilket indikerar konventionalisering av sammansättningens betydelse i språkgemenskapen (hypotes 1c), medan låg frekvens innebär en mer förhand-

lingsbar betydelse där fler alternativ i meningspotentialen kan komma att aktualiseras. Jag jämför också svaren för de olika informantgrupperna, och ställer dem i relation till de antaganden som tillsammans utgör hypotes 3 om skillnader mellan första- och andraspråkstalarna vid etablerade och frekventa sammansättningar (hypotes 3c), etablerade men mindre frekventa sammansättningar (hypotes 3d) och tillfälliga sammansättningar (hypotes 3e). Utifrån mitt teoretiska perspektiv förväntas sådana skillnader föreligga och vara resultatet av skillnader i befästning av specifika sammansättningar och analogibaser på olika schematicitetsnivåer.

Vilka semantiska relationer som förekommer mellan leden och hur föremål conceptualiseras undersöks i en grupp av sammansättningar som betecknar behållare av olika slag. Jag har genom enkäten undersökt om sammansättningar som *fruktorg*, *olivkorg* och *plastlåda* conceptualiseras med exempelvis ÄNDAMÅL, INNEHÅLL eller MATERIAL som semantisk relation mellan leden, i anslutning till tidigare forskning om dessa semantiska relationer (se avsnitt 2.2).

En annan fråga som uppmärksammas i sammansättningsforskningen är styrkeförhållandet mellan jämförande egenskapsprojicering och tematisk relationslänkning vid tolkning av sammansatta ord. För att bidra med empiri till denna fråga analyserar jag tolkningen av ett antal sammansättningar som skulle kunna tolkas antingen egenskapsprojicerande (ett *zebralejon* är 'ett lejon som är randigt som en zebra') eller med en tematisk relation mellan leden (ett *zebralejon* är 'ett lejon som jagar zebra').

Frågan om jämförande egenskapstolkning kontra tematisk relationslänkning aktualiseras i ytterligare en kategori av nominala sammansättningar, nämligen personbetecknande sammansättningar där förledet betecknar genus och efterledet betecknar antingen yrke eller genus, t.ex. *kvinnopräst* och *kvinnokarl*. När det gäller den här typen av sammansättningar undersöker jag om informanterna gör en identitetstolkning (t.ex. en *kvinnopräst* är 'en präst som är kvinna') eller en jämförande egenskapstolkning (t.ex. en *kvinnokarl* är 'en karl som är som en kvinna'). Dessa alternativ ställs mot tematisk relationslänkning, dvs. en relation mellan leden som bygger på att ledens referenter intar olika funktionella roller i förhållande till varandra (t.ex. en *kvinnopräst* är 'en präst som tar emot och hjälper kvinnor' eller en *kvinnokarl* är 'en man som har många relationer med kvinnor').

I en annan grupp av sammansättningar undersöker jag agens vid verbavledning. Frågan jag analyserar vid den här typen av sammansättningar är om förledet tolkas som AGENT eller PATIENT i sammansättningar som *kvinnomisshandel*, *mansmisshandel* och *polismisshandel*, vilka är strukturellt ambigua.

Den sista gruppen av nominala sammansättningar i enkäten är betydelsesammansmälta AN-sammansättningar, där helheten betecknar något

mer specifikt än vad som framgår av de ingående leden, t.ex. *finkläder*. Jag undersöker hur informanterna hanterar denna typ av icke-kompositionalitet och vilka skillnader som kan iakttas härvidlag mellan informantgruppernas tolkningar.

I urvalet av adjektiviska sammansättningar uppmärksammas en strukturell aspekt hos AA-sammansättningar, nämligen förekomsten och placeringen av det semantiska huvudledet (eller de semantiska huvudleden). Huvudledet och dess placering antas vara mindre etablerat i adjektiviska sammansättningar än i nominala sammansättningar (hypotes 1b). Detta antas visa sig genom en större spridning i informantgruppens tolkningar när det gäller att identifiera det semantiska huvudledet i AA-sammansättningar än i nominala sammansättningar (hypotes 2b). Dessa hypoteser är formulerade utifrån det teoretiska antagandet om att närvaron av två konkurrerande strukturer (i det här fallet ett kopulativt och ett determinativt schema) leder till en lägre grad av befästning och etablering jämfört med huvudets etablering i nominala sammansättningar, som följer *en* struktur. I anslutning till detta testar jag hypotes 3b, dvs. hypotesen om att L1-talarna i högre grad än L2-talarna följer gängse principer för huvudledets placering i adjektiviska sammansättningar, dvs. att de i högre grad undviker vänstertolkningar. I analysen diskuterar jag vad eventuella skillnader betyder och hur de kan förklaras.

Mer yt nära, semantiska mönster undersöks i några typer av NA-sammansättningar. I en grupp av sammansättningar analyserar jag tolkningarna av NA-sammansättningar som är potentiellt ambigüosa mellan en jämförelse- och en orsaksrelation (t.ex. *iskall* och *solvarm*). Graden av samstämmighet i informantgruppen som helhet ger information om graden av befästning (hypotes 2c) och sammansättningarnas etableringsgrad (hypotes 1c). De eventuella skillnaderna mellan L1- och L2-talarnas tolkningar sätts därefter i relation till frekvenserna för de enskilda sammansättningarna (hypotes 3c–d), tillgången till analogibaser (hypotes 3e) och frekvenserna för det jämförande respektive orsaksrelaterade mönstret hos NA-sammansättningar.

I en mindre grupp av NA-sammansättningar med efterledet *-rädd* analyseras valet mellan jämförelse [rädd som N] och bundet adverbial [rädd för N]. Jag undersöker vilken inverkan frekvensen för den enskilda sammansättningen respektive lågnivåsschemat har på informanternas tolkningar.

Ytterligare en kategori av sammansättningar i enkäten är adjektiviska sammansättningar med efterledet *-fri*. I detta fall analyseras valet mellan ett produktivt avledningsschema (*fri* som avledningssuffix) och ett mindre produktivt sammansättningsmönster (*fri* som självständigt led) i sammansättningar av olika frekvens.

Slutligen analyserar jag adjektiviska sammansättningar med efterledet *-färdig*, för att se om informanterna väljer tolkningen [färdig att börja

VERBA] eller [färdig med att VERBA]. Kategorin innehåller en lexikaliserad sammansättning – *gråtfärdig* – och två tillfälliga bildningar: *skrattfärdig* och *skrikfärdig*. Genom detta urval kan jag testa betydelsen av frekvens och graden av analogisk tolkning hos L1- respektive L2-talarna.

Med två undantag (sammansättningarna med *-färdig* respektive *-fri*) är alla kategorier av sammansättningar representerade både i fritext- och flervalsdelen av enkäten. Därigenom kan jag jämföra hur sammansättningstyperna hanteras med och utan kontext. Genom fritextdelen kan jag på så sätt undersöka reliabiliteten i resultaten från flervalsdelen av enkäten (se vidare avsnitt 4.12).

Jag har genomgående i enkäten valt att inkludera flera ord med samma efterled, eller samma förled, för att kunna jämföra tolkningen av strukturellt lika ord (t.ex. *polismisshandel*, *kvinnomisshandel* och *barnmisshandel*). Detta gäller även de tillfälliga sammansättningarna, vilka är konstruerade med etablerade sammansättningar som mönster: *skrattfärdig* är exempelvis bildad utifrån den etablerade sammansättningen *gråtfärdig*, och *manspräst* är bildad utifrån *kvinnopräst*. Syftet är att se i vilken grad informanterna använder sig av analogi när de tolkar okända sammansättningar (se hypotes 3e).

Jag har också valt ut icke-parallella fall, där strukturellt lika sammansättningar uppvisar olika relation mellan för- och efterled, såsom i ordparet *polismisshandel* – *kvinnomisshandel*, där den förra sammansättningen har konventionaliserats i en betydelse där förledet pekar ut verbhandlingens AGENT och den senare har konventionaliserats i en betydelse där förledet pekar ut verbhandlingens PATIENT.

I urvalet av sammansättningar till de olika kategorierna har jag valt ut både frekventa, mindre frekventa och tillfälliga sammansättningar. Detta urval är gjort för att kunna testa hypotes 1c om sambandet mellan frekvens och konventionalisering, hypotes 2c om sambandet mellan frekvens och befästning, samt hypoteserna 3c–e om sambandet mellan frekvens, befästning och L1/L2-skillnader. Sammansättningarnas frekvens har undersökts med hjälp av en referenskorpus (se vidare avsnitt 4.4).

4.3.1. Nominala sammansättningar

I detta avsnitt presenteras studiens sju kategorier av nominala sammansättningar.

4.3.1.1. Huvudled i nominala sammansättningar

Som diskuterats i avsnitt 2.4 antas svenska sammansättningar följa en universell preferens, där huvudledet i sammansättningen är placerat sist

i sammansättningen. Hur informanterna förhåller sig till detta schema testas genom tolv nominala sammansättningar i enkäten. Förklaringar som följer det gängse schemat med efterledet som sammansättningens huvud, har kompletterats med distraktorer som pekar ut förledet som sammansättningens huvudled (jfr ett *båthus* är en 'båt som man kan bo i'). I den här kategorin av sammansättningar testas hypotes 1a och 2a om huvudets status i nominala sammansättningar samt hypotes 3a om skillnader mellan informantgrupperna. Utöver tolkningen av lexikaliserade sammansättningar analyseras hur informanterna tolkar tillfälliga sammansättningar (t.ex. *vagnbarn*), där förståelsen är mer beroende av en strukturell analys. I fritextdelen av enkäten möter informanterna sammansättningarna *vagnbarn*, *husbil* och *skalpotatis*. I flervaldsdelen möter de nio sammansättningar som är etablerade i olika grad (se tabell 4:1). Som framgår är merparten av dem NN-sammansättningar, men *spelkort* skulle kunna betraktas som en VN-sammansättning.

TABELL 4:1. *Undersökta sammansättningar i kategorin huvudled i nominala sammansättningar.*

Fritextdel	Flervaldsdel
vagnbarn	vagnbarn stövelbarn
husbil	husbåt båthus båtsommar
skalpotatis	burköl boktext spelkort flodstrand

Sammansättningen *vagnbarn* finns representerad både i fritext- och flervaldsdelen av enkäten, eftersom jag vill undersöka om informanternas svar är konsistenta, eller om de påverkas av tillgången till kontext. Utöver förmågan att identifiera huvudledet vill jag undersöka hur informanterna närmare förklarar relationen mellan för- och efterled, t.ex. om *vagnbarn* förklaras med PLATS 'ett barn som sitter i vagn' eller med PREFERENS 'ett barn som gillar att åka vagn' eller med ytterligare andra alternativ.

4.3.1.2. Huvudled i treledade sammansättningar

I sammansättningar där förledet utgörs av en sammansättning och denna slutar med en konsonant, är det brukligt att skjuta in ett fogeelement före det sista ledet, dvs. före sammansättningens huvudled (jfr *vinbär-s-buske*). Denna fogeregeln kan komma att markera en semantisk skillnad mellan treledade sammansättningar med och utan foge-s (se avsnitt 2.4.1). Hur informanterna förhåller sig till fogeregeln undersöks genom fyra sammansättningar i enkäten, som framgår av tabell 4:2.

TABELL 4:2. *Undersökta sammansättningar i kategorin huvudled i treledade nominala sammansättningar.*

Fritextdel	Flervalsdel
barnbokhylla	barnbokhylla barnbokshylla
	skolboksförsäljning
	sommardagdröm

Barnbokhylla utan foge-s och *barnbokshylla* med foge-s är ett sammansättningspar av två lexikaliserade men lågfrekventa sammansättningar. Hur informanterna segmenterar och tolkar *barnbokhylla* undersöks både genom fritextdelen och flervalsdelen. I flervalsdelen ingår två ännu mer lågfrekventa, tillfälliga bildningar, nämligen *skolboksförsäljning* och *sommardagdröm* (vilka har sina tänkta motsvarigheter i *skolboksförsäljning* och *sommardagsdröm*).

4.3.1.3. Semantisk relation i sammansättningar som betecknar behållare

Som beskrivits i avsnitt 3.1.2 antar man inom kognitiv semantik att betydelse uppstår genom konceptualisering. Hur informanterna konceptualiserar potentiellt ambigüosa sammansättningar som betecknar behållare av olika slag undersöks genom tre sammansättningar med efterledet *-korg* och två sammansättningar med efterledet *-låda* i flervalsdelen av enkäten (se tabell 4:3). Informanterna får i enkäten välja mellan konkurrerande mönster såsom INNEHÅLL, ÄNDAMÅL och MATERIAL. I fritextdelen av enkäten möter informanterna sammansättningarna *glaslåda* och *blomkruka*.

TABELL 4:3. *Undersökta sammansättningar i kategorin semantisk relation i sammansättningar som betecknar behållare.*

Fritextdel	Flervalsdel
glaslåda	plastlåda blomlåda
blomkruka	svampkorg fruktkorg olivkorg

Urvalet av ord är gjort med tanke på att kunna jämföra frekventa sammansättningar med mindre frekventa. De frekventa sammansättningarna i tabell 4:3 förväntas ge mer samstämmiga tolkningar än de tillfälliga bildningarna (*olivkorg*, *glaslåda*), där flera alternativ i meningspotentialen kan antas vara öppna för såväl L1- som L2-talare av svenska.

4.3.1.4. Jämförande egenskapsprojicering eller tematisk relationslänkning

Två övergripande tolkningsstrategier som identifierats för nominala sammansättningar, och som beskrivs i avsnitt 2.3, är jämförande egenskapsprojicering och tematisk relationslänkning. Hur informanternas tolkningar fördelar sig mellan dessa testas genom två sammansättningar i fritextdelen och sex sammansättningar i flervalsdelen av enkäten (se tabell 4:4)

TABELL 4:4. *Undersökta sammansättningar i kategorin jämförande egenskapsprojicering eller tematisk relationslänkning.*

Fritextdel	Flervalsdel
zebraväska	zebralejon
issaft	isvatten citronvatten
	skräpkruka
	apörn korpörn

I enkäten förekommer *zebra* som förled i de tillfälliga bildningarna *zebraväska* (i fritextdelen) och *zebralejon* (i flervalsdelen). Informanterna möter också den tillfälliga bildningen *issaft* i fritextdelen av enkäten. I flervalsdelen av enkäten förekommer sammansättningar som är etablerade i olika grad, bl.a. *apörn* som är en belagd men mindre frekvent sammansättning och *corpörn* som är en tillfällig bildning som jag lånat från Svanlund (2013).

Även i denna kategori undersöks betydelsen av etablering och frekvens. När det gäller de tillfälliga bildningarna vill jag undersöka om det föreligger skillnader mellan L1- och L2-talarna med avseende på preferenser för det ena eller det andra tolkningsmönstret.

4.3.1.5. *Identitet/jämförande egenskapstolkning eller tematisk relationslänkning i personbetecknande sammansättningar*

Till den här kategorin har jag valt ut etablerade och tillfälliga sammansättningar där ett av sammansättningens led, eller båda, betecknar en persons genus. I enkäten ställs identitetstolkning och jämförande egenskapstolkning mot tematisk relationslänkning, dvs. någon annan typ av relation mellan leden. Jag har valt att behandla identitetstolkning och jämförande egenskapstolkning tillsammans i analysen, eftersom typerna påminner om varandra: i det här fallet markerar de att en person antingen *är* eller *är som* den referent som förledet pekar ut – i kontrast till en tematisk relationslänkning där ledens referenter intar olika funktionella roller gentemot varandra. Den tillfälliga sammansättningen *manspräst* kan exempelvis tolkas som 'man som är präst' (identitetstolkning) eller en 'präst som tar emot och hjälper män' (relationslänkning). Flera av sammansättningarna rör genus i relation till profession. I några av sammansättningarna betecknar båda leden genus, som framgår av sammanställningen i tabell 4:5.

TABELL 4:5. *Undersökta sammansättningar i kategorin identitet/jämförande egenskapstolkning eller tematisk relationslänkning i personbetecknande sammansättningar.*

Fritextdel	Flervalsdel
kvinnoläkare	mansläkare
	kvinnopräst
	manspräst
	herrfrisör
	flickpojke
kvinnoman	kvinnokarl

För sammansättningarna i tabell 4:5 ingår både identitets-/egenskapstolkning och relationslänkning i meningspotentialen. Frågan jag vill undersöka är vilket av dessa mönster som är det prefererade i informanternas tolkningar, och vilken betydelse frekvens har i sammanhanget.

4.3.1.6. *Agens vid verbavledning*

Enkäten innehåller nominala verbavledningar där sammansättningens semantiska roller – AGENT och PATIENT – inte är a priori förutsägbara. I fritextdelen förekommer *barnmisshandel* och *politikerförakt*. I flervalsdelen ingår också sammansättningar med *-misshandel* och *-förakt* som efterled samt sammansättningen *kvinnovård* (se tabell 4:6).

TABELL 4:6. *Undersökta sammansättningar i kategorin agens vid verbavledning.*

Fritextdel	Flervalsdel
barnmisshandel	polismisshandel kvinnomisshandel mansmisshandel
politikerförakt	studentförakt kvinnovård

Med sammansättningarna på *-misshandel* avser jag att testa hur informanterna förhåller sig till till synes jämförbara sammansättningar med olika semantisk struktur. *Kvinnomisshandel* och *barnmisshandel* är konventionaliserade med förledets referent som verbhandlingens PATIENT, medan *polismisshandel* är konventionaliserad med förledets referent som verbhandlingens AGENT.

Den tillfälliga bildningen *studentförakt* kan inte antas finnas befäst i det mentala lexikonet hos vare sig L1- eller L2-talarna. I detta fall vill jag undersöka i vilken mån informantgrupperna använder sig av analogi som tolkningsprincip och tolkar *studentförakt* i analogi med den etablerade sammansättningen *politikerförakt* (se hypotes 3e).

4.3.1.7. *Betydelsesammansmälta AN-sammansättningar*

Bakkens term *betydelsesammansmältning* refererar, som diskuteras i avsnitt 3.2.2, till sammansättningar som har lexikaliserats i en snävare, mer specifik betydelse än vad som kan förutses utifrån de ingående leden.

När det gäller den här typen av sammansättningar är det intressant att undersöka hur informanterna hanterar ordens icke-kompositionalitet. I fritextdelen av enkäten tolkar informanterna sammansättningen *grön-område*. I flervalsdelen möter de fyra etablerade sammansättningar, som framgår av sammanställningen i tabell 4:7.

TABELL 4:7. *Undersökta sammansättningar i kategorin betydelssammansmälta AN-sammansättningar.*

Fritextdel	Flervalsdel
grönområde	grönyta
	småkakor
	sötsaker
	finkläder

Betydelsesammansmälta sammansättningar som de i tabell 4:7 kan inte till fullo förstås enbart utifrån de ingående leden, utan måste läras in som helheter. Jag vill undersöka om L1-talarna är mer samstämmiga i sina svar än L2-talarna vid tolkningen av denna typ av sammansättningar (hypotes 3c–d), som en följd av skillnader med avseende på befästning.

4.3.2. Adjektiviska sammansättningar

I detta avsnitt beskrivs de fem kategorier av adjektiviska sammansättningar som ingår i studien.

4.3.2.1. Huvudled i AA-sammansättningar

Hur informanterna identifierar huvudledet i adjektiviska AA-sammansättningar testas genom åtta sammansättningar (se tabell 4:8). Ordet *fulsnygg* finns representerat både i fritext- och flervalsdelen, eftersom jag vill undersöka kontextens inverkan på tolkningen.

AA-sammansättningarna är lexikaliserade enligt två skilda mönster (determinativt respektive kopulativt; se avsnitt 1.2). Denna strukturella ambiguitet gäller i hög grad de sammansättningar som anger färg. *Svartvit* följer det kopulativa mönstret 'svart och vit', medan *gulvit* enligt traditionella lexikologiska principer skulle räknas som en determinativ sammansättning 'vit med en nyans av gult'. Genom den här kategorin av sammansättningar testas hypotes 1b, dvs. hypotesen om att huvudledets

placering sist i determinativa sammansättningar är en mindre etablerad struktur för adjektiviska än för nominala sammansättningar, vilket skulle kunna yttra sig i mindre samstämmiga tolkningar (i enlighet med hypotes 2b). Jag undersöker vidare om strukturell kunskap om huvudets placering i sammansättningar ger en lägre grad av vänstertolkningar hos L1-talarna än hos L2-talarna (se hypotes 3b).

TABELL 4:8. *Undersökta sammansättningar i kategorin huvudled i AA-sammansättningar.*

Fritextdel	Flervalsdel
fulsnygg	fulsnygg
dumsnäll	sötsur
	gulvit
	svartvit
	vitgul
	gröngul
	blågul
	gulgrön

4.3.2.2. Jämförelse eller orsak i NA-sammansättningar

Ytterligare en kategori av sammansättningar som ingår i enkäten är adjektiviska NA-sammansättningar som skulle kunna tolkas jämförande eller med en orsaksrelation mellan leden. Kategorin är representerad av flera sammansättningar som rör perception, alltifrån lexikaliserade ord som *kryddstark* och *iskall* till tillfälliga bildningar som *lökstark* och *glasskall*. En översikt över sammansättningarna i denna kategori ges i tabell 4:9.

Urvalet av ambigüosa NA-sammansättningar är gjort med syfte att kunna jämföra sammansättningar av olika frekvens. Avsikten är att sätta sammansättningarnas frekvens i relation till graden av befästning (hypotes 2c) och konventionalisering (hypotes 1c), samt att undersöka skillnader i tolkning mellan de informanter som har svenska som förstaspråk och de informanter som har svenska som andraspråk (hypotes 3c–e). I flera fall har jag låtit samma efterled representeras av en etablerad och en mindre etablerad sammansättning (t.ex. *stenhård* och *berghård*). Syftet med detta är att kunna analysera i vilken mån informanterna använder sig av analogi och tolkar nya sammansättningar i analogi med befästa.

TABELL 4:9. *Undersökta sammansättningar i kategorin jämförelse eller orsak i NA-sammansättningar.*

Fritextdel	Flervalsdel
	kryddstark lökstark
honungssött	honungssöt sockersöt
	stenhård berghård
sandmjukt	snömjukt
	iskall glasskall snökall
	solvarm spisvarm
blodrött	
dammgrått	
	solgul
	snövit

4.3.2.3. *Sammansättningar med -rädd*

I sammansättningar med strukturen [N-rädd] betecknar förledet vanligen den företeelse man är rädd för, som i *hundrädd*. Det finns emellertid en annan fullt möjlig tolkning, som uttrycker en jämförande relation, 'rädd som en N'. I enkäten undersöks informanternas hantering av denna potentiella ambiguitet genom en sammansättning i fritextdelen (*musrädd*) och två sammansättningar i flervalsdelen (*fågelrädd* och *hundrädd*), som framgår av tabell 4:10.

TABELL 4:10. *Undersökta sammansättningar i kategorin NA-sammansättningar med -rädd.*

Fritextdel	Flervalsdel
musrädd	fågelrädd hundrädd

4.3.2.4. *Sammansättningar med -fri*

I sammansättningar med *-fri* kan efterledet representera det självständiga ordet *fri*, eller fungera mer som ett avledningssuffix, med betydelsen 'fri från', dvs. 'utan'. Det förra bildningsmönstret är sällsynt, och representeras av *valfri* 'fri att väljas' i flervalsdelen av enkäten. Bildningsmönstret med avledningssuffix är däremot produktivt och representeras i enkäten av två lexikaliserade sammansättningar i flervalsdelen, nämligen *helgfri* och *bilfri* (se tabell 4:11). I samtliga fall råder potentiell ambiguitet: *valfri* skulle kunna betyda 'fri från val', och *helgfri* respektive *bilfri* skulle kunna betyda 'fri på helgen', 'fri på grund av bil'. Jag vill undersöka hur informanterna förhåller sig till denna ambiguitet.

TABELL 4:11. *Undersökta sammansättningar i kategorin sammansättningar med -fri.*

Fritextdel	Flervalsdel
	valfri
	helgfri
	bilfri

4.3.2.5. *Sammansättningar med -färdig*

I sammansättningar med *-färdig* kan *-färdig* betyda 'färdig, klar (med)' som i *halvfärdig*, men vanligen (och mindre genomskinligt) betyder efterledet 'färdig att börja'. I enkäten ingår tre sammansättningar av den senare typen, samtliga i flervalsdelen. En av dem är lexikaliserad (*gråtfärdig*) och två är tillfälliga bildningar (*skratzfärdig* och *skrikfärdig*) (se tabell 4:12). I den här kategorin av sammansättningar vill jag undersöka i vilken mån L1- respektive L2-talarna använder sig av analogi som tolkningsstrategi och tolkar de tillfälliga bildningarna i analogi med den lexikaliserade sammansättningen *gråtfärdig*. Frågan jag ställer är om L1-talarna är mer samstämmiga i sina tolkningar än L2-talarna på grund av en större kunskap om analogibasen *gråtfärdig* (se hypotes 3e).

TABELL 4:12. *Undersökta sammansättningar i kategorin sammansättningar med -färdig.*

Fritextdel	Flervalsdel
	gråtfärdig
	skratzfärdig
	skrikfärdig

4.4. Referenskorpus

För att kunna relatera informanternas svar till sammansättningarnas frekvens (och indirekt till förväntad språklig input), har jag sammanställt en referenskorpus. Korpusfrekvenser är visserligen ett trubbigt verktyg för att komma åt information om språklig input hos språkbrukare, eftersom de ger information om språket, inte om språkbrukarna. Det som korpusen kan bidra med är ett mått på ordets frekvens i den typ av språk som korpusen innehåller. Om korpusen är representativ för det språk som språkbrukarna möter är det statistiskt sett mer sannolikt att frekventa ord i korpusen är bekanta för språkbrukarna, än mindre frekventa ord, men korpusrepresentationen säger ingenting om vilken språklig input som den enskilde språkbrukaren fått.

Till referenskorpusen har jag valt ut fem av Språkbankens korpusar, som är tänkta att representera sådana texttyper som ungdomarna i min enkätstudie kan förväntas möta (Språkbanken/Korp; se även Borin et al. 2012). Modernt tidningspråk representeras av korpusarna Göteborgs-Posten (GP) 2013 och helgmagasinet *Två Dagar* (bilaga till GP). Relativt modern skönlitteratur representeras av korpusen SUC-romaner med 58 svenska romaner från tidigt 1990-tal samt korpusen Norstedts-romaner med 23 svenska romaner från 1999. Mer vardagsnära språk är rikligt representerat, i form av bloggspråk från *Bloggmix* 2014. Av praktiska skäl har jag däremot inte använt mig av någon talspråkskorpus. Utöver dessa fem korpusar, som finns tillgängliga i Korp (se Språkbanken/Korp), innehåller mitt referensmaterial en korpus med läroböcker från högstadiet och gymnasiet. Lärobokskorpusen har tidigare använts som primärmaterial för en studie om ämnesspråken i skolan (Ribeck 2015).³ Korpusen innehåller dels material som samlats in inom ramen för projektet *Ord i läromedel*, *OrdiL* (Lindberg & Johansson Kokkinakis 2007), dels nyare material insamlat av Ribeck (se Ribeck 2015). *OrdiL*-korpusen omfattar cirka en miljon token och består av två läroböcker vardera från ämnena biologi, fysik, kemi, matematik, geografi, historia, religion och samhällskunskap, samtliga från högstadiet. Ribeck (2015) har kompletterat denna med fler och nyare läroböcker för högstadiet och gymnasiet. Tillsammans med materialet från *OrdiL* omfattar lärobokskorpusen 31 läroböcker från högstadiet och 32 läroböcker från gymnasiet, motsvarande drygt 5 miljoner token (se Ribeck 2015:appendix för en sammanställning av lärobokskorpusens innehåll). De fem korpusarna i Språkbanken och lärobokskorpusen (Ribeck 2015)

3 Stort tack till Judy Ribeck Nyström, fil. dr i språkvetenskaplig databehandling, som tillgängliggjort och hjälpt mig med sökningar i lärobokskorpusen.

behandlas tillsammans i studien, såsom en referenskorpus. Delkorpusarnas storlek framgår av tabell 4:13.⁴

TABELL 4:13. *Fördelning mellan studerade korpusar.*

Korpus	Token	Andel
GP 2013	16 872 043	24,0 %
Två Dagar	1 033 747	1,5 %
Bloggmix 2014	40 133 589	57,0 %
Norstedtsromaner	2 533 209	3,6 %
SUC-romaner	4 651 200	6,6 %
Lärobokskorpus	5 221 788	7,4 %
Totalt	70 445 576	

Man kan naturligtvis diskutera fördelningen mellan delkorpusarna, och fråga sig hur väl min korpus motsvarar ungdomarnas språkliga input – är det t.ex. adekvat att ge Bloggmix så stort utrymme? Jag har dock bedömt det som rimligt att ge det vardagsnära språket en stark representation, eftersom sammansättningar i hög grad uppstår och utvecklas i vardags-språket.

När jag i det följande anger med vilken frekvens enskilda sammansättningar från enkäten är representerade i referenskorpusen avses frekvensen för sammansättningen inklusive alla dess böjningsformer. Även förekomster där sammansättningen återfinns i en större sammansättning (såsom *båthus* i *båthusbygge*) är medräknade. En sammanställning över de enskilda sammansättningarnas absoluta och relativa tokenfrekvenser ges i bilaga 3.

4.5. Enkätens orddel

Som nämnts i föregående avsnitt är mitt undersökningsobjekt kompositionellt uppbyggda sammansättningar som ändå inte är självklart genomskinliga eller analyserbara. I avsnitt 4.1 redovisades urvalskriterierna för sammansättningarna, vilka upprepas här:

⁴ Sökningarna utfördes i Språkbankens korpusar med hjälp av sökverktyget Korp i juni 2017.

- de kan förstås utan metaforisk tolkning
- helhetens betydelse är potentiellt kompositionell
- de är inte självklart genomskinliga eller analyserbara, på grund av relationell polysemi och (i vissa fall) på grund av polysemi hos de ingående leden

Enkäten innehåller totalt 50 nominala sammansättningar och 35 adjektiviska sammansättningar. Sammansättningarna är etablerade i olika grad. Del 1 är en fritextdel där informanterna själva formulerar förklaringen till var och en av 20 sammansättningar, genom att fullfölja en parafras (se exempel 4:1). Del 2 är en flervalsdel med 69 testord, där informanterna väljer den definition, av två eller tre alternativ, som de tycker bäst speglar testordets betydelse. Definitionerna följs av en eller flera kontextmeningar, som framgår av exempel (4:2). Enkäten finns tillgänglig i sin helhet i bilaga 1.

(4:1) En *zebraväska* är _____

- barn som sitter i vagn

T.ex. ”Barnen är fortfarande små. De är vagnbarn båda två.”

(4:2) Vagnbarn

- vagn som man har det lilla barnet i

T.ex. ”Det är dyrt att köpa vagnbarn.”

Genom fritextdelen kommer jag åt förklaringar som inte är styrda av någon erbjuden kontext från min sida. Det primära syftet med denna del är att kontrollera reliabiliteten i resultatet från flervalsdelen. I flervalsdelen av enkäten väljer informanterna mellan färdiga svarsalternativ med kontextmeningar. Denna begränsning i djup innebär samtidigt en vinst vad gäller bredd. Med fasta tolkningsalternativ kan fler sammansättningar och sammansättningskategorier testas. Informanternas svar i denna del analyseras med statistiska metoder i en kvantitativ analys.

I den pilotstudie som genomfördes hösten 2013 testade jag olika definitionsformat för flervalsdelen av enkäten, bl.a. COBUILD (se Svensén 2004:290–291), där ordet som ska förklaras ingår i definitionen enligt mönstret i exempel (4:3).

(4:3) En kvinnopräst är en präst som är kvinna

Efter utvärdering av pilotstudien beslutade jag mig för att använda ett annat definitionsformat som jag också testat i pilotstudien, nämligen en lättillgänglig förklaring eller parafras följt av en eller flera illustrerande

kontextmeningar (se exempel 4:2). Jag har bedömt det som angeläget att använda ett enkelt språk i förklaringarna, så att de är tillgängliga även för andraspråkstalare med en kortare exponeringstid för svenskan, jämför exempel (4:4) där substantivet *polismisshandel* förklaras med en omskrivning.

- (4:4) Polismisshandel
- När polisen slår eller misshandlar en person
T.ex. "När polisen gripit mannen slog och sparkade de honom. De blev senare åtalade för polismisshandel."
 - När polisen blir misshandlad
T.ex. "Killarna slog och sparkade polismannen. De blev senare åtalade för polismisshandel."

Vid sidan av de lättillgängliga parafraserna har kontextmeningarna en viktig funktion i enkäten genom att ge orden ett sammanhang. Enligt Svanlund (2009:28) är det annars ett problem inom sammansättningsforskningen att testorden i informantundersökningarna normalt saknar både sammanhang och språklig kontext. I vissa fall i enkäten är det dessutom kontextmeningen som tydligt visar om en parafras fungerar eller inte, som framgår av exempel (4:5).

- (4:5) Finkläder
- Kläder som är fina
T.ex. "Hon hade köpt finkläder: ett par mjukisbyxor och tre par sockar."
 - Kläder som passar vid finare, högtidligare tillfällen
T.ex. "Hon hade köpt finkläder: en långklänning och en kavaj."

I detta exempel illustrerar den första kontextmeningen att parafrasen 'kläder som är fina' inte täcker in alla nödvändiga semantiska komponenter hos *finkläder*.

4.6. Enkätens språkliga bakgrundsfrågor

Enkäten avslutas med 30 bakgrundsfrågor. Förutom allmänna uppgifter om ålder, kön, gymnasieprogram och årskurs, efterfrågas här information om informanternas språk och språkanvändning.

Bakgrundsdelens inleds med ikryssningsfrågor om informantens kön (med svarsalternativen tjej/kille/transperson) och födelseår. I fråga 3 får informanten ange om hen är född i Sverige eller inte, eftersom den frågan bidrar till bilden av informantens språkliga bakgrund. För de informanter som uppger att de inte är födda i Sverige följer en fråga om åldern vid ankomsten till Sverige. Fråga 5 är en allmänt hållen fråga som inleder språkdelens och som efterfrågar vilket/vilka språk informanten kan tala. Därefter följer en fråga om informantens modersmål (se exempel 4:6).

(4:6) Vilket/vilka språk lärde du dig tala först (ditt/dina modersmål)?

I fråga 7 kryssar informanten för vid vilken ålder hen började lära sig svenska, och i fråga 8 efterfrågas platsen för denna tidiga inlärning (se exempel 4:7).

(4:7) Var började du först
lära dig svenska?

- Hemma
- På dagis/förskola
- Hos dagmamma
- I skolan
- Med kompisar
- Annat, nämligen _____

Därefter ställs frågan om informanten under sin skoltid läst svenska som andraspråk (fråga 9), med följdfrågorna 10–13 om huruvida informanten gjort detta under lågstadiet, mellanstadiet, högstadiet respektive gymnasiet. I fråga 14 efterfrågas om informanten under sin skoltid läst ämnet modersmål/hemspråk, följt av fråga 15–18 om huruvida informanten gjort detta under lågstadiet, mellanstadiet, högstadiet respektive gymnasiet. Fråga 19 handlar om vilken typ av svenskundervisning informanten deltar i för närvarande (svenska, svenska som andraspråk, ingen). Frågorna 20–21 rör gymnasieprogram och årskurs. I fråga 22–23 efterfrågas om mamma respektive pappa är född i Sverige eller inte. Fråga 24 rör föräldrarnas modersmål (se exempel 4:8).

(4:8) Föräldrars
modersmål

- Mina föräldrar har svenska som modersmål
- En av mina föräldrar har svenska som modersmål
- Mina föräldrar har inte svenska som modersmål

Kommentar _____

I fråga 25–26 efterfrågas vilket/vilka modersmål mamma respektive pappa har. Frågorna 27–29 handlar om vilka språk informanten talar hemma i familjen, med vänner respektive med eventuell partner. Under punkt 30 ges möjlighet att lämna övriga kommentarer.

4.7. Etiska överväganden: information, samtycke och anonymitet

För att undvika hantering av känsliga personuppgifter valde jag att låta deltagandet i enkäten vara anonymt, dvs. inga personuppgifter – såsom namn eller personnummer – samlades in. På enkätens första blad presenteras en formulering om samtycke till deltagande, där det framgår att deltagandet i enkäten är frivilligt och anonymt men att forskningsresultat från enkätstudien är avsedda att användas för publicering och för presentation i vetenskapliga och andra publika sammanhang (se bilaga 1). Denna information presenterades även i den muntliga introduktion som jag genomförde i samtliga klasser och undervisningsgrupper.

4.8. Instruktion till och genomförande av enkäten

I mötet med eleverna inledde jag med en muntlig genomgång, innan enkäten delades ut. För att alla informanter i studien skulle få samma information valde jag att genomföra denna enligt ett på förhand skrivet manus. Detta var ett sätt att uppfylla kravet på reliabilitet.

Introduktionen som tog cirka tio minuter innehöll en övergripande beskrivning av institutionens ordforskning och ordboksarbete. Jag berättade att mitt avhandlingsarbete befinner sig inom denna ordforskningstradition och handlar om ”strategier vid tolkning av ord”, samt betonade att enkäten inte är ett test, utan syftar till att belysa hur (olika) vi tolkar ord. Jag informerade om frivilligheten i att delta i studien och framhöll att enkätsvaren skulle behandlas anonymt. Samtidigt upplystes informanterna om att de genom att lämna in enkäten godkänner att jag använder forskningsresultaten i min avhandling, i tidskrifter och i presentationer i olika sammanhang. Slutligen gavs information om själva genomförandet av enkäten.

Den skriftliga instruktionen till fritextdelen har medvetet hållits kortfattad, eftersom jag ville undvika att styra informanternas förklaringar.

I den muntliga instruktionen gav jag emellertid tre exempel på tavlan på hur man skulle kunna fullfölja parafaserna (se exempel 4:9). Med dessa exempel ville jag täcka in variationer i fritextdelen vad gäller ordklass och parafasformat.

- (4:9) En *badboll* är _____
 En *morgontrött* person är _____
Värlängtan är när _____

För exemplet *badboll* angav jag (och fyllde i på tavlan) att en badboll skulle kunna förklaras vara 'en boll som man kan leka med när man badar' eller 'en boll som man kan ha i vattnet'. Jag bad informanterna att förklara ordet så som de skulle förklara det för någon som inte känner till ordet och sa att det inte räcker med att säga att det är 'en sorts boll'. Jag framhöll att det inte är konstigt om de inte skulle känna igen alla ord, eftersom enkäten innehåller både vanliga och ovanliga ord. Informanterna uppmanades att förklara så gott de kunde och att gissa om de kände sig osäkra på betydelsen.

Därefter instruerade jag informanterna om att först svara på del 1, dvs. fritextdelen (vilken utgjordes av ett separat pappersark), och att vända ner papperet för insamling innan de gick vidare till del 2–3 (som utgjordes av ett häfte). Anledningen till detta var att jag ville undvika att informanterna skulle påverkas av kontexterna i del 2 och gå tillbaka och ändra svar i det öppna formatet i del 1. Detta var särskilt viktigt med tanke på att några ord ingår i både del 1 och del 2. För att kompensera för en potentiell utmattningseffekt distribuerades fyra versioner av enkäten där orden förekommer i olika, framslumpad ordning. Alla versioner distribuerades med jämn fördelning i samtliga klasser. Varje utdelad enkät var märkt med en kod som kopplade samman del 1 med del 2–3, som gav möjlighet att koppla ihop enkäterna med rätt skola. Kodens första två bokstäver anger enkätversion – BC, BF, KV och PS – där bokstäverna står för begynnelsebokstaven i det första ordet i del 1 respektive 2. Bokstavskombinationen följs av en siffra mellan 1 och 6 som anger från vilken skola enkäten samlats in, följt av en löpande numrering för informanterna i respektive klass.

Jag visade på tavlan hur flervalsdelen i del 2 ser ut och gav instruktionen att informanterna endast skulle kryssa i *ett* alternativ, även om de skulle tycka att flera alternativ var möjliga, samt informerade om att både förklaringen och kontextmeningen skulle stämma. Jag uppmanade informanterna att svara på orden i tur och ordning. Språkbakgrundsfrågorna i del 3 presenterades som "frågor om vilka språk ni kan och använder". Angående startålder för svenskan kommenterade jag att språkbakgrunden

innehåller ”frågor om hur gamla ni var när ni började lära er svenska”. Jag beskrev att jag tänker att man börjar lära sig språket när man hör det, och fortsatte (i princip ordagrant):

Om man har haft svenska i familjen från början är man alltså noll år när man börjar lära sig svenska. Men om föräldrarna talat andra språk än svenska med en så kanske det var först på dagis eller i skolan man började lära sig svenska och då fyller man i hur gammal man var då.

Jag avslutade med att öppna för frågor, och uppmanade informanterna att be mig om hjälp om de fick problem med att tolka enkätens bakgrundsfrågor.

Den erhållna tiden för att fylla i enkäten varierade något mellan klasserna, beroende på hur klassernas schema såg ut. Den muntliga instruktionen tog tio minuter, och därefter återstod 50–60 minuter för att svara på enkäten (60 minuter i samtliga svenska som andraspråksgrupper), förutom i en klass (med stor övervikt av L1-elever), där schemat endast tillät 45 minuter. Min bedömning är att det i samtliga grupper fanns gott om tid för informanterna att fylla i enkäten, vilket avspeglas i i huvudsak tidigt inlämnade och fullständigt ifyllda enkäter.

I många klasser jag besökte var frånvaron relativt hög. Den genomsnittliga frånvaron för samtliga klasser var drygt 25 %. Med något undantag var dock frånvaron vid mitt besök inte högre än normalt, enligt lärarna. Av de elever jag mötte var det endast en person som valde att inte delta i studien.

4.9. Skolor och gymnasieprogram

Enkäten har samlats in på sex gymnasieskolor fördelade på olika typer av skolor: friskolor respektive kommunala skolor, innerstadsskolor respektive förortsskolor på större ort, samt skolor på mindre ort. Databasinsamlingen skedde under vårterminen 2015 och höstterminen 2015. Vid tidpunkten för databasinsamlingen gick informanterna i årskurs 1, 2 eller 3 på något av de sex studieförberedande gymnasieprogrammen, dvs. samhällsvetenskapsprogrammet (SAM), ekonomiprogrammet (EK), humanistiska programmet (HUM), estetiska programmet (ES), naturvetenskapsprogrammet (NA) eller teknikprogrammet (TE).

I min analys studerar jag mönster i informanternas sammansättnings- tolkningar i relation till variabeln språklig bakgrund (svenska som L1 eller L2). Informanternas språkliga bakgrund kartläggs genom ett antal frågor i enkäten (se avsnitt 4.6). För att kunna fokusera kartläggningen av språklig bakgrund har jag valt att inte efterfråga uppgifter om socioeko-

nomiska förhållanden i enkäten. I valet av skolor och gymnasieprogram har dock sociala aspekter beaktats. Avgränsningen till studieförberedande program innebär att informanterna har en mer likartad bakgrund än om samtliga gymnasieprogram hade inkluderats. På de studieförberedande programmen läsåret 2015/2016 hade cirka två tredjedelar av eleverna högutbildade föräldrar, vilket kan jämföras med cirka en tredjedel av eleverna på de yrkesförberedande programmen (Skolverket 2015). Snittet i avgångsbetyg var dessutom högre för de studieförberedande programmen än för de yrkesförberedande programmen (Skolverket 2016). Genom min avgränsning har jag säkerställt en konstant; studierna på ett studieförberedande program är en gemensam nämnare för samtliga informanter oavsett språklig bakgrund (L1 eller L2), och jag har därmed eliminerat en variabel som skulle kunna påverka informanternas enkätresultat.

I urvalet av skolor har jag strävat efter en god representation av både modersmålstalare och andraspråkstalare av svenska i samtliga områden och skolor. Som forskare är det emellertid inte möjligt att göra något åt de demografiska förhållanden som gör att jag mött en betydligt större andel flerspråkiga informanter i storstadens förortsskolor än i dels storstadens innerstadsskolor, dels skolorna på mindre ort, medan förhållandet varit det omvända för de informanter som har svenska som sitt enda L1. Likväl har jag mött en betydande andel elever med flerspråkig bakgrund i samtliga grupper, vilket framgår i den följande karakteristiken av skolorna som sammanfattas i tabell 4:14 på nästa sida. Tabellen ger en översikt över de skoltyper, gymnasieprogram och klasser som ingår i studien.

Skola 1 är en liten friskola med inriktning mot teknik och design på en mindre ort. Andelen flerspråkiga elever på skolan är liten. I de två klasser jag besökte (tekniskt och estetiskt program) hade 4 av 33 informanter (eller 12 %) en flerspråkig bakgrund (dvs. de uppger att de har ett annat modersmål än svenska eller att de har svenska och ytterligare ett eller flera modersmål).

Skola 2 är en kommunal skola på en ort som ligger i en kranskommun till en större stad. Orten ligger nära denna större stad och fungerar i vissa avseenden som förort till storstaden. Andelen flerspråkiga elever är relativt stor. I den klass jag besökte (ekonomiprogrammet) hade 7 av 16 informanter (dvs. 44 %) en flerspråkig bakgrund.

Skola 3 är en kommunal skola på en mindre ort i en kranskommun till en större stad. Orten har inte samma närhet till storstaden som skola 2, och saknar därmed den förortskaraktär som präglar skola 2. Andelen flerspråkiga elever på skola 3 är relativt liten, men jag besökte tre undervisningsgrupper i ämnet svenska som andraspråk (med elevrepresentation från naturvetenskapsprogrammet, teknikprogrammet, samhällsprogram-

met och ekonomiprogrammet), så alla 15 informanter jag träffade hade en flerspråkig bakgrund.

Skola 4 är en friskola i innerstaden i en större stad. Andelen flerspråkiga elever är relativt liten. Jag besökte två klasser (naturvetenskapsprogrammet och humanistiska programmet) och en undervisningsgrupp i engelska med representation från teknikprogrammet och det estetiska programmet. Av de 40 elever jag mötte hade 8 informanter (dvs. 20 %) en flerspråkig bakgrund.

Skola 5 är en kommunal skola i innerstaden i en större stad. I den klass jag besökte (sambhällsvetenskapsprogrammet) hade 9 av 31 informanter (dvs. 29 %) en flerspråkig bakgrund.

Skola 6 är en kommunal skola i förorten till en större stad. I de fem klasser jag besökte (fyra klasser på samhällsvetenskapsprogrammet och en klass på naturvetenskapsprogrammet) hade 62 av 67 informanter (dvs. 92,5 %) en flerspråkig bakgrund.

TABELL 4:14. *Beskrivning av ingående skolor och klasser i enkätstudien.*

	Typ av skola	Besökta gymnasieprogram	Årskurs	Antal besökta klasser / antal informanter	Antal (andel) flerspråkiga informanter i de besökta klasserna
Skola 1	Friskola på mindre ort	TE, ES	3	2/33	4 (12 %)
Skola 2	Kommunal skola på ort i kranskommun till större stad	EK	3	1/16	7 (44 %)
Skola 3	Kommunal skola på ort i kranskommun till större stad	Integrerade grupper från NA, TE, EK, SAM	1–3	3/15	15 (100 %)
Skola 4	Friskola i större stad, innerstad	NA, HUM, Integrerad grupp från TE/ES	3	3/40	8 (20 %)
Skola 5	Kommunal skola i större stad, innerstad	SAM	1	1/31	9 (29 %)
Skola 6	Kommunal skola i förort till större stad	SAM (4 klasser), NA	2–3	5/67	62 (92,5 %)

4.10. Informanterna och deras språkliga bakgrund

Studien omfattar 202 informanter på sex gymnasieskolor. Knappt hälften av informanterna har svenska som sitt enda L1 (n=94), och drygt hälften av dem (n=105) har en flerspråkig bakgrund. Tre informanter har inte uppgett några bakgrundsuppgifter om modersmål. I det följande redogör jag för studiens informantgrupper, samt beskriver bakomliggande kategoriseringsprinciper och diskuterar metodologiska svårigheter vad gäller gränsdragning.

4.10.1. Informantgrupper och kategoriseringsprinciper

Utifrån deras språkliga bakgrundsförhållanden har jag delat in informanterna i fem överordnade grupper (SVE, SVE+, START<6, START≥6 och INLÄRARE; se tabell 4:15 på nästa sida). Gruppen SVE+ har vidare delats in i två undergrupper; information om dessa ges i kursiv stil i tabellen.

Till gruppen SVE räknar jag de informanter som uppger svenska som sitt enda modersmål. Med några få undantag har de minst en förälder som har svenska som modersmål. 94 informanter i studien hör till denna grupp.

Till gruppen SVE+ räknar jag de informanter som anger svenska och ytterligare minst ett språk som modersmål och som anger en startålder för svenskan före tre års ålder, eftersom treårsåldern i tidigare forskning lyfts fram som en brytpunkt mellan första- och andraspråksinläring (McLaughlin 1978:99, Håkansson 2003:145, jfr Ortega 2013:5). 30 av informanterna i min studie räknas till denna informantgrupp.⁵ Eftersom jag antar att mängden svenskspråkig input har betydelse för befästningen av svenskan har gruppen SVE+ delats in i två undergrupper. De flerspråkiga informanter som har svenska som L1 och som anger att (minst) en av föräldrarna har svenska som L1 räknar jag till gruppen SVE+/S. Detta gäller 15 av informanterna. De flerspråkiga informanter som anger att de har svenska som L1 men att ingen av föräldrarna har svenska som L1 räknar jag till undergruppen SVE+/A, där A står för att föräldrarna har annat/andra modersmål än svenskan. Även denna undergrupp består av 15 informanter. Jag vill undersöka om denna skillnad i språklig bakgrund ger utslag i mitt material (se vidare information om indelningen i avsnitt 4.10.2).

5 Om informanten uppger svenska och ytterligare ett L1 men en startålder för svenskan vid tre år eller senare har jag ändå följt uppgiften om svenska som modersmål om informanten har minst en förälder med svenska som modersmål.

TABELL 4:15. *Antal informanter, fördelade efter informantgrupp.*

Beteckning	Informantgrupp	n	Ålder	Angiven startålder för svenskan
			Medel (SD) Median (Min; Max)	Medel (SD) Median (Min; Max)
SVE	Svenska = enda L1	94	18,4 (0,98) 19 (16; 20)	
SVE+	Flerspråkig i svenska och ytterligare (minst) ett L1	30	18,1 (0,91) 18 (17; 20)	0,7 (1,19) 0 (0; 4,5)
SVE+/S	<i>Flerspråkig i svenska och ytterligare (minst) ett L1 samt har (minst) en förälder som har svenska som L1</i>	(15)	18,3 (1,03) 19 (17; 20)	0,8 (1,41) 0 (0; 4,5)
SVE+/A	<i>Flerspråkig i svenska och ytterligare (minst) ett L1. Föräldrarna har andra modersmål än svenska</i>	(15)	17,9 (0,74) 18 (17; 19)	0,7 (0,96) 0 (0; 2,5)
START<6	Flerspråkig med svenska som L2, startålder för svenskan <6 år	34	17,9 (0,81) 18 (16; 19)	2,6 (1,32) 2,75 (0; 5)
START≥6	Flerspråkig med svenska som L2, startålder för svenskan ≥6 år, samt minst 4 års erfarenhet av svenska språket	32	19,1 (1,23) 19 (17; 22)	11,6 (3,02) 12 (6; 17)
INLÄRARE	Flerspråkig med svenska som L2, startålder för svenskan ≥6 år och mindre än 4 års exponering för svenska språket	9	18,8 (1,64) 18 (17; 22)	15,7 (1,48) 16,0 (14,0; 18,5)

Gruppen START<6 består av informanter som uppger ett annat (eller flera andra) modersmål än svenska och som börjat lära sig svenska före sex års ålder. Gränsen mellan tidig och sen startålder för andraspråket har dragits vid sex år dels för att tidigare studier påvisat en (skenbart) kritisk inlärningsperiod för ett andraspråk före 6–7 års ålder (se Abrahamsson & Hyltenstam 2013:245–246), dels för att sexårsåldern sammanfaller med skolstarten i förskoleklass.⁶ Informanterna i START<6 har därmed fått hela sin skolgång på svenska (förutsatt att de bott i Sverige hela tiden). Till denna informantgrupp hör 34 av studiens informanter.

Informanterna i gruppen START≥6 har annat modersmål än svenska och uppger en startålder sex år eller äldre för svenskan. Gruppen avgränsas åt andra hållet av en miniminivå av fyra års exponering för svenskan. Denna minimigräns är satt baserat på tidigare forskningsrön om utvecklingen av olika färdighetsnivåer i ett andraspråk. Enligt Cummins (2008) kan det i ett andraspråk ta cirka två år att utveckla ett vardagsnära språk, Basic Interpersonal Communication Skills (BICS), med vilket man kan kommunicera om enklare saker. Denna färdighetsnivå skiljer han från Cognitive Academic Language Proficiency (CALP) som står för det mer kognitivt krävande språk som behövs för formellt lärande. Studier visar att det kan ta mellan fyra och tio år att utveckla CALP (se Thomas & Collier 2002, Cummins 2008). Eftersom fokus i min studie ligger på tolkningen av betydelserelationen mellan för- och efterled, är det angeläget att de ingående ordleden är kända för informanterna, även om sammansättningen som helhet inte är bekant. Merparten av sammansättningarna i min enkät innehåller led som betecknar basala företeelser och egenskaper, såsom *sten* och *sol*, *hård* och *varm*. För dessa kan BICS möjligen räcka, men vissa sammansättningar innehåller abstrakta, svårare led som möjligen kräver CALP, t.ex. *politiker* och *förakt*. Min slutsats blir att avkodningen av leden kräver en färdighetsnivå någonstans mellan BICS och CALP, och jag har därför valt att sätta gränsen för START≥6 vid minst fyra års exponering för svenskan. Gruppen START≥6 består av 32 informanter.

Informanterna med en kortare tids exponering för svenskan än fyra år betraktar och betecknar jag som mer utpräglade ”inlärare” (9 informanter). I deras fall kan enkätresultatet ha mer att göra med en kort exponeringstid för svenskan och medföljande svårigheter att avkoda leden än en flerspråkighetssituation som medför andra associationer när det gäller relationen mellan för- och efterled. Av detta skäl kommer inlärnarnas enkäter inte att vidare behandlas i analysen.

6 Förskoleklass var från 1998–2018 en frivillig skolform som majoriteten av sexåringarna deltog i (Ackesjö & Persson 2010:142–143). Från och med hösten 2018 är den obligatorisk (Skolverket 2018).

Att skapa adekvata kategorier för att beskriva informanternas språkliga bakgrund är en grannliga uppgift, eftersom en informants språkliga bakgrund alltid är komplex och svårligen låter sig fångas (se t.ex. Fraurud & Boyd 2011). Även tillämpningen av mina kategoriseringsprinciper har i vissa fall varit problematisk eftersom enkätsvar alltid innehåller ett visst mått av oklarheter, utelämnningar och tvetydigheter. Några informanternas språkliga bakgrundsuppgifter är sådana att informanterna inte självklart låter sig inordnas i någon av kategorierna. I det följande (avsnitt 4.10.2–4.10.6) diskuteras dessa metodologiska svårigheter: dels övergripande svårigheter vad gäller gränsdragningen mellan kategorier, dels tillämpningen av mina kategoriseringsprinciper och svårbedömda informantfall.

4.10.2. Flerspråkighetsbakgrund och L2

De flerspråkiga informanter i studien som uppger svenska och ytterligare (minst) ett språk som modersmål har jag räknat till gruppen SVE+, där SVE markerar svenska som L1 och plustecknet markerar informantens ytterligare L1. Detta gäller 30 informanter. De flerspråkiga informanter som anger ett annat (eller flera andra) modersmål än svenska har jag räknat till START<6 (34 informanter) eller START≥6 (32 informanter) beroende på deras startålder för svenskan. Informanterna i START<6 och START≥6 betraktar jag som flerspråkiga med svenska som L2.

När jag i det följande hänvisar till *flerspråkiga* avser jag informanterna i SVE+, som har (minst) två modersmål, och informanterna i START<6 och START≥6, som förutom sitt (eller sina) modersmål dagligen använder svenskan. När jag hänvisar till *L2-talare* avses informanterna i START<6 och START≥6, dvs. de informanter som uppger annat modersmål än svenska.

När det gäller informanterna i SVE+ framgår det av bakgrundsuppgifterna för dessa informanter att deras flerspråkighetssituationer skiljer sig åt när det gäller svenskans status som L1 i familjen. Hälften av informanterna i SVE+ har (minst) en förälder som har svenska som L1, medan hälften av dem har föräldrar som har andra modersmål än svenska. Den senare gruppen har således inte haft stöd för svenskan från en förälder som har svenska som L1. För att kunna skilja dessa flerspråkighetssituationer åt i analysen har jag valt att också använda mig av en finare indelning i två undergrupper.

Till gruppen SVE+/S räknar jag de informanter som anger att (minst) en av föräldrarna har svenska som L1. Tolv av informanterna i denna grupp har en förälder med svenska som modersmål och en förälder med annat modersmål. Det innebär att de representerar en flerspråkighetssituation med ett modersmål från vardera föräldern. För tre informanter i SVE+/S

är situationen sådan, att båda föräldrarna har svenska som modersmål, men informanterna har under sina tidiga levnadsår bott i ett annat land och förvärvat detta lands språk som ett av sina modersmål. Sammanfattningsvis har dessa 15 informanter svenska som modersmål via förälders modersmål, och de kan förväntas ha påbörjat sin inläring av svenska från födseln.

De övriga 15 informanterna i gruppen representerar en annan flerspråkighetssituation. De har (minst) ett annat modersmål än svenska med sig från föräldrarna. Trots att ingen av föräldrarna har svenska som modersmål uppges svenskan ha kommit in tidigt i informanternas liv – före tre års ålder – och de betraktar svenska som ett av sina modersmål. Dessa informanter räknar jag till undergruppen SVE+/A. Tolv av dessa informanter är födda i Sverige, och de övriga tre har kommit till Sverige under det första eller andra levnadsåret. Det är svårt att veta hur den språkliga situationen och den svenskspråkiga inputen sett ut för informanterna med denna språkbakgrund. Viss information kan dock inhämtas från enkätfråga 8 som rör platsen för den tidiga svenskinläringen och enkätfråga 27 om använda språk i familjen. Nio av informanterna i SVE+/A anger att de lärt sig svenska hemma och att svenska används i familjen. Det är naturligtvis fullt tänkbart att föräldrarna talat svenska med sina barn från en tidig ålder, även om föräldrarna själva inte har svenska som modersmål. Dessutom kan informanterna ha fått svenskspråkig input från (äldre) syskon. Förutom samtal på svenska i familjen kan man tänka sig att ”hemma” för informanten även inbegriper samtal på svenska med kamraterna hemma på gården eller i området. Två av informanterna anger att den tidiga svenskinläringen skett hemma och på dagis/förskola, och anger svenska som ett använt språk i familjen. Fyra av informanterna anger endast dagis/förskola som den tidiga svenskspråkiga inlärningsarenan, och tre av dessa anger att svenska är ett språk som används i familjen.

När jag i det följande hänvisar till SVE+ refererar jag till den överordnade gruppen av 30 flerspråkiga informanter som uppger svenska och ytterligare ett (eller flera) språk som L1. När jag vill särredovisa uppgifter från undergrupperna hänvisar jag till SVE+/S respektive SVE+/A.

Eftersom förstaspråksinläring anses kunna ske upp till tre års ålder (se t.ex. McLaughlin 1978:99, Håkansson 2003:145) har jag valt att följa informanternas uppgift om svenska som modersmål om de angett en startålder lägre än tre år, oavsett om den tidiga inläringen skett i hemmet eller på förskolan (se vidare diskussion om modersmålsbegreppet, startålder och språklig profilering i Fraurud & Boyd 2011). Denna kategorisering innebär att två flerspråkiga informanter som båda har föräldrar med annat L1 än svenska och som båda börjat med svenska vid två års ålder på förskola, kan hamna i två olika informantgrupper (SVE+/A eller

START<6), beroende på hur de tolkat frågan om modersmål som presenterades i exempel (4:6) och upprepas i exempel (4:10).

(4:10) Vilket/vilka språk lärde du dig tala först (ditt/dina modersmål)?

Den informant som associerar modersmål till ”bäst behärskade språk” kan tänkas ange svenska och ytterligare ett modersmål p.g.a. fullgod behärskning i dessa två språk, medan den informant som tar fasta på ursprungskriteriet ”första språk” kan tänkas ange ett annat modersmål än svenska – oavsett färdighetsnivå för svenskan – eftersom informanten börjat med ett annat språk före svenskan (se Skutnabb-Kangas 1981:22–23). Därför kan det föreligga en överlappning vad gäller språkliga bakgrundsförhållanden mellan informanter som jag (utifrån modersmålsfrågan) hänfört till SVE+/A respektive START<6.

Om informanten angett svenska som modersmål (eller svenska och ytterligare ett modersmål) men uppger en startålder för svenskan vid tre års ålder eller senare har jag valt att inte betrakta svenskan som ett modersmål, utan har räknat informanten till kategorin START<6. Detta har skett för fyra informanter i studien (BF611, PS602, PS606 och KV611).

4.10.3. Svenska som L1 (SVE) trots andra modersmål hos föräldrarna

I den grupp av informanter som anger svenska som sitt enda modersmål (dvs. SVE) har 90 av informanterna minst en förälder med svenska som modersmål. Fyra av informanterna uppger dock att de har svenska som enda modersmål (med en startålder noll år), trots att ingen av föräldrarna har svenska som modersmål. I dessa fall har jag av olika skäl ändå valt att följa deras uppgifter. För informant BC411 och KV503 gäller, att föräldrarna har två sinsemellan olika (och obesläktade) modersmål. Föräldrarna till BC411 har två afrikanska språk som modersmål, men de kommunicerar med varandra på ett tredje språk och med barnen på svenska. Föräldrarna till KV503 har ett afrikanskt språk respektive ett finskugriskt språk som modersmål, men i familjen används svenska och i viss mån engelska som kommunikationsspråk. Inte i något av fallen anges föräldrarnas modersmål vara informantens modersmål (enkätfråga 6) eller ett språk som informanten kan tala (enkätfråga 5). Informant BC411 uppger visserligen att hon kan tala lite på det språk som föräldrarna använder som kommunikationsspråk med varandra, men modifie-

ringen ”lite” gör att det inte finns anledning att tro att detta tredje språk i familjen är att betrakta som ytterligare ett modersmål för informanten. Även informanten KV107 har två föräldrar med två sinsemellan olika, och obesläktade, modersmål (ett finskugriskt språk och ett nordiskt), men den ena föräldern är född i Sverige. Informanten uppger visserligen att han kan tala det nordiska språket, men detta anges efter engelska i behärskade språk (vilket tyder på att det inte är likställt svenskan), och språket anges inte som använt i familjen (bara svenska anges). Därför har jag valt att hålla fast vid informantens uppgifter om svenska som enda modersmål. Informant KV401, slutligen, har två föräldrar med samma modersmål (kurdiska). För båda föräldrarna gäller dock att svenskan kommit in tidigt i deras liv, vid sex respektive tolv års ålder. Föräldrarna anges tala sitt modersmål med varandra men i övrigt används svenska som familjespråk. Informanten själv anger inte föräldrarnas modersmål som ett behärskat språk. Därmed förefaller uppgiften om svenska som enda modersmål även i detta fall rimlig. Sammanfattningsvis har alla fyra informanter som anger svenska som modersmål, trots att ingen av föräldrarna har svenska som modersmål, på redovisade grunder hänförs till kategorin SVE (svenska som enda modersmål).

4.10.4. Flerspråkig (SVE+) trots ett angivet modersmål

I några fall har jag haft anledning att lägga till ett modersmål till det som informanten angett, vilket inneburit att informanten hänförs till gruppen SVE+. Sex informanter har angett svenska som enda modersmål, men det framgår av andra bakgrundsfrågor att informanten har ytterligare ett språk med sig från start. Informanterna BF617, BC617 och BC605 har alla uppgett svenska som enda modersmål (i kombination med en startålder lägre än tre år). Informanternas föräldrar har emellertid andra modersmål än svenska (kurdiska; somaliska; somaliska), vilka informanterna uppger som behärskade språk (fråga 5), och som använda språk i familjen (fråga 27). Det förefaller således rimligt att lägga till dem som informanternas modersmål, och informanterna har därmed hänförs till gruppen SVE+ (undergruppen SVE+/A). Informanterna BC601 och PS403 anger svenska som modersmål, och detta modersmål har erhållits via den ena föräldern. Den andra förälderns L1 (arabiska respektive engelska) uppges dock vara behärskat av informanten och använt i familjen. Jag har därför lagt till detta som ett modersmål och hänfört informanterna till gruppen SVE+ med svenska och ytterligare ett modersmål via förälders modersmål (dvs. SVE+/S). PS505 uppger svenska som enda modersmål, men hon behärskar dessutom ett slaviskt språk, vilket också används i familjen och med vänner. Ena föräldern är född i Sverige och är tvåspråkig i

svenska och detta slaviska språk, medan den andra föräldern har ett annat slaviskt språk som modersmål. Utifrån dessa bakgrundsuppgifter har jag kategoriserat informanten som tvåspråkig med svenska och ytterligare ett modersmål via föräldrarna (dvs. SVE+/S), även om båda språken i det här fallet kommer från den ena föräldern.

Även informanten BF612 uppger endast ett modersmål, men i detta fall engelska. Hon anger en start för svenskan vid två års ålder på dagis. Informanten är född i Sverige, den ena föräldern har engelska och den andra föräldern har svenska som modersmål, och engelska anges som använt språk i familjen. Eftersom den ena föräldern har svenska som modersmål och majoritetsspråket svenska kommit in i informantens liv så tidigt som vid två års ålder, har jag bedömt det som rimligt att betrakta svenskan som modersmål, trots att informanten inte själv anger svenska som modersmål. Informanten har således hänförs till gruppen SVE+ med svenska via förälders modersmål, dvs. SVE+/S.

4.10.5. SVE trots svenska och ytterligare modersmål i familjen

Som redan anförts har 90 av 94 informanter i SVE minst en förälder med svenska som modersmål. Några av dessa lever emellertid i en familj där ytterligare modersmål finns representerade hos föräldrarna. Tre av informanterna i den enspråkiga L1-gruppen (SVE) har en förälder med svenska som modersmål och en förälder med annat modersmål. Ytterligare tre informanter har en förälder med svenska som modersmål och en med svenska och ytterligare ett modersmål. En informant har två föräldrar som båda har svenska och ytterligare ett språk som modersmål. I det följande förklarar jag varför jag, i likhet med dem själva, inte betraktar dessa sju informanter som flerspråkiga med två modersmål. Jag börjar med de tre informanter som har en förälder med svenska som modersmål och en förälder med annat modersmål.

Informant BC105 har en förälder med annat modersmål än svenska (finska), men informanten själv anger inte finska som ett behärskat språk och det används inte i hans familj. Informanten uppger om att föräldern kom till Sverige redan som liten och endast talar finska med en äldre släkting. Även informant BF101 har en förälder med finska som modersmål, men inte heller i detta fall anges finska som ett behärskat språk av informanten, eller som ett använt språk i familjen. Informanten anger vidare att föräldern är född i Sverige. Eftersom varken BC105 eller BF101 talar finska räknar jag dem till kategorin svenska som enda modersmål. I fallet med informant PS501 har gränsdragningen varit svårare. Informanten

uppger att den ena föräldern har två andra modersmål än svenska (holländska, tyska). På frågan om vilka språk informanten kan tala nämns ett av dem (holländska), men det nämns sist, efter både svenska och engelska. Informanten gör dessutom en modifiering när han på frågan om använda språk i familjen skriver ”svenska och *lite* holländska”. Det faktum att holländskan nämns efter engelskan på frågan om behärskade språk, i kombination med att det bara talas lite holländska i familjen, gör att jag bedömer informantens uppgifter om svenska som enda modersmål som trovärdiga och placerar informanten i kategorin svenska som enda modersmål.

PS203 och BC606 har båda *en* förälder med svenska som modersmål och *en* förälder med svenska och ytterligare ett modersmål, men dessa ytterligare språk (finska respektive ett austronesiskt språk) är inte använda i familjerna och anges inte som behärskade av informanterna. KV505 har en förälder med svenska som modersmål och en förälder med svenska och danska som modersmål. Informanten uppger visserligen danska som ett behärskat språk, men han nämner det efter både svenska och engelska och han anger det inte som modersmål eller som ett använt språk i familjen. Därför har jag valt att följa informantens uppgifter om att endast svenska är informantens modersmål.

Ytterligare en informant i kategorin SVE, nämligen BF104, har två föräldrar som båda har svenska och ytterligare ett språk som modersmål (ett nordiskt språk respektive finska), men även i detta fall är det endast svenska som används i familjen. Endast svenska uppges som modersmål, och på frågan om behärskade språk svarar informanten svenska, engelska och ”väldigt lite finska”. Utifrån svaret på modersmålsfrågan och modifieringen ”väldigt lite” om behärskningen av finska, bedömer jag det som rimligt att följa informantens uppgifter om att endast svenska är informantens modersmål.

4.10.6. Tvetydiga uppgifter om modersmål och startålder för svenskan

I några fall är informanternas uppgifter tvetydiga. BF601 uppger arabiska som modersmål men anger samtidigt att hon börjat lära sig svenska vid noll års ålder hemma. Med en så låg startålder är det förväntat att informanten skulle betrakta språket som ett modersmål. BF602 svarar inte på frågan om modersmål men uppger arabiska som ett behärskat språk och anger en start för svenskan vid ett års ålder hemma. I båda dessa fall har jag valt att låta fritextsvaret på modersmålsfrågan väga tyngre än ikryssningsfrågan om startålder. Eftersom informanterna inte anger svenska som

modersmål har jag valt att hänföra dem till gruppen START<6. En annan informant, KV606, anger finska som modersmål, men uppger samtidigt att inlärnigen av svenska börjat hemma vid noll års ålder. Informanten har också kryssat i att han har en förälder med svenska som modersmål, men följdfrågorna om mammas respektive pappas modersmål besvaras inte, så språksituationen förblir oklar. Även i detta fall har jag låtit svaret på fritextfrågan om modersmål väga tyngre än ikryssningsfrågan om föräldrars modersmål. Därmed hänförs även KV606 till gruppen START<6. BF502 anger fyra modersmål, varav ett är svenska, och båda föräldrarna uppges ha persiska som modersmål. Informanten kan möjligen anses ha en något vidlyftig syn på vad som kan betraktas som ett modersmål. Samtidigt anges startåldern 2–3 år för svenskan (vilket jag avrundat till före tre år), varför jag valt att följa informantens uppgifter och räknat informanten till SVE+/A. Informanten anger att hon börjat lära sig svenska på förskolan, och hon uppger svenska och persiska som använda språk i familjen.

4.11. Kvantitativ analys av resultaten i enkätens flervalsdel

I avhandlingen används både deskriptiv och analytisk statistik. Valet av statistisk metod och statistiska test är gjort i samråd med en expertkonsult inom statistik.⁷

4.11.1. Statistisk metod

Deskriptiv statistik har angivits med medelvärde (M), standardavvikelse (SD), median, minimum och maximum för kontinuerliga variabler, med antal och procent för kategoriska variabler, och med Spearmans rangkorrelationskoefficient för korrelationer.

Först analyserades variablerna med ett övergripande test mellan de fyra informantgrupperna, som analyserar om det finns några övergripande skillnader mellan dem. Om p -värdet för det övergripande testet var signifikant, gick jag vidare och tolkade de parvisa analyserna.

För de övergripande jämförelserna mellan informantgrupperna har Pearsons exakta χ^2 -test använts för dikotoma variabler (dvs. i analysen

⁷ Jag vill rikta ett innerligt tack till Nils-Gunnar Pehrsson, senior biostatistiker, B.A.; CEO Statistiska konsultgruppen.

av svaren på de enskilda sammansättningarna) och Kruskal-Wallis rangtest för kontinuerliga variabler (dvs. i analysen av medelvärden för en kategori av sammansättningar). För de parvisa jämförelserna mellan två informantgrupper har Fishers exakta test använts för dikotoma variabler och Mann-Whitney U-test för kontinuerliga variabler. På grund av att de kontinuerliga variablerna inte var normalfördelade har jag genomgående använt icke-parametriska test.

I två kategorier av sammansättningar förekommer tre svarsalternativ, i stället för två. För att renodla analysen har två av svarsalternativen slagits samman i den statistiska analysen. Därmed blev även dessa variabler dikotoma, och samma procedur kunde följas i dessa kategorier som i de övriga. Ett övergripande Pearsons χ^2 -test med exakt beräkning följdes upp av parvis analys med Fishers exakta test.

Vid alla signifikanstester har jag använt ett tvåsidigt test med signifikansnivån $p < 0,05$. Den exakthet med vilken p-värden anges följer en statistisk konvention enligt följande. Värden mindre än 0,01 anges med fyra decimalers noggrannhet. Värden mellan 0,01 och 0,10 anges med tre decimaler. Värden större än 0,10 anges med två decimalers noggrannhet.

4.11.2. Bortfall

Av de elever jag träffade på plats i skolorna var det bara en person (i skola 4) som valde att inte delta i studien, vilket betyder att det externa bortfallet är mycket litet. (Jag bortser här från frånvaron i klasserna; se avsnitt 4.8.)

Även det interna bortfallet är mycket begränsat med enstaka missade svar för enstaka informanter. Bortfall av detta slag skulle kunna bero på att informanterna (visuellt) inte uppfattat frågan, men det skulle också kunna bero på att informanterna hoppat över särskilt svåra ord.

Det förekommer emellertid en annan typ av bortfall i enkäten. Tre informanter (BC202, PS201 och BC611) har tidigt avbrutit sitt deltagande. När det gäller orddelen har de besvarat fritextdelen men bara de första två-tre sidorna (av tolv) i flervalsdelen. De tre informanterna har svarat noggrant på de första sidorna och därefter ingenting på de följande sidorna. De har med andra ord bara tolkat en bråkdel av sammansättningarna i flervalsdelen, innan de avbrutit ifyllandet av enkäten. När jag gjort beräkningar av informantgruppernas medelvärden på en hel grupp av sammansättningar har jag räknat bort resultatet för dessa informanter i de fall där deras bortfall överstigit 50 %, eftersom deras medelvärde inte kan sägas spegla deras kunskap om sammansättningstypen i fråga. Det tidigt avbrutna deltagandet innebär att de inte visat hur de tolkar sammansättningstypen.

4.12. Analys av fritexttolkningarna i relation till svaren i enkätens flervalsdel

Fritextdelen av enkäten används huvudsakligen som en kontrolldel. Till flervalsdelen har jag ur sammansättningarnas meningspotentialer valt ut två eller tre alternativa förslag, som är försedda med en kontextualiserande mening, dvs. informanternas tolkningar är delvis styrda. I fritextdelen formulerar informanterna själva en förklaring. För att kunna jämföra hur informanterna tolkar sammansättningar med och utan kontext är några sammansättningar representerade både i fritextdelen (som informanterna möter först) och i flervalsdelen (se vidare avsnitt 4.3). I övrigt har jag valt ut sammansättningar av samma typ till båda delar av enkäten, t.ex. *kvinnoläkare* i fritextdelen och *mansläkare* i flervalsdelen. Således används fritextdelen som ett referensmaterial till flervalsdelen av enkäten. Utöver detta görs en noggrannare semantisk analys av enstaka exempel i fritextdelen som belyser spännvidden i sammansättnings meningspotentialer.

5

Tolkningar i enkätens flervalssdel

I det här och det följande kapitlet redovisas resultaten från den enkätstudie om sammansatta ord som föreliggande avhandling baseras på. Resultaten från flervalssdelen av enkäten redovisas i detta kapitel, medan resultat från fritextdelen av enkäten presenteras i det nästföljande kapitlet.

Datainsamlingen genomfördes på ett antal gymnasieskolor under våren och hösten 2015, och studien omfattar 202 informanter. Informanterna i studien har delats in i fyra huvudgrupper efter språklig bakgrund. Grupp SVE består av informanter med svenska som enda L1 (n=94). Grupp SVE+ består av informanter som är flerspråkiga i svenska och ytterligare (minst) ett L1 (n=30). Grupp START<6 består av flerspråkiga informanter med svenska som andraspråk och en låg startålder för svenskan, före sex års ålder (n=34). Grupp START≥6 består av flerspråkiga informanter med svenska som andraspråk och en startålder för svenskan vid sex års ålder eller senare, samt minst fyra års exponering för svenskan (n=32). Analysen grundar sig på enkätsvaren från dessa 190 informanter.⁸ Informantgrupperna finns närmare beskrivna i metodavsnittet 4.10.

⁸ Därutöver har en mindre grupp om nio informanter med kortare exponeringstid än fyra år för svenskan uteslutits ur analysen, liksom tre informanter som inte svarat på frågorna om språklig bakgrund.

Detta kapitel innehåller en redogörelse för resultaten från enkätens flervalsdel som innehåller 69 sammansättningar med flervalsalternativ. Resultaten presenteras efter den indelning jag gjort av sammansättningarna i tolv kategorier (se avsnitt 4.3). I avsnitt 5.1 presenteras resultaten för de nominala sammansättningarna, och i avsnitt 5.2 presenteras resultaten för de adjektiviska sammansättningarna. Inom kategorierna av sammansättningar presenterar jag i tabellform hur informanterna tolkar de enskilda sammansättningarna, dvs. hur de väljer mellan de två eller tre svarsalternativen. Sammansättningarna presenteras i en fallande frekvensordning i tabellerna, från sammansättningen med högst till sammansättningen med lägst antal förekomster i en referenskorpus (se avsnitt 4.4). De enskilda sammansättningarnas frekvens i referenskorpusen finns tillgängliga i bilaga 3. Kontextmeningarna för svarsalternativen är inte representerade i tabellerna men även dessa finns tillgängliga i appendix, se bilaga 1.

Resultaten i de olika kategorierna av sammansättningar presenteras för informantgruppen som helhet och för de olika informantgrupperna. Därefter jämförs informantgruppernas svar på de enskilda sammansättningarna med varandra. Signifikanta skillnader mellan grupperna är markerade med symboler i tabellerna. Eftersom informanterna i SVE+ skiljer sig åt när det gäller svenskans status som L1 i familjen (se avsnitt 4.10.2) har jag särredovisat resultatet för undergruppen SVE+/S, dvs. de informanter i SVE+ som har en förälder som har svenska som L1. Resultatet för SVE+/S anges i kursiv stil under resultatet för den överordnade gruppen SVE+. Där anges också om resultatet för SVE+/S skiljer sig signifikant från resultatet för SVE.

Efter genomgången av de enskilda orden riktas fokus mot mer generella likheter och skillnader mellan informantgruppernas tolkningar med avseende på sammansättningskategorierna som helhet.

5.1. Nominala sammansättningar

I en kategori av nominala sammansättningar undersöks hur informanterna tolkar en sammansättnings semantiska huvudled. I en annan kategori av sammansättningar analyseras hur informanterna konceptualiserar behållare av olika slag, och vilka semantiska relationer mellan leden som aktualiseras i dessa tolkningar. Dessutom undersöker jag i enkäten vilket som är det prefererade mönstret i valet mellan egenskapstolkning och relationslänkning. I ytterligare en kategori av sammansättningar studeras informanternas tolkningar av agens vid verbavledning. I en annan kategori av nominala sammansättningar analyseras tolkningen av bety-

delsesammansmälta AN-sammansättningar (för en närmare beskrivning av sammansättningskategorierna; se avsnitt 4.3).

5.1.1. Huvudled i nominala sammansättningar – typen *spelkort* och *vagnbarn*

Som framgår i avsnitt 1.2 och avsnitt 2.4 utgår man från att nominala sammansättningar i svenskan är högerställda, dvs. att efterledet i sammansättningen utgör det semantiska och grammatiska huvudet, medan förledet utgör underordnad modifierare. Distraktorerna i den enkät som informanterna besvarade pekar i stället ut förledet som sammansättningens morfologiska och semantiska huvud, och speglar därmed en struktur med huvudledet till vänster i sammansättningen.

I flervalsdelen av enkäten ingår nio sammansättningar som är avsedda att ge information om vilket led informanterna identifierar som sammansättningens semantiska huvud. Till de mer frekventa sammansättningarna i kategorin hör *husbåt* och *båthus* med förhållandevis god representation i referenskorpusen. Sammansättningarna *spelkort* och *flodstrand* har låg representation i referenskorpusen, och för *burköl* och *båtsommar* noteras bara enstaka förekomster. De tillfälliga bildningarna *boktext*, *vagnbarn* och *stövelbarn* saknar representation i referenskorpusen.

I tabell 5:1 redovisas, för informantgruppen som helhet och per informantgrupp, hur svaren på de enskilda sammansättningarna i kategorin huvud fördelar sig mellan höger- och vänstertolkningar. För varje ord redovisas andelen högertolkningar (i fet stil), dvs. de tolkningar som följer schemat för huvudets placering i svenska sammansättningar, följt av andelen vänstertolkningar. Som tidigare påpekats är sammansättningarna i tabellen sorterade i frekvensordning, i en fallande skala från sammansättningen med högst till sammansättningen med lägst antal förekomster i referenskorpusen.

TABELL 5:1. Informanternas svar på nominala sammansättningar som testar identifiering av huvudled.⁹

	Totalt, alla informanter n=190	SVE n=94	SVE+ n=30 SVE+/S n=15	START<6 n=34	START≥6 n=32	P-värde övergripande (overall)
Båt som man kan bo i	166 (88 %)	84 (90 %)	28 (97 %) 13 (93 %)	32 (94 %)	22 (69 %)	** □□ ◇
Husbåt	22 (12 %)	9 (10 %)	1 (3 %) 1 (7 %)	2 (6 %)	10 (31 %)	p=0,0018 ** □□ ◇
Hus där man kan lägga båten	102 (55 %)	69 (75 %)	10 (33 %) 7 (47 %) *	10 (31 %) ***	13 (41 %) ***	
Båthus	84 (45 %)	23 (25 %)	20 (67 %) 8 (53 %) *	22 (69 %) ***	19 (59 %) ***	p<0,0001
Speciella kort (i en kortlek) som man använder för att spela med	136 (74 %)	76 (83 %)	22 (76 %) 11 (73 %)	22 (69 %)	16 (52 %)	**
Spelkort	48 (26 %)	16 (17 %)	7 (24 %) 4 (27 %)	10 (31 %)	15 (48 %)	** p=0,0066
Strand vid floden	129 (69 %)	75 (81 %)	19 (63 %) 10 (67 %)	17 (52 %) **	18 (56 %) **	
Flodstrand	59 (31 %)	18 (19 %)	11 (37 %) 5 (33 %)	16 (48 %) **	14 (44 %) **	p=0,0035

⁹ Övergripande p-värde beräknat med Pearsons exakta chi²-test, parvisa jämförelser beräknade med Fishers exakta test.

* p<0,05, ** p<0,01, *** p<0,001

□ p<0,05, □□ p<0,01, □□□ p<0,001

◇ p<0,05, ◇◇ p<0,01, ◇◇◇ p<0,001

Siffror i kursiv representerar gruppen SVE+/S.

Parvis jämförelse SVE mot var och en av de övriga grupperna

Parvis jämförelse SVE+ mot START<6, START≥6

Parvis jämförelse START<6 mot START≥6

	Totalt, alla informanter n=190	SVE n=94	SVE+ n=30 SVE+S n=15	START<6 n=34	START≥6 n=32	P-värde övergripande (overall)
Burköl	Öl i burk	157 (84 %)	88 (96 %)	25 (86 %) 13 (93 %)	25 (74 %) ***	19 (59 %) *** <input type="checkbox"/>
	Burk för öl	30 (16 %)	4 (4 %)	4 (14 %) 1 (7 %)	9 (27 %) ***	13 (41 %) *** <input type="checkbox"/>
Båt-sommar	Sommar i båt	131 (71 %)	82 (90 %)	17 (59 %) *** 11 (79 %)	22 (67 %) **	10 (31 %) *** <input type="checkbox"/> ∅∅
	Båt som man bara kan använda på sommaren	54 (29 %)	9 (10 %)	12 (41 %) *** 3 (21 %)	11 (33 %) **	22 (69 %) *** <input type="checkbox"/> ∅∅
Boktext	Text ur/i en bok	134 (72 %)	79 (86 %)	24 (80 %) 13 (87 %)	17 (52 %) *** <input type="checkbox"/>	14 (45 %) *** <input type="checkbox"/> <input type="checkbox"/>
	Bok med texter	52 (28 %)	13 (14 %)	6 (20 %) 2 (13 %)	16 (49 %) *** <input type="checkbox"/>	17 (55 %) *** <input type="checkbox"/> <input type="checkbox"/>
Stövel-barn	Barn som gillar att ha stövlar	150 (82 %)	83 (91 %)	23 (79 %) 14 (93 %)	25 (78 %)	19 (59 %) ***
	Stövel som passar ett barn	34 (18 %)	8 (9 %)	6 (21 %) 1 (7 %)	7 (22 %)	13 (41 %) ***
Vagn-barn	Barn som sitter i vagn	145 (78 %)	83 (90 %)	19 (63 %) ** 10 (67 %) *	25 (76 %)	18 (60 %) ***
	Vagn som man har det lilla barnet i	40 (22 %)	9 (10 %)	11 (37 %) ** 5 (33 %) *	8 (24 %)	12 (40 %) ***

5.1.1.1. Tolkningar i informantgruppen som helhet

Av tabell 5:1 framgår att informanterna i stor utsträckning identifierar efterledet som sammansättningens huvud, men även det alternativ där förledet utgör huvudled aktualiseras i många fall. Frekvenserna för de enskilda sammansättningarna i referenskorpusen har viss betydelse men får inget entydigt genomslag i tolkningarna från informantgruppen som helhet. *Husbåt* som är mest välrepresenterad i referenskorpusen har störst

andel högertolkningar i enkäten (88 %), men *båthus* på plats 2 i den frekvensordnade tabellen är den sammansättning som får lägst grad av högertolkningar, endast 55 %. De mindre frekventa sammansättningarna *spelkort* och *flodstrand* får en relativt stor andel högertolkningar i enkäten (74 % respektive 69 %). *Burköl* och *båtsommar*, som bara har enstaka träffar i referenskorpusen, får likväl en stor andel högertolkningar, 84 % respektive 71 %. De tillfälliga sammansättningarna *boktext*, *stövelbarn* och *vagnbarn*, som saknar representation i referenskorpusen, får ändå så många som 72 %, 82 % respektive 78 % högertolkningar.

Sammanfattningsvis är tolkningarna av de enskilda sammansättningarna ingen direkt avspeglning av frekvensförhållandena för de enskilda sammansättningarna. Mönstret med efterledet som sammansättningens huvudled slår över lag igenom starkt i informanternas tolkningar, oavsett de enskilda sammansättningarnas representation i referenskorpusen.

5.1.1.2. Jämförelse mellan informantgruppernas tolkningar – enskilda sammansättningar

Av tabell 5:1 kan man utläsa att L1-talarna (SVE och SVE+) generellt har en större andel högertolkningar av de nominala sammansättningarna än L2-talarna ($START < 6$ och $START \geq 6$). Det finns vid alla sammansättningar i kategorin ett signifikant samband mellan vilken informantgrupp man tillhör och hur man svarar, dvs. det föreligger statistiskt signifikanta skillnader mellan informantgruppernas tolkningar av de enskilda orden.

De sammansättningar som genererar störst skillnader mellan grupperna är *båtsommar*, *båthus* och *boktext*. Differensen mellan den största och minsta andelen högertolkningar är 59, 44 respektive 41 procentenheter vid dessa sammansättningar. Svagast är sambandet mellan informantgrupp och svar vid *flodstrand* och *husbåt* (29 respektive 28 procentenheter i differens mellan största och minsta andel högertolkningar). I det följande redovisas mellan vilka grupper det föreligger signifikanta skillnader i tolkningen av de enskilda sammansättningarna.

Vid den lågfrekventa sammansättningen *båtsommar* (ord 6 i tabellen) är skillnaderna mellan grupperna stora. 90 % av informanterna i SVE väljer högertolkningen 'sommar i båt', vilket är en signifikant större andel än i gruppen SVE+ (59 %, $p=0,0003$), $START < 6$ (67 %, $p=0,0042$) och $START \geq 6$ (31 %, $p < 0,0001$). SVE skiljer sig däremot inte signifikant från SVE+S (79 % högertolkning $p=0,20$). Gruppen SVE+ har signifikant fler högertolkningar än informanterna i $START \geq 6$ ($p=0,041$). Tolkningarna från informanterna i $START < 6$ skiljer sig signifikant från de tolkningar som informanterna i $START \geq 6$ gör ($p=0,0063$).

Ytterligare ett ord med stora mellangrupsliga skillnader är den mindre frekventa sammansättningen *boktext* (ord nummer 7 i tabellen). 86 % av

informerterna i grupp SVE och 80 % av informerterna i grupp SVE+ väljer en tolkning med huvudledet till höger: 'text ur/i en bok', medan motsvarande siffra är 52 % i START<6 och 45 % i START≥6. Övriga väljer distraktorn 'bok med texter'. Tolkningarna från informerterna i SVE skiljer sig signifikant från tolkningarna i START<6 ($p=0,0002$) och START≥6 ($p<0,0001$). Även informerterna i SVE+ skiljer sig signifikant från START<6 ($p=0,033$) och START≥6 ($p=0,0079$) i sina tolkningar av *boktext*.

För sammansättningen *båthus* (ord 2 i tabellen) föreligger också ett relativt starkt samband mellan informantgrupp och svar. Informanterna i SVE har en signifikant större andel högertolkningar – 'hus där man kan lägga båten' – än övriga grupper. I SVE väljer 75 % denna tolkning, att jämföra med 33 % i SVE+ ($p<0,0001$), 31 % i START<6 ($p<0,0001$) och 41 % i START≥6 ($p=0,0009$). Resultatet för SVE skiljer sig signifikant också från resultatet för SVE+/S, dvs. de flerspråkiga informanter som har en förälder som har svenska som L1 (47 % högertolkning, $p=0,034$). Informanterna i SVE är således mer benägna än övriga grupper att betrakta 'hus' som det semantiska huvudet hos *båthus*. Samtidigt är det så många som en fjärdedel av SVE-informerterna som väljer distraktorn 'båt som man kan bo i'.

Vid *burköl* är sambandet mellan informantgrupp och svar svagare, men informerterna i SVE har signifikant fler högertolkningar (96 %) än informerterna i START<6 (74 %, $p=0,0010$). Tolkningarna från SVE skiljer sig också signifikant från tolkningarna i START≥6 (59 % högertolkningar, $p<0,0001$). Gruppen SVE+ har signifikant fler högertolkningar (86 %) än gruppen START≥6 (59 %, $p=0,024$).

Vid den tillfälliga bildningen *vagnbarn* har grupp SVE en signifikant större andel högertolkningar (90 %) än informerterna i gruppen START≥6 (60 %, $p=0,0004$). SVE-gruppen skiljer sig signifikant i sina tolkningar även från SVE+ som har 63 % högertolkningar av *vagnbarn* ($p=0,0013$). Om man avgränsar analysen till de informanter i SVE+ som har svenska som L1 via förälders L1 (SVE+/S, $n=15$) är andelen högertolkningar av *vagnbarn* 67 %, vilket är en signifikant skillnad gentemot SVE (90 %, $p=0,026$).

Informanterna i SVE har en signifikant större andel högertolkningar (91 %) än informerterna i START≥6 (59 %) även vid nybildningen *stövelbarn* ($p=0,0001$).

Vid *flodstrand* har SVE-informerterna en signifikant större andel högertolkningar (81 %) än informerterna i START<6 (52 %, $p=0,0025$). Tolkningarna i SVE-gruppen skiljer sig också signifikant från tolkningarna i START≥6 (56 % högertolkningar, $p=0,0096$).

SVE-informerternas tolkningar av *spelkort* skiljer sig signifikant från tolkningarna som informerterna i START≥6 gör. 83 % av informerterna i grupp SVE och 52 % av informerterna i START≥6 gör högertolkningar med *kort* som det semantiska huvudledet ($p=0,0014$).

Vid *husbåt* har informanterna i $START \geq 6$ en betydligt lägre grad av högertolkningar (69 %) än informanterna i SVE (90 %), SVE+ (97 %) och $START < 6$ (94 %). Resultatet för $START \geq 6$ skiljer sig signifikant från resultatet för SVE ($p=0,0079$), SVE+ ($p=0,0064$) och $START < 6$ ($p=0,010$).

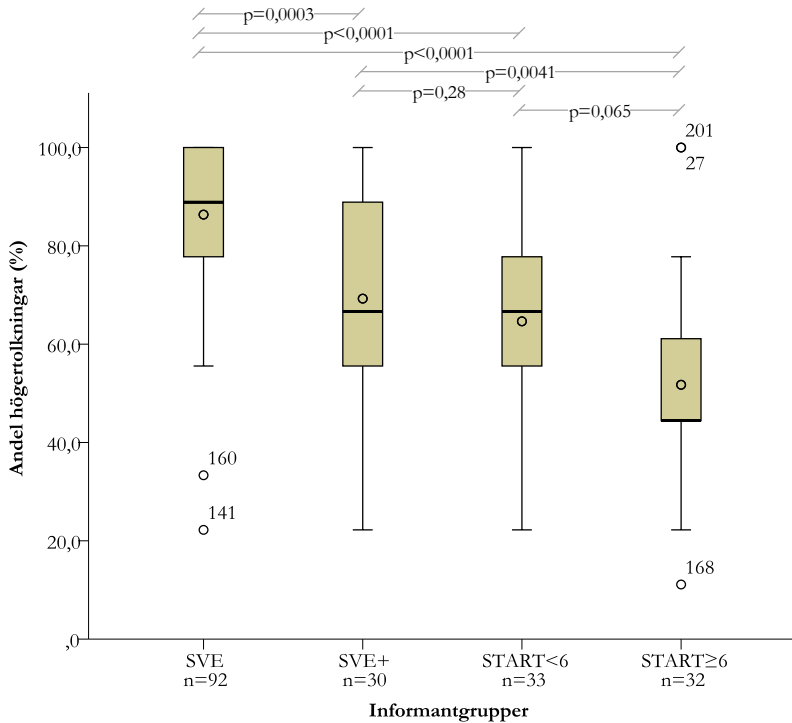
Sammanfattningsvis föreligger signifikanta skillnader mellan informantgrupperna vid alla enskilda sammansättningar. L1-talarna har generellt fler högertolkningar av sammansättningarna än L2-talarna. Störst skillnader mellan L1- och L2-talarna förekommer vid de sammansättningar som är lågfrekventa i referenskorpusen (*båtsommar* och *boktext* är på plats 6 respektive 7 i den frekvensordnade tabellen), även om *båthus* utgör ett undantag. De största skillnaderna föreligger mellan L1-talarna i grupp SVE och L2-talarna i $START \geq 6$. Mellan dessa föreligger signifikanta skillnader vid alla ord ($p < 0,01$). Vid hälften av sammansättningarna skiljer sig SVE-informanternas tolkningar signifikant även från gruppen $START < 6$. Vidare har grupp SVE+ signifikant fler högertolkningar än $START < 6$ och $START \geq 6$ på några sammansättningar. $START < 6$ har en signifikant större andel högertolkningar än gruppen $START \geq 6$ vid två sammansättningar (*båtsommar* och *båthus*). SVE-informanterna har signifikant fler högertolkningar än SVE+ vid tre sammansättningar (*båthus*, *båtsommar* och *vagnbarn*). Om man avgränsar analysen till de informanter i SVE+ som har svenska som L1 via förälders L1 är skillnaden mellan SVE och SVE+/S bara statistiskt signifikant vid *båthus* och *vagnbarn*.

5.1.1.3. Övergripande jämförelse mellan informantgruppernas tolkningar

Av tabell 5:1 framgår att grupp SVE genomgående har en större andel högertolkningar än övriga grupper på de sammansättningar som är ämnade att ge information om etableringen av sammansättningens huvudled. Tolkningarna från gruppen SVE+ och gruppen $START < 6$ följs i stor utsträckning åt och speglar en struktur med huvudledet till höger i lägre grad än SVE-informanternas tolkningar. $START \geq 6$ är den informantgrupp som har minst andel högertolkningar av sammansättningarna i tabell 5:1.

I det följande zoomar jag in högertolkningarna av de nio sammansättningar i flervalsdelen av enkäten som var avsedda att testa identifiering av semantiskt huvudled. I boxplotten i figur 5:1 visas andelen högertolkningar (i procent) per informantgrupp.¹⁰

10 I boxplotten visas kvartil 1 och 3 (boxen), median (strecket), medelvärde (ringen i boxen). Whiskers uppåt och nedåt visar största värdet inom $1.5 * (K3-K1)$ från boxen och uppåt/nedåt. Ringarna utanför whiskers indikerar outliers, dvs. extremvärden.



FIGUR 5:1. *Fördelning av tolkningar med huvudledet till höger, redovisat per informantgrupp.*

Genom en övergripande analys med Kruskal-Wallis påvisas signifikanta skillnader mellan gruppernas medelvärden, när det gäller andelen höger-tolkningar ($p < 0,0001$, $n = 187$). Boxplotten i figur 5:1 visar att resultatet för SVE-gruppen skiljer sig signifikant från resultatet för de övriga grupperna (se p-värden i figuren). Informanterna i gruppen SVE har en signifikant större andel högertolkningar ($M = 86,4$ %, $SD = 16,5$) än informanterna i SVE+ ($M = 69,3$ %, $SD = 23,6$)¹¹, START<6 ($M = 64,6$ %, $SD = 17,9$) och START≥6 ($M = 51,7$ %, $SD = 20,3$). Gruppen SVE+ har en signifikant större andel högertolkningar än gruppen START≥6, men skiljer sig däremot inte signifikant från gruppen START<6. Skillnaden mellan informanterna med låg och hög startålder för svenskan, dvs. informanterna i grupperna

11 Skillnaden mellan informantgruppen SVE och SVE+/S är däremot inte signifikant (se vidare avsnitt 5.1.1.4).

START<6 respektive START≥6, når inte statistisk signifikans.¹² I informantgruppen som helhet (n=187) är medelvärdet 73,9 % (SD=22,9) högertolkningar av NN-sammansättningarna.

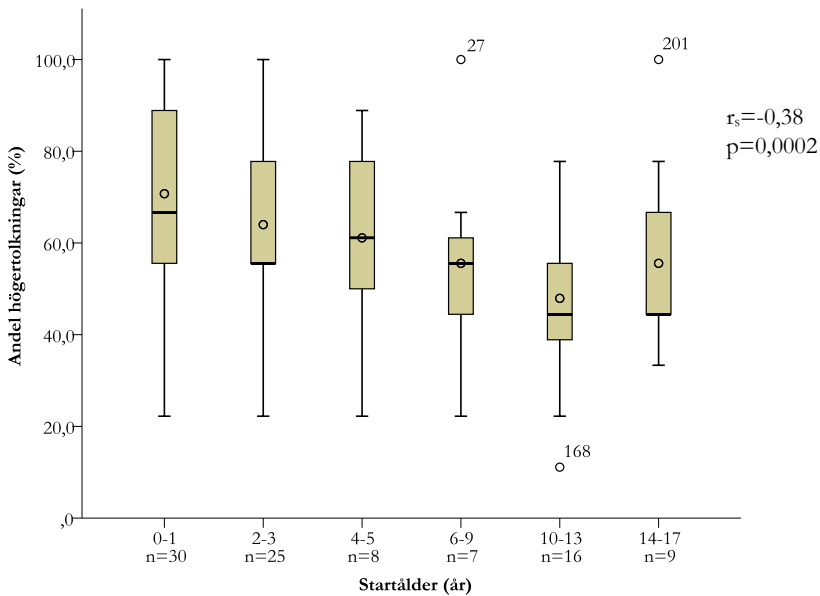
5.1.1.4. *Tolkningar i relation till språkbakgrund i SVE+ och i relation till startålder*

Som framgått uppvisar de L1-talare som har svenska och ytterligare ett modersmål, dvs. informanterna i SVE+, en signifikant lägre grad av högertolkningar än informanterna i SVE, som har svenska som enda modersmål. I avsnitt 4.10.2 diskuterades variationen i flerspråkighetsbakgrund i den informantgrupp som uppger svenska och ytterligare ett modersmål, och undergrupperna SVE+/S och SVE+/A introducerades. Informanterna i SVE+/S har (minst) en förälder som har svenska som L1, till skillnad från informanterna i SVE+/A, vars föräldrar har andra modersmål. I det följande jämförs dessa undergrupper med varandra och med SVE.

I en uppföljande analys med ett (independent samples) Mann-Whitney U-test konstateras inga signifikanta skillnader, när det gäller andelen högertolkningar, mellan SVE+/S (n=15; M=75,6 %, SD=26,0) och SVE+/A (n=15; M=63,0 %, SD=20,0; p=0,15). Samplen är dock små och standardfelet därmed relativt stort. I ett uppföljande Mann-Whitney U-test där medelvärdet för SVE (n=92) M=86,4 % (SD=16,5) jämförs med medelvärdet för SVE+/S (75,6 %) konstateras inga signifikanta skillnader mellan grupperna (p=0,17). Ett motsvarande Mann-Whitney U-test där SVE jämförs med SVE+/A (63,0 % högertolkning) uppvisar däremot signifikanta skillnader mellan grupperna (p<0,0001).

Som diskuterats i metodavsnittet (avsnitt 4.10.2) föreligger en överlappning mellan informanterna i SVE+ och START<6, eftersom två flerspråkiga informanter med samma startålder för svenskan kan hamna i olika informantgrupper beroende på hur de tolkat modersmålsfrågan. För att komma åt variabeln startålder (och i förlängningen exponeringstiden för svenskan) har en uppföljande korrelationsanalys genomförts, med Spearmans korrelationstest, mellan andel högertolkningar och angiven startålder för svenskan. Denna analys är oberoende av informanternas egen modersmålsrapportering. Svaren från alla informanter med flerspråkig bakgrund (dvs. informanter som tillhör någon av informantgrupperna SVE+, START<6 och START≥6) har jämförts enbart efter variabeln startålder. I boxplotten i figur 5:2 åskådliggörs relationen mellan andel högertolkningar och informanternas startålder.

12 De höga extremvärdena i START≥6 kommer från två informanter med modersmål som strukturellt liknar svenska språket, närmare bestämt isländska respektive engelska.



FIGUR 5:2. *Fördelning av tolkningar med huvudet till höger, redovisat efter startålder.*

Korrelationsanalysen påvisar en signifikant korrelation mellan startålder och andel högerställda tolkningar ($r_s = -0,38$, $p = 0,0002$; $n = 95$). För att närmare undersöka betydelsen av startålder inom de lägre startåldrarna genomfördes en korrelationsanalys inom ett mindre åldersspann. För startåldrar lägre än sju år påvisas ett signifikant samband mellan startålder och andel högertolkningar ($r_s = -0,29$, $p = 0,016$; $n = 67$). För startåldrar lägre än sex år påvisas ingen signifikant korrelation (men en icke signifikant trend) mellan startålder och andel högertolkningar ($r_s = -0,24$, $p = 0,058$; $n = 63$). Som framgår av boxplotten, figur 5:2, är fördelningen mellan olika startåldrar mycket ojämn. Det är tänkbart att signifikanta skillnader hade kunnat påvisas också mellan de lägsta startåldrarna om åldersfördelningen varit jämnare och om studien innefattat fler informanter som angett en startålder för svenskan vid fyra eller fem års ålder.

5.1.2. Huvudled i treledade sammansättningar – typen *barnbokshylla* och *sommardagdröm*

I följande avsnitt presenteras svaren på de sammansättningar i enkäten som testar segmenteringen av treledade, nominala sammansättningar med och utan fog. Som redovisats tidigare (avsnitt 2.4.1) finns det en skrivregel om fog framför det tredje ledet i en sammansättning om detta utgör sammansättningens huvudled. Regelen markerar en strukturell och semantisk skillnad mellan å ena sidan *barnbokshylla* {barnbok + s + hylla} 'hylla för barnböcker' och, å den andra, *barnbokhylla* {barn + bokhylla} 'bokhylla för barn'. I det följande undersöks i vilken grad denna regel avspeglar sig i informanternas segmentering av treledade sammansättningar, dvs. om förekomst eller avsaknad av fog påverkar hur de uppfattar det semantiska huvudets räckvidd.

I enkätens flervalsdel ingår fyra treledade sammansättningar. Två av dem förekommer i vardagsspråket, *barnbokhylla* och *barnbokshylla*, och två av dem är tillfälliga bildningar: *sommardagdröm* och *skolboksförsäljning*.

I tabell 5:2 redovisas, för informantgruppen som helhet och per informantgrupp, i vilken mån informanternas segmentering av de treledade sammansättningarna sammanfaller med vad fogeregeln säger om det semantiska huvudets räckvidd i sammansättningar med och utan fog. För varje ord redovisas först andelen tolkningar som överensstämmer med fogeregeln (fetmarkerade i tabellen), följt av andelen tolkningar som avviker från fogeregeln. Sammansättningarna presenteras frekvensordnade efter antal förekomster i referenskorpusen. Alla sammansättningarna är lågfrekventa, och det är bara *barnbokhylla* som finns belagd i referenskorpusen.

5.1.2.1. Tolkningar i informantgruppen som helhet

Tabell 5:2 visar att tolkningarna av de treledade sammansättningarna följer fogeregeln i varierande grad. Sammansättningen *skolboksförsäljning* är den sammansättning som i högst grad (84 %) tolkas enligt fogeregeln (dvs. med *försäljning* som semantiskt huvud). Tolkningarna av *barnbokshylla*, följer fogeregeln i 71 % av fallen (*hylla* som semantiskt huvud). *Sommardagdröm* tolkas i enlighet med fogeregeln och med *dagdröm* som semantiskt huvud i 46 % av fallen. *Barnbokhylla* segmenteras enligt fogeregeln (*barn-bokhylla*) i 40 % av fallen. Denna hierarkiska

TABELL 5:2. Informanternas svar vid treledade sammansättningar med och utan fog.¹³

	Totalt, alla infor- manter n=190	SVE n=94	SVE+ n=30 SVE+/S n=15	START<6 n=34	START≥6 n=32	P-värde över- gripande (overall)	
Barn- bokhylla	Liten hylla som passar i ett barnrum	73 (39 %)	40 (44 %)	12 (40 %) 5 (33 %)	8 (24 %)	13 (42 %)	
	Hylla med barnböcker, t.ex. på ett bibliotek	112 (61 %)	51 (56 %)	18 (60 %) 10 (67 %)	25 (76 %)	18 (58 %)	p=0,26
Barnboks- hylla	Hylla med barnböcker, t.ex. på ett bibliotek	133 (71 %)	67 (73 %)	22 (73 %) 10 (67 %)	24 (73 %)	20 (63 %)	p=0,71
	Liten hylla som passar i ett barnrum	54 (29 %)	25 (27 %)	8 (27 %) 5 (33 %)	9 (27 %)	12 (38 %)	
Skolboks- försälj- ning	Försäljning av skolböcker	155 (84 %)	80 (88 %)	27 (90 %) 15 (100 %)	25 (76 %)	23 (74 %)	p=0,13
	Försäljning av olika typer av böcker, som äger rum i skolan	30 (16 %)	11 (12 %)	3 (10 %) 0 (0 %)	8 (24 %)	8 (26 %)	
Sommar- dagdröm	Dagdröm som man har på sommaren	87 (46 %)	42 (45 %)	12 (40 %) 5 (33 %)	17 (52 %)	16 (50 %)	p=0,79
	Dröm om en sommardag	101 (54 %)	51 (55 %)	18 (60 %) 10 (67 %)	16 (48 %)	16 (50 %)	

13 Övergripande p-värde beräknat med Pearsons exakta chi²-test, parvisa jämförelser beräknade med Fishers exakta test.

* p<0,05, ** p<0,01, *** p<0,001

□ p<0,05, □□ p<0,01, □□□ p<0,001

◇ p<0,05, ◇◇ p<0,01, ◇◇◇ p<0,001

Siffror i kursiv representerar gruppen SVE+/S.

Parvis jämförelse SVE mot var och en av de övriga grupperna

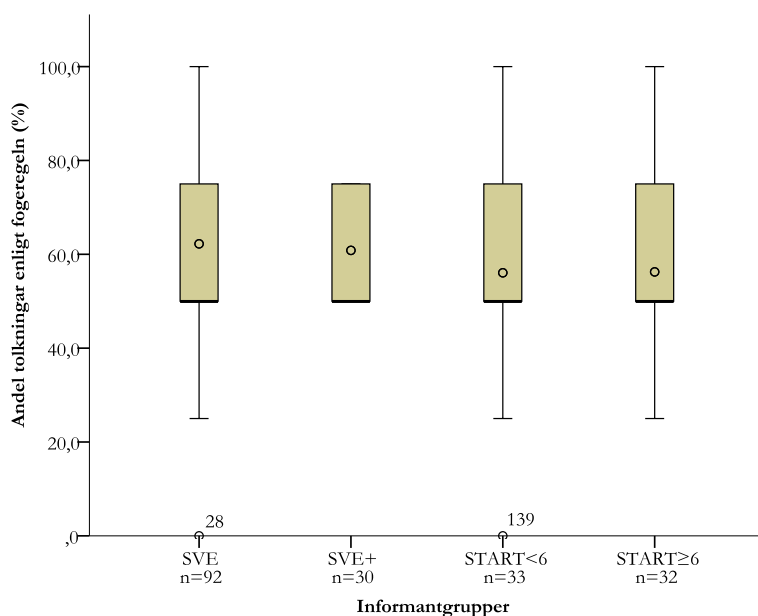
Parvis jämförelse SVE+ mot START<6, START≥6

Parvis jämförelse START<6 mot START≥6

ordning sammanfaller inte med sammansättningarnas frekvensordning i referenskorpusen, där endast *barnbokhylla* finns belagd. Siffrorna visar att informanterna i hög grad segmenterar sammansättningarna med den stora ordgränsen mellan led två och led tre, oavsett om sammansättningen har utsatt fog eller inte. Detta blir tydligt när man jämför ordparet *barnbokhylla*–*barnbokshylla*. 71 % av informanterna tolkar *s*-strukturen *barnbokshylla* som en 'hylla med barnböcker, t.ex. på ett bibliotek'. Så många som 61 % av informanterna väljer dock denna tolkning även vid den *s*-lösa strukturen *barnbokhylla*, där regeln om fog i treledad sammansättning predicerar tolkningen 'liten (bok)hylla som passar i ett barnrum'.

5.1.2.2. Jämförelse mellan informantgruppernas tolkningar

När det gäller de treledade sammansättningarna är medelvärdet för tolkningar som följer fogeregeln 59,9 % (SD=19,6; n=187) för informantgruppen som helhet. Det är endast knappt 10 procentenheter bättre än ett slumpmässigt utfall. Detta tyder på att fogeregeln som ansetts åtskilja sammansättningspar med och utan fog inte i nämnvärd utsträckning aktiveras hos studiens informanter. I figur 5:3 redovisas hur de tolkningar som följer fogeregeln fördelar sig, redovisat per informantgrupp.



FIGUR 5:3. Fördelning av tolkningar i enlighet med fogeregeln, redovisat per informantgrupp.

Ett Kruskal-Wallis-test visar att skillnaderna i medelvärde mellan de olika informantgrupperna i kategorin som helhet inte är statistiskt signifikanta ($p=0,38$). Inget signifikant samband påvisas heller med χ^2 -test mellan informantgruppstillhörighet och svar på de enskilda orden.

5.1.3. Semantisk relation i sammansättningar som betecknar behållare – typen *fruktkorg* och *blomlåda*

Som beskrivs i avsnitt 4.3.1.3 har NN-sammansättningar som betecknar behållare flera potentiella betydelserelationer mellan för- och efterled. Några av dessa är INNEHÅLL, ÄNDAMÅL och MATERIAL. I flervalsdelen av enkäten ingår fem sammansättningar av denna typ: *blomlåda* och *plastlåda* är väl representerade i referenskorpusen, medan *fruktkorg* och *svampkorg* har få förekomster i korpusen. *Olivkorg* utgör en mer tillfällig bildning och saknar representation i referenskorpusen.

Sammansättningarna i denna kategori har tre svarsalternativ, i stället för två som i de tidigare redovisade kategorierna. I tabell 5:3 presenteras hur informantgruppernas svar fördelar sig mellan dessa. Eftersom jag i den övergripande χ^2 -analysen, och i den parvisa jämförelsen med Fishers exakta test, har valt att ställa det alternativ som samlat flest tolkningar mot de övriga två (se vidare avsnitt 5.1.3.2), följer också presentationen i tabellen denna ordning. Det innebär att jag för varje sammansättning först presenterar det alternativ av de semantiska funktionerna INNEHÅLL, ÄNDAMÅL, MATERIAL och DEKORATION, som samlat flest tolkningar, följt av de övriga alternativen. Sammansättningarna i tabellen är ordnade efter grad av förekomst i referenskorpusen.

TABELL 5:3. *Informanternas val av semantisk relation vid sammansättningar som betecknar behållare.*¹⁴

	Totalt, alla informanter n=190	SVE n=94	SVE+ n=30 SVE+/S n=15	START<6 n=34	START≥6 n=32	P-värde över- gripande (overall)
Blom- låda	ÄNDAMÅL	112 (61 %)	64 (70 %)	15 (50 %)	18 (56 %)	15 (47 %)
	Speciell låda för blommor som man kan hänga upp på balkongen				*	*
	INNEHÅLL	57 (31 %)	23 (25 %)	11 (37 %)	10 (31 %)	13 (41 %)
	Låda/kartong som det är blommor i			8 (53 %)		
	DEKORATION	16 (9 %)	4 (4 %)	4 (13 %)	4 (13 %)	4 (13 %)
	Låda som är blom- mig			1 (7 %)		
Plast- låda	MATERIAL	156 (84 %)	88 (96 %)	26 (87 %)	25 (76 %)	17 (57 %)
	Låda som är gjord av plast			14 (93 %)	**	***
	ÄNDAMÅL	17 (9 %)	3 (3 %)	2 (7 %)	5 (15 %)	7 (23 %)
	Speciell låda som man använder för att samla plast i			1 (7 %)		
	INNEHÅLL	12 (6 %)	1 (1 %)	2 (7 %)	3 (9 %)	6 (20 %)
Låda som det ligger plast i			0 (0 %)			

14 Övergripande p-värde beräknat med Pearsons exakta χ^2 -test, parvisa jämförelser beräknade med Fishers exakta test. Två svarsalternativ har slagits samman till ett i den statistiska analysen. Det alternativ som samlat flest tolkningar (rad 1) har ställts mot de övriga två. Signifikansvärdena på rad 1 avser denna analys där alternativet på rad 1 jämförs med de övriga två tillsammans.

* $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$

Parvis jämförelse SVE mot var och en av de övriga grupperna

□ $p < 0,05$, □□ $p < 0,01$, □□□ $p < 0,001$

Parvis jämförelse SVE+ mot START<6, START≥6

◇ $p < 0,05$, ◇◇ $p < 0,01$, ◇◇◇ $p < 0,001$

Parvis jämförelse START<6 mot START≥6

† Den uppmätta signifikansen med Fishers exakta test bör ej räknas som ett signifikant värde, eftersom det övergripande p-värdet (χ^2 overall) inte är signifikant.

Siffror i kursiv representerar gruppen SVE+/S.

	Totalt, alla informanter n=190	SVE n=94	SVE+ n=30 SVE+S n=15	START<6 n=34	START≥6 n=32	P-värde övergripande (overall)	
Frukt- korg	INNEHÅLL	140 (75 %)	75 (82 %)	23 (77 %)	23 (70 %)	19 (59 %)	
	Korg som det ligger frukt i			13 (87 %)		*† p=0,059	
	ÄNDAMÅL	45 (24 %)	16 (18 %)	7 (23 %)	9 (27 %)	13 (41 %)	
	Speciell korg som man använder när man plockar frukt			2 (13 %)			
Svamp- korg	DEKORATION	1 (1 %)	0 (0 %)	0 (0 %)	1 (3 %)	0 (0 %)	
	Korg som är dekorerad med frukter på utsidan			0 (0 %)			
	ÄNDAMÅL	128 (68 %)	69 (74 %)	21 (70 %)	19 (57 %)	19 (61 %)	
	Speciell korg som man använder när man plockar svamp			10 (67 %)			p=0,27
Oliv- korg	INNEHÅLL	49 (26 %)	23 (25 %)	8 (27 %)	9 (27 %)	9 (29 %)	
	Korg som det ligger svamp i			5 (33 %)			
	DEKORATION	10 (5 %)	1 (1 %)	1 (3 %)	5 (15 %)	3 (10 %)	
	Korg som är dekorerad med svampar på utsidan			0 (0 %)			
Oliv- korg	INNEHÅLL	79 (43 %)	38 (41 %)	11 (37 %)	15 (47 %)	15 (52 %)	
	Korg som det ligger oliver i			5 (33 %)			p=0,64
	ÄNDAMÅL	63 (34 %)	31 (34 %)	12 (40 %)	10 (31 %)	10 (34 %)	
	Speciell korg som man använder när man plockar oliver			5 (33 %)			
Oliv- korg	MATERIAL	41 (22 %)	23 (25 %)	7 (23 %)	7 (22 %)	4 (14 %)	
	Korg som är flätad av kvistar från olivträd			5 (33 %)			

5.1.3.1. Tolkningar i informantgruppen som helhet

Av tabell 5:3 framgår att sammansättningarnas frekvens över lag inverkar på graden av samstämmighet. Vid *blomlåda*, som har flest belägg i referenskorpusen, samlar ett av svarsalternativen (ÄNDAMÅL) 60 % av tolkningarna, även om relationen INNEHÅLL också förekommer frekvent i informanternas tolkningar (31 %). *Plastlåda*, som i princip är lika frekvent som *blomlåda* i referenskorpusen, tolkas mest samstämmigt: 84 % av informanterna väljer relationen MATERIAL och tolkar *plastlåda* som en 'låda av plast'. Sammansättningen *fruktkorg* tolkas i hög grad (75 %) som en 'korg med frukt i', dvs. informanterna väljer relationen INNEHÅLL, men även tolkningsalternativet 'korg för fruktplockning' (dvs. relationen ÄNDAMÅL) aktualiseras för en fjärdedel av informanterna. *Svampkorg* tolkas med den semantiska relationen ÄNDAMÅL i 68 % av fallen, medan en dryg fjärdedel av informanterna tolkar sammansättningen med relationen INNEHÅLL. Den tillfälliga bildningen *olivkorg* ger också lägst grad av samstämmighet: 43 % tolkar *olivkorg* som en 'korg som det ligger oliver i' (dvs. INNEHÅLL), 34 % väljer tolkningen 'speciell korg som man använder när man plockar oliver' (dvs. ÄNDAMÅL), medan 22 % väljer tolkningen 'korg som är flätad av kvistar från olivträd' (dvs. MATERIAL).

Sammanfattningsvis är det tydligt att det finns ett prefererat svarsalternativ vid de frekventa sammansättningarna, medan det spretiga resultatet för den tillfälliga bildningen *olivkorg* illustrerar att flera alternativ i meningspotentialen är öppna.

5.1.3.2. Jämförelse mellan informantgruppernas tolkningar

Sammansättningarna som betecknar behållare har tre svarsalternativ. För att göra analysen mer överblickbar har jag slagit samman svarsalternativen till två i de statistiska analyserna. Jag har ställt det svarsalternativ som samlat flest tolkningar mot de övriga två.

För orden *svampkorg* och *olivkorg* påvisas inget signifikant samband mellan vilken informantgrupp man tillhör och hur man tolkar den aktuella sammansättningen ($p=0,27$ respektive $p=0,64$). Skillnaderna vid *frukt-korg* balanserar på gränsen, men når inte statistisk signifikans ($p=0,059$). Vid sammansättningen *blomlåda* föreligger dock signifikanta skillnader ($p=0,049$), liksom vid *plastlåda* ($p<0,0001$).

Av tabell 5:3 framgår att SVE är den grupp som är mest samstämmig i sin tolkning av *blomlåda*, nämligen 70 % tolkningar med ledrelationen ÄNDAMÅL. Resultatet för SVE skiljer sig signifikant från SVE+ (50 % ändamålstolkningar, $p=0,049$) och från $START \geq 6$ (47 % ändamålstolkningar, $p=0,031$). Tolkningarna från informanterna i SVE skiljer sig signifikant också från undergruppen SVE+/S (40 % ändamålstolkningar, $p=0,037$).

Det förekommer också tydliga skillnader mellan grupperna vid *plastlåda*. Informantgruppen SVE har 96 % materialtolkningar, vilket är en signifikant större andel än i gruppen $START < 6$ (76 % materialtolkningar, $p = 0,0025$) och gruppen $START \geq 6$ (57 % materialtolkningar, $p < 0,0001$). Även SVE+ (87 % materialtolkningar) skiljer sig signifikant från $START \geq 6$ ($p = 0,020$).

Sammanfattningsvis har informanterna i gruppen SVE en högre grad av samstämmiga tolkningar än de andra informantgrupperna vid de frekventa sammansättningarna *blomlåda* och *plastlåda*, medan de övriga etablerade sammansättningarna – *fruktkorg* och *svampkorg* – inte ger några signifikanta skillnader mellan grupperna. Inte heller tolkningarna av den tillfälliga bildningen *olivkorg* uppvisar signifikanta mellangrupsliga skillnader.

5.1.4. Jämförande egenskapsprojicering eller tematisk relationslänkning – typen *corpörn* och *citronvatten*

I avsnitt 2.3 presenteras de två konkurrerande tolkningsmönstren egenskapsprojicering och relationslänkning vid nominal sammansättning. Jämförande egenskapsprojicering innebär att en framträdande egenskap hos förledets referent projiceras på huvudledets referent (t.ex. *tigerhaj* = 'randig haj'), dvs. det rör sig om en jämförande relation. Tematisk relationslänkning innebär att det finns någon annan typ av relation mellan leden, där ledens referenter intar olika funktionella roller i relation till varandra.

I den enkät som informanterna besvarade ingår åtta sammansättningar där informanterna kunde välja mellan egenskapsprojicerande och relationslänkande tolkningar. Sex av dessa förekommer i delen med flervalsalternativ. *Citronvatten* och *isvatten* är frekventa sammansättningar. Sammansättningen *apörn* betecknar ett djur som visserligen förekommer i världen, men sammansättningen är lågfrekvent och saknar representation i referenskorpusen. Inte heller *skräpkruka* har några förekomster i referenskorpusen. Detsamma gäller de tillfälliga bildningarna *corpörn* och *zebralejon*.

I tabell 5:4 redovisas för informantgruppen som helhet ($n = 190$), och per informantgrupp, hur svaren på de enskilda sammansättningarna fördelar sig mellan egenskaps- och relationstolkning. För varje ord redovisas först andelen egenskapsprojicerande och jämförande tolkningar, t.ex. 'vatten som smakar surt som citron', och därefter andelen relationslänkande tolkningar, t.ex. 'vatten med citron i'. Sammansättningarna i tabellen är ordnade efter deras frekvens i referenskorpusen, i en fallande skala från sammansättningen med högst till sammansättningen med lägst antal belägg i korpusen.

TABELL 5:4. Informanternas svar på sammansättningar som testar valet mellan egenskapsprojicering och relationslänkning.¹⁵

	Totalt, alla informanter n=190	SVE n=94	SVE+ n=30 SVE+/S n=15	START<6 n=34	START≥6 n=32	P-värde övergripande (overall)	
Citron- vatten	EGENSKAP Vatten som smakar surt som citron	18 (10 %)	5 (5 %)	0 (0 %) 0 (0 %)	7 (21 %) *	6 (19 %) *	p=0,0042
	RELATION Vatten med citron i	171 (90 %)	89 (95 %)	30 (100 %) 15 (100 %)	27 (79 %) *	25 (81 %) *	
Is- vatten	EGENSKAP Vatten som är kallt som is	42 (23 %)	14 (16 %)	5 (17 %) 2 (13 %)	9 (27 %)	14 (48 %) ***	p=0,0024
	RELATION Vatten med is i	140 (77 %)	76 (84 %)	25 (83 %) 13 (87 %)	24 (73 %)	15 (52 %) ***	
Apörn	EGENSKAP Örn som låter eller tjattrar som en apa	124 (67 %)	62 (67 %)	22 (73 %) 14 (93 %)	22 (67 %)	18 (60 %)	p=0,76
	RELATION Örn som jagar apor	61 (33 %)	30 (33 %)	8 (27 %) 1 (7 %)	11 (33 %)	12 (40 %)	

15 Övergripande p-värden har beräknats med Pearsons exakta chi²-test, parvisa jämförelser har genomförts med Fishers exakta test.

* p<0,05, ** p<0,01, *** p<0,001

Parvis jämförelse SVE mot var och en av de övriga grupperna

□ p<0,05, □□ p<0,01, □□□ p<0,001

Parvis jämförelse SVE+ mot START<6, START≥6

◇ p<0,05, ◇◇ p<0,01, ◇◇◇ p<0,001

Parvis jämförelse START<6 mot START≥6

† Den uppmätta signifikansen med Fishers exakta test bör ej räknas som ett signifikant värde, eftersom det övergripande p-värdet (chi² overall) inte är signifikant.

Siffror i kursiv representerar gruppen SVE+/S.

		Totalt, alla informanter n=190	SVE n=94	SVE+ n=30 SVE+/S n=15	START<6 n=34	START≥6 n=32	P-värde övergripande (overall)
Skräpkruka	EGENSKAP (Blom)-kruka som inte är så fin utan ser lite skräpig ut	92 (50 %)	52 (57 %)	10 (34 %) 7 (47 %)	19 (59 %)	11 (35 %)	p=0,043
	RELATION Kruka med skräp i	92 (50 %)	40 (43 %)	19 (66 %) 8 (53 %)	13 (41 %)	20 (65 %)	
Zebra-lejon	EGENSKAP Lejon som är randigt som en zebra	162 (87 %)	83 (90 %)	27 (90 %) 13 (87 %)	28 (85 %)	24 (75 %) *†	p=0,16
	RELATION Lejon som äter zebror	25 (13 %)	9 (10 %)	3 (10 %) 2 (13 %)	5 (15 %)	8 (25 %) *†	
Korp-örn	EGENSKAP Örn som är svart som en korp	127 (71 %)	68 (77 %)	18 (62 %) 9 (60 %)	23 (74 %)	18 (58 %)	p=0,14
	RELATION Örn som jagar korp-fåglar	52 (29 %)	20 (23 %)	11 (38 %) 6 (40 %)	8 (26 %)	13 (42 %)	

5.1.4.1. Tolkningar i informantgruppen som helhet

Tabell 5:4 visar att det råder stor samstämmighet i informanternas tolkningar av de frekventa sammansättningarna *citronvatten* och *isvatten*. I informantgruppen som helhet ges sammansättningarna 90 respektive 77 % relationslänkande tolkningar. Lägst grad av samstämmighet noteras för de mindre frekventa sammansättningarna *skräpkruka* och *apörn*. Vid *skräpkruka* fördelar sig informanternas svar lika mellan den egenskapsprojicerande parafrasen '(blom)kruka som inte är så fin utan ser lite skräpig ut' och den relationslänkande 'kruka med skräp i'. *Apörn* ges en stor andel jämförande egenskapstolkningar, dvs. 'örn som låter eller tjattrar som en apa' (67 %). De övriga djursammansättningarna, *zebralejon* och *korpörn*, är tillfälliga bildningar som betecknar fenomen som torde vara obekanta för informanterna. Vid dessa väljer en majoritet av informanterna det egenskapsprojicerande alternativet. I informantgruppen som helhet

är det 86 % som anser att ett *zebralejon* är ett 'randigt lejon', och 71 % av informanterna uppger att en *korpörn* är en 'svart örn'.

Sammanfattningsvis ger de frekventa sammansättningarna samstämmiga tolkningar, men också de tillfälliga bildningarna av två djurlad tenderar att tolkas relativt samstämmigt, och egenskapsprojicerande. Det egenskapsprojicerande och jämförande mönstret slår igenom starkt även vid den tredje, lågfrekventa djursammansättningen *apörn*.

5.1.4.2. Jämförelse mellan informantgruppernas tolkningar

För orden i tabell 5:4 påvisas ett signifikant samband mellan informantgrupp och svar vid de frekventa sammansättningarna *citronvatten* ($p=0,0042$) och *isvatten* ($p=0,0024$) samt vid den tillfälliga sammansättningen *skräpkruka* ($p=0,043$). Inga signifikanta skillnader mellan grupperna påvisas vid den lågfrekventa sammansättningen *apörn* eller de tillfälliga bildningarna *zebralejon* och *korpörn*.

Vid *isvatten* väljer 84 % av informanterna i grupp SVE det relationslänkande alternativet 'vatten med is i', liksom 83 % av informanterna i grupp SVE+, medan motsvarande siffra är 73 % för gruppen $START < 6$ och 52 % för $START \geq 6$. Gruppen SVE har en signifikant större andel relationslänkande tolkningar än gruppen $START \geq 6$ ($p=0,0007$). Även SVE+ skiljer sig signifikant från $START \geq 6$ ($p=0,013$).

Vid *citronvatten* har de båda L1-grupperna en större andel relationslänkande tolkningar än de båda L2-grupperna. Svaren från informantgruppen SVE (95 % relationslänkande tolkningar) skiljer sig signifikant från svaren i gruppen $START < 6$ (79 % relationslänkning, $p=0,015$) och $START \geq 6$ (81 % relationslänkning, $p=0,027$). Även svaren från SVE+ skiljer sig signifikant från L2-gruppernas svar. Samtliga informanter i SVE+ gör en relationstolkning av *citronvatten*, och gruppens tolkningar skiljer sig därmed signifikant från de tolkningar som görs av informanterna i $START < 6$ ($p=0,012$) och informanterna i $START \geq 6$ ($p=0,024$).

För *skräpkruka* uppvisas också ett signifikant samband mellan informantgrupp och svar i den överordnade χ^2 -analysen. Tabell 5:4 visar att informanterna i grupperna SVE och $START < 6$ har 57 % respektive 59 % egenskapsprojicerande tolkningar, att jämföra med 35 % egenskapsprojicerande tolkningar bland informanterna i $START \geq 6$ och 34 % egenskapsprojicering i SVE+, men i de uppföljande, parvisa jämförelserna med Fishers Exakt nås inte statistisk signifikans.

Sammanfattningsvis tolkas de förhållandevis frekventa sammansättningarna *citronvatten* och *isvatten* samstämmigt och relationslänkande i högre grad av L1-informanterna än av L2-informanterna. Samstämmigheten är begränsad i alla informantgrupper vid den lågfrekventa sammansättningen *skräpkruka*, men en icke signifikant tendens är att den

tolkas mer samstämmigt, och egenskapsprojicerande, av informanterna i SVE och START<6 än av informanterna i SVE+ och START≥6. De minst frekventa sammansättningarna – *apörn*, *corpörn* och *zebralejon* – uppvisar inga signifikanta skillnader mellan grupperna. Dessa djursammansättningar tolkas i hög grad egenskapsprojicerande, dvs. jämförande.

5.1.5. Identitet/jämförande egenskapstolkning eller tematisk relationslänkning – typen *kvinnopräst* och *herrfrisör*

I enkäten ingår en grupp av sammansättningar som är avsedda att ge information om hur informanterna väljer mellan å ena sidan identitets- eller jämförande egenskapstolkning, å andra sidan relationslänkande tolkning av personbetecknande sammansättningar.

I flervalsdelen ingår sex sammansättningar som skulle kunna ges identitets- eller egenskapstolkningar – 'en Y som är (som en) X' – eller tolkas med tematisk relationslänkning – 'en Y som har med X att göra'. Sammansättningarna *kvinnopräst*, *kvinnokarl* och *herrfrisör* är sparsamt belagda i referenskorpusen, och *mansläkare* saknar belägg. *Flickpojke* och *manspräst* är mer att betrakta som tillfälliga bildningar. *Flickpojke* har två träffar i referenskorpusen, medan *manspräst* inte förekommer alls.

I den följande resultatpresentationen i tabell 5:5 redovisas för varje sammansättning först andelen identitetstolkning (t.ex. 'präst som är kvinna') eller egenskapstolkning (t.ex. 'kvinnlig man'), och därefter andelen relationslänkande tolkningar (t.ex. 'präst som tar emot och hjälper kvinnor'). Sammansättningarna presenteras i en ordningsföljd som speglar deras frekvens i referenskorpusen.

TABELL 5:5. Informanternas svar på sammansättningar som testar valet mellan identitets-/egenskapstolkning och relationslänkning.¹⁶

	Totalt, alla informanter n=190	SVE n=94	SVE+ n=30 SVE+/S n=15	START<6 n=34	START≥6 n=32	P-värde övergripande (overall)
Kvinno- präst	IDENTITET Präst som är kvinna	141 (75 %)	76 (83 %)	23 (77 %) 11 (73 %)	26 (79 %)	16 (50 %) *** □ ◇
	RELATION Präst som tar emot och hjälper kvinnor	46 (25 %)	16 (17 %)	7 (23 %) 4 (27 %)	7 (21 %)	16 (50 %) *** □ ◇
Kvinno- karl	EGENSKAP Kvinnlig man	41 (22 %)	14 (15 %)	6 (20 %) 2 (13 %)	11 (33 %) *†	10 (31 %)
	RELATION Man som är förtjust i och har många kärleksrela- tioner med kvinnor	146 (78 %)	78 (85 %)	24 (80 %) 13 (87 %)	22 (67 %) *†	22 (69 %) ***
Herr- frisör	IDENTITET Frisör som är man	72 (38 %)	40 (43 %)	9 (30 %) 5 (33 %)	16 (47 %)	7 (22 %) *† ◇†
	RELATION Frisör som bara klipper män	116 (62 %)	52 (57 %)	21 (70 %) 10 (67 %)	18 (53 %)	25 (78 %) *† ◇†

16 Övergripande p-värde beräknat med Pearsons exakta chi²-test, parvisa jämförelser beräknade med Fishers exakta test.

* p<0,05, ** p<0,01, *** p<0,001

□ p<0,05, □□ p<0,01, □□□ p<0,001

◇ p<0,05, ◇◇ p<0,01, ◇◇◇ p<0,001

† Den uppmätta signifikansen med Fishers exakta test bör ej räknas som ett signifikant värde, eftersom det övergripande p-värdet (chi² overall) inte är signifikant.

Siffror i kursiv representerar gruppen SVE+/S.

	Totalt, alla informanter n=190	SVE n=94	SVE+ n=30 SVE+S n=15	START<6 n=34	START≥6 n=32	P-värde övergripande (overall)
Flickpojke	EGENSKAP Pojke som har intressen som anses vara "flickiga"	161 (87 %)	85 (92 %)	25 (89 %) 12 (86 %)	29 (85 %)	22 (69 %) **
	RELATION Pojke som gillar flickor	25 (13 %)	7 (8 %)	3 (11 %) 2 (14 %)	5 (15 %)	10 (31 %) **
Mansläkare	IDENTITET Läkare som är man	92 (49 %)	51 (55 %)	15 (50 %) 6 (40 %)	16 (48 %)	10 (31 %) *†
	RELATION Läkare som arbetar med sjukdomar som bara män kan få	95 (51 %)	41 (45 %)	15 (50 %) 9 (60 %)	17 (52 %)	22 (69 %) *†
Manspräst	IDENTITET Präst som är man	133 (72 %)	76 (83 %)	21 (70 %) 11 (73 %)	20 (61 %) *	16 (52 %) **
	RELATION Präst som tar emot och hjälper män	53 (28 %)	16 (17 %)	9 (30 %) 4 (27 %)	13 (39 %) *	15 (48 %) **

5.1.5.1. Tolkningar i informantgruppen som helhet

Av tabell 5:5 framgår att *kvinnopräst* och *kvinnokarl* ges relativt samstämmiga tolkningar, trots att dessa sammansättningar har låg representation i referenskorpusen. 75 % av informanterna tolkar *kvinnopräst* med en identitetsrelation mellan leden, och 78 % av informanterna gör en relationslänkande tolkning av *kvinnokarl*. *Herrfrisör*, som har enstaka belägg i referenskorpusen, tolkas relationslänkande i 62 % av fallen. Vid sammansättningen *mansläkare*, som inte är belagd i referenskorpusen, förefaller båda alternativen lika öppna, sett till tolkningarna från informantgruppen som helhet. Svaren fördelar sig jämnt mellan identitetstolkning och relationslänkande tolkning. *Flickpojke* och *manspräst* tolkas samstämmigt, med en jämförande egenskapstolkning respektive identitetstolkning, av så många som 87 % respektive 72 % av informanterna, trots att det rör sig om tillfälliga bildningar.

Det föreligger ett signifikant samband mellan hur *kvinnopräst* och *manspräst* tolkas i informantgruppen som helhet, som framgår av tabell 5:6.

TABELL 5:6. *Samband mellan identitetstolkningar av kvinnopräst och manspräst i informantgruppen som helhet.*

	Kvinnopräst IDENTITET	Kvinnopräst RELATION	P-värde Fishers exakta test
Manspräst IDENTITET	122 (87 %)	11 (24 %)	p<0,0001
Manspräst RELATION	18 (13 %)	35 (76 %)	

Av tabell 5:6 framgår att den grupp av informanter som väljer en identitetstolkning av *kvinnopräst* ('präst som är kvinna') har en signifikant större andel identitetstolkningar av *manspräst* (87 %) än den grupp som väljer en relationslänkande tolkning av *kvinnopräst* (24 % identitetstolkning av *manspräst*, $p < 0,0001$, $n = 186$).

Sammanfattningsvis när det gäller de personbetecknande sammansättningarna får de mest frekventa sammansättningarna i tabell 5:5 också förhållandevis samstämmiga tolkningar, medan spridningen är stor vid tolkningen av de lågfrekventa sammansättningarna *herrfrisör* och *mansläkare*. Vid de tillfälliga bildningarna *flickpojke* och *manspräst* är samstämmigheten ändå stor, och jämförande egenskaps- respektive identitetstolkning slår igenom starkt. *Manspräst* tolkas i hög grad i analogi med *kvinnopräst*, som framgår av tabell 5:6.

5.1.5.2. Jämförelse mellan informantgruppernas tolkningar

För orden i tabell 5:5 påvisas ett signifikant samband mellan informantgrupp och svar vid orden *kvinnopräst* ($p = 0,0028$), *manspräst* ($p = 0,0034$) och *flickpojke* ($p = 0,0080$), dvs. vid dessa ord föreligger signifikanta skillnader mellan informantgruppernas svar. Vid *kvinnokarl*, *herrfrisör* och *mansläkare* föreligger inget signifikant samband mellan informantgrupp och svar.

Störst skillnader mellan informantgrupperna föreligger vid sammansättningen *kvinnopräst* där differensen mellan högsta och lägsta värde för identitetsrelation är 33 procentenheter, följt av *manspräst* med 31 procentenheters motsvarande differens.

Informanterna i grupperna SVE, SVE+ och START<6 har en väsentligt större andel identitetstolkningar av *kvinnopräst* (83, 77 respektive 79 %) än informanterna i START≥6 (52 %). Resultatet för START≥6 skiljer sig signifikant från resultatet för SVE ($p = 0,0007$), SVE+ ($p = 0,038$) och START<6 ($p = 0,020$).

Den tillfälliga bildningen *manspräst* tolkas med en identitetsrelation mellan leden i högre grad av L1-informanterna i SVE (83 %) än av L2-informanterna i START<6 (61 %) och START≥6 (52 %). Skillnaden mellan SVE och START<6 är signifikant ($p=0,016$), liksom skillnaden mellan SVE och START≥6 ($p=0,0014$).

De enspråkiga L1-informanterna (grupp SVE) har signifikant fler jämförande egenskapstolkningar av *flickpojke* (92 %) än informanterna i START≥6 (69 %, $p=0,0019$).

I samtliga informantgrupper föreligger ett signifikant samband mellan hur man tolkar *kvinnopräst* och *manspräst*. De informanter som väljer en identitetstolkning av *kvinnopräst* ('präst som är kvinna') har en signifikant större andel identitetstolkningar av *manspräst* än de informanter som väljer en relationslänkande tolkning av *kvinnopräst*, som framgår av tabell 5:7.

TABELL 5:7. Samband mellan identitetstolkningar av kvinnopräst och manspräst, redovisat per informantgrupp.

Informantgrupp		Kvinnopräst IDENTITET	Kvinnopräst RELATION	P-värde Fishers exakta test
SVE	Manspräst IDENTITET	70 (92 %)	6 (37,5 %)	$p<0,0001$
	Manspräst RELATION	6 (8 %)	10 (62,5 %)	
SVE+	Manspräst IDENTITET	21 (91 %)	0 (0 %)	$p<0,0001$
	Manspräst RELATION	2 (9 %)	7 (100 %)	
START<6	Manspräst IDENTITET	19 (73 %)	1 (14 %)	$p=0,0084$
	Manspräst RELATION	7 (27 %)	6 (86 %)	
START≥6	Manspräst IDENTITET	12 (80 %)	4 (25 %)	$p=0,0038$
	Manspräst RELATION	3 (20 %)	12 (75 %)	

I tabell 5:7 åskådliggörs signifikansen mellan andelen identitetstolkningar av den etablerade sammansättningen *kvinnopräst* och andelen identitetstolkningar av den tillfälliga bildningen *manspräst* i samtliga informantgrupper. Sambandet är något starkare i L1-grupperna (SVE och SVE+)

än i L2-grupperna (START<6 och START≥6). Informanterna tolkar med andra ord *manspräst* i analogi med *kvinnopräst*.

Sammanfattningsvis påvisas i kategorin av personbetecknande sammansättningar signifikanta skillnader mellan grupperna vid tolkningen av den etablerade sammansättningen *kvinnopräst* och vid tolkningen av de tillfälliga bildningarna *manspräst* och *flickpojke*, som ges en högre grad av identitets- respektive egenskapstolkningar i L1-grupperna än i L2-grupperna (se tabell 5:5 ovan). Inga signifikanta skillnader mellan informantgrupperna påvisas vid de etablerade (men mindre frekventa) sammansättningarna *kvinnokarl*, *herrfrisör* och *mansläkare*. I samtliga grupper tolkar informanterna *manspräst* i analogi med *kvinnopräst* (se tabell 5:7).

5.1.6. Agens vid verbavledning – typen *polismisshandel* och *studentförakt*

Vid verbavledda sammansättningar med personbetecknande förled kan förledet tolkas som PATIENT eller AGENT (se vidare i avsnitt 4.3.1.6).

Fem sammansättningar i flervalsdelen av enkäten är avsedda att testa tolkningen av agens vid verbavledning. *Kvinnomisshandel* och *polismisshandel* finns belagda i referenskorpusen, medan *mansmisshandel* och *kvinnovård* saknas. *Studentförakt* är en tillfällig bildning som inte heller förekommer i referenskorpusen.

I tabell 5:8 redovisas fördelningen av informanternas svar på de sammansättningar som testar agens vid verbavledning. För varje sammansättning redovisas först andelen patienttolkningar och därefter andelen agenttolkningar. Sammansättningarna är ordnade efter deras frekvens i referenskorpusen.

5.1.6.1. Tolkningar i informantgruppen som helhet

Av tabell 5:8 kan man utläsa att för samtliga sammansättningar är patienttolkningen prefererad framför agenttolkningen. Den mest frekventa sammansättningen, *kvinnomisshandel*, får mest samstämmiga tolkningar, 93 % patienttolkningar i informantgruppen som helhet. Samstämmigheten är dock stor även vid de mindre frekventa sammansättningarna *mansmisshandel* och *kvinnovård*, vilka i hög grad tolkas med förledet som PATIENT (85 respektive 91 % patienttolkningar). Även den tillfälliga bildningen *studentförakt* ges en patienttolkning i 85 % av fallen. För *polismisshandel*, som bara har ett belägg i referenskorpusen, fördelar sig tolkningarna jämnt mellan agenttolkning och patienttolkning, så båda alternativen förefaller lika relevanta för informantgruppen.

TABELL 5:8. Informanternas svar på verbavledda NN-sammansättningar som testar tolkningen av agens.¹⁷

		Totalt, alla informanter n=190	SVE n=94	SVE+ n=30 SVE+/S n=15	START<6 n=34	START≥6 n=32	P-värde övergripande (overall)
Kvinnomisshandel	PATIENT När kvinnor blir miss-handlade	177 (93 %)	93 (99 %)	28 (93 %) 14 (93 %)	28 (82 %) **	28 (88 %) *	p=0,0058
	AGENT När kvinnor misshandlar	13 (7 %)	1 (1 %)	2 (7 %) 1 (7 %)	6 (18 %) **	4 (12 %) *	
Polismisshandel	PATIENT När polisen blir miss-handlad	93 (50 %)	55 (60 %)	11 (37 %) * 6 (40 %)	15 (44 %)	12 (39 %) *	p=0,043
	AGENT När polisen slår eller misshandlar en person	93 (50 %)	36 (40 %)	19 (63 %) * 9 (60 %)	19 (56 %)	19 (61 %) *	
Mansmisshandel	PATIENT När män blir misshandlade	156 (85 %)	87 (96 %)	24 (83 %) * 13 (93 %)	26 (79 %) **	19 (61 %) ***	p<0,0001
	AGENT När män misshandlar	28 (15 %)	4 (4 %)	5 (17 %) * 1 (7 %)	7 (21 %) **	12 (39 %) ***	
Kvinnovård	PATIENT Vård som kvinnor får	168 (91 %)	86 (96 %)	27 (90 %) 13 (87 %)	29 (88 %)	26 (84 %) *†	p=0,19
	AGENT Vård som kvinnor ger	16 (9 %)	4 (4 %)	3 (10 %) 2 (13 %)	4 (12 %)	5 (16 %) *†	

17 Övergripande p-värde beräknat med Pearsons exakta chi²-test, parvisa jämförelser beräknade med Fishers exakta test.

* p<0,05, ** p<0,01, *** p<0,001

Parvis jämförelse SVE mot var och en av de övriga grupperna

□ p<0,05, □□ p<0,01, □□□ p<0,001

Parvis jämförelse SVE+ mot START<6, START≥6

◇ p<0,05, ◇◇ p<0,01, ◇◇◇ p<0,001

Parvis jämförelse START<6 mot START≥6

† Den uppmätta signifikansen med Fishers exakta test bör ej räknas som ett signifikant värde, eftersom det övergripande p-värdet (chi² overall) inte är signifikant.

Siffror i kursiv representerar gruppen SVE+/S.

5 | Tolkningar i enkätens flervalsdel

	Totalt, alla informanter n=190	SVE n=94	SVE+ n=30 SVE+/S n=15	START<6 n=34	START≥6 n=32	P-värde övergripande (overall)
PATIENT Förakt som är riktat mot	156 (85 %)	83 (91 %)	23 (77 %)	25 (81 %)	25 (81 %)	p=0,14
Student- studenter			13 (87 %)			
förakt						
AGENT Förakt som studenter visar	27 (15 %)	8 (9 %)	7 (23 %)	6 (19 %)	6 (19 %)	
			2 (13 %)			

Sammanfattningsvis är de enskilda sammansättningarnas frekvens inte avgörande för graden av samstämmighet vid sammansättningarna som testar bedömningen av agens. Mönstret patienttolkning slår igenom starkt vid alla orden utom *polismisshandel*, där båda de tolkningsperspektiv som presenterats förefaller lika tillgängliga för informanterna.

5.1.6.2. Jämförelse mellan informantgruppernas tolkningar

Inget signifikant samband mellan informantgrupp och svar påvisas vid orden *kvinnovård* och *studentförakt*, vilka har en stor andel patienttolkningar i samtliga grupper. Signifikanta skillnader föreligger dock mellan grupperna vid *kvinnomisshandel* ($p=0,0058$), *mansmisshandel* ($p<0,0001$) och *polismisshandel* ($p=0,043$).

Vid sammansättningen *kvinnomisshandel* väljer informanterna med svenska som L1 i högre grad än informanterna med svenska som L2 tolkningen där kvinnan utgör verbhandlingens PATIENT: i gruppen SVE 99 % och i gruppen SVE+ 93 %. Motsvarande siffra är 82 % för informanterna i gruppen START<6 och 88 % för START≥6. SVE skiljer sig signifikant från START<6 ($p=0,0014$) och START≥6 ($p=0,015$).

Vid *mansmisshandel* anger 96 % av informanterna i gruppen SVE en patientrelation, 83 % i SVE+, 79 % i START<6 och 61 % i START≥6. Resultatet för informanterna i SVE skiljer sig signifikant från resultatet för START<6 ($p=0,0078$) och START≥6 ($p<0,0001$). SVE skiljer sig signifikant även från SVE+ ($p=0,037$), men skillnaden är inte signifikant om analysen avgränsas till de informanter i SVE+ som har svenska som L1 via förälders modersmål, dvs. undergruppen SVE+/S (93 % patienttolkning, $p=0,52$).

Signifikanta skillnader mellan grupperna förekommer också vid *polismisshandel*, där informanterna i grupp SVE utmärker sig. Bara 40 % av

informerarna i grupp SVE väljer det konventionaliserade alternativet, dvs. parafrazen som pekar ut polisen som agent, vilket är en signifikant lägre andel än i gruppen SVE+ (63 %, $p=0,034$) och i gruppen $START \geq 6$ (61 %, $p=0,040$). I gruppen $START < 6$ är det 56 % som väljer agenttolkningen. Informanterna i SVE+/S följer med sina 60 % agenttolkningar de övriga flerspråkiga informanterna, men på grund av liten sampelstorlek ($n=15$) skiljer sig resultatet för SVE+/S inte signifikant från resultatet för SVE ($p=0,17$).

Sammanfattningsvis föreligger det inga signifikanta skillnader mellan grupperna i deras tolkningar av *kvinnovård* och den tillfälliga bildningen *studentförakt*, vilka båda i hög grad tolkas med förledet som patient. Den förhållandevis frekventa sammansättningen *kvinnomisshandel* och den lågfrekventa sammansättningen *mansmisshandel* tolkas i hög grad med förledet som patient i alla grupper, men andelen patienttolkningar är högre i de grupper som har svenska som L1 än i de grupper som har svenska som L2. Vid *polismisshandel* har grupp SVE en signifikant lägre grad av konventionella agenttolkningar än SVE+ och $START \geq 6$.

5.1.7. Betydelsesammansmälta AN-sammansättningar – typen *grönyta* och *småkakor*

I avsnitten 3.2.2 och 4.3.1.7 redogörs för AN-sammansättningar som konventionaliserats i en mer specifik betydelse än vad leden antyder, t.ex. *blåbär* vars betydelse är mer specifik än vad parafrazen 'bär som är blått' anger. Bakken kallar sådana sammansättningar icke-kompositionella och betydelsesammansmälta. Fyra sammansättningar av detta slag testas i flervalsdelen av enkäten: *sötsaker*, *småkakor*, *finkkläder* samt *grönyta*. Dessa har god representation i referenskorpusen.

I tabell 5:9 redovisas hur informanternas svar i denna kategori fördelar sig mellan de specifika förklaringarna (fetmarkerade i tabellen) och de strikt kompositionella som utgör distraktorer. Sammansättningarna presenteras i en ordning som speglar deras frekvens i referenskorpusen, i en fallande skala från sammansättningen med högst till sammansättningen med lägst antal förekomster.

TABELL 5:9. Informanternas svar vid betydelsesammansmälta AN-sammansättningar.¹⁸

	Totalt, alla informanter n=190	SVE n=94	SVE+ n=30 SVE+S n=15	START<6 n=34	START≥6 n=32	P-värde övergripande (overall)	
Sötsaker	Söta saker som man kan äta	154 (83 %)	86 (92 %)	25 (86 %) 12 (80 %)	26 (79 %) *	17 (57 %) *** □	p=0,0001
	Saker som är söta	31 (17 %)	7 (8 %)	4 (14 %) 3 (20 %)	7 (21 %) *	13 (43 %) *** □	
Småkakor	Kakor som är små och hårda	121 (64 %)	75 (80 %)	20 (69 %) 11 (79 %)	17 (50 %) **	9 (28 %) *** □□	p<0,0001
	Kakor som är små	68 (36 %)	19 (20 %)	9 (31 %) 3 (21 %)	17 (50 %) **	23 (72 %) *** □□	
Finkläder	Kläder som passar vid finare, högtidligare tillfällen	156 (84 %)	86 (95 %)	28 (93 %) 15 (100 %)	27 (82 %)	15 (47 %) *** □□□ ◇◇	p<0,0001
	Kläder som är fina	30 (16 %)	5 (5 %)	2 (7 %) 0 (0 %)	6 (18 %)	17 (53 %) *** □□□ ◇◇	
Grönyta	Naturområde med stor gräsplan	132 (71 %)	75 (81 %)	17 (57 %) *	20 (61 %) *	20 (65 %)	p=0,023
	Yta som är grön	55 (29 %)	18 (19 %)	13 (43 %) *	13 (39 %) *	11 (35 %)	
				7 (47 %) *			

18 Övergripande p-värde beräknat med Pearsons exakta chi²-test, parvisa jämförelser beräknade med Fishers exakta test.

* p<0,05, ** p<0,01, *** p<0,001

□ p<0,05, □□ p<0,01, □□□ p<0,001

◇ p<0,05, ◇◇ p<0,01, ◇◇◇ p<0,001

Siffror i kursiv representerar gruppen SVE+S.

Parvis jämförelse SVE mot var och en av de övriga grupperna

Parvis jämförelse SVE+ mot START<6, START≥6

Parvis jämförelse START<6 mot START≥6

5.1.7.1. Tolkningar i informantgruppen som helhet

Av tabell 5:9 framgår att de specificerande parafraserna är prefererade över de strikt kompositionella i informantgruppen som helhet. Mest samstämmiga tolkningar ges de frekvent förekommande sammansättningarna *finkläder* och *sötsaker*, där det specificerade alternativet samlar 84 % respektive 83 % av informanternas svar. *Grönyta*, som har förhållandevis få belägg i referenskorpusen, samlar 71 % tolkningar som speglar en specificerad relation mellan leden, snarare än en strikt kompositionell relation. Sammansättningen *småkakor* har förvånansvärt god representation i referenskorpusen men samstämmigheten i tolkningarna är något lägre. I informantgruppen som helhet är det 64 % som väljer det specifika alternativet 'kakor som är små och hårda'.

5.1.7.2. Jämförelse mellan informantgruppernas tolkningar – enskilda sammansättningar

Signifikanta skillnader mellan informantgruppernas svar påvisas vid alla betydelsesammansmälta AN-sammansättningar i enkäten ($p < 0,0001$ för *finkläder*; $p = 0,0001$ för *sötsaker*; $p = 0,023$ för *grönyta*, $p < 0,0001$ för *småkakor*).

Som framgår av tabell 5:9 väljer L1-informanterna i högre grad än L2-informanterna de mer specificerade tolkningsalternativen. Vidare har L2-informanterna i gruppen $START < 6$ en högre grad av specifika tolkningar än L2-informanterna i gruppen $START \geq 6$. Tydligast samband mellan informantgrupp och svar föreligger vid sammansättningarna *småkakor* och *finkläder*, där differensen mellan högsta och lägsta andel specifika tolkningar är 52 respektive 48 procentenheter.

Vid *småkakor* väljer 80 % av informanterna i SVE den specifika parafrasen ('kakor som är små och hårda'), vilket är signifikant fler än i gruppen $START < 6$, där 50 % väljer detta alternativ ($p = 0,0016$). Resultatet för SVE skiljer sig också signifikant från resultatet för $START \geq 6$, där 28 % av informanterna väljer det specificerade alternativet ($p < 0,0001$). I SVE+ samlar den specificerande parafrasen 69 % av informanternas svar, vilket är en signifikant större andel än i $START \geq 6$ ($p = 0,0020$).

Vid *finkläder* väljer informanterna i SVE och SVE+ det mer specifika svarsalternativet ('kläder som passar vid finare, högtidligare tillfällen') i signifikant högre grad (95 % respektive 93 %) än informanterna i $START \geq 6$ (47 %, $p < 0,0001$ respektive $p < 0,0001$). Informanterna i $START < 6$ skiljer sig med 82 % specifika tolkningar från informanterna i $START \geq 6$ (47 %, $p = 0,0044$).

Sammansättningen *sötsaker* tolkas konventionellt i högre grad av modersmålstalarna än av andraspråkstalarna. 92 % av informanterna i SVE

väljer tolkningen 'söta saker som man kan äta'. Därmed skiljer sig deras tolkningar signifikant från de tolkningar som informanterna i $START < 6$ gör (79 % konventionell tolkning, $p=0,049$) och från de tolkningar som informanterna i $START \geq 6$ gör (57 % konventionell tolkning, $p < 0,0001$). Även tolkningarna från informanterna i SVE+ (86 % konventionell tolkning) skiljer sig signifikant från tolkningarna i $SVE \geq 6$ ($p=0,020$).

I tolkningarna av *grönyta* skiljer sig informanterna i SVE (81 % konventionell tolkning) signifikant från informanterna i $START < 6$ (61 % konventionell tolkning, $p=0,033$). Tolkningarna i SVE skiljer sig också signifikant från tolkningarna i SVE+ (57 % konventionell tolkning, $p=0,014$) och från tolkningarna i undergruppen SVE+/S (53 % konventionell tolkning, $p=0,042$).

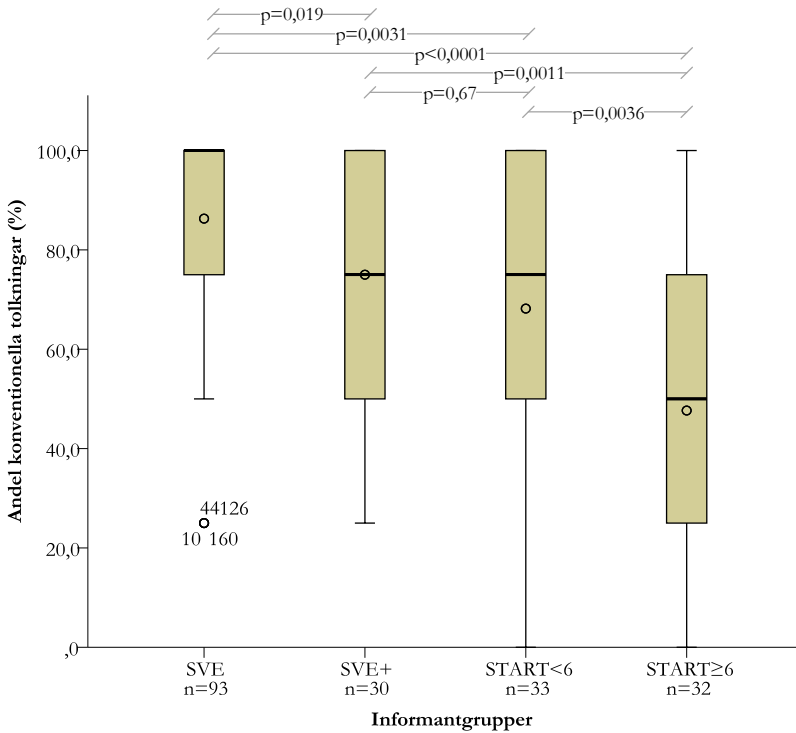
Sammanfattningsvis förekommer det signifikanta skillnader mellan grupperna vid alla de betydelsesammansmälta sammansättningarna. L1-talarna väljer i högre grad än L2-talarna de konventionella alternativen. Störst är skillnaderna mellan informanterna med svenska som enda modersmål (SVE) och informanterna med svenska som andraspråk och en högre startålder ($START \geq 6$). Skillnader mellan dessa gruppers tolkningar föreligger vid tre av fyra sammansättningar ($p < 0,001$). Skillnaderna är störst vid *småkakor* och *finkläder*.

5.1.7.3. Övergripande jämförelse mellan informantgruppernas tolkningar

I boxplotten i figur 5:4 åskådliggörs fördelningen av konventionella, specifika tolkningar i kategorin betydelsesammansmälta AN-sammansättningar, redovisat per informantgrupp. I informantgruppen som helhet är medelvärdet 74,7 % konventionella tolkningar, men med Kruskal-Wallis påvisas signifikanta skillnader mellan grupperna ($p < 0,0001$, $n=188$).

Grupp SVE har högst medelvärde ($M=86,3$ %, $SD=20,0$) och skiljer sig i sina tolkningar signifikant från SVE+ ($M=75,0$ %, $SD=22,7$), $START < 6$ ($M=68,2$ %, $SD=33,2$) och $START \geq 6$ ($M=47,7$ %, $SD=29,3$). P-värden framgår av figur 5:4. Medelvärdet för SVE skiljer sig däremot inte signifikant från medelvärdet för SVE+/S ($n=15$) ($M=76,7$ %, $SD=22,1$, $p=0,076$). Informanterna i grupp SVE+ har ett signifikant högre medelvärde än informanterna i gruppen $START \geq 6$, men skiljer sig inte signifikant från informanterna i $START < 6$. Informanterna i gruppen $START < 6$ har ett signifikant högre medelvärde än informanterna i gruppen $START \geq 6$.

Sammanfattningsvis är L1-talarna mer samstämmiga (och konventionella) i sina tolkningar än L2-talarna. L2-talarna med en längre exponeringstid för svenskan är mer samstämmiga än L2-talarna med en kortare exponeringstid för svenskan.



FIGUR 5:4. Fördelning av konventionella tolkningar av betydelse-sammansmälta AN-sammansättningar.

5.2. Adjektiviska sammansättningar

I det följande redovisas hur informanterna tolkar de adjektiviska sammansättningarna i enkätens flervalsdel.

5.2.1. Huvudled i AA-sammansättningar – typen *gulvit* och *svartvit*

Som framhållits i avsnitt 1.2 och 4.3.2.1 har adjektiviska sammansättningar med adjektiviskt förled, dvs. AA-sammansättningar, traditionellt ansetts fördela sig på två strukturella mönster: determinativa sammansättningar med efterledet som sammansättningens huvud (t.ex. *gulvit*)

och kopulativa sammansättningar med två semantiska huvudled (t.ex. *svartvit*). I flervalsdelen av enkäten ingår åtta AA-sammansättningar: *sötsur* och *fulsnygg*, samt färgsammansättningarna *gulvit*, *vitgul*, *gröngul*, *gulgrön*, *blågul* och *svartvit*. Informanternas svar redovisas i tabell 5:10 på nästa sida. Förutom de etablerade mönstren kopulativ relation och determinativ relation med efterledet som huvudled, erbjuds informanterna i denna kategori ett tredje svarsalternativ, nämligen determinativ relation med förledet som huvud (dvs. vänsterställd sammansättningsstruktur).

För varje sammansättning presenteras här först andelen svar som följer det förmodat mest konventionaliserade alternativet (kopulativ relation vid *sötsur*, *blågul* och *svartvit*; determinativ relation vid *gulvit*, *vitgul*, *gröngul*, *gulgrön* och *fulsnygg*), därefter andelen tolkningar som följer det andra strukturellt tänkbara mönstret, och slutligen andelen svar med förledet som huvudled, dvs. andelen vänstertolkningar. Sammansättningarna presenteras i en fallande frekvensordning efter deras representation i referenskorpusen.

För att renodla analysen har svarsalternativen slagits samman i den statistiska analysen. I en sammanslagningsmodell studeras betydelsen av de enskilda sammansättningarnas konventionaliseringsgrad. Det förmodat mest konventionaliserade svarsalternativet (rad 1) ställs mot de övriga två, och signifikansvärdena för denna analys markeras på rad 1.

I en annan sammanslagningsmodell undersöks om det föreligger skillnader mellan informanterna med avseende på andelen vänstertolkningar. Jag redovisar signifikansuppgifter efter en sammanslagning där det alternativ som speglar en vänstertolkning (rad 3) ställs mot de övriga två, strukturellt tänkbara alternativen, och signifikansvärdena av denna analys redovisas på rad 3.

5.2.1.1. Tolkningar i informantgruppen som helhet

Tabell 5:10 visar att det är stor spridning i tolkningarna av AA-sammansättningarna. Informanterna är mest samstämmiga i sina tolkningar av *svartvit*, *blågul* och *sötsur*, som har störst representation i referenskorpusen, och i tolkningarna av *fulsnygg*, vilka alla har en övervikt av kopulativ tolkning. *Svartvit* ges 83 % kopulativa tolkningar, *fulsnygg* 64 %, *sötsur* 63 % och *blågul* 60 % kopulativa tolkningar. Färgsammansättningarna *gröngul*, *vitgul*, *gulvit* och *gulgrön* tolkas påfallande divergerande. Vid *gulvit* och *gulgrön* fördelar sig svaren ganska jämnt över de tre svarsalternativen, och i valet av determinativ tolkning är vänstertolkning lika vanligt som högertolkning. Vid *vitgul* och *gröngul* väljer närmare hälften av informanterna förklaringen att *vitgul* är mer vit än gul och *gröngul* mer grön än gul, dvs. förledet tolkas som sammansättningens huvud.

TABELL 5:10. Informanternas svar vid AA-sammansättningar.¹⁹

	Totalt, alla informanter n=190	SVE n=94	SVE+ n=30 SVE+/S n=15	START<6 n=34	START≥6 n=32	P-värde övergripande (overall)
Svart-vit	KOPULATIV Både svart och vit	154 (83 %)	81 (89 %)	28 (93 %) 14 (93 %)	24 (73 %) * □	21 (68 %) * □ p=0,0061
	DET. HÖGER Vit färg som samtidigt har en lite svart nyans	19 (10 %)	7 (8 %)	1 (3 %) 1 (7 %)	4 (12 %)	7 (23 %)
	DET. VÄNSTER Svart färg som samtidigt har en lite vit nyans	12 (6 %)	3 (3 %)	1 (3 %) 0 (0 %)	5 (15 %) *†	3 (10 %) p=0,080
Blågul	KOPULATIV Både blå och gul	111 (60 %)	56 (62 %)	21 (70 %) 13 (87 %)	15 (45 %)	19 (59 %) p=0,23
	DET. HÖGER Gul färg som samtidigt har en lite blå nyans	25 (14 %)	10 (11 %)	3 (10 %) 0 (0 %)	6 (18 %)	6 (19 %)
	DET. VÄNSTER Blå färg som samtidigt har en lite gul nyans	49 (26 %)	24 (27 %)	6 (20 %) 2 (13 %)	12 (36 %)	7 (22 %) p=0,47

19 Övergripande p-värde beräknat med Pearsons exakta χ^2 -test, parvisa jämförelser beräknade med Fishers exakta test. I den statistiska analysen har två alternativ slagits samman till ett, i två olika sammanslagningsmodeller. I en sammanslagningsmodell ställs det förmodat mest konventionaliserade alternativet (rad 1) mot de övriga två, och p-värden för denna analys anges på rad 1. I en annan modell ställs det minst konventionaliserade alternativet (rad 3) mot de övriga två alternativen i den statistiska analysen. P-värden för denna analys redovisas på rad 3.

* $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$

Parvis jämförelse SVE mot var och en av de övriga grupperna

□ $p < 0,05$, □□ $p < 0,01$, □□□ $p < 0,001$

Parvis jämförelse SVE+ mot START<6, START≥6

◇ $p < 0,05$, ◇◇ $p < 0,01$, ◇◇◇ $p < 0,001$

Parvis jämförelse START<6 mot START≥6

† Den uppmätta signifikansen med Fishers exakta test bör ej räknas som ett signifikant värde, eftersom det övergripande p-värdet (χ^2 overall) inte är signifikant.

Siffror i kursiv representerar gruppen SVE+/S.

5 | Tolkningar i enkätens flervalsdel

	Totalt, alla informanter n=190	SVE n=94	SVE+ n=30 SVE+/S n=15	START<6 n=34	START≥6 n=32	P-värde övergripande (overall)	
Gul-grön	DET. HÖGER Grön färg som samtidigt har en lite gul nyans eller färgton	47 (26 %)	20 (22 %)	9 (31 %) 3 (20 %)	11 (34 %)	7 (22 %)	p=0,47
	KOPULATIV Både grön och gul [sic!]	71 (39 %)	28 (31 %)	14 (48 %) 9 (60 %)	11 (34 %)	18 (56 %)	
Gulvit	DET. VÄNSTER Gul färg som samtidigt har en lite grön nyans eller färgton	66 (36 %)	43 (47 %)	6 (21 %) 3 (20 %)	10 (31 %)	7 (22 %) *	p=0,011
	KOPULATIV Både gul och vit	82 (44 %)	29 (32 %)	16 (53 %) 11 (73 %)	16 (48 %)	21 (66 %)	
Sötsur	DET. HÖGER Vit färg med en lite gul nyans eller färgton	53 (28 %)	32 (35 %)	7 (23 %) 2 (13 %)	7 (21 %)	7 (22 %)	p=0,28
	DET. VÄNSTER Gul färg med en lite vit nyans eller färgton	51 (27 %)	30 (33 %)	7 (23 %) 2 (13 %)	10 (30 %)	4 (13 %) *†	p=0,15
Sötsur	KOPULATIV Lika söt som sur (eller lika sur som söt)	116 (63 %)	59 (64 %)	17 (59 %) 9 (64 %)	21 (68 %)	19 (61 %)	p=0,89
	DET. HÖGER Både söt och sur, men mest sur	27 (15 %)	11 (12 %)	3 (10 %) 1 (7 %)	5 (16 %)	8 (26 %)	
	DET. VÄNSTER Både söt och sur, men mest söt	40 (22 %)	22 (24 %)	9 (31 %) 4 (29 %)	5 (16 %)	4 (13 %)	p=0,30

	Totalt, alla informanter n=190	SVE n=94	SVE+ n=30 SVE+/S n=15	START<6 n=34	START≥6 n=32	P-värde övergripande (overall)
Fullsnygg	DET. HÖGER Mer snygg än ful	39 (21 %)	17 (19 %)	5 (17 %) 3 (20 %)	4 (12 %)	13 (42 %) * □ ◇ p=0,015
	KOPULATIV Mittemellan ful och snygg	119 (64 %)	58 (64 %)	23 (77 %) 11 (73 %)	23 (70 %)	15 (48 %)
	DET. VÄNSTER Mer ful än snygg	27 (15 %)	16 (18 %)	2 (7 %) 1 (7 %)	6 (18 %)	3 (10 %) p=0,38
Grön-gul	DET. HÖGER Gul färg som samtidigt har en lite grön nyans eller färgton	31 (17 %)	17 (19 %)	3 (10 %) 2 (13 %)	6 (19 %)	5 (16 %) p=0,73
	KOPULATIV Både grön och gul	68 (37 %)	26 (29 %)	13 (43 %) 8 (53 %)	12 (38 %)	17 (55 %)
	DET. VÄNSTER Grön färg som samtidigt har en lite gul nyans eller färgton	85 (46 %)	48 (53 %)	14 (47 %) 5 (33 %)	14 (44 %)	9 (29 %) *† p=0,15
Vitgul	DET. HÖGER Gul färg med en lite vit nyans eller färgton	30 (16 %)	15 (16 %)	3 (10 %) 1 (7 %)	5 (15 %)	7 (22 %) p=0,67
	KOPULATIV Både vit och gul	71 (38 %)	28 (31 %)	13 (43 %) 9 (60 %)	14 (42 %)	16 (50 %)
	DET. VÄNSTER Vit färg med en lite gul nyans eller färgton	85 (46 %)	48 (53 %)	14 (47 %) 5 (33 %)	14 (42 %)	9 (28 %) *† p=0,11

Sammanfattningsvis visar tabell 5:10 att det är de kopulativa sammansättningarna, som också är de mest frekventa i referenskorpusen, som tolkas med störst – men likväl begränsad – samstämmighet i informantgruppen som helhet. En kopulativ tolkning är prefererad även vid *fulsnygg*, där det kopulativa svarsalternativet samlar närmare två tredjedelar av tolkningarna. De determinativa färgsammansättningarna tolkas mycket

divergerande, och vänstertolkning är ett lika tillgängligt alternativ i meningspotentialen som högertolkning.

5.2.1.2. Jämförelse mellan informantgruppernas tolkningar

AA-sammansättningarna har tre svarsalternativ. För att förenkla analysen och göra den mer överblickbar har jag slagit samman svarsalternativen till två.

Utifrån hypotes 3b är det särskilt intressant att se om det föreligger skillnader mellan informanterna med avseende på andelen vänstertolkningar. I tabell 5:10 redovisas därför signifikansuppgifter efter en sammanslagning där det alternativ som speglar en vänstertolkning ställs mot de övriga två strukturellt tänkbara alternativen. Ett signifikant samband mellan informantgrupp och svar föreligger vid *gulgrön* ($p=0,011$). Påfallande många av informanterna i gruppen SVE tolkar *gulgrön* med förledet som huvud: 'gul färg som samtidigt har en lite grön nyans eller färgton'. Informanterna i SVE har en signifikant större andel vänstertolkningar (47 %) än informanterna i SVE+ (21 %, $p=0,016$) och START \geq 6 (22 %, $p=0,013$). På grund av liten sampelstorlek skiljer sig resultatet för informanterna i SVE+/S (20 % vänstertolkningar) inte signifikant från resultatet för SVE ($p=0,055$). Det finns en svag, icke-signifikant tendens att informanterna i SVE väljer en vänstertolkning i högre grad än informanterna i START \geq 6 också vid de övriga sammansättningarna som brukar anses determinativa (se tabell 5:10).

I en alternativ modell för sammanslagning undersöks betydelsen av de enskilda sammansättningarnas konventionaliseringsgrad. Där ställs det förmodat mest konventionaliserade svarsalternativet mot de övriga två. Signifikanta skillnader konstateras vid *svartvit* ($p=0,0061$) och *fulsnygg* ($p=0,015$).

Vid *svartvit* har de båda L1-grupperna en större andel kopulativa tolkningar (ca 90 %) än de båda L2-grupperna (ca 70 %) (SVE–START $<$ 6, $p=0,045$; SVE–START \geq 6, $p=0,010$; SVE+–START $<$ 6, $p=0,046$; SVE+–START \geq 6, $p=0,022$).

Vid *fulsnygg* väljer informanterna i gruppen START \geq 6 svarsalternativet med huvudet till höger 'mer snygg än ful' i betydligt större utsträckning (42 %) än övriga informantgrupper, där det kopulativa alternativet är prefererat. START \geq 6 skiljer sig därmed signifikant från SVE (19 % högertolkning, $p=0,015$), SVE+ (17 % högertolkning, $p=0,049$) och START $<$ 6 (12 % högertolkning, $p=0,010$).

I ytterligare en sammanslagningsmodell (som inte redovisas i tabellen) undersöks tendensen till mer kopulativa tolkningar i L2-grupperna än i L1-grupperna vid de presumtvt determinativa sammansättningarna. Signifikanta skillnader mellan informantgrupperna konstateras vid *gulgrön*

($p=0,046$) och *gulvit* ($p=0,0047$), men inte vid *gröngul* ($p=0,057$), *vitgul* ($p=0,21$) och *fulsnygg* ($p=0,12$).

Vid *gulgrön* har informanterna i $\text{START} \geq 6$ en signifikant större andel kopulativa tolkningar (56 %) än informanterna i SVE (31 %, $p=0,019$). Även informanterna i SVE+/S har en signifikant större andel kopulativa tolkningar (60 %) än informanterna i SVE (31 %, $p=0,040$).

Samma förhållande gäller vid *gulvit* där informanternas tolkningar i $\text{START} \geq 6$ (66 % kopulativa tolkningar) skiljer sig signifikant från tolkningarna i SVE (32 % kopulativa tolkningar, $p=0,0015$). SVE+ har signifikant fler kopulativa tolkningar av *gulvit* än SVE (53 % mot 32 % kopulativa tolkningar, $p=0,049$). Även informanterna i undergruppen SVE+/S har en signifikant större andel kopulativa tolkningar (73 %) än informanterna i SVE (32 % kopulativa tolkningar, $p=0,0034$).

Sammanfattningsvis noteras en svag tendens till att L2-informanterna med en sen startålder (grupp $\text{START} \geq 6$) och de flerspråkiga informanterna i SVE+ är mer benägna än de enspråkiga informanterna i gruppen SVE att välja en kopulativ tolkning av de färgsammansättningar som brukar betraktas som determinativa. Informanterna i SVE väljer i högre grad determinativ tolkning vid dessa sammansättningar, men tendensen är att vänstertolkning är prefererad framför högertolkning. Vid kopulativa sammansättningar är det bara *svartvit* som har signifikanta skillnader mellan grupperna. Sammansättningen tolkas mer kopulativt och konventionellt av L1- än av L2-grupperna.

5.2.2. Jämförelse eller orsak i NA-sammansättningar – typen *iskall* och *solvarm*

Många NA-sammansättningar är ambigüosa på grund av relationell polysemi (se avsnitt 3.2.1). I det följande redovisas svaren på NA-sammansättningar som skulle kunna aktualisera en jämförelserelation eller en orsaksrelation.

Kategorin med NA-sammansättningar som är potentiellt ambigüosa mellan jämförelse och orsak representeras av 14 sammansättningar i flervalsdelen. Sju av dessa är frekventa sammansättningar: *stenhård*, *iskall*, *sockersöt*, *solvarm*, *snövit*, *solgul* och *kryddstark*. De övriga sju är mindre frekventa eller tillfälliga bildningar. *Snömjuk*, *snökall*, *glasskall*, *lökstark*, *spisvarm* och *berghård* saknar helt representation i referenskorpusen, och *honungssöt* förekommer med ett fåtal träffar.

I tabell 5:11 redovisas fördelningen av informanternas svar på de NA-sammansättningar som skulle kunna aktualisera en jämförande relation eller en orsaksrelation mellan för- och efterled. För varje sam-

mansättning redovisas först andelen jämförande och därefter andelen orsaksrelaterade tolkningar. Sammansättningarna är ordnade efter deras frekvens i referenskorpusen.

TABELL 5:11. *Informanternas val av jämförelse- eller orsaksrelation vid NA-sammansättningar.*²⁰

	Totalt, alla informanter n=190	SVE n=94	SVE+ n=30 SVE+/S n=15	START<6 n=34	START≥6 n=32	P-värde övergripande (overall)	
Stenhård	JÄMFÖRELSE Väldigt hård (som sten)	177 (94 %)	92 (99 %)	27 (90 %) * 13 (87 %)	28 (85 %) **	30 (94 %)	p=0,016
	ORSAK Hård av sten	11 (6 %)	1 (1 %)	3 (10 %) * 2 (13 %)	5 (15 %) **	2 (6 %)	
Iskall	JÄMFÖRELSE Mycket kall (som is)	172 (92 %)	89 (96 %)	28 (93 %) 15 (100 %)	31 (94 %)	24 (80 %) *	p=0,040
	ORSAK Kall (eller kyld) av is	14 (8 %)	4 (4 %)	2 (7 %) 0 (0 %)	2 (6 %)	6 (20 %) *	
Socker-söt	JÄMFÖRELSE Mycket söt (som socker)	160 (87 %)	81 (90 %)	26 (87 %) 15 (100 %)	31 (94 %)	22 (71 %) * ∅	p=0,026
	ORSAK Söt av socker	24 (13 %)	9 (10 %)	4 (13 %) 0 (0 %)	2 (6 %)	9 (29 %) * ∅	
Solvarm	JÄMFÖRELSE Mycket varm (som solen)	70 (38 %)	23 (25 %)	15 (52 %) * 6 (40 %)	15 (45 %) *	17 (53 %) **	p=0,0046
	ORSAK Varm (eller uppvärmd) av solen	116 (62 %)	69 (75 %)	14 (48 %) * 9 (60 %)	18 (55 %) *	15 (47 %) **	

20 Övergripande p-värde beräknat med Pearsons exakta chi²-test, parvisa jämförelser beräknade med Fishers exakta test.

* p<0,05, ** p<0,01, *** p<0,001

Parvis jämförelse SVE mot var och en av de övriga grupperna

□ p<0,05, □□ p<0,01, □□□ p<0,001

Parvis jämförelse SVE+ mot START<6, START≥6

∅ p<0,05, ∅∅ p<0,01, ∅∅∅ p<0,001

Parvis jämförelse START<6 mot START≥6

† Den uppmätta signifikansen med Fishers exakta test bör ej räknas som ett signifikant värde, eftersom det övergripande p-värdet (chi² overall) inte är signifikant.

Siffror i kursiv representerar gruppen SVE+/S.

		Totalt, alla informanter n=190	SVE n=94	SVE+ n=30 SVE+S n=15	START<6 n=34	START≥6 n=32	P-värde övergripande (overall)
Kryddstark	JÄMFÖRELSE Mycket stark (som kryddor)	39 (21 %)	12 (13 %)	7 (24 %) 1 (7 %)	14 (42 %) ***	6 (19 %)	p=0,0043
	ORSAK Stark av kryddor	147 (79 %)	81 (87 %)	22 (76 %) 13 (93 %)	19 (58 %) ***	25 (81 %)	
Solgul	JÄMFÖRELSE Gul som solen	162 (87 %)	81 (88 %)	28 (93 %) 14 (93 %)	30 (91 %)	23 (72 %) * □	p=0,046
	ORSAK Gul av solen	25 (13 %)	11 (12 %)	2 (7 %) 1 (7 %)	3 (9 %)	9 (28 %) * □	
Snövit	JÄMFÖRELSE Vit som snö	171 (91 %)	86 (92 %)	29 (97 %) 14 (93 %)	29 (88 %)	27 (87 %)	p=0,49
	ORSAK Vit av snö	16 (9 %)	7 (8 %)	1 (3 %) 1 (7 %)	4 (12 %)	4 (13 %)	
Honungsöt	JÄMFÖRELSE Mycket söt (som honung)	144 (78 %)	69 (75 %)	24 (80 %) 12 (80 %)	25 (78 %)	26 (84 %)	p=0,77
	ORSAK Söt av honung	41 (22 %)	23 (25 %)	6 (20 %) 3 (20 %)	7 (22 %)	5 (16 %)	
Snökall	JÄMFÖRELSE Mycket kall (som snö)	119 (65 %)	55 (60 %)	21 (75 %) 12 (86 %)	23 (70 %)	20 (65 %)	p=0,50
	ORSAK Kall av snö	64 (35 %)	36 (40 %)	7 (25 %) 2 (14 %)	10 (30 %)	11 (35 %)	
Spisvarm	JÄMFÖRELSE Mycket varm (som spisen)	108 (58 %)	48 (52 %)	18 (62 %) 8 (57 %)	20 (61 %)	22 (71 %)	p=0,30
	ORSAK Varm (eller uppvärmd) av spisen	77 (42 %)	44 (48 %)	11 (38 %) 6 (43 %)	13 (39 %)	9 (29 %)	
Snömjuk	JÄMFÖRELSE Mycket mjuk (som snö)	156 (83 %)	73 (78 %)	28 (97 %) 14 (100 %)	27 (82 %) *+	28 (90 %)	p=0,069
	ORSAK Mjuk av snö	31 (17 %)	21 (22 %)	1 (3 %) 0 (0 %)	6 (18 %) *+	3 (10 %)	

5 | Tolkningar i enkätens flervalsdel

	Totalt, alla informanter n=190	SVE n=94	SVE+ n=30 SVE+/S n=15	START<6 n=34	START≥6 n=32	P-värde övergripande (overall)	
Berghård	JÄMFÖRELSE Väldigt hård (som berg)	160 (86 %)	81 (87 %)	28 (97 %) 14 (100 %)	26 (79 %)	25 (78 %)	p=0,13
	ORSAK Hård av berg	27 (14 %)	12 (13 %)	1 (3 %) 0 (0 %)	7 (21 %)	7 (22 %)	
Lökstark	JÄMFÖRELSE Mycket stark (som lök)	60 (32 %)	20 (22 %)	11 (37 %) 6 (40 %)	13 (41 %) *	16 (50 %) **	p=0,012
	ORSAK Stark av lök	127 (68 %)	73 (78 %)	19 (63 %) 9 (60 %)	19 (59 %) *	16 (50 %) **	
Glasskall	JÄMFÖRELSE Mycket kall (som glass)	134 (73 %)	71 (77 %)	23 (82 %) 13 (87 %)	24 (73 %)	16 (52 %) *	p=0,027
	ORSAK Kall av glass	50 (27 %)	21 (23 %)	5 (18 %) 2 (13 %)	9 (27 %)	15 (48 %) *	

5.2.2.1. Tolkningar i informantgruppen som helhet

Tabell 5:11 visar att det generellt råder stor samstämmighet i tolkningen av NA-sammansättningarna. Störst är samstämmigheten vid de frekventa sammansättningarna. Sammansättningarna *stenhård*, *iskall*, *sockersöt*, *solgul* och *snövit* tolkas lika – och jämförande – av cirka 90 % av informantgruppen som helhet. Samstämmigheten vid *solwarm* är betydligt lägre, trots att sammansättningen har relativt god representation i referenskorpusen; 62 % av informanterna ger en orsaksrelaterad tolkning. Sammansättningen *kryddstark*, som är en förhållandevis frekvent sammansättning, tolkas relativt samstämmigt: 80 % av informanterna väljer en orsaksrelaterad tolkning. *Honungssöt* har bara tre träffar i referenskorpusen men tolkas ändå med stor övervikt för det ena alternativet: 78 % av informanterna väljer en jämförande tolkning. Resterande sammansättningar saknar representation i referenskorpusen. För dessa överväger jämförande tolkningar: *snökall* (65 %), *spiswarm* (58 %), *snömjuk* (83 %), *berghård* (86 %), *glasskall* (73 %). *Lökstark* utgör ett undantag och tolkas i 68 % av fallen med en orsaksrelation mellan leden, 'stark av lök'.

Sammanfattningsvis är informanterna relativt samstämmiga i sina tolkningar av NA-sammansättningarna. Störst är samstämmigheten vid de frekventa sammansättningarna, men även vid de mindre frekventa eller

tillfälliga sammansättningarna samlar ett av svarsalternativen – i de flesta fall det jämförande – en stor andel av informanternas tolkningar.

5.2.2.2. Jämförelse mellan informantgruppernas tolkningar

Som framgår av tabell 5:11 föreligger ett signifikant samband mellan informantgrupp och svar för sammansättningarna *stenhård* ($p=0,016$), *iskall* ($p=0,040$), *sockersöt* ($p=0,026$) *solvarm* ($p=0,0046$), *kryddstark* ($p=0,0043$), *solgul* ($p=0,046$), *lökstark* ($p=0,012$) och *glasskall* ($p=0,027$). Inga signifikanta skillnader föreligger vid sammansättningarna *berghård*, *snövit*, *honungssöt*, *snökall*, *spisvarm* och *snömjuk*. Sambandet mellan informantgrupp och svar – dvs. skillnaderna mellan informantgrupperna – är tydligast vid *kryddstark*, *lökstark* och *solvarm*. Differensen mellan högsta och lägsta andel tolkningar med ett orsakssamband är 29, 28 respektive 28 procentenheter

Vid *solvarm* gör informanterna i grupp SVE orsakrelaterade tolkningar 'varm av solen' i signifikant högre grad (75 %) än grupp SVE+ (48 %, $p=0,011$), grupp START<6 (55 %, $p=0,046$) och grupp START≥6 (47 %, $p=0,0046$). Tolkningarna från SVE skiljer sig däremot inte signifikant från tolkningarna i SVE+/S (60 % orsakstolkning, $p=0,23$).

Vid *kryddstark* har informanterna i grupp SVE en signifikant större andel orsakrelaterade tolkningar 'stark av kryddor' (87 %) än gruppen START<6 (58 %, $p=0,0008$).

Den tillfälliga bildningen *lökstark* tolkas mer orsakrelaterat 'stark av lök' i gruppen SVE (78 %) än i START<6 (59 %, $p=0,040$) och START≥6 (50 %, $p=0,0033$).

Även *glasskall* tolkas jämförande i större utsträckning av informanterna i SVE och SVE+ (77 respektive 82 %) än i gruppen START≥6 (52 % jämförande tolkningar). Tolkningarna från informanterna i gruppen START≥6 skiljer sig signifikant från tolkningarna i SVE ($p=0,011$) och SVE+ ($p=0,026$).

I tolkningarna av *stenhård* skiljer sig informanterna i SVE (99 % jämförande tolkning) från informanterna i SVE+ (90 % jämförande tolkning, $p=0,045$). Skillnaden mellan SVE och SVE+/S balanserar på signifikansgränsen, men når inte statistisk signifikans (87 % jämförande tolkning, $p=0,050$). SVE skiljer sig signifikant även från START<6 (85 % jämförelse, $p=0,0047$).

De parvisa jämförelserna när det gäller tolkningen av *iskall* påvisar endast en signifikant skillnad: informanterna i SVE har en större andel jämförande tolkningar (96 %) än informanterna i START≥6 (80 % jämförelse, $p=0,013$).

Sammansättningen *sockersöt* tolkas i högre grad jämförande av informanterna i gruppen SVE (90 %) än av informanterna i START≥6 (71 %, $p=0,0047$).

$p=0,017$). Även skillnaden mellan informanterna i $START < 6$ (94 % jämförelse) och informanterna i $START \geq 6$ (71 % jämförelse) är signifikant ($p=0,021$).

I tolkningarna av *solgul* skiljer sig SVE (88 % jämförande tolkning) från informanterna i $START \geq 6$ (72 % jämförelse, $p=0,049$). Även SVE+ har en signifikant större andel jämförande tolkning (93 %) än informanterna i $START \geq 6$ ($p=0,044$).

Sammanfattningsvis föreligger signifikanta skillnader mellan grupperna i tolkningen av åtta av de fjorton NA-sammansättningar som presenterades med ett jämförande respektive ett orsakrelaterat svarsalternativ. Störst skillnader konstateras vid de mindre frekventa eller tillfälliga sammansättningarna (*kryddstark*, *solvarm* och *lökstark*). De grupper som skiljer sig mest åt är SVE och $START \geq 6$. Dessa skiljer sig åt vid sex sammansättningar ($p < 0,05$); vid de etablerade sammansättningarna är informanterna i SVE mer konventionella i sina tolkningar än andraspråkstalarna. Tolkningarna från SVE skiljer sig också signifikant från tolkningarna i $START < 6$ vid fyra sammansättningar och från SVE+ (men inte SVE+/S) vid två sammansättningar. SVE+ skiljer sig signifikant från $START \geq 6$ vid sammansättningarna *glasskall* och *solgul*, som tolkas jämförande i högre grad av L1-talarna. I tolkningen av *sockersöt* väljer informanterna i $START < 6$ i signifikant högre grad än informanterna i $START \geq 6$ det jämförande alternativet.

5.2.3. Sammansättningar med *-rädd*

I avsnitt 4.3.2.3 diskuteras NA-sammansättningar med efterledet *-rädd*. Dessa är potentiellt ambigüosa mellan tolkningsmönstren [rädd för N] och [rädd som N]. I flervalsdelen av enkäten undersöks tolkningarna av sammansättningarna *hundrädd* och *fågelrädd*, och i fritextdelen av enkäten tolkar informanterna sammansättningen *musrädd*.

I tabell 5:12 redovisas hur informanternas tolkningar av *hundrädd* och *fågelrädd* fördelar sig mellan tolkning med bundet adverbial [rädd för N] och jämförelse [rädd som N]. Sammansättningarna har ingen representation i referenskorpusen.

Av tabellen framgår att informanterna är relativt samstämmiga i sina tolkningar. 90 % av informanterna väljer tolkningen [rädd för N] vid *hundrädd*, och 85 % av informanterna väljer [rädd för N] vid *fågelrädd*. Det finns dock en tendens att de informanter som har svenska som modersmål är mer benägna än de som har svenska som andraspråk att välja tolkningarna med [rädd för N]. Vid *hundrädd* når skillnaden inte statistisk signifikans. Vid *fågelrädd* däremot föreligger en signifikant skillnad mellan SVE och $START \geq 6$. I gruppen SVE väljer 95 % av informanterna tolkningen 'rädd för en fågel' medan bara 71 % av informanterna i

START \geq 6 väljer samma tolkning ($p=0,0012$). Informanterna i SVE skiljer sig även signifikant från informanterna i SVE+ (73 % tolkning 'rädd för', $p=0,0032$) vid *fågelrädd*. SVE-gruppen skiljer sig däremot inte signifikant från SVE+/S (87 % 'rädd för', $p=0,26$).

TABELL 5:12. Informanternas svar vid sammansättningar med -rädd.²¹

		Totalt, alla informanter n=190	SVE n=94	SVE+ n=30 SVE+/S n=15	START<6 n=34	START \geq 6 n=32	P-värde övergripande (overall)
Hund- rädd	Rädd för hundar	168 (90 %)	88 (95 %)	26 (90 %) 12 (86 %)	29 (88 %)	25 (81 %)	*†
	Rädd som en hund	18 (10 %)	5 (5 %)	3 (10 %) 2 (14 %)	4 (12 %)	6 (19 %)	*†
Fågel- rädd	Rädd för fåglar	157 (85 %)	86 (95 %)	22 (73 %) 13 (87 %)	27 (82 %)	22 (71 %)	**
	Rädd som en fågel	28 (15 %)	5 (5 %)	8 (27 %) 2 (13 %)	6 (18 %)	9 (29 %)	**

5.2.4. Sammansättningar med -fri

Efterledet *-fri* i sammansättningar utgör vanligen avledningssuffix med betydelsen 'fri från', men *-fri* kan också utgöra självständigt sammansättningsled med betydelsen 'fri (att)' som i *valfri* (se vidare avsnitt 4.3.2.4). I flervalsdelen av enkäten efterfrågas informanternas tolkning av de potentiellt ambigüsa sammansättningarna *bilfri*, *valfri* och *helfri*.

I tabell 5:13 redovisas informanternas svar i kategorin *-fri*. För varje sammansättning redovisas först andelen tolkningar med relationen 'fri från'. Sammansättningarna är ordnade efter deras frekvens i referenskorpusen.

21 Övergripande p-värde beräknat med Pearsons exakta chi²-test, parvisa jämförelser beräknade med Fishers exakta test.

* $p<0,05$, ** $p<0,01$, *** $p<0,001$

Parvis jämförelse SVE mot var och en av de övriga grupperna

□ $p<0,05$, □□ $p<0,01$, □□□ $p<0,001$

Parvis jämförelse SVE+ mot START<6, START \geq 6

◇ $p<0,05$, ◇◇ $p<0,01$, ◇◇◇ $p<0,001$

Parvis jämförelse START<6 mot START \geq 6

† Den uppmätta signifikansen med Fishers exakta test bör ej räknas som ett signifikant värde, eftersom det övergripande p-värdet (chi² overall) inte är signifikant.

Siffror i kursiv representerar gruppen SVE+/S.

TABELL 5:13. *Informanternas svar vid sammansättningar med -fri.*²²

	Totalt, alla informanter n=190	SVE n=94	SVE+ n=30 SVE+/S n=15	START<6 n=34	START≥6 n=32	P-värde övergripande (overall)
Fri från val = utan val	29 (16 %)	7 (8 %)	5 (17 %) 3 (20 %)	6 (18 %)	11 (34 %) ***	
Valfri	Som kan väljas fritt 158 (84 %)	85 (92 %)	25 (83 %) 12 (80 %)	27 (82 %)	21 (66 %) ***	p=0,0040
Bilfri	Fri från bilar = utan bilar 23 (13 %)	11 (12 %)	2 (7 %) 1 (7 %)	7 (22 %)	3 (10 %)	p=0,31
Helgfri	När det inte är helg (= fri från helg) 79 (42 %)	37 (39 %)	13 (43 %) 7 (47 %)	11 (33 %)	18 (56 %)	p=0,26

I tabell 5:13 påvisas stor samstämmighet i tolkningarna av *bilfri* och *valfri*. *Valfri* är den sammansättning i enkäten som har den i särklass största representationen i referenskorpusen. Sammansättningen tolkas i stor utsträckning (84 %) konventionellt med *-fri* som självständigt led 'fri att'. Även *bilfri* med måttlig representation i referenskorpusen tolkas mycket

22 Övergripande p-värde beräknat med Pearsons exakta chi²-test, parvisa jämförelser beräknade med Fishers exakta test.

* p<0,05, ** p<0,01, *** p<0,001

Parvis jämförelse SVE mot var och en av de övriga grupperna

□ p<0,05, □□ p<0,01, □□□ p<0,001

Parvis jämförelse SVE+ mot START<6, START≥6

◇ p<0,05, ◇◇ p<0,01, ◇◇◇ p<0,001

Parvis jämförelse START<6 mot START≥6

† Den uppmätta signifikansen med Fishers exakta test bör ej räknas som ett signifikant värde, eftersom det övergripande p-värdet (chi² overall) inte är signifikant.

Siffror i kursiv representerar gruppen SVE+/S.

samstämmigt: 88 % av informanterna väljer den semantiska relationen 'fri från'. Vid *helgfri* är samstämmigheten däremot begränsad: 58 % av informanterna väljer relationen 'fri från', medan 42 % väljer tolkningen 'när man är fri på helgen' med *-fri* som självständigt sammansättningsled.

Det föreligger inga signifikanta skillnader mellan informantgruppernas tolkningar av *bilfri* och *helgfri*. Vid *valfri* har informanterna i SVE signifikant fler tolkningar 'som kan väljas fritt' (92 %) än informanterna i START \geq 6 (66 %, $p=0,0006$).

5.2.5. Sammansättningar med *-färdig*

Sammansättningar med efterledet *-färdig* är potentiellt ambigüosa mellan parafraserna 'färdig att (börja) göra' den verbhandling som förledet markerar, dvs. 'färdig att börja VERBA' och 'färdig, klar med att göra' den verbhandling som förledet anger (se vidare 3.1.1.12).

I fritextdelen av enkäten representeras denna kategori av en relativt frekvent förekommande sammansättning – *gråtfärdig* – och två tillfälliga bildningar: *skratfärdig* och *skrikfärdig*.

I tabell 5:14 på nästa sida presenteras informanternas svar på sammansättningarna i kategorin *-färdig*. Vid varje enskild sammansättning presenteras först svarsalternativet som följer schemat [färdig att börja VERBA] och därefter svarsalternativet som följer schemat [färdig/klar med att VERBA]. Endast *gråtfärdig* är representerad i referenskorpusen.

5.2.5.1. Tolkningar i informantgruppen som helhet

Tabell 5:14 visar att informanterna i relativt stor utsträckning (83 %) väljer tolkningsmönstret [färdig att börja VERBA] vid den etablerade sammansättningen *gråtfärdig*. De tillfälliga bildningarna *skratfärdig* och *skrikfärdig* tolkas båda enligt detta mönster i 81 % av fallen.

Det finns ett tydligt samband mellan informanternas tolkningar av *gråtfärdig* och deras tolkningar av å ena sidan *skratfärdig*, å andra sidan *skrikfärdig*. Som framgår av tabell 5:15 är det en signifikant större andel av dem som följer schemat [färdig att börja VERBA] vid *gråtfärdig* som följer detta schema också för *skratfärdig* och *skrikfärdig* jämfört med de informanter som väljer [färdig med att VERBA] vid *gråtfärdig* ($p<0,0001$ för *skratfärdig*, $p<0,0001$ för *skrikfärdig*).

TABELL 5:14. *Informanternas svar vid sammansättningar med -färdig.*²³

	Totalt, alla informanter n=190	SVE n=94	SVE+ n=30 SVE+/S n=15	START<6 n=34	START≥6 n=32	P-värde övergripande (overall)
Gråt-färdig	När man är så ledsen att man nästan börjar gråta	152 (83 %)	88 (96 %)	26 (93 %) 13 (93 %)	25 (76 %) **	13 (42 %) *** □□□ ◇
	När man gråtit färdigt	32 (17 %)	4 (4 %)	2 (7 %) 1 (7 %)	8 (24 %) **	18 (58 %) *** □□□ ◇
Skratt-färdig	När man tycker något är så roligt att man nästan börjar skratta	152 (81 %)	85 (92 %)	26 (90 %) 14 (100 %)	27 (79 %)	14 (44 %) *** □□□ ◇◇
	När man skrattat färdigt	35 (19 %)	7 (8 %)	3 (10 %) 0 (0 %)	7 (21 %)	18 (56 %) *** □□□ ◇◇
Skrik-färdig	När man är så arg och upprörd att man nästan börjar skrika	150 (81 %)	88 (96 %)	24 (83 %) * 13 (93 %)	27 (82 %) *	11 (34 %) *** □□□ ◇◇◇
	När man skrikat färdigt	36 (19 %)	4 (4 %)	5 (17 %) * 1 (7 %)	6 (18 %) *	21 (66 %) *** □□□ ◇◇◇

23 Övergripande p-värde beräknat med Pearsons exakta chi²-test, parvisa jämförelser beräknade med Fishers exakta test.

* p<0,05, ** p<0,01, *** p<0,001

□ p<0,05, □□ p<0,01, □□□ p<0,001

◇ p<0,05, ◇◇ p<0,01, ◇◇◇ p<0,001

† Den uppmätta signifikansen med Fishers exakta test bör ej räknas som ett signifikant värde, eftersom det övergripande p-värdet (chi² overall) inte är signifikant.

Siffror i kursiv representerar gruppen SVE+/S.

TABELL 5:15. Samband mellan tolkningar av -färdig i informantgruppen som helhet.

	Gråtfärdig Färdig att börja VERBA	Gråtfärdig Färdig med att VERBA	P-värde Fishers exakta test
Skrattfärdig Färdig att börja VERBA	139 (91 %)	12 (37,5 %)	p<0,0001
Skrattfärdig Färdig med att VERBA	13 (9 %)	20 (62,5 %)	
Skrikfärdig Färdig att börja VERBA	137 (91 %)	11 (34 %)	p<0,0001
Skrikfärdig Färdig med att VERBA	14 (9 %)	21 (66 %)	

Sammanfattningsvis visar tabell 5:14 och 5:15 att den etablerade sammansättningen *gråtfärdig* i stor utsträckning tolkas på konventionellt sätt. De tillfälliga bildningarna *skrattfärdig* och *skrikfärdig* tolkas i analogi med betydelsen hos *gråtfärdig*.

5.2.5.2. Jämförelse mellan informantgruppernas tolkningar

Av tabell 5:14 framgår att det råder ett signifikant samband mellan informantgruppstillhörighet och svar vid samtliga ord i kategorin -färdig ($p<0,0001$ för *gråtfärdig*, $p<0,0001$ för *skrattfärdig*, $p<0,0001$ för *skrikfärdig*).

Störst skillnader mellan grupperna noteras för *skrikfärdig* där differensen mellan största och minsta andel tolkningar 'färdig att börja VERBA' är 62 procentenheter. Skillnaderna är stora även vid *gråtfärdig* och *skrattfärdig* där motsvarande siffra är 54 respektive 48 procentenheters differens mellan högsta och lägsta värde.

Tabell 5:14 visar att informanterna i de båda L1-grupperna har en större andel svar som följer mönstret 'färdig att börja VERBA' än informanterna i de båda L2-grupperna. En betydande skillnad föreligger också mellan grupperna $START<6$ och $START\geq 6$.

Vid *gråtfärdig* väljer 96 % av informanterna i SVE den konventionaliserade betydelsen 'färdig att börja gråta'. Därmed skiljer sig deras resultat signifikant från resultatet för å ena sidan informanterna i $START<6$ (76 % konventionella tolkningar, $p=0,0025$), å andra sidan informanterna i $START\geq 6$ (42 % konventionella tolkningar, $p<0,0001$). Informanterna i

gruppen SVE+ har 93 % konventionella tolkningar och skiljer sig signifikant från START \geq 6 ($p < 0,0001$). Även skillnaden mellan START $<$ 6 och START \geq 6 är statistiskt signifikant ($p = 0,010$).

Vid *skrattfärdig* har informantgruppen SVE 92 % tolkningar 'färdig att', att jämföra med 90 % i SVE+, 79 % i START $<$ 6 och 44 % i START \geq 6. Resultatet för SVE skiljer sig signifikant från resultatet för informantgruppen START \geq 6 ($p < 0,0001$). Även resultatet för SVE+ skiljer sig signifikant från START \geq 6 ($p < 0,0001$). Informanterna med låg startålder för svenskan (START $<$ 6) har en signifikant större andel tolkningar enligt mönstret 'färdig att' än informanterna med en högre startålder för svenskan, dvs. gruppen START \geq 6 ($p = 0,0048$).

Sammanställningen *skrikfärdig* tolkas enligt mönstret 'färdig att' i högre grad av de informanter som har svenska som L1 (96 % i SVE och 83 % i SVE+) än av de informanter som har svenska som L2 (82 % i START $<$ 6 och 34 % i START \geq 6). SVE skiljer sig signifikant från START $<$ 6 ($p = 0,021$) och START \geq 6 ($p < 0,0001$). SVE+ skiljer sig signifikant från START \geq 6 ($p = 0,0002$). Informanterna i START $<$ 6 har en betydligt högre grad av tolkningar 'färdig att' än informanterna i START \geq 6 ($p = 0,0001$). Även skillnaden mellan SVE och SVE+ är signifikant ($p = 0,035$). Skillnaden mellan SVE och SVE+/S är däremot inte signifikant ($p = 0,52$).

I tabell 5:16 presenteras sambandet för de olika informantgrupperna mellan tolkningarna av *gråtfärdig*, *skrattfärdig* och *skrikfärdig*.

Ett signifikant samband mellan tolkningarna av dessa sammansättningar påvisas för informantgrupperna START $<$ 6 och START \geq 6, men inte för SVE och SVE+. Den grupp av informanter i START $<$ 6 som följer schemat [färdig att VERBA] vid *gråtfärdig* har en signifikant större andel tolkningar 'färdig att' vid *skrattfärdig* (92 %) och *skrikfärdig* (92 %) än den grupp informanter i START $<$ 6 som väljer 'färdig med' vid *gråtfärdig* (50 % 'färdig att börja skratta/skrika', $p = 0,020$ respektive $p = 0,020$). Även i gruppen START \geq 6 har den grupp som väljer 'färdig att' vid *gråtfärdig* en signifikant större andel tolkningar 'färdig att' vid *skrattfärdig* (77 %) och *skrikfärdig* (69 %) än den grupp av informanter som väljer 'färdig med' vid *gråtfärdig* (22 % 'färdig att börja skratta', $p = 0,0040$ respektive 11 % 'färdig att börja skrika', $p = 0,0017$).

I SVE och SVE+ råder ingen signifikant skillnad mellan dem som väljer [färdig att VERBA] och dem som väljer [färdig med att VERBA] vid *gråtfärdig*, när det gäller hur de tolkar *skrattfärdig* och *skrikfärdig*. Det beror på att den grupp som väljer alternativet [färdig med att VERBA] vid *gråtfärdig* är så liten att inga signifikanta skillnader kan detekteras. Som framgår av tabell 5:16 väljer nästan alla informanterna [färdig att VERBA] vid alla sammansättningarna, dvs. *skrattfärdig* och *skrikfärdig* tolkas i analogi med *gråtfärdig*.

TABELL 5:16. Samband mellan tolkningar av -färdig, redovisat per informantgrupp.

		Gråtfärdig 'färdig att'	Gråtfärdig 'färdig med'	P-värde Fishers exakta test
SVE	Skrattfärdig 'färdig att'	82 (93 %)	3 (75 %)	p=0,28
	Skrattfärdig 'färdig med'	6 (7 %)	1 (25 %)	
	Skrikfärdig 'färdig att'	83 (95 %)	4 (100 %)	p=1,00
	Skrikfärdig 'färdig med'	4 (5 %)	0 (0 %)	
SVE+	Skrattfärdig 'färdig att'	24 (92 %)	1(50 %)	p=0,21
	Skrattfärdig 'färdig med'	2 (8 %)	1 (50 %)	
	Skrikfärdig 'färdig att'	22 (85 %)	1 (50 %)	p=0,33
	Skrikfärdig 'färdig att'	4 (15 %)	1 (50 %)	
START<6	Skrattfärdig 'färdig att'	23 (92 %)	4 (50 %)	p=0,020
	Skrattfärdig 'färdig med'	2 (8 %)	4 (50 %)	
	Skrikfärdig 'färdig att'	23 (92 %)	4 (50 %)	p=0,020
	Skrikfärdig 'färdig med'	2 (8 %)	4 (50 %)	
START≥6	Skrattfärdig 'färdig att'	10 (77 %)	4 (22 %)	p=0,0040
	Skrattfärdig 'färdig med'	3 (23 %)	14 (78 %)	
	Skrikfärdig 'färdig att'	9 (69 %)	2 (11 %)	p=0,0017
	Skrikfärdig 'färdig med'	4 (31 %)	16 (89 %)	

Sammanfattningsvis är det stora skillnader mellan informantgruppernas tolkningar av sammansättningarna med *-färdig*. Den stora skiljelinjen går mellan informanterna med svenska som L1 och informanterna med svenska som L2 och en hög startålder för svenskan. Informanterna med svenska som L1 följer i betydligt högre grad än informanterna i START \geq 6 mönstret [färdig att VERBA]. Bland informanterna med svenska som andraspråk är det en betydligt större andel tolkningar [färdig att VERBA] hos de informanter som har en låg startålder för svenskan än hos de informanter som har en högre startålder för svenskan. I samtliga grupper föreligger ett signifikant samband mellan tolkningarna av *gråtfärdig*, *skrattfärdig* och *skrikfärdig*.

5.3. Sammanfattning

Resultaten från flervalsdelen visar att det råder stora skillnader mellan informanternas tolkningar, beroende på om de har svenskan som sitt första- eller andraspråk. Störst skillnader föreligger mellan de enspråkiga L1-informanterna i SVE och de flerspråkiga L2-informanterna i START \geq 6, vars tolkningar med något undantag skiljer sig åt i samtliga kategorier och för en betydande del av de enskilda sammansättningarna. SVE skiljer sig även signifikant från START $<$ 6 vid många (men färre) sammansättningar. Det förekommer skillnader också mellan de enspråkiga L1-talarna i gruppen SVE och de flerspråkiga i gruppen SVE+ vid en del av de undersökta sammansättningarna. Om gruppen SVE+ avgränsas till de informanter som har en förälder som har svenska som modersmål (SVE+/S) förekommer skillnader endast vid ett fåtal sammansättningar. Skillnader mellan informanternas tolkningar i grupperna START $<$ 6 och START \geq 6 noteras vid några sammansättningar.

Störst skillnader mellan grupperna konstateras för de sammansättningar som (indirekt) testar hur informanterna förhåller sig till huvudledets placering i NN-sammansättningar. Resultatet visar att L1-talarna är mer benägna än L2-talarna att tolka nominala sammansättningar med huvudledet till höger. Även betydelsesammansmälta AN-sammansättningar tolkas mer konventionellt, med en specifik relation mellan leden, av L1- än av L2-talarna. Skillnaderna är också stora i tolkningarna av agens i verbavledning. Där är L1-talarna mer benägna än L2-talarna att välja patienttolkning av sammansättningar av typen [N-misshandel]. De tre sammansättningarna med *-färdig* tolkas enligt mönstret [färdig att VERBA] i högre grad av L1- än av L2-talarna, och informanterna i START $<$ 6 väljer detta mönster i högre grad än informanterna i START \geq 6.

Även i tolkningarna av de NA-sammansättningar som presenteras med en jämförande eller en orsaksrelation föreligger signifikanta skillnader mellan grupperna vid hälften av sammansättningarna. Också i denna kategori är skillnaderna störst mellan de enspråkiga informanterna i SVE och de flerspråkiga informanterna i START \geq 6. Sammantaget kan man konstatera att exponeringstiden för svenskan ger utslag: ju längre exponeringstid desto större samstämmighet.



6

Tolkningar i enkätens fritextdel

I detta kapitel diskuteras svaren i fritextdelen, där informanterna fått formulera egna förklaringar till var och en av 20 sammansättningar genom att fullfölja parafrafer av typen ”En *zebraväska* är ...”. Här diskuteras i vilken grad dessa svar tyder på tolkningar som följer resultaten i flervalsdelen, där informanterna fått välja mellan färdiga svarsalternativ, som var försedda med en eller flera kontextualiserande meningar.

Alla sammansättningar i fritextdelen tillhör någon av de kategorier som också undersöks i flervalsdelen av enkäten. Detta urval är gjort med tanke på att kunna jämföra informanternas tolkningar i den fria och den mer styrda delen av enkäten. Fyra sammansättningar förekommer i enkätens båda delar. I övrigt återfinns i stor utsträckning sammansättningar med samma för- eller efterled i båda delarna av enkäten, t.ex. *zebra-* som förled och *-misshandel* som efterled. Sålunda används fritextdelen som ett referensmaterial till flervalsdelen.

Eftersom fokus ligger på att presentera samband och skillnader mellan informanternas helt fria och deras delvis styrda tolkningar, får resultatredovisningen i detta kapitel en mer syntetiserande karaktär än redovisningen av resultaten i flervalsdelen av enkäten (kapitel 5). Av samma skäl är presentationen inriktad på de övergripande dragen i materialet. För att ändå belysa bredden och de semantiska detaljerna i informanternas fritexttolkningar presenteras mot slutet av kapitlet en noggrannare se-

mantisk genomgång av en sammansättning i enkäten, nämligen *vagnbarn* (som också behandlas med avseende på andra aspekter tidigare i kapitlet). Tolkningarna av *vagnbarn* får därmed tjäna som ett illustrativt exempel på sammansättningsars vida meningspotentialer.

Vid insamlandet av materialet fick informanterna besvara och lämna in fritextdelen innan de besvarade flervalsdelen av enkäten, för att svaren i fritextdelen inte skulle påverkas av tillgången till kontext i flervalsdelen, särskilt vid de sammansättningar som förekommer i enkätens båda delar.

6.1. Sammansättningar som förekommer i enkätens båda delar

De sammansättningar som förekommer i enkätens båda delar är *vagnbarn*, *barnbokhylla*, *fulsnygg* och *honungssöt*. Det rör sig om tre sammansättningar som kan ge närmare information om informanternas uppfattning om sammansättningens struktur med avseende på det semantiska huvudledet (*vagnbarn*, *barnbokhylla*, *fulsnygg*), och en sammansättning (*honungssöt*) där styrkeförhållandet mellan (exempelvis) jämförande och orsaksrelaterade relationer testas.

När informanterna möter den tillfälliga bildningen *vagnbarn* i fritextdelen av enkäten ges de i uppgift att fullfölja en parafra (se exempel 6:1). När ordet återkommer i flervalsdelen ges två alternativa parafrafer med en eller två kontextmeningar (se exempel 6:2).

(6:1) Ett *vagnbarn* är _____

- barn som sitter i vagn
T.ex. ”Barnen är fortfarande små. De är vagnbarn båda två.”

(6:2) Vagnbarn

- vagn som man har det lilla barnet i
T.ex. ”Det är dyrt att köpa vagnbarn.”

Det är relativt många av informanterna som inte ger någon förklaring till sammansättningen *vagnbarn* i fritextdelen av enkäten. 25 personer, eller 13 % av informantgruppen, avger inget svar, vilket är ett betydligt större bortfall än bortfallet för de flesta av de andra sammansättningar som förekommer i fritextdelen. Detta kan ses som ett tecken på en osäkerhet om sammansättningens betydelse. En av informanterna ger explicit uttryck

för sin osäkerhet: ”ingen aning, har ingen gissning”. Av dem som svarar ger två tredjedelar av informanterna ett svar som speglar en struktur med huvudet till höger, dvs. ’en sorts barn’. Detta kan jämföras med resultatet i flervalsdelen av enkäten, där bortfallet är mycket begränsat (5 informantsvar), och där 78 % av de avgivna svaren speglar en högerstruktur. Merparten av informanterna ger en förklaring till *vagnbarn* i fritextdelen som strukturellt överensstämmer med det alternativ som de väljer i flervalsdelen av enkäten, men några av informanterna gör alltså olika tolkingar i de olika delarna. Dessa informanter gör i typfallet en avvikande tolkning med huvudledet till vänster (dvs. ett *vagnbarn* är en ’vagn för ett barn’) när de först möter ordet i fritextdelen. När ordet återkommer i flervalsdelen av enkäten och kontexten finns där som stöd väljer de det alternativ som speglar en högerstruktur.

I gruppen SVE är det i princip ingen skillnad mellan svaren i fritextdelen och flervalsdelen: drygt 90 % högertolkingar, ’en sorts barn’ i båda delar. För informanterna i SVE+ med svenska och ytterligare ett modersmål är det större skillnad mellan svaren i enkätens olika delar. I fritextdelen ger 44 % av informanterna en högertolkning (’en sorts barn’), men i flervalsdelen av enkäten är det så många som 63 % av informanterna som väljer alternativet med huvudet till höger, dvs. ’barn som sitter i vagn’. I gruppen SVE+/S följs däremot tolkningarna i enkätens båda delar åt: 69 % högertolkning i fritextdelen och 67 % i flervalsdelen. När det gäller tolkningarna från informanterna i gruppen START<6 skiljer sig dessa åt mellan fritext- och flervalsdelen. Drygt hälften av dem ger i fritextdelen en förklaring som följer mönstret med huvudet till höger. I flervalsdelen väljer tre fjärdedelar av informanterna denna betydelse. I gruppen START≥6 är skillnaderna mellan fritext- och flervalsdelen ännu större: bara 22 % ger i produktion en förklaring som ryms inom en högerstruktur, dvs. ’en sorts barn’, medan 60 % av dem följer högermönstret i flervalsdelen av enkäten. Siffrorna tyder på att de flerspråkiga informanterna är mer kontextberoende än L1-talarna. De flerspråkiga informanterna (med undantag av gruppen SVE+/S) har svårigheter att förklara den oetablerade sammansättningen *vagnbarn*, men med stöd av färdiga alternativ och en (visserligen begränsad) kontext når de en större andel högertolkingar i flervalsdelen än i fritextdelen. L1-talarna med svenska som enda modersmål tolkar sammansättningen enligt det gängse strukturella språkmönstret utan stöd av kontext.

Sammansättningen *barnbokhylla*, dvs. en treledad sammansättning utan foge-s, finns också representerad i båda delar av enkäten. I flervalsdelen förekommer den med svarsalternativen ”liten hylla som passar i ett barnrum” och ”hylla med barnböcker, t.ex. på ett bibliotek”, medan den i fritextdelen tolkas fritt. Utifrån fogeregeln kunde man förvänta sig en seg-

mentering *barn-bokhylla*, i kontrast till segmenteringen *barnbok-s-hylla* vid förekomst av *foge-s*, men så ser det inte ut i materialet. Liksom i flervalsdelen finns det i fritextdelen ingen tydlig preferens för tolkningen där *bokhylla* utgör huvudled och *barn* modifierare. I informantgruppen som helhet är det endast knappt hälften som ger sådana förklaringar, vilket överensstämmer med resultatet i flervalsdelen av enkäten. Resterande förklaringar fördelar sig lika mellan tolkningar som kan inordnas under parafrasen 'hylla för/med barnböcker' (dvs. den förklaring som motsvaras av uttrycket *barnbokshylla*) och tolkningar som ryms under parafrasen 'bokhylla för barnböcker'. I det senare fallet utnyttjas det mellersta ledet, *bok*, både till huvudledet (*bokhylla*) och till förledet (*barnbok*). Förklaringarna tyder på att förekomsten eller avsaknaden av fog inte nämnvärt påverkar det semantiska huvudets räckvidd.

Det föreligger inte några stora skillnader mellan informantgrupperna när det gäller hur *barnbokshylla* segmenteras. Om man däremot granskar de semantiska detaljerna i förklaringarna, har svaren från informanterna med svenska som andraspråk mindre semantisk precision än svaren från modersmålstalarna, och de är svårare att inordna i kategorier, t.ex. "ligger alla barns leksaker någonting (kanske)" eller "där man lägger olika saker i".

I tolkningarna av *fulsnygg* aktualiseras många delar i meningspotentialen. I detta fall blir det ännu tydligare att tolkningarna inte enkelt kan sorteras in i de kategorier som informanterna möter i flervalsdelen av enkäten, nämligen "mer snygg än ful", "mer ful än snygg" respektive "mitemellan ful och snygg". I kontrast till dessa svarsalternativ förekommer många fritexttolkningar där *ful* och *snygg* betecknar olika dimensioner, förklaringar av typen 'snyggt yttre, fult inre'. Den semantiska bredden och kreativiteten i tolkningarna är stor, och det är uppenbart att informanternas tolkningar tvingas in i för snäva alternativ i flervalsdelen av enkäten, dvs. alternativen i denna del fångar inte komplexiteten i informanternas konceptualiseringar och ger därmed inte en rättvisande bild av deras tolkningar. Tolkningarna av AA-sammansättningar som *fulsnygg* återkommer i avsnitt 6.3 som handlar om identifiering av semantiskt huvudled med och utan kontext. I det avsnittet diskuteras det kunskapsbidrag som tolkningarna i fritextdelen ger på så sätt att de illustrerar svårigheten med att fastställa det semantiska huvudledet i AA-sammansättningar.

Fritextförklaringarna av *fulsnygg* ger närmare information om en skillnad i tolkning mellan grupperna. Informanterna i gruppen START \geq 6 väljer i större utsträckning än de övriga det mest konventionaliserade alternativet "mer snygg än ful" i flervalsdelen. I fritexttolkningarna har de en hög andel tolkningar där förledet fungerar förstärkande, t.ex. "jättesnygg" och "för snygg", dvs. tolkningar där ledet *ful* inte har någon re-

presentation i tolkingen. Detta tyder på att de tolkar sammansättningen som *full-snygg*.

Även sammansättningen *honungssöt* förekommer i enkätens båda delar. I flervalsdelen kunde informanterna välja mellan ett jämförande alternativ ”mycket söt (som honung)” och ett orsakssamband mellan leden ”söt av honung”. Kontextmeningen till det jämförande alternativet omtalar en ”honungssöt klänning” (dvs. *honungssöt* i den metaforiska betydelsen ’sockersöt’), medan alternativet med ett orsakssamband mellan leden illustreras med en kontextmening som omtalar ”en kopp honungssött te”. I informantgruppen som helhet valde 78 % det jämförande (metaforiska) alternativet. I fritextsvaren, där informanterna fick fullfölja parafrasen ”något som är *honungssött* är ...”, gjorde nästan alla informanter en jämförande tolkning. I några av dem förekommer metaforik, när en person uppges vara ’söt som honung’ (t.ex. ”att personen är för söt!”) eller när *honungssött* beskrivs som något ’alltför sött’ (t.ex. ”något som är så sött att man gör grimaser”), eller när *honungssött* likställs med ”sockersöt”.²⁴ I de flesta av svaren anges *honungssött* rätt och slätt betyda ’sött (som honung)’. Detta visar att jämförelse ligger nära till hands när informanterna formulerar egna förklaringar.

Sammanfattningsvis illustrerar de testord som förekommer både i fritextdelen och flervalsdelen tre saker. För det första synliggör de att kontexten har potential att påverka förståelsen av sammansättningen. För det andra indikerar de en större semantisk spännvidd (vid alla sammansättningar) och en lägre grad av semantisk precision (vid de etablerade sammansättningarna) i L2-talarnas fritexttolkingar än i L1-talarnas motsvarande tolkingar. Därmed blir skillnaderna mellan svaren i fritextdelen och flervalsdelen större för andraspråkstalarna än för modersmålstalarna. För det tredje illustreras svårigheterna med att fastställa det semantiska huvudledets räckvidd i treledade nominala sammansättningar och huvudledet i AA-sammansättningar.

6.2. Val av semantisk relation mellan leden

I föregående avsnitt redovisades skillnaderna mellan fritext- och flervalsdelen för de sammansättningar som förekommer i båda delar av enkäten. I det följande analyseras vilka semantiska relationer som är prefererade i svaren vid de övriga sammansättningarna i fritextdelen, och resultatet

24 Informanternas förklaringar återges här och i det följande oredigerade, inklusive konguens- och stavfel.

jämförs med tolkningarna av motsvarande sammansättningar i flervalsdelen av enkäten.

6.2.1. Material, innehåll eller ändamål

Den semantiska relationen <gjord av MATERIAL/SUBSTANS> har i tidigare studier visat sig vara en lättillgänglig semantisk relation vid mindre etablerade sammansättningar (såsom *leather needle* hos Maguire et al. 2010a; se avsnitt 2.2). Så ser det också ut i enkättolkningarna. MATERIAL som semantisk relation mellan leden slår igenom starkt i tolkningarna av den mindre etablerade sammansättningen *glaslåda* i fritextdelen av enkäten. Nästan fyra femtedelar av informantgruppen tolkar *glaslåda* som en 'låda tillverkad av materialet glas', snarare än en 'låda för glas av något slag' (dvs. ÄNDAMÅL) eller 'låda med glas av något slag' (dvs. INNEHÅLL). Resultatet följer resultatet i flervalsdelen av enkäten, där 84 % av informanterna väljer MATERIAL-tolkningen av *plastlåda*. Tendensen i enkäten är att informanterna som har svenska som L1 är mer benägna att välja MATERIAL-tolkningar än de informanter som har svenska som L2, både i de fria tolkningarna av *glaslåda* i fritextdelen och i de styrda tolkningarna av *plastlåda* i flervalsdelen.

Tolkningar som speglar den semantiska relationen ÄNDAMÅL slår igenom starkt när informanterna förklarar den etablerade sammansättningen *blomkruka* i fritextdelen. Fyra femtedelar av de avgivna svaren speglar denna relation, dvs. 'kruka för blomma', trots att även 'kruka med blomma i' (dvs. relationen INNEHÅLL) är en vedertagen betydelse i sammansättningens meningspotential. ÄNDAMÅL slår igenom starkare vid *blomkruka* i fritextdelen än vid *blomlåda* i flervalsdelen, vilket är förväntat med tanke på frekvensförhållanden. Liksom i flervalsdelen har informanterna med svenska som L1 en högre grad av ändamålstolkningar än informanterna i gruppen START_{≥6}.

6.2.2. Jämförande egenskapsprojicering eller tematisk relationslänkning – djur och artefakter

Styrkeförhållandet mellan jämförande egenskapstolkning (XY är en Y som i något avseende är som X) och tematisk relationslänkning (XY är en Y som har med X att göra) har diskuterats i tidigare forskning (se avsnitt 2.3).

I mitt material får jämförande egenskapstolkning stort genomslag i de sammansättningar där ett eller båda leden betecknar ett djur. Detta gäller inte bara tolkningarna av *zebralejon* (och de andra djursammansättningar) i flervalsdelen av enkäten, utan också tolkningarna av *zebraväska* i

fritextdelen. Närmare 80 % av informanterna formulerar egenskapsprojicerande förklaringar till sammansättningen *zebraväska*, dvs. tolkningar av typen 'randig väska', trots att det rör sig om en tillfällig sammansättning. Majoriteten av de övriga svaren speglar tolkningar där förledets referent betecknar väskans material, dvs. 'väska av zebraskinn'. Några informanter konceptualiserar i stället *zebraväska* som en 'väska med en zebra på'. Ytterligare några informanter ger tolkningar av typen 'väska av märket zebra' respektive 'väska för zebror'. Sammansättning med *zebra* som förled ger sammanfattningsvis en hög grad av egenskapsprojicerande tolkningar både i de fria förklaringarna och i flervalsdelen. Detta gäller samtliga informantgrupper.

Vid den tillfälliga sammansättningen *issaft* är sådana tolkningar som kan kategoriseras under paraplytermen tematisk relationslänkning prefererade framför egenskapsprojicerande tolkningar. Bara en fjärdedel av informanterna gör tolkningar som tydligt faller inom en egenskapsprojicerande ram, dvs. tolkningar där en egenskap från modifierarens referent projiceras på huvudledets referent. I dessa fall ger förledet *is* associationer till isens kyla, varvid *issaft* tolkas som 'kall saft'. I övrigt görs relationslänkande tolkningar av typen 'saft med is i'. I vissa fall är saften och isen så integrerade i informanternas förklaringar att det är svårt att avgöra om det rör sig om en saft med is i (relationslänkning), en saft som har vissa egenskaper från isen (egenskapsprojicering) eller en saft som också är is (identitet). Som exempel förekommer tolkningar av typen 'fryst saft' och 'isig saft'. Ett fåtal informanter tolkar *issaft* som 'saft från is', dvs. som 'smältvatten'. Resultatet för *issaft* följer samma mönster som den etablerade sammansättningen *isvatten* i flervalsdelen av enkäten, som i tre fjärdedelar av fallen tolkas med en tematisk relationslänkning mellan leden, dvs. 'vatten med is i'. Skillnaderna mellan informantgruppernas tolkningar av *issaft* är obetydliga, till skillnad från tolkningarna av *isvatten* där informanterna som börjat med svenska vid sex års ålder eller senare har en betydligt högre grad av egenskapsprojicering än de övriga grupperna.

6.2.3. Jämförande egenskapstolkning, identitet eller tematisk relationslänkning – personbetecknande sammansättningar

Hur informanternas tolkningar fördelar sig mellan jämförande egenskapstolkning, identitetstolkning och tematisk relationslänkning testas i en kategori av personbetecknande sammansättningar. I fritextdelen av enkäten

möter informanterna *kvinnoman* och *kvinnoläkare*, och i flervalsdelen möter de *kvinnokarl* och *mansläkare*.

När informanterna formulerar förklaringar till *kvinnoläkare* beskriver de i hög grad *kvinnoläkare* som en 'kvinnlig läkare', i strid med den konventionaliserade betydelsen 'läkare för kvinnors sjukdomar'. 60 % av informanterna gör en sådan identitetstolkning. Resten av informanterna gör, med något undantag, tolkningar som speglar sammansättningens konventionaliserade betydelse. Resultatet följer därmed resultatet i flervalsdelen, där tolkningarna av motsvarigheten *mansläkare* fördelar sig lika mellan identitetstolkning ('läkare som är man') och tematisk relationslänkning ('läkare som arbetar med sjukdomar som bara män kan få'), trots att sammansättningen är lexikaliserad enligt det senare mönstret.

Vid *kvinnoman* fördelar sig tolkningarna över fler alternativ än de som förekommer som svarsalternativ till *kvinnokarl* i flervalsdelen av enkäten. En fjärdedel av informanterna gör relationslänkande tolkningar av *kvinnoman*, där mannen beskrivs ha många intima relationer med kvinnor. I en del av dessa svar är mannen tydlig agent, såsom i förklaringen "en man som jagar kvinnor", dvs. betydelsen motsvarar betydelsen hos *kvinnokarl*. I en del av svaren konstrueras mannen snarare som mottagare, t.ex. "en man som kvinnor dras till". En tredjedel av informanterna gör jämförande egenskapstolkningar, som fördelar sig jämnt mellan 'man som är kvinnlig' och 'kvinna som är manlig'. I det senare fallet rör det sig således om tolkningar där sammansättningens vänstra led utgör det semantiska huvudledet. Utöver tematisk relationslänkning och jämförande egenskapstolkning framträder ett tredje mönster som får stort genomslag i tolkningarna av *kvinnoman*, nämligen ett mönster som rör könsöverskridande identitet. I så många som 40 % av tolkningarna ges förklaringar till *kvinnoman* som skulle kunna sammanfattas under paraplytermen *transperson*. I vissa av tolkningarna utgör efterledet det kön som personen identifierar sig som, medan förledet markerar biologiska drag, t.ex. "en person som fysiskt är en tjej men mentalt en kille". I andra fall markerar efterledet biologiska drag och förledet identitet, t.ex. "en kvinna i en mans kropp". I den här typen av tolkningar överskrider gränsen mellan semantiskt huvudled och semantisk modifierare. Det blir irrelevant och missvisande att vikta sammansättningens delar, och det går inte heller att kategorisera tolkningarna efter etablerade semantiska relationer. Tendensen i materialet är att informanterna med svenska som L2 är något mer benägna än informanterna med svenska som L1 att välja tolkningen 'transperson', medan L1-talarna är mer benägna att göra relationslänkande tolkningar av typen 'man som tjuvar kvinnor', där *kvinnoman* får en betydelse som ungefär motsvarar betydelsen hos den etablerade sammansättningen *kvinnokarl*.

6.2.4. Agens vid verbavledning

Vid verbavledningar med personbetecknande förled är det semantiskt ambiguöst om förledet utgör verbhandlingens PATIENT eller AGENT (se avsnitt 4.3.1.6). I enkätens fritextdel representeras denna sammansättningstyp av *barnmisshandel* och *politikerförakt*.

Samstämmigheten i tolkningarna av *barnmisshandel* är mycket stor. Alla informanter utom en gör en tolkning där det rör sig om ett barn som är drabbat av misshandel, fysiskt och/eller psykiskt. Bara en informant (i gruppen START \geq 6) tolkar barnet som agent, i en krigskontext: ”barn har vapen, hjälp land gå till krig”. Patienttolkningen är prefererad även vid *kvinnomisshandel* och *mansmisshandel* i flervalsdelen av enkäten, medan tolkningarna av *polismisshandel* fördelar sig jämnt mellan agent- och patienttolkning.

I kontrast till resultatet för *barnmisshandel* är osäkerheten stor i informantgruppen när det gäller den semantiska relationen mellan leden i den etablerade sammansättningen *politikerförakt* (som motsvaras av den tillfälliga bildningen *studentförakt* i flervalsdelen av enkäten). Nästan en tredjedel av informanterna ger inget förslag på vad *politikerförakt* skulle kunna betyda. Av dem som ger en förklaring är det så många som 15 % som missuppfattar *förakt*. I flertalet av dessa tolkningar uppfattas *förakt* som en politisk akt som ’äger rum i förväg eller i ett inledningsskede’ (jfr *för-akt*), t.ex. ”uppvärmningen inför en politisk debatt”, ”politiker agerar i förhand” och ”en inledning en politiker håller”. I knappt två tredjedelar av de avgivna svaren ges en tolkning som någorlunda väl speglar den etablerade betydelsen där politikern utgör sammansättningens PATIENT mot vilken förakt riktas, medan en knapp tiondel betraktar politikern som AGENT. Detta är en lägre andel patienttolkning än i tolkningarna av *studentförakt* i flervalsdelen där 85 % av informanterna betraktar förledet som PATIENT. I detta fall visar jämförelsen av tolkningarna i enkätens delar att de färdiga alternativen och kontextmeningen, där *förakt* indirekt förklaras, hjälper informanterna att tolka sammansättningen och styr in dem mot större samstämmighet och konventionalitet. Detta är särskilt tydligt i grupperna av andraspråkstalare, som är betydligt mer samstämmiga i tolkningarna i flervalsdelen än i fritextdelen. Skillnaderna mellan tolkningarna av agens i fritext- och flervalsdelen är särskilt stora för de informanter som börjat med svenskan vid sex års ålder eller senare.

6.2.5. Betydelsesammansmälta AN-sammansättningar

Tolkningen av betydelsesammansmälta AN-sammansättningar som konventionaliserats i en mer specifik betydelse än vad leden anger (t.ex. *småkakor*; se avsnitt 3.2.2) visar hur informanterna förhåller sig till icke-kompositionalitet. I fritextdelen av enkäten tolkar informanterna *grönområde* (som motsvaras av *grönyta* i flervalsdelen av enkäten).

I mötet med *grönyta* i flervalsdelen väljer 71 % av informanterna alternativet ”naturområde med stor gräsplan”, där exempelmeningen omtalar en person som uppsöker stadens grönytor för att promenera, medan de övriga väljer alternativet ”yta som är grön”, där exempelmeningen beskriver hur en person målar väggarna gröna för att hon älskar grönytor. Medan kontextmeningarna i flervalsdelen tydliggör skillnaden mellan det specifika tillika konventionaliserade alternativet kontra det strikt kompositionella, förblir det i många av fritextsvaren oklart huruvida informanterna uppfattar denna distinktion. En femtedel av informanterna ger förklaringar av typen ’grönt område’, utan att närmare specificera vad det rör sig om för område, så det framgår inte om de är medvetna om den semantiska preciseringen att ett *grönområde* är ’grönt av växtlighet’. I drygt två tredjedelar av svaren förklaras det explicit att området är ’grönt av växtlighet’. I vissa av dessa svar ingår naturvård som en semantisk komponent hos sammansättningen, t.ex. i förklaringar som ”ett naturreservat” och ”ett ekovänligt område”.

Fritexttolkningarna följer i stora drag resultatet i flervalsdelen. I flervalsdelen av enkäten har informanterna i gruppen SVE signifikant fler konventionella tolkningar än de flerspråkiga informanterna. I fritextdelen fördelar sig tolkningarna på jämförbart sätt över de övergripande kategorierna, oavsett vilken informantgrupp man tillhör. Det finns dock en tendens till större semantisk precision i förklaringarna från modersmålstalarna av svenska (t.ex. ”en park eller skog i en stad/tätort”) än i andraspråkstalarnas förklaringar, som i vissa fall innehåller mindre konventionaliserade semantiska komponenter (såsom i svaret ”en bonde gård fylld med grönt gräs” eller ”en rabbat/gräsplan”). Ett intressant inslag i fritextdelen är det tiotal svar där sammansättningen tolkas metaforiskt; *grönområde* kommer då att beteckna ett område (konkret eller abstrakt) som är grönt i betydelsen ’ofarligt’ eller ’okej’, t.ex. ”en zon man är trygg i” eller ”något som man kan snacka om i en diskussion”.

6.2.6. Jämförelse

Liksom i flervalsdelen slår jämförelse igenom starkt i informanternas fritexttolkningar av NA-sammansättningarna *honungssött*, *sandmjukt*,

blodrött och *dammgrått* i fritextdelen av enkäten. Som redan framgått tolkas *honungssött* nästan uteslutande jämförande i fritexttolkningarna (se avsnitt 6.1), vilket innebär att sammansättningen tolkas jämförande i högre grad i fritext- än i flervalsdelen där tolkningarna fördelar sig 78/22 % mellan jämförelse och orsak. Det föreligger inte heller några nämnvärda skillnader mellan informantgruppernas tolkningar. *Sandmjukt* i fritextdelen motsvaras av *snömjuk* i flervalsdelen. Vid *snömjuk* är 83 % av tolkningarna jämförande, 'mjuk som snö', medan 17 % väljer alternativet 'mjuk av snö'. I de fria tolkningarna är informanterna ännu mer benägna att göra en jämförande tolkning där någonting är 'mjukt som sand' eller 'mjukt och strävt som sand'. I några tolkningar antas en person vara 'mjuk som sand'. Inte i någon tolkning lyfts alternativet att *sandmjuk* skulle kunna beteckna något som är 'mjukt av sand', såsom i fallet med *sandstrand*. Även färgsammansättningarna ges i stor utsträckning jämförande parafraaser, på samma sätt som jämförelse slår igenom starkt i tolkningarna av *snövit* och *solgul* i flervalsdelen (91 % respektive 87 % jämförande relation).

Blodrött är en etablerad och frekvent sammansättning. Tre fjärdedelar av tolkningarna är direkt jämförande 'rött som blod' medan bara fem tolkningar eller 3 % är orsakrelaterade 'röd av blod'. Fyra av dessa formuleras av informanter i $START \geq 6$. En knapp fjärdedel av tolkningarna är tolkningar av typen 'mörkrött'. Dessa speglar troligen en jämförande konceptualisering ('mörkrött som blod'), men man kan inte utesluta en konceptualisering där ett orsakssamband föreligger ('mörkrött av blod'). Även färgsammansättningen *dammgrått* tolkas i hög grad jämförande. I två tredjedelar av tolkningarna uttrycks jämförelseledet explicit, 'grått som damm'. I övrigt förekommer tolkningar av typen 'nyans av grått' eller 'tråkigt'. I 12 tolkningar (dvs. 8 % av tolkningarna), uttrycks en orsakrelation mellan eller en samförekomst av ledens referenter 'grått av/med damm'.

När informanterna formulerar egna förklaringar till den tillfälliga bilden *musrädd* (som motsvaras av *hundrädd* och *fågelrädd* i flervalsdelen av enkäten) får mönstret [rädd för N] större genomslag i tolkningarna än ett jämförande mönster. Nästan tre fjärdedelar formulerar förklaringar där en *musrädd* person är rädd för förledets referent (i de flesta tolkningar djuret, i ett fåtal tolkningar kvinnans könsorgan). Tio tolkningar (6 %) är direkt jämförande 'rädd som en mus'. I en femtedel av tolkningarna formuleras förklaringar av typen 'väldigt rädd, lättskrämd'. I dessa fall är det oklart om informanten gör en jämförelse ('väldigt rädd som en mus') eller om informanten konceptualiserar en person som är så rädd och nervös att den till och med låter sig skrämmas av en mus. Informanterna i gruppen $START \geq 6$ är mer benägna att välja jämförelse än informanterna i de andra

grupperna. Tolkningarna i fritextdelen följer i stora drag tolkningarna av *hundrädd* och *fågelrädd* i flervalsdelen, där 90 % respektive 85 % av informanterna väljer [rädd för N], i stället för jämförelsemönstret [rädd som N].

6.3. Sammansättningens huvudled i informanternas fritexttolkningar

När det gäller identifieringen av det semantiska huvudledet i sammansättningar varierar tolkningarna av de enskilda sammansättningarna i fritextdelen. Detta indikerar att inte bara det strukturella schemat utan också semantiska aspekter och frekvens har betydelse för informanternas tolkningar. Informanterna är samstämmiga i tolkningarna av den frekventa sammansättningen *husbil*, medan osäkerheten är större om betydelsen hos en lågfrekvent sammansättning som *skalpotatis* och en tillfällig bildning som *vagnbarn*, särskilt bland de informanter som har svenska som L2. Den stora skiljelinjen går dock mellan tolkningarna av de nominala och de adjektiviska sammansättningarna. Tolkningarna av *fulsnygg* och *dumsnäll* illustrerar att det vid AA-sammansättningar är svårt att överhuvudtaget fastställa vilket (eller vilka) led som utgör semantiskt huvudled.

Samstämmigheten i tolkningarna av *husbil* är mycket stor, vilket dels kan ha med frekvens och semantik att göra, dels med struktur. *Husbil* är en frekvent sammansättning, och referenten, hybriden av ett hus och en bil, kan betraktas som semantiskt lättillgänglig och salient. Även sammansättningens struktur kan inverka. Eftersom sammansättningen är predikativ, en *husbil* är 'en bil som är som ett hus', faller inte bara 'en sorts bil'-tolkningarna utan även 'en sorts hus'-tolkningarna inom ramen för högertolkningar, t.ex. förklaringen "ett hus på hjul". Därmed blir inte informanternas behärskning av schemat [efterled=huvud] utslagsgivande. I enkätmaterialiet återfinns endast två svar som avviker från högerstrukturen, nämligen två tolkningar av *husbil* som ett 'hus för en bil', dvs. ett 'garage'. Samstämmigheten är därmed större än vid tolkningen av *husbåt* i flervalsdelen av enkäten, där 12 % av tolkningarna speglar en struktur med huvudet till vänster. Skillnaden kan sammanhånga med att *husbil* är en mer frekvent sammansättning än *husbåt*. Skillnaderna mellan tolkningarna i fritext- och flervalsdelen är stora för informanterna i gruppen $START \geq 6$, som i nästan en tredjedel av fallen tolkar *husbåt* i flervalsdelen som ett hus där man lägger båten.

Skalpotatis är en mindre frekvent sammansättning, och den har också en lägre andel högertolkningar. Fyra femtedelar av informanterna tolkar

sammansättningen som 'en sorts potatis'. Resterande tolkningar speglar en struktur med huvudet till vänster. I en del av dessa utgör 'skal' det semantiska huvudledet, men många av svaren tyder på att informanterna konceptualiserar en 'potatisskalare'. Informanterna i gruppen START \geq 6 skiljer ut sig från de övriga med drygt två tredjedelar tolkningar där 'skal' eller 'skalare' utgör semantiskt huvudled, medan bara en knapp tredjedel av tolkningarna är av typen 'en sorts potatis'. Också SVE+ har en relativt stor andel vänstertolkningar (en tredjedel), och i hälften av dessa konceptualiserar informanterna en 'potatisskalare'.

Som redan redovisats är det för de flerspråkiga informanterna stor skillnad mellan hur de tolkar *vagnbarn* när de formulerar sig fritt och när de väljer mellan två färdiga svarsalternativ. De har en betydligt lägre grad av högertolkningar när de själva formulerar förklaringen. Informanterna i gruppen SVE har samma (höga) andel högertolkningar i båda delarna av enkäten. Sammansättningen torde vara okänd för de flesta av informanterna. De som har svenska som modersmål identifierar ändå 'barn' som det semantiska huvudledet redan i fritextdelen, medan de övriga grupperna är mer beroende av den kontext som ges i flervalsdelen.

Fritexttolkningarna av AA-sammansättningarna *fulsnygg* och *dumsnäll* visar att betydelsen hos sådana sammansättningar är svår att fånga genom en enkel kategorisering i å ena sidan kopulativa, å andra sidan determinativa sammansättningar. Relationen mellan sammansättningens led är betydligt mer komplex. I vissa av tolkningarna samexisterar det fula och det snygga, utan att det blir tydligt vad som överväger, såsom i svaret "på gränsen mellan ful och snygg, båda samtidigt" eller "en som inte ser bra ut men ändå gör det". I vissa förklaringar tycks den ena egenskapen slå över i dess motsats, t.ex. "så snygg att hen blir ful, för 'perfekt'" eller – tvärtom – "så ful så att personen blir snygg igen". I flertalet svar råder ett motsatsförhållande mellan det yttre och det inre, såsom i "fint utseende, ful personlighet eller tvärtom". I några fall tycks materiell rikedom väga upp utseendet, såsom i "en person med mycket pengar". Några informanter tolkar den *fulsnygge* som attraktiv trots viss fulhet, t.ex. "otroligt attraktiv pga. ful med stil". I andra fall har fulheten kontra snyggheten med tillfället att göra: "ibland snygg ibland ful" eller "snygg när han vill". I åter andra förklaringar förekommer ett spänningsförhållande mellan objektivt och subjektivt, t.ex. "egentligen ful men man tycker ändå personen är snygg" eller "snygg trots att den i mångas ögon anses vara ful". Några av dessa tolkningar kommer nära det som brukar anses som en gängse definition av *fulsnygg*, t.ex. "speciell i sitt utseende och det blir attraktivt" eller "någon som inte är traditionellt attraktiv (inom mediabildens ramar) men ändå är snygg". I dryga tiotalet tolkningar har ledet *ful* ingen representation i tolkningen, t.ex. "jättesnygg/vacker". Dessa förekommer bara i andra-

språkstalarnas tolkningar och tyder på en läsning *full-snygg* (se vidare i avsnitt 6.1).

Sammantaget framträder en bild där det råder ett spänningsförhållande mellan det fulla och det snygga i *fulsnygg*. Motsatserna kan gälla olika dimensioner, t.ex. yttre och inre egenskaper. I andra fall gäller spänningsförhållandet den ”objektiva” kontra den ”subjektiva” bilden. Det blir närmast omöjligt – för att inte säga irrelevant och missvisande – att försöka fastställa vilket av sammansättningens led som är att betrakta som det semantiska huvudledet.

Tolkningarna av *dumsnäll* illustrerar delvis samma problematik som tolkningarna av *fulsnygg*. Men eftersom leden inte står i direkt motsatsförhållande till varandra framstår det som något enklare att fastställa det semantiska huvudledet i *dumsnäll*. I 40 % av tolkningarna formuleras parafraaser på temat ’(överdrivet) snäll på ett dumt eller naivt sätt’ t.ex. ”naivt snäll, lätt att lura”, ”för snäll för sitt eget bästa” eller ”en idiot med ett stort hjärta”. Här skulle *snäll* kunna ses som semantiskt huvudled, man är ’snäll på ett dumt sätt’. Men i många av förklaringarna framgår inte relationen mellan leden, såsom i ”snäll fast dum” eller ”en person som visar upp två sidor”. Dessutom ställer polysemin hos *dum* till det, vilket gör att många informanter ändå laborerar med en dikotomi, nämligen motsatsförhållandet mellan elak och snäll, t.ex. ”elak ibland men ändå snäll” eller ”en person som retas, men på ett vänskapligt sätt”. Därmed blir det svårt att avgöra vilket som är det semantiska huvudledet. Liksom i fallet med *fulsnygg* kan den ena egenskapen i sammansättningen slå över i dess motsats, t.ex. ”så elak man blir snäll”.

Tolkningarna av AA-sammansättningarna *fulsnygg* och *dumsnäll* illustrerar att sådana sammansättningar svårligen låter sig inordnas i dikotomin determinativa kontra kopulativa sammansättningar. När det gäller AA-sammansättningar, som betecknar egenskaper, är det semantiska innehållet ofta alltför komplext för att beskrivas i termer av att det finns ett semantiskt huvudled (som i determinativa sammansättningar) eller att leden är likställda (som i kopulativa sammansättningar).

6.4. Meningspotentialen hos en sammansättning – exemplet *vagnbarn*

I det följande görs en mer ingående semantisk analys av en sammansättning i fritextdelen, nämligen *vagnbarn*, som får tjäna som illustrerande exempel på sammansättningsars vida meningspotentialer och möjliga konnotationer.

Tolkningarna av *vagnbarn* i fritextdelen av enkäten fördelar sig mellan en tredjedel tolkningar av typen 'en sorts vagn' och två tredjedelar av typen 'en sorts barn'. I nästa steg av analysen är det intressant att se vilken tematisk relationslänkning mellan 'barn' och 'vagn' som åberopas av de 110 informanter som väljer tolkningen att det rör sig om en sorts barn. Av dessa 110 gör 44 en tolkning där relationen mellan leden är PLATS, t.ex. "ett barn i vagn" eller "ett barn som ligger i en vagn". 28 av informanterna ger en förklaring som relaterar till TID, t.ex. "ett barn som fortfarande ligger/sitter i barnvagn" eller "ett barn som inte lärt sig att gå". I ytterligare sju informanttolkningar från både L1- och L2-informanter ses vagnsittandet som ett uttryck för något annat än att barnet är för litet för att gå själv; vagnsittandet blir i dessa fall en symbol för bortskämdhet, såsom i förklaringen "ett barn som fått allt gjort för hen", "någon som haft hönsiga föräldrar i sin uppväxt" eller rätt och slätt "jobbigt barn". 13 av informanterna tolkar vagnbarn HABITUELLT, t.ex. "ett barn som alltid sitter i barnvagn". Ytterligare 12 informanter ger en förklaring med relationen PREFERENS, t.ex. "ett barn som gillar att färdas i vagn" eller "ett barn som kan gå men hellre åker vagn".

En informant gör (sannolikt med glimten i ögat) en jämförande koppling mellan leden och tolkar ett *vagnbarn* som ett "barn som ser ut som vagn". I två förklaringar anger vagnen platsen för barnets tillblivelse: "ett namn på ett barn som tillkommit på en vagn" eller barnets födelse "ett barn som fötts i en spårvagn". Sex svar har hänförs till kategorin "övrigt". I ett av dessa är vagnen en nödvändighet p.g.a. funktionsvariation hos barnet: "ett barn som måste ha vagn, pga handikap". I två fall är barnet upptaget av vagnar i sitt arbete eller på sin fritid, jämför "barn som arbetar med vagnar" och "ett barn som hänger i vagnar". Två tolkningar kan väl uppfattas som något skämtsamma: "ett barn till en vagnfamilj" och "ett barn på hjul". En tolkning från en informant med kurdiska som modersmål lyder "Abdirizag", vilket enligt obekräftade källor är ett förekommande pojknamn i Mellanöstern.

6.5. Skillnader mellan L1- och L2-talarnas tolkningar i fritext- och flervalssdelen

Liksom i flervalssdelen förekommer i fritextdelen skillnader mellan de informanter som har svenska som modersmål och de som har svenska som andraspråk. När det gäller de etablerade sammansättningarna visar sig dessa skillnader i en högre grad av konventionalitet i L1-talarnas än i L2-talarnas tolkningar. Vid de mindre konventionaliserade eller tillfälliga

bildningarna är skillnaden en fråga om en högre respektive lägre grad av samstämmighet.

Informanterna som har svenska som andraspråk har en betydligt större andel konventionella respektive samstämmiga tolkningar i flervalsdelen än i fritextdelen av enkäten, medan skillnaderna mellan fritext- och flervalsdelen är små för de informanter som har svenska som förstaspråk. I det här fallet tycks andraspråkstalarna i högre grad än förstaspråkstalarna dra nytta av de färdiga alternativen och kontextmeningarna i flervalsdelen. Detta gäller i synnerhet informanterna i $START \geq 6$. L2-talarna har vidare en större semantisk spridning och en lägre grad av semantisk precision i sina fritextsvar än L1-talarna. När de i flervalsdelen möter de färdigformularade svarsalternativen behöver informanterna inte processa alla potentiella betydelser, till skillnad från i fritextdelen där mångfalden av aktualiserade betydelser i svaren påvisar en sådan process. Informanterna kan också få hjälp att hitta ett alternativ som de inte skulle ha kommit på själva. Den styrning som informanterna får i flervalsdelen visar sig särskilt gynna de informanter som inte har svenska som modersmål.

Skillnaderna mellan de fria förklaringarna och de styrda tolkningarna kan förstås mot bakgrund av den diskussion som förts om meningspotentialer. Orden i sig bär potentiella betydelser som först i en kontext kan övergå i aktualiserad betydelse.

6.6. Sammanfattning

Fritextdelen av enkäten fyller två syften. Å ena sidan tjänar den som ett kontrollinstrument eller ett referensmaterial till flervalsdelen av enkäten, där informanterna endast kunde välja mellan två eller tre alternativa betydelser. Genom att jämföra tolkningar av samma eller strukturellt lika ord i fritext- och flervalsdelen kan jag kontrollera huruvida ett mönster som slår igenom starkt i flervalsdelen också är starkt i fritextdelen, där informanterna själva formulerar förklaringar. Å andra sidan tjänar fritextdelen syftet att belysa den semantiska spännvidd som uppstår när två ord sätts ihop.

När det gäller det första syftet styrker resultatet i fritextdelen över lag resultatet i flervalsdelen. De mönster som slår igenom starkt i flervalsdelen av enkäten, t.ex. jämförande egenskapstolkningar och materialtolkningar, är starkt representerade också i fritextdelen av enkäten där informanterna formulerar sina förklaringar helt fritt. Konsistensen i tolkningarna är särskilt stor för de informanter som har svenska som modersmål, medan skillnaderna mellan tolkningarna i enkätens båda delar är större för de

informeranter som inte har svenska som modersmål, i synnerhet för dem med en startålder för svenskan vid sex års ålder eller senare.

När det gäller det andra syftet belyser fritexttolkningarna att kategorierna i flervalsdelen är övergripande, och i vissa fall alltför grova. I fritextdelen aktualiseras andra potentiella betydelser. Tolkningarna av AA-sammansättningar i fritextdelen visar exempelvis att det är svårt att fastställa AA-sammansättnings semantiska huvudled. I tolkningarna av NN-sammansättningen *kvinnoman* aktualiseras fler alternativ i meningspotentialen än de som gavs som svarsalternativ till *kvinnokarl* i flervalsdelen av enkäten. En stor andel tolkningar där *kvinnoman* får en könsöverskridande betydelse visar att det inte alltid är enkelt – eller adekvat – att fastställa det semantiska huvudledet heller i NN-sammansättningar. Sålunda kompletterar och nyanserar informanternas egenformulerade förklaringar de resultat som framkommer i enkätens flervalsdel.



7

Analys

I det här kapitlet analyseras informanternas tolkningar och sätts i relation till de hypoteser som presenteras i avsnitt 4.2 i metod- och materialkapitlet. Resultatet analyseras mot bakgrund av den tidigare forskning som presenteras i kapitel 2 och med hjälp av det teoretiska ramverk som presenteras i kapitel 3. Med hjälp av kognitiv teori och lexikal semantik analyseras faktorer som inverkar på informanternas tolkningar av sammansättningarna i fritextdelen och flervalsdelen av enkäten.

Med utgångspunkt i kognitiv teori undersöks inledningsvis vilket genomslag det teoretiska antagandet om en korrelation mellan frekvens och befästning får i mitt material (avsnitt 7.1). I denna del av analysen studeras informanternas tolkningar kategorivis efter en indelning i tolv kategorier av sammansättningar (avsnitt 7.1.1–7.1.12). I genomgången aktualiseras även flerspråkighetsaspekter, såsom interferens och transfer av frekventa mönster från L1 till L2. Därefter följer ett avsnitt där det sammanfattas hur betydelsen av frekvens och befästning avspeglas i materialet (avsnitt 7.1.13). I de följande avsnitten diskuteras genomslaget för jämförelse som kognitiv princip (avsnitt 7.2) och betydelsen av saliens för befästning (avsnitt 7.3). Avsnitt 7.4 behandlar principen om blockering. I avsnitt 7.5 analyseras inverkan från omvärldskunskap, dvs. relationen mellan uttryck och referent, vilket är en viktig faktor i såväl lexikala som kognitiva perspektiv. I avsnitt 7.6 görs en sammanfattande konklusion.

När jag i det följande skriver om *L1-talarna* åsyftas informanterna i SVE, som uppger svenska som sitt enda modersmål, och informanterna i SVE+ som uppger svenska som ett av sina modersmål. *L2-talarna* syftar på de informanter som har svenska som andraspråk, dvs. informanterna i $START < 6$ och $START \geq 6$. *Flerspråkiga informanter* syftar på informanterna i SVE+, $START < 6$ och $START \geq 6$.

7.1. Frekvens och befästning av schematiska mönster och specifika instanser

Inom det kognitiva paradigmet utgår man från att språkbruksfrekvenser påverkar den mentala representationen av språket. Hög tokenfrekvens leder till att specifika instanser, dvs. enskilda ord och uttryck, blir befästa som enheter eller exemplar. Typfrekvens sammanhänger med schematisering och befästning av scheman (se avsnitt 3.1.4).

I det följande visas att de enskilda sammansättningarnas tokenfrekvens har betydelse för informanternas tolkningar av deras betydelse, eftersom frekvens korrelerar med konventionalisering (i en språkgemenskap) och befästning (hos individen). Även typfrekvens har betydelse, eftersom förekomsten av flera typer kan leda till uppkomst och befästning av ett schema på lägre eller högre abstraktionsnivå, som kan fungera vägledande i tolkningen av en mindre känd sammansättning. En okänd sammansättning kan också tolkas med en befäst sammansättning som analogibas, såsom när *berghård* tolkas med samma relation mellan leden som *stenhård*. Språkliga mallar som kan fungera som analogibaser förekommer således på olika schematicitetsnivåer: som enstaka instanser, som lågnivåscheman och som högnivåscheman (Ryder 1994:79–80). Frekventa enheter i språket har större potential att fungera som analogibaser än mindre frekventa (se avsnitt 3.1.5).

7.1.1. Identifiering av huvudled i nominala sammansättningar

I det följande analyseras vilken inverkan schemat för sammansättningsstruktur å ena sidan, och befästningen av enskilda instanser å den andra, har på informanternas bedömning av sammansättningens semantiska huvudled. När det (i detta och följande avsnitt) hänvisas till enskilda sammansättnings frekvens är utgångspunkten frekvensuppgifter från den

referenskorpus som beskrivs i avsnitt 4.4. Sammansättningarnas frekvens i referenskorpusen redovisas i bilaga 3.

7.1.1.1. *Betydelsen av enskilda instanser och scheman vid identifieringen av huvudled*

I resultatkapitlet, avsnitt 5.1.1, redovisas informanternas tolkningar av nominala sammansättningar som är avsedda att ge information om huvudledets status i svenska sammansättningar. Av redovisningen framgår att sammansättningarna i stor utsträckning tolkas på konventionellt sätt med efterledet som huvudled, i enlighet med antagandet i hypotes 2a. Resultatet är intressant med tanke på att kategorin innehåller tillfälliga sammansättningar, vid sidan av de etablerade. Därmed bekräftar empirin hypotes 1a om att huvudets placering sist i nominala sammansättningar är en etablerad struktur i svenskan.²⁵

Det finns en svag tendens i materialet att de sammansättningar som har många förekomster i referenskorpusen (t.ex. *husbåt*) tolkas med huvudet till höger i större utsträckning än de sammansättningar som har få förekomster (t.ex. *båtsommar*), dvs. högre tokenfrekvens leder till större samstämmighet i informantgruppen (såsom antas i hypotes 2c). Resultatet indikerar skillnader avseende befästning mellan frekventa och mindre frekventa instanser (se avsnitt 3.1.4), och befästa sammansättningar behöver inte analyseras utifrån det övergripande schemat för sammansättningsstruktur.

Det finns ett tydligt samband i materialet mellan exponeringstiden för svenskan och andelen högertolkningar. De informanter som har längre exponeringstid för svenskan (dvs. L1-talarna i SVE och SVE+) har en större andel tolkningar som följer sammansättningsschemat med efterledet som sammansättningens huvud än de informanter som har svenska som L2 och en kortare exponeringstid för svenskan (dvs. grupperna START<6 och START≥6; se resultatredovisningen i avsnitt 5.1.1.3). Det finns signifikanta skillnader mellan informantgruppernas svar vid alla enskilda sammansättningar i kategorin (se tabell 5:1). I fritextdelen av enkäten är skillnaden mellan grupperna ännu större. Sammantaget ges empiriska belegg för hypotes 3a, dvs. det föreligger skillnader mellan informanternas tolkningar av sammansättningens huvudled beroende på om svenskan är

²⁵ Sammansättningen *båthus* utmärker sig emellertid med en synnerligen liten andel högertolkningar (55 % i informantgruppen som helhet). En bidragande orsak till detta skulle kunna vara att distraktorn för *båthus* faktiskt är kompatibel med regeln om huvudet som efterled om man tänker sig en predikativ tolkning av ordet, nämligen att ett båthus är ett 'hus som är som en båt'. Denna parafras ligger betydelsemässigt nära distraktorn 'båt som man kan bo i'.

deras första- eller andraspråk. Störst skillnader mellan informantgruppernas svar noteras för de lexikaliserade men mindre frekventa sammansättningarna *båthus*, *båtsommar* och *boktext*.²⁶ Utifrån kognitiv teori är det ett förväntat resultat att skillnaden mellan L1- och L2-talarna generellt ökar med avtagande frekvens (såsom antas i hypotes 3d), eftersom befästningen av mindre frekventa sammansättningar kräver en längre exponeringstid än mer frekventa instanser. Sammanfattningsvis har L1-talarna ett försprång framför L2-talarna med avseende på befästning av specifika instanser. Tendensen är att detta tilltar vid lexikaliserade men mindre frekventa sammansättningar, dvs. att vid den typen av sammansättningar är den relativa skillnaden mellan grupperna störst.

Skillnaderna mellan L1- och L2-talarna när det gäller tolkningen av de etablerade sammansättningarna indikerar skillnader avseende lexikal befästning. Sammansättningar som blivit befästa som enheter eller exemplar behöver inte analyseras (Langacker 1987:57, Bybee 2013:61; se vidare avsnitt 3.1.4). Vid tillfälliga sammansättningar kan informanterna däremot inte konsultera sin kunskap om specifika enheter, utan är hänvisade till analogibaser på olika schematicitetsnivåer. Därför är det intressant att notera att de L1-talare som har svenska som enda modersmål (dvs. inte de tvåspråkiga informanterna i SVE+) har en större andel högertolkningar än L2-talarna även vid tillfälliga sammansättningar som *stövelbarn* och *vagnbarn*.

Skillnaderna mellan L1- och L2-talarnas tolkningar av tillfälliga sammansättningar indikerar en skillnad avseende tillgång till det strukturella schemat för svenska sammansättningar. Det strukturella schemat [efterledet = huvudled] är mer befäst hos L1- än hos L2-talarna av svenska på grund av längre exponeringstid för instanser som aktiverar detta schema. Tidigare studier bekräftar att barn tidigt kan identifiera huvudledet i sammansättningar på sitt modersmål (se avsnitt 2.4.1). De enspråkiga L1-talarna har, till skillnad från L2-talarna, bara ett språkligt mönster att beakta, medan L2-talarnas tolkningar i andraspråket kan påverkas av interferens från ett befäst mönster i deras modersmål, som konkurrerar med L2-mönstret (se avsnitt 2.4.2). Även skillnader i tillgång till mer yt-nära scheman, såsom [N-barn] i tolkningen av *vagnbarn*, kan inverka på tolkningarna (se vidare avsnitt 7.6).

Sammanfattningsvis kan skillnaderna i resultat mellan de informanter som har svenska som förstaspråk och de informanter som har svenska

26 *Båtsommar* och *boktext* har låg respektive ingen representation i referenskorpusen. *Båthus* är visserligen relativt frekvent förekommande i referenskorpusen, men det finns skäl att anta att sammansättningen är överrepresenterad, eftersom förekomsterna är koncentrerade till särskilda material i korpusen. Min bedömning är att *båthus* liksom *båtsommar* och *boktext* är lågfrekventa i den typ av språk som informanterna möter.

som andraspråk, med avseende på andelen tolkningar med huvudledet till höger, förklaras med skillnader i mängden svenskspråkig input. Dessa skillnader i input resulterar i skillnader när det gäller befästning av å ena sidan de enskilda sammansättningarna, å andra sidan mönster, såsom det övergripande schemat $[NN_{\text{huvud}}]_N$ och mer ytnära scheman, såsom $[N\text{-barn}]$.

7.1.1.2. Betydelsen av språkbakgrund och startålder vid identifieringen av huvudled

De redovisade skillnaderna i andel högertolkningar mellan informanter som har svenska som förstaspråk och informanter som har svenska som andraspråk är förväntade utifrån skillnader i input. Att de flerspråkiga informanter som har svenska och ytterligare ett språk som L1 (grupp SVE+) har en signifikant lägre grad av högertolkningar än de informanter som har svenska som enda L1 (grupp SVE) är ett mer oväntat resultat. Resultatet sammanhänger med den stora spännvidden avseende språklig bakgrund i informantgruppen SVE+. Till denna grupp räknas informanter som har en flerspråkighetssituation med svenska och ytterligare ett modersmål från vardera föräldern, men också informanter som räknar svenska som sitt modersmål trots att ingen av föräldrarna har detta språk som modersmål (se avsnitt 4.10.2). I gruppen av informanter som uppger svenska och ytterligare ett modersmål (grupp SVE+) är det bara de informanter som *inte* har någon förälder med svenska som modersmål (dvs. informanterna i undergruppen SVE+/A) som skiljer sig signifikant från SVE med avseende på andel högertolkningar i enkäten. Informanterna i SVE+/S, som har en förälder med svenska som modersmål, har sannolikt fått en större mängd svenskspråkig input, och därmed mer gynnsamma förhållanden för utvecklingen av svenskan, än de informanter som inte haft samma språkliga stöd för svenskan från en förälder.

Startåldern för svenskan torde ha betydelse för befästningen av å ena sidan specifika sammansättningar, å andra sidan sammansättningsschemat med efterledet som huvud, eftersom en lägre startålder innebär en längre exponeringstid för svenska språket. Ett metodologiskt problem i min studie är att två flerspråkiga informanter med samma startålder skulle kunna hamna i olika informantgrupper (SVE+/A eller START<6) beroende på hur de tolkar modersmålsfrågan (se avsnitt 4.10.2). I resultatkapitlet (5.1.1.4) redovisas resultatet av en korrelationsanalys mellan startålder och andel högertolkningar för alla informanter som har en flerspråkig bakgrund (dvs. informanter som tillhör någon av informantgrupperna SVE+, START<6 och START≥6). Därmed studeras variabeln startålder isolerat från informanternas egen modersmålsrapportering. Korrelations-

analysen påvisar en signifikant korrelation mellan startålder och andel högertolkningar. En inzoomning på lägre startåldrar visar att korrelationen kvarstår om man avgränsar korrelationsanalysen till startåldrar i intervallet 0–6 år, men inte om man sänker den övre gränsen till fem år. I studien ingår dock få informanter med en startålder fyra eller fem år. Man kan inte utesluta att signifikanta skillnader hade kunnat påvisas också mellan de lägsta startåldrarna om studien innefattat fler informanter med en start för svenskan vid fyra eller fem års ålder. I denna studie påvisas ingen generell, signifikant skillnad mellan flerspråkiga ungdomars identifiering av sammansättningens semantiska huvudled beroende på om de börjat lära sig svenska vid 0, 1, 2 eller 3 års ålder. Som framgått tidigare (avsnitt 5.1.1.4) visar studien däremot att det har betydelse för behärskningen av sammansättningsstruktur om den flerspråkiga individen haft stöd för svenskan från en förälder som har svenska som modersmål.

Sammanfattningsvis tyder resultatet på att skillnaderna mellan SVE och SVE+, med avseende på andelen högertolkningar, primärt är ett resultat av den stora spännvidden inom gruppen SVE+, där bara hälften av informanterna växt upp med stöd för svenskan från en förälder som har svenska som modersmål.

7.1.2. Segmentering av treledade sammansättningar

Av resultatredovisningen i avsnitt 5.1.2 framgår att informanternas segmentering av treledade sammansättningar med och utan fog inte följer den fogeregeln som har potential att särskilja strukturer i sammansättningspar som *barnbokhylla*–*barnbokshylla*. Det finns en svag preferens bland informanterna att välja tolkningen där det semantiska huvudet är begränsat till det sista ledet, oavsett om sammansättningen har utsatt fog eller inte. Inget signifikant samband föreligger heller mellan informantgrupp och svar vare sig för de enskilda sammansättningarna eller för sammansättningskategorin som helhet.

De treledade sammansättningarna i enkäten är förhållandevis lågfrekventa. Det är därför rimligt att anta att de endast i låg grad är befästa som enheter. Av resultatredovisningen för enkätens flervaldsdel (avsnitt 5.1.2) framgår att informanterna har en större andel svar som följer fogeregeln vid de sammansättningar som har utsatt fog (*skolboks försäljning* och *barnbokshylla*) än vid de sammansättningar som saknar fog (*sommardagdröm* och *barnbokhylla*). Detta förhållande är svårt att förklara enbart utifrån de enskilda sammansättningarnas frekvensförhållanden (jfr hypotes 2c). Resultatet tyder snarare på att informanterna följer en annan segmenteringsprincip än vad fogeregeln anger. Informanterna har en benägenhet att segmentera mellan led 2 och 3, oavsett om det finns en

utsatt fog eller inte, dvs. de föredrar ett enledat huvud framför ett tvåledat. Detta resulterar i att tolkningarna av sammansättningar med utsatt foge-s överensstämmer med fogeregeln, medan sammansättningarna utan fog tolkas i strid med fogeregeln.

Det är troligt att språkbrukare på grund av frekvensförhållanden omedvetet läser in ett foge-s och tolkar s-lösa strukturer på samma sätt som motsvarande s-strukturer. Merparten av treledade, nominala sammansättningar har utsatt fog och segmenteras med den stora ordgränsen mellan det andra och det tredje ledet jämför [XXsY]. Detta schema kan därför antas vara mer befast hos språkbrukare än schemat där huvudet är tvåledat [XYY].

Ett noterbart inslag i fritextdelen av enkäten är de tolkningar där *barnbokhylla* förklaras vara en 'bokhylla för barnböcker', dvs. där det mel-lersta ledet nyttjas både till huvudledet och till modifieraren.

Sammanfattningsvis slås de lågfrekventa sammansättningarnas förväntade betydelser ut av ett frekvent använt mönster [XX(s)Y], som väger tyngre än regeln om fog. Enkätresultatet tyder således på att denna skrivregel, som går tillbaka på Tamms fogeregeln, inte är kognitivt motiverad (se avsnitt 2.4.1). Huvudledet kan visserligen omfatta två led, men det finns en preferens för en segmentering där huvudledets räckvidd är begränsad till ledet längst till höger, oavsett om sammansättningen har utsatt fog eller inte.

7.1.3. Semantisk relation i sammansättningar som betecknar behållare

I avsnitt 5.1.3 redovisas informanternas tolkningar av relationen mellan leden i nominala sammansättningar som betecknar behållare (t.ex. *plastlåda* och *olivkorg*). Informanterna kunde välja mellan de semantiska relationerna MATERIAL, INNEHÅLL, ÄNDAMÅL och DEKORATION. Den här typen av semantiska relationer har varit föremål för tidigare forskning om relationspreferenser och lagrade mönster (se avsnitt 2.2).

De sammansättningar som har god representation i referenskorpusen ges förhållandevis samstämmiga tolkningar, vilket överensstämmer med hypotes 2c, dvs. antagandet om en korrelation mellan frekvens och samstämmighet i informantgruppen som helhet. Skillnader mellan L1- och L2-talarna noteras vid hälften av de lexikaliserade sammansättningarna, vilka tolkas mer samstämmigt (och konventionellt) av L1- än av L2-talarna (i enlighet med hypotes 3c–d). Även i fritextdelen av enkäten är de informanter som har svenska som förstaspråk mer samstämmiga i sina tolkningar än de som har svenska som andraspråk. Resultatet markerar

skillnaden mellan möjligt och brukligt. Det mesta är möjligt, men några alternativ i sammansättningarnas meningspotentialer har kommit att konventionaliseras och bli de brukliga. Dessa språkspecifika konventioner är befästa i högre grad hos dem som haft svenskan med sig från början (L1-talarna) än hos dem som börjat med svenskan först efter att begrepps-bildning skett på ett annat språk (L2-talarna).

De sammansättningar där förledets referent kan tolkas som ett material tolkas i stor utsträckning med den semantiska relationen <gjord av MATERIAL>. Detta gäller *plastlåda* i flervalsdelen såväl som *glaslåda* i fritextdelen av enkäten. Förledet i *glaslåda* är polysemt, men en majoritet av informanterna tolkar *glaslåda* som en låda 'gjord av materialet glas', snarare än som en låda 'avsedd för glas av något slag'. Detta är anmärkningsvärt, eftersom lådor av glas vid tiden för undersökningen (2015) var ett ovanligt fenomen, innan de lanserades som alternativ till lunchlådor i plast. I det här fallet följer informanternas tolkningar det produktiva mönstret <gjord av MATERIAL/SUBSTANS>. Detta mönster har i informantstudier visat sig vara lättillgängligt för språkbrukare (se vidare avsnitt 2.2). Resultaten från tidigare studier och resultaten från föreliggande studie kan relateras till CARIN-modellens antagande om lagrade relationspreferenser. Liksom *choklad* som modifierare ofta bildar en sammansättning där efterledets referent är gjord av choklad, förefaller andra substanser eller material trigga en tolkning där sammansättningens referent är gjord av det material som modifieraren betecknar. Huruvida tolkningen utgår från lagrade relationspreferenser för enskilda modifierare, såsom *plast-* och *glas-* (såsom antas i CARIN-modellen) eller om de utgår från semantiska kategorier för sammansättningens båda led [substans/material–objekt] (såsom Maguire et al. 2010a, 2010b föreslår) kan inte avgöras här (se avsnitt 2.2).

I den här kategorin av sammansättningar avviker mina resultat från resultaten hos Ryder (1994). I hennes experiment ges sammansättningar där efterledet betecknar en behållare i huvudsak tolkningar där förledet anger vad behållaren innehåller, t.ex. *bean-barrel* 'tunna med bönor'. Utifrån dessa resultat drar Ryder slutsatsen att INNEHÅLL är den semantiskt mest rimliga relationen för sammansättningar som betecknar behållare. Det är emellertid värt att notera att modifierarna i Ryders sammansättningar inte är av sådan art att de kan beteckna MATERIAL eller SUBSTANS (1994:232). Det förekommer exempelvis inga förled *trä-*, *glas-* eller *plast-* i hennes undersökning. Min bedömning är att Ryders analys därför inte ger en rättvisande bild av sammansättningar där referenten är en behållare. Mina empiriska resultat talar för att det finns flera starka scheman för den här typen av sammansättningar, t.ex. [material–behållare] 'behållare gjord av material', [sak–behållare] 'behållare som innehåller sak' eller [sak–behål-

lare] ’behållare som är avsedd för sak’. Vissa modifierare passar in i flera scheman, och schemana konkurrerar i dessa fall med varandra. Tolkningarna av *glaslåda* och *plastlåda* i enkäten visar att [material–behållare] ligger nära till hands om förledets referent kan tolkas som material/substans.

Sammanfattningsvis får frekvensen för de enskilda sammansättningarna visst genomslag i informanternas tolkningar, vilket yttrar sig i mer samstämmiga tolkningar av de frekventa än de lågfrekventa sammansättningarna i informantgruppen som helhet. Mindre frekventa sammansättningar tolkas relativt samstämmigt om de instantierar ett frekvent mönster, såsom i fallet med *glaslåda*. Samstämmigheten i tolkningarna är större i de grupper av informanter som har svenska som förstaspråk än i de grupper som har svenska som andraspråk.

7.1.4. Jämförande egenskapsprojicering eller tematisk relationslänkning

I avsnitt 5.1.4 redovisas informanternas tolkningar av sammansättningar med potential att tolkas egenskapsprojicerande, dvs. jämförande, eller relationslänkande (t.ex. *zebralejon* och *isvatten*). På en övergripande mönsternivå står valet mellan två abstrakta, semantiska mönster – jämförande egenskapsprojicering och tematisk relationslänkning (se avsnitt 2.3). I kontrast till detta har man i tidigare kognitiv forskning betonat vikten av ytnära analogibaser, såsom specifika instanser och lågnivåscheman, vid produktion och reception (se avsnitt 3.1.5).

Störst samstämmighet i tolkningarna från informantgruppen som helhet påvisas vid sammansättningarna med störst representation i referenskorpusen, vilket överensstämmer med antagandet om en korrelation mellan frekvens och konventionalisering. Informanterna är dock förvånansvärt samstämmiga även i sina tolkningar av de lågfrekventa respektive oetablerade djursammansättningarna (*apörn*, *zebralejon* och *corpörn*) i strid med hypotes 2c. Dessa tolkas i stor utsträckning egenskapsprojicerande, liksom också *zebraväska* i fritextdelen av enkäten. *Issaft* i fritextdelen av enkäten tolkas i hög grad relationslänkande, som ’saft med is i’. Vid den här typen av sammansättningar har informanterna inte tillgång till sammansättningarna som befästa enheter, utan samstämmigheten förutsätter tillgång till befästa mönster eller liknande sammansättningar som kan utgöra analogibaser.

När djursammansättningarna i enkäten tolkas egenskapsprojicerande går tolkningen sannolikt via ett ytnära, egenskapsprojicerande och jämförande schema [djur1–djur2] ’djur 2 är (i något avseende) som djur 1’. Detta schema konkurrerar av pragmatiska skäl ut ett relationslänkande

jaktschema (se vidare avsnitt 7.5). Det har belagts i tidigare forskning att det egenskapsprojicerande mönstret är frekvent förekommande i etablerade sammansättningar som betecknar djur. Det har också visat sig vara ett tillgängligt mönster vid tolkning av nybildade djursammansättningar (se avsnitt 2.3, samt vidare diskussion i avsnitt 7.2–7.3).

Som Ryder (1994) framhåller kan språkbrukare dessutom använda liknande sammansättningar som analogibaser för tolkningen. I mitt material kan instanser som *zebrafisk* och *zebrafink*, liksom andra jämförande sammansättningar med det aktuella förledet, fungera som analogibaser vid tolkningen av *zebralejon* i flervalsdelen och *zebraväska* i fritextdelen av enkäten. Korpens svärta och zebrans randighet åberopas exempelvis i adjektiviska sammansättningar som *korpsvart* och *zebrarandig*. Även om mönstren bakom dessa är mer perifera i relation till enkätorden har de potential att påverka språkbrukarna i egenskapsprojicerande riktning när de tolkar sammansättningarna med *korp* och *zebra* som förled. Om man relaterar tolkningarna till CARIN-modellen tyder resultatet på att modifierarna *korp-* och *zebra-* är lagrade som enheter i det mentala lexikonet med informationen om att de i sammansättningar brukar markera utmärkande egenskap, svärta respektive randighet. Det rör sig om lågfrekventa instanser som är befästa som enheter p.g.a. saliens snarare än frekvens (se vidare avsnitt 7.3 om saliens).

Den höga andelen relationslänkande tolkningar av *issaft* i fritextdelen av enkäten kan också sammanhånga med analogisk tolkning som strategi. När informanterna tolkar *issaft* som 'saft med is' förefaller tolkningen göras i analogi med *isvatten*.

Avslutningsvis är samstämmigheten i informantgruppernas tolkningar stor vid de frekventa sammansättningarna, såsom förväntat utifrån hypotes 2c. Detta indikerar i sin tur att sammansättningarna har en konventionaliserad betydelse i språkgemenskapen (i linje med hypotes 1c). Dessa frekventa sammansättningar tolkas mer konventionellt av L1-talarna än av L2-talarna (i enlighet med hypotes 3c). Samstämmigheten i informanternas tolkningar är dock förvånansvärt stor även vid de lågfrekventa och tillfälliga sammansättningarna (jfr hypotes 2c), där inga signifikanta skillnader påvisas mellan informantgrupperna (i strid med hypoteserna 3d–e). De tillfälliga bildningarna finns inte befästa som enheter hos informanterna. Samstämmigheten kan i detta fall förklaras med tillgång till frekventa och befästa språkliga mallar, i form av specifika sammansättningar och lågnivåmönster, som fungerar som analogibaser i tolkningen av det nya uttrycket (se avsnitt 3.1.5).

7.1.5. Identitet/jämförande egenskapstolkning eller tematisk relationslänkning i personbetecknande sammansättningar

I avsnitt 5.1.5 redovisas tolkningar av personbetecknande sammansättningar där informanterna kunde välja mellan å ena sidan en identitetstolkning (XY är en Y som också är X) eller en jämförande egenskapstolkning (XY är en Y som är som X), å andra sidan någon typ av tematisk relationslänkning mellan leden. Sammansättningarna betecknar genus, i flera fall i relation till yrke. Den här gruppen av sammansättningar representeras inte av några högfrekventa sammansättningar i enkäten, men graden av samstämmighet i tolkningarna av de lexikaliserade sammansättningarna följer likväl sammansättningarnas frekvens. Av resultatredovisningen framgår att sammansättningarna med störst representation i referenskorpusen (*kvinnopräst* och *kvinnokarl*) tolkas relativt samstämmigt i informantgruppen som helhet: det mest konventionaliserade alternativet samlar en stor del av tolkningarna. Spridningen är däremot stor vid de sammansättningar som har få eller inga förekomster i referenskorpusen (*herrfrisör*, *mansläkare* och *kvinnoläkare*).

Sammansättningarna *flickpojke* och *manspräst* tolkas relativt samstämmigt, med en jämförande egenskapstolkning respektive en identitetstolkning, trots att det rör sig om tillfälliga bildningar som inte kan antas vara befästa som enheter eller exemplar. I det här fallet förefaller analogibaser på den mest ytnära nivån inverka. *Flickpojke* tolkas sannolikt i analogi med betydelsen hos den etablerade sammansättningen *pojkflicka*. Enkätsvaren visar ett signifikant samband mellan tolkningarna av *kvinnopräst* och *manspräst* i alla informantgrupper, dvs. alla grupper använder sig av analogi som strategi vid tolkningen av *manspräst*. I fritextdelen av enkäten tolkas *kvinnoman* i relativt stor utsträckning i analogi med betydelsen hos den etablerade sammansättningen *kvinnokarl*, särskilt av de informanter som har svenska som modersmål.

L1-talarna har en högre grad av identitets- och egenskapstolkningar än L2-talarna vid sammansättningarna *kvinnopräst*, *manspräst* och *flickpojke*. I fallet med *kvinnopräst* kan skillnaden förstås som en skillnad avseende lexikal befästning av en etablerad sammansättning (enligt hypotes 3d). När den tillfälliga bildningen *manspräst* tolkas i analogi med *kvinnopräst* resulterar detta i en högre grad av identitetstolkningar hos L1-talarna än hos L2-talarna, eftersom L1-talarna är mer bekanta med den konventionaliserade betydelsen hos analogibasen (vilket överensstämmer med hypotes 3e). Samma resonemang torde kunna tillämpas på tolkningarna av *flickpojke*. Inga signifikanta skillnader mellan första- och andraspråkstalarna föreligger vid de lexikaliserade sammansättningarna

herrfrisör och *mansläkare* som har få respektive inga förekomster i referenskorpusen (jfr hypotes 3d). Vid dessa sammansättningar utnyttjas identitetstolkningen i meningspotentialen i nästan lika hög grad som den lexikaliserade tematiska relationslänknings i alla informantgrupper.

Sammanfattningsvis råder det stor osäkerhet i informantgruppen om vad sammansättningarna *mansläkare*, *kvinnoläkare* och *herrfrisör* betecknar, trots att dessa är lexikaliserade. Samstämmigheten är större vid *kvinnopräst* och *kvinnokarl*, som har något större representation i referenskorpusen. Men också vid de tillfälliga bildningarna *manspräst* och *flickpojke* (utan belagda betydelser) är samstämmigheten förvånansvärt stor, och identitets- respektive jämförelsemönstret är prefererat. Denna preferens – som är särskilt tydlig hos L1-talarna – är ett resultat av analogisk tolkning. Den sammantagna bilden är att identitets- och jämförande egenskapstolkning ligger nära till hands för informanterna vid obekanta sammansättningar som relaterar till genus.

7.1.6. Agens vid verbavledning

I avsnitt 5.1.6 i resultatkapitlet behandlas informanternas tolkningar av agens i verbavledningar som *kvinnomisshandel* och *studentförakt* i flervalssdelen av enkäten.

De sammansättningar som har god representation i referenskorpusen tolkas som förväntat synnerligen samstämmigt och konventionellt (i linje med hypotes 2c respektive 1c). Men även vid mindre frekventa sammansättningar och nybildningar är informanterna förvånansvärt överens om att efterledets verbhandling är riktad mot det personbetecknande förledets referent. I det här fallet är samstämmigheten inte ett resultat av de specifika sammansättningarnas frekvens och etablering (jfr hypotes 2c). Informanterna förefaller här göra bruk av en språklig mall som fungerar som analogibas vid tolkningen. I sammansättningar där efterledet är avlett från ett transitivt verb är det vanligare att förledets referent fyller verbhandlingens patientroll, snarare än verbhandlingens agentroll (Nesset 2016:89; se även Selkirk 1982:34). Informanterna följer således ett etablerat schema [PATIENT-verbavledning] i sina tolkningar.

Informanterna i gruppen SVE har generellt en högre grad av patienttolkningar än övriga grupper. Vid en frekvent sammansättning som *kvinnomisshandel* kan förhållandet förklaras som ett försprång avseende befestning av en specifik instans (i enlighet med hypotes 3c), och den synnerligen lågfrekventa sammansättningen *mansmisshandel* tolkas sannolikt med *kvinnomisshandel* som analogibas (se hypotes 3e). Vid en tillfällig sammansättning som *studentförakt* speglar tolkningarna skillnader i befest-

ning av ett frekvent förekommande schema [PATIENT–verbavledning], till följd av skillnader i exponering.

Befästningen av schemat [PATIENT–verbavledning] ger generellt ett stöd i tolkningen av verbavledda sammansättningar, men i fallet med *polismisshandel* leder det i fel riktning. *Polismisshandel* har konventionaliserats i en betydelse där förledets referent utgör verbhandlingens agent, dvs. den etablerade betydelsen avviker från schemat [PATIENT–verbavledning]. *Polismisshandel* är den av verbavledningarna som tolkas minst samstämmigt. I informantgruppen som helhet fördelar sig tolkningarna 50/50 mellan agent- och patienttolkning. I det här fallet pekar ledtrådar på den specifika sammansättningens nivå i *en* riktning, dvs. att förledet är AGENT, och ledtrådar på mönsternivå i en annan riktning, dvs. att förledet är PATIENT. Resultatet indikerar att sammansättningens konventionaliserade betydelse inte är befäst hos flertalet informanter. Det är värt att notera att informanterna i SVE har en signifikant lägre grad av konventionella tolkningar än de flerspråkiga informanterna (grupperna SVE+, START<6, START≥6) i tolkningen av *polismisshandel*. Bara 40 % av informanterna i SVE tolkar polisen som AGENT i sammansättningen, medan 60 % ser polisen som verbhandlingens PATIENT. Resultatet sammanhänger med den generella preferensen bland L1-talarna att välja det etablerade och frekventa mönstret [PATIENT–verbavledning].

Sammanfattningsvis tolkas de verbavledda sammansättningarna relativt samstämmigt och konventionellt, i stor utsträckning med en relation mellan leden där förledets referent intar en patientroll. Vid de etablerade sammansättningarna är informanternas tolkningar ett resultat av befästning av enskilda instanser. Vid tillfälliga sammansättningar kan det frekventa schemat [PATIENT+verbavledning], eller mer ytnära analogibaser, leda informanterna till en tolkning där förledets referent utgör verbhandlingens patient. Stor spridning i tolkningarna av *polismisshandel* illustrerar en konflikt vid denna sammansättning mellan ytnära (lexikal) information och språklig information på en högre schematicitetsnivå. Informanternas låga grad av konventionella agenttolkningar av *polismisshandel* tyder på att de inte har sammansättningen befäst i sitt mentala lexikon. L1-talarna i SVE gör i högre grad än de flerspråkiga informanterna en patienttolkning av *polismisshandel*, vilket sammanhänger med en allmän preferens bland L1-talarna (särskilt informanterna i SVE) att välja patientmönstret. Sammantaget tyder resultatet på att schemat [PATIENT–verbavledning] är mer befäst hos L1-talarna än hos L2-talarna.

7.1.7. Betydelsesammansmälta AN-sammansättningar

I resultatkapitlets avsnitt 5.1.7 redogörs för tolkningar av betydelse-sammansmälta AN-sammansättningar, såsom *sötsaker* och *finkläder*. Kategorin innehåller sammansättningar som är icke-kompositionella eftersom de lexikaliseras i en snävare betydelse än vad leden anger (se avsnitt 3.2.2). De sammansättningar som finns i enkäten är vardagliga och har god representation i referenskorpusen. Sett till informantgruppen som helhet är de förklaringar som speglar den konventionaliserade betydelsen prefererade, men skillnaderna mellan informantgrupperna är betydande. De informanter som har svenska som förstaspråk väljer i betydligt högre grad än de informanter som har svenska som andraspråk specificerade och konventionella alternativ. L2-talarna med låg startålder (dvs. gruppen $START < 6$) väljer de konventionella alternativen i betydligt större utsträckning än L2-talarna med högre startålder (dvs. $START \geq 6$; se figur 5:4).

Eftersom de sammansättningar som valts ut till denna kategori är frekventa är skillnaderna mellan L1- och L2-grupperna förvånansvärt stora (jfr hypotes 3c). Det är dock värt att notera att de konventionella alternativen och distraktorerna i enkäten ligger betydelsemässigt nära varandra vid den här typen av sammansättningar. Om man är väl bekant med begreppet *finkläder* känner man till att det innefattar fler betydelsekomponenter än dem som är explicit uttryckta i sammansättningen. Om man däremot är mer ytligt bekant med begreppet är det rimligt att välja det mindre markerade, tillika kompositionella, alternativet, 'kläder som är fina'. De stora skillnaderna mellan första- och andraspråkstalarnas tolkningar speglar en skillnad avseende djup och precision i ordförståelsen (Henriksen 1999:311; se vidare avsnitt 3.2.3).

Vikten av lexikal befästning illustreras tydligt i den här kategorin av sammansättningar. Eftersom AN-sammansättningarna inte är kompositionella förutsätter tillgången till de konventionaliserade betydelsena att sammansättningarna är befästa som enheter. Skillnaderna mellan L1- och L2-talarna vid dessa sammansättningar tyder på att de finns lagrade i större utsträckning hos L1- än hos L2-talarna. L2-talarna förefaller i högre grad förlita sig på kompositionalitetsprincipen och pussla ihop betydelsen utifrån de ingående leden. Därmed faller valet på de kompositionella men underspecificerade alternativen. Denna slutsats överensstämmer med Wrays (2002, 2008) konstaterande om att andraspråkstalare lagrar mindre enheter och är mer beroende av kompositionell analys (se avsnitt 3.2.3).

Sammanfattningsvis visar resultatet att L1-talarna har en mer precis kunskap om den här typen av sammansättningar än L2-talarna. Dessutom gör L2-talarna med låg startålder fler konventionella tolkningar än

L2-talarna med en högre startålder. Den sammantagna bilden är att de informanter som haft en längre tid av exponering för svenskan också är mer bekanta med de konventionaliserade betydelseerna, medan informanterna med en kortare exponeringstid för svenskan väljer de kompositionella, mindre specificerade och mindre konventionella alternativen. Med andra ord korrelerar graden av befästning av de betydelsesammansmälta AN-sammansättningar tydligt med exponeringstiden för svenskan.

7.1.8. Identifiering av huvudledet i AA-sammansättningar

När det gäller enkätens AA-sammansättningar möter informanterna tre svarsalternativ i stället för två (se avsnitt 5.2.1). Därför är det rimligt med en större spridning i tolkningarna. Icke desto mindre är spridningen anmärkningsvärt stor, särskilt med tanke på att sammansättningarna, med något undantag, har många förekomster i referenskorpusen. De sammansättningar som har konventionaliserats med en kopulativ relation mellan leden (*svartvit*, *blågul*, *sötsur*) tolkas mer samstämmigt än de sammansättningar som brukar anses determinativa (*gulvit*, *gröngul*). En bidragande orsak skulle kunna vara de enskilda sammansättningarnas frekvens. *Svartvit* och *blågul* har en mycket god representation i referenskorpusen. Det kan alltså vara en fråga om skillnad i konventionaliseringsgrad mellan mer och mindre frekventa sammansättningar (såsom antas i hypotes 1c). En större andel konventionella, kopulativa tolkningar av *svartvit* hos informanter som har svenska som förstaspråk än hos informanter som har svenska som andraspråk tyder på skillnader mellan grupperna med avseende på befästning av den enskilda instansen (i linje med hypotes 3c).

Utifrån frekvensförhållanden kunde man förvänta sig en större samstämmighet även vid de determinativa sammansättningarna, eftersom också dessa har förhållandevis många belägg i referenskorpusen. Man kan konstatera att sammansättningarnas konventionaliserade betydelse inte är befäst hos språkbrukarna i min informantgrupp. I förlängningen kan man, med utgångspunkt i resultatet, sätta frågetecken för om man över huvud taget kan tala om en konventionaliserad betydelse hos AA-sammansättningar av denna typ. Resultatet tyder på att sammansättningarnas uttryckssida är etablerad (vilket avspeglas i användningsfrekvens) men inte deras innehållssida (se Hannesdóttir 2012:54–55; se även avsnitt 3.1.7).

Det är inte så förvånande att informanterna är öppna för både determinativ och kopulativ tolkning av sammansättningar som *gulvit* och *sötsur*, så att en traditionellt determinativ sammansättning ges en kopulativ tolkning och vice versa. Däremot är det mycket intressant att notera att många informanter – och i synnerhet de informanter som har svenska som

sitt enda modersmål – gör determinativa tolkningar där förledet uppfattas som sammansättningens huvud, dvs. de följer ett mönster som har ansetts vara främmande i svenska sammansättningar. Nästan hälften av informanterna i SVE anger exempelvis att *gulgrön* är 'mer gul än grön' och att *gröngul* är 'mer grön än gul'. Som framgått tidigare förefaller det etablerade schemat för huvudets placering i svenska sammansättningar vara starkt närvarande hos L1-talarna när de tolkar nominala sammansättningar (i enlighet med hypotes 2a), men så verkar alltså inte vara fallet när de tolkar adjektiviska sammansättningar. Ur ett kognitivt perspektiv kan detta förklaras med att förekomsten av två etablerade och konkurrerande mönster för AA-sammansättningar (determinativt respektive kopulativt) leder till en lägre grad av befestning (såsom antas i hypotes 2b). En annan förklaring kan vara att olika ordning mellan leden i en adjektivisk sammansättning resulterar i förhållandevis små betydelskillnader. I de flesta sammanhang är det inte av avgörande betydelse om en nyans mellan grönt och gult är mest gul (som i *gröngul*) eller mest grön (som i *gulgrön*). Framför allt möter talaren sällan negativ evidens som visar att vederbörande tolkar en AA-sammansättning på ett annat sätt än samtalspartnern. AA-sammansättningarna är ofta kopplade till perception, och alla människor har sina individuella upplevelser av exempelvis synintryck (jfr färgsammansättningarna och sammansättningen *fulsnygg* i enkäten) och smakintryck (jämför *sötsur*).

Informanternas förklaringar av *fulsnygg* i fritextdelen av enkäten (se avsnitt 6.1 och 6.3) illustrerar att dikotomin kopulativ–determinativ sammansättning är för grov vid den här typen av AA-sammansättningar. Informanterna ger svar som inte kan inordnas i en enkel kategorisering 'mer ful än snygg', 'mer snygg än ful' eller 'mittemellan ful och snygg'. Flera informanter gör tolkningar av typen 'så snygg att man blir ful' eller 'så ful att man blir snygg', och det framstår därmed som oklart vilken av de kontrasterande egenskaperna som överväger. I andra svar lyfts variation efter tillfälle in som en aspekt, i förklaringar av typen 'ful ibland, snygg ibland'. I ytterligare andra betecknar komponenterna *ful* och *snygg* skilda dimensioner 'fult yttre, snyggt inre' och vice versa, dvs. det är inte en fråga om att vara mest ful eller mest snygg. I flervalsdelen av enkäten, där informanterna väljer mellan fasta alternativ, har informanterna i START \geq 6 signifikant fler konventionella tolkningar av *fulsnygg* – 'mer snygg än ful' – än informanterna i SVE. Resultatet kan i förstone framstå som något förvånande, men min hypotes är att flera av L2-talarna förväxlar de grafiskt snarlika orden *ful* och *full*, dvs. att de uppfattar sammansättningen som *fullsnygg*, vilket leder fram till tolkningen 'mycket snygg'. L2-informanternas förklaringar i fritextdelen av enkäten styrker denna hypotes.

Av jämförelsen mellan första- och andraspråkstalarnas svar i flervalsdelen av enkäten framgår det att L2-talarna i gruppen START \geq 6 är mer

benägna än L1-talarna att göra kopulativa tolkningar av de färgsammansättningar som består av två varandra närliggande färger – t.ex. *gulvit* – medan L1-talarna i högre grad väljer determinativ tolkning. Dessa L2-talare väljer med andra ord det minst markerade, i det här fallet något underspecificerade, alternativet, vilket kan relateras till resonemanget om partiell ordförståelse (se avsnitt 3.2.3). Att komplexa begrepp i en inlärningsfas tenderar att tolkas sida vid sida snarare än integrerade med varandra (såsom när *book magazine* ges förklaringen ”a big magazine and a little book”) har också konstaterats i barnspråksstudier (se avsnitt 2.4.1). Det är mer förvånande att också de flerspråkiga informanterna med svenska som ett av sina förstaspråk (både informanterna i SVE+/S och informanterna i SVE+/A) föredrar en kopulativ tolkning av de presumtivt determinativa sammansättningarna. De enspråkiga L1-talarna är mer förtrogna med mönstret att använda sammansättning för att modifiera en färgnyans. Men intressant nog är benägenheten stor bland dem att tolka förledet som huvudled och efterledet som modifierare, tvärtemot den struktur som antas vara etablerad för sammansättningar. De gör alltså i högre grad än L2-talarna vänstertolkningar. Därmed kan hypotes 3b avvisas. Mot bakgrund av informanternas tolkningar finns det anledning att ifrågasätta hur etablerat schemat $[AA_{\text{huvud}}]_A$ med efterledet som huvud egentligen är (se vidare avsnitt 7.6 och 8.3).

7.1.9. Jämförelse eller orsak vid NA-sammansättningar

I avsnitt 5.2.2 presenteras informanternas tolkningar av de NA-sammansättningar där relationen mellan leden kan tolkas antingen som jämförande eller som ett orsakssamband. Samstämmigheten i tolkningarna från informantgruppen som helhet är betydande och det jämförande mönstret prefererat, vilket framgår av tabell 5:11. De sammansättningar som har god representation i referenskorpusen (t.ex. *iskall*, *sockersöt* och *kryddstark*) tolkas anmärkningsvärt samstämmigt, vilket tyder på konventionisering och befestning på grund av hög tokenfrekvens (i enlighet med hypotes 1c och 2c). Spridningen är större vid *solvarm*, där mindre än två tredjedelar av informanterna ger ett orsakrelaterat svar i enlighet med den konventionaliserade betydelsen hos sammansättningen. *Solvarm* har en förhållandevis god representation i referenskorpusen, men min bedömning är att sammansättningen är överrepresenterad i de utvalda korpuserna. *Solvarma* frukter och tomater är mycket frekvent förekommande i vissa av delkorpuserna i min referenskorpus (se avsnitt 4.4). Den relativt låga graden av konventionella tolkningar i informantgruppen skulle kunna sammanhånga med en blygsam representation för *solvarm* i det språk som ungdomarna möter.

Vid de oetablerade sammansättningarna utan representation i referenskorpusen är samstämmigheten ändå förvånansvärt stor – i strid med hypotes 2c – och det jämförande mönstret slår igenom starkt vid sammansättningar som *berghård*, *snömjuk* och *glasskall*. *Lökstark* utgör ett undantag och tolkas övervägande med en orsaksrelation mellan leden, dvs. 'stark av lök'.

Även i fritextdelen av enkäten, där informanterna själva formulerar förklaringarna, dominerar jämförande förklaringar, inte bara av de frekvent förekommande NA-sammansättningarna *blodröd* och *dammgrått*, utan också av de mer tillfälliga sammansättningarna *honungssött* och *sandmjukt*.

Resultatet från de båda delarna av enkäten tyder på att det jämförande mönstret ligger nära till hands för informanterna. Detta kan relateras till allmänna frekvensmönster för NA-sammansättningar. I en tidigare korpusstudie om NA-sammansättningar (Loenheim 2016) konstaterades att den jämförande relationen [A som N] generellt är stark för den här typen av sammansättningar, medan NA-sammansättningar med en orsaksrelation mellan leden [A av N] är mer blygsamt representerade (se avsnitt 2.1). Benägenheten hos informanterna att välja jämförande tolkningar av de obekanta sammansättningarna i enkäten kan tolkas som att de baserar tolkningarna på ett produktivt schema [A som N].

Å andra sidan har det uppmärksammats i tidigare forskning att språkbrukare snarare än abstrakta mönster gör bruk av lågnivåmönster och specifika enheter eller exemplar när de tolkar obekanta ord (se avsnitt 3.1.5). Språkbrukaren kan på en ytnära nivå finna stöd i att det jämförande mönstret, tillsammans med det semantiskt närliggande förstärkande mönstret, är betydligt vanligare än det orsaksrelaterade mönstret vid NA-sammansättningar med *-varm* och *-kall* som efterled (se avsnitt 2.1). Utifrån de många jämförande sammansättningarna med *-varm* och *-kall* är det förväntat att informanterna gör en schematisering [N-kall] 'kall som N' eller [N-varm] 'varm som N', som får ligga till grund för deras tolkningar av obekanta ord med det aktuella efterledet (t.ex. *glasskall*). Förekomsten av ett sådant ytnära schema kan förklara varför så många informanter i min enkätundersökning – både L2- och L1-talare – väljer tolkningen 'varm som solen' i stället för 'varm av solen' vid *solvarm*.

En annan förklaring, som ryms inom den kognitiva ramen, är att informanternas tolkningar baseras på analogi utifrån specifika exemplar (Bybee 2013), snarare än ett produktivt schema. *Glasskall* och *snökall* skulle exempelvis kunna tolkas med *iskall* som analogibas, *honungssöt* förefaller tolkas jämförande (och metaforiskt) i analogi med *sockersöt*, och *lökstark* tolkas i förhållandevis stor utsträckning i analogi med betydelsen hos *kryddstark*.

Som framgått är samstämmigheten i informanternas tolkningar påfallande stor, och de jämförande tolkningarna dominerar vid merparten av de tillfälliga sammansättningarna. De lexikaliserade sammansättningarna tolkas emellertid mer konventionellt av L1-talarna än av L2-talarna, vilket indikerar skillnader i befästning av specifika instanser. Störst är skillnaderna vid lexikaliserade men mindre frekventa sammansättningar som *solvarm* och *kryddstark* (i linje med hypotes 3d). Samstämmigheten är större bland L1-talarna än bland L2-talarna även vid de tillfälliga bildningarna *glasskall* och *lökstark*. Skillnaden vid *lökstark* kan relateras till en större kunskap om analogibasen *kryddstark* (såsom antas i hypotes 3e).

7.1.10. Sammansättningar med *-rädd*

Informanternas tolkningar av sammansättningar med *-rädd* som efterled, i flervalsdelen av enkäten, redovisas i avsnitt 5.2.3. Tolkningen [rädd för N] dominerar i informantgruppen som helhet. Samstämmigheten vid *hundrädd* skulle kunna bero på att ordet är etablerat (även om det saknar representation i referenskorpusen). Samstämmigheten är emellertid förhållandevis stor även vid den mindre etablerade sammansättningen *fågelrädd*, liksom vid *musrädd* i fritextdelen av enkäten. Resultatet tyder på att informanterna stöder sig på ett frekvent mönster [rädd för N] eller ett något mer abstrakt mönster [rädd för X], där X kan instantieras av ett substantiv (som i *höjdrädd*) eller ett verb (som i *flygrädd*). [N-rädd]/[X-rädd] i jämförande betydelse är mindre vanligt.

Det finns en svag tendens i båda delar av enkäten till att informanterna som har svenska som modersmål i högre grad än L2-talarna gör tolkningar med relationen [rädd för N]. Detta tyder på att L1-talarna är något mer benägna än L2-talarna att ta stöd i ett etablerat, frekvent mönster, såsom [rädd för N] eller [rädd för X], när de möter specifika instanser utan befästning i deras mentala lexikon. L2-informanterna kan tänkas vara mindre bekanta med detta lågnivåmönster och i stället stödja sig på ett mer övergripande, jämförande mönster i språket (se vidare avsnitt 7.2).

7.1.11. Sammansättningar med *-fri*

I avsnitt 5.2.4 redovisas tolkningar av enkätens sammansättningar med efterledet *-fri*. *Fri* som efterled har kommit att fungera suffixliknande och markera att något är fritt från den referent som förledet betecknar (jfr *bilfri*, *fettfri*, *alkoholfri*). Från detta vanliga mönster avviker *valfri* där *fri* fungerar som självständigt led 'fri (att)'. *Valfri* är en mycket frekvent använd sammansättning. Av alla sammansättningar i enkäten är *valfri* – värt

att notera – den sammansättning som har flest förekomster i referenskorpusen. Därmed är samstämmigheten i tolkningarna från informantgruppen som helhet väntad (enligt hypotes 2c), trots att sammansättningen avviker från ett etablerat och frekvent mönster. Det är också väntat att *valfri* tolkas något mer samstämmigt av L1-talarna än av L2-talarna (enligt hypotes 3c).

Spridningen i tolkningarna är dock stor vid *helgfri*. Bara drygt hälften av informanterna väljer den lexikaliserade betydelse som följer det produktiva mönstret [fri från N], medan övriga informanter väljer tolkningen 'när man är fri på helgen' med *fri* som självständigt led. Den låga graden av befästning hos informanterna sammanhänger med en förhållandevis låg frekvens för sammansättningen (i linje med hypotes 2c). Dessutom kan semantiska aspekter inverka. Många sammansättningar med *-fri* som efterled betecknar avsaknad av något i sammanhanget icke önskvärt (se Teleman 1970:44). I sammansättningen *helgfri* betecknar förledet snarare något som i många sammanhang framstår som önskvärt, vilket skulle kunna leda till att tolkningen med *fri* som självständigt led blir mer tillgänglig än en tolkning som instantierar den semantiska relationen 'fri från'. Den konventionaliserade betydelsen hos *helgfri* är inte etablerad i någon av informantgrupperna till skillnad från *valfri* och *bilfri*, där de lexikaliserade betydelseerna är starkt befästa i samtliga grupper.

7.1.12. Sammansättningar med *-färdig*

Tolkningarna av sammansättningar med efterledet *-färdig* (*gråtfärdig*, *skratzfärdig*, *skrikfärdig*) presenteras i avsnitt 5.2.5. I den här kategorin av sammansättningar kunde informanterna välja mellan ett alternativ där verbhandlingen är på väg att inledas och ett alternativ där verbhandlingen är avslutad, dvs. mönstren [färdig att börja VERBA] och [färdig med att VERBA]. Samstämmigheten i informantgruppen som helhet är som väntat stor vid den etablerade sammansättningen *gråtfärdig*, som i hög grad ges en konventionell tolkning (i enlighet med hypotes 2c). Men samstämmigheten är påfallande stor även vid de tillfälliga bildningarna *skratzfärdig* och *skrikfärdig*, i strid med hypotes 2c. I samtliga fall är tolkningen [färdig att börja VERBA] det prefererade valet.

Det föreligger emellertid stora skillnader mellan grupperna. Vid *gråtfärdig* väljer L1-talarna i signifikant högre grad än L2-talarna den konventionaliserade betydelsen (i linje med hypotes 3c). Vid de tillfälliga bildningarna föreligger skillnader mellan L1-informanterna och L2-informanterna i gruppen START \geq 6 (i linje med hypotes 3e), men också mellan L2-informanterna med lägre respektive högre startålder. L1-talarna och

L2-talarna i gruppen START<6 väljer betydelsen 'färdig att' i betydligt högre grad än L2-informanterna med sen startålder.

Intressant nog följer samtliga informantgrupper ett mönster där samma relation väljs för alla sammansättningarna, dvs. de förefaller använda sig av analogi som tolkningsprincip. I L1-grupperna yttrar sig detta i drygt 90 % tolkningar som följer mönstret [färdig att börja VERBA]. I START<6 är motsvarande andel 80 %. I L2-gruppen med sen startålder ges sammansättningarna bara cirka 40 % tolkningar av detta slag. I alla informantgrupper följer informanterna således ett analogiskt mönster, men skillnader i bekantskapen med den etablerade sammansättningen *gråtfärdig* yttrar sig i skillnader med avseende på genomslaget för mönstret [färdig att börja VERBA] (såsom antas i hypotes 3e).

7.1.13. Sammanfattning

På de enskilda sammansättningarnas nivå kan man förvänta sig att frekventa sammansättningar är lättare att tolka än mindre frekventa sammansättningar, eftersom hög frekvens leder till konventionalisering i språkgemenskapen och befästning hos de enskilda språkbrukarna. Det är således större chans att språkbrukarna hört och lagrat information om de högfrekventa än de lågfrekventa sammansättningarna i enkäten, och att de är förhållandevis eniga om sammansättningarnas betydelse. Till stor del ser det också ut så i materialet. Därmed bekräftas hypotes 2c: frekventa sammansättningar ges mera samstämmiga tolkningar.

Men det finns också exempel på att tillfälliga bildningar, som inte kan antas vara befästa som enheter, ges förvånansvärt samstämmiga svar. Detta förhållande illustreras exempelvis av orden *vagnbarn* och *zebralejon* i enkäten. Jag har argumenterat för att samstämmigheten i dessa fall beror på tillgång till scheman, eller sammansättningar av liknande slag, som kan fungera som analogibaser i tolkningen. *Vagnbarn* kan exempelvis tolkas utifrån ett övergripande schema $[NN_{\text{huvud}}]_N$, ett lågnivåschema $[N\text{-barn}]$ eller en specifik instans som exempelvis *skolbarn*. Det är principiellt svårt att avgöra vilken schematicitetsnivå som språkbrukare stöder sig på, men jag finner empiriska belägg för att ytnära analogi är en tillgänglig tolkningsstrategi. Detta diskuteras vidare i avsnitt 7.6.

I enkäten förekommer också exempel på frekventa sammansättningar som tolkas divergerande i informantgruppen som helhet, i strid med de förväntningar som uttrycks i hypotes 2c. Detta gäller framför allt de förmodat determinativa AA-sammansättningarna, där oenigheten är stor. Mycket tyder på att bara uttrycken är etablerade, inte deras semantiska innehåll (Hannesdóttir 2012:54–55; se avsnitt 7.1.8). Många av dessa sammansättningar tolkas i betydande omfattning på ett sätt som indikerar

att förledet betraktas som huvud. Resultatet visar därmed att schemat $[AA_{\text{huvud}}]_A$ med efterledet som huvud inte är befäst hos informanterna. Det i sin tur sätter frågetecken för om detta schema, eller ett mera abstrakt, allomfattande sammansättningsschema $[XX_{\text{huvud}}]_X$, med efterledet som huvud, verkligen är en etablerad struktur i den svenska språkgemenskapen. Med en kognitiv förklaringsmodell kan förekomsten av två strukturella mönster för AA-sammansättningar (ett kopulativt och ett determinativt) antas resultera i en lägre grad av etablering i språkgemenskapen (vilket bekräftar hypotes 1b) och en lägre grad av befästning hos individerna (vilket bekräftar hypotes 2b).

Schemat för huvudledet i nominala sammansättningar slår däremot igenom förhållandevis starkt i informantgruppen som helhet (såsom antas i hypotes 2a), vilket indikerar att det är etablerat i den svenska språkgemenskapen (såsom stipuleras i hypotes 1a). Detta är förväntat, eftersom det rör sig om ett produktivt schema med synnerligen hög typfrekvens. Det föreligger dock stora skillnader mellan informantgrupperna. Första-språkstalarnas tolkningar speglar den etablerade strukturen i högre grad än andraspråkstalarnas tolkningar (helt i linje med hypotes 3a). Ju längre exponeringstid för svenskan, desto mer högertolkningar gör informanterna. Vänstertolkningar är däremot inte uteslutna för någon grupp av informanter.

Genomgången av frekvens och befästning för de olika typerna av sammansättningar visar att det föreligger skillnader mellan första- och andraspråkstalarna av svenska. L1-talarna uppvisar generellt större samstämmighet i sina svar och väljer i högre grad än L2-informanterna de mer konventionella alternativen, både vad gäller strukturella mönster (se hypotes 3a; jfr dock hypotes 3b) och mer semantiska mönster (se hypotes 3c–e). Bland L2-talarna har de informanter som har en låg startålder för svenskan mer samstämmiga tolkningar och följer i högre grad konventionella mönster än L2-talarna i gruppen $START \geq 6$. Ett viktigt undantag från denna generella bild utgör AA-sammansättningarna, där modersmåls-talarna i SVE i högre grad än de flerspråkiga informanterna med svenska som första- eller andraspråk väljer (okonventionella) vänstertolkningar av sammansättningarna. Detta resultat strider mot hypotes 3b.

Gruppen av flerspråkiga informanter som har svenska och ytterligare (minst) ett modersmål (dvs. SVE+) är intressant. I denna grupp finns en tendens till en lägre grad av konventionalitet i tolkningarna än i tolkningarna från de enspråkiga informanterna i gruppen SVE. Det visar sig emellertid att skillnaderna mellan de enspråkiga och flerspråkiga L1-informanterna är mycket begränsade om analysen avgränsas till de informanter i SVE+ som haft stöd för svenskan från en förälder som har svenska som modersmål (dvs. informanterna i SVE+/S). Dessa kan tänkas ha fått en

större mängd av svenskspråkig input från låg ålder än de flerspråkiga informanter som uppger att föräldrarna har andra modersmål än svenska (SVE+/A). Flerspråkigheten är en gemensam nämnare för informanterna i SVE+, men de skiljer sig alltså åt när det gäller variabeln förälders modersmål. Detta tyder på att skillnader i konventionalitet mellan SVE och SVE+ inte primärt beror på flerspråkigheten i sig utan är ett utslag av de skillnader i mängden svenskspråkig input som följer av skillnaden mellan att ha eller inte ha svenska som ett L1 i familjen.

Det finns en tendens till att skillnaderna mellan första- och andraspråkstalarnas tolkningar är större vid etablerade men lågfrekventa sammansättningar, än vid etablerade och högfrekventa sammansättningar i enkäten (se hypotes 3c–d), även om resultatet härvidlag inte är entydigt. Frekventa sammansättningar har en större potential än mindre frekventa att nå fram till och bli befästa hos informanter med en kortare exponeringstid för svenska språket (dvs. L2-talare). Så länge orden inte är lexikalt befästa är flera alternativ i meningspotentialen öppna vid tolkningen.

Vid tillfälliga bildningar har första- och andraspråkstalarna av svenska mer lika förutsättningar på den lexikala nivån. L1-talarna tolkar dock även dessa mer samstämmigt än L2-talarna, vilket sammanhänger med skillnader i tillgången till analogibaser. I hypotes 3e antas att L1-talarna i högre grad än L2-talarna tolkar tillfälliga bildningar i analogi med betydelsen hos etablerade sammansättningar. Denna hypotes bekräftas av min empiri: L1-talarna tolkar i högre grad än L2-talarna *skrattfärdig* med relationen [färdig att VERBA] i analogi med betydelsen hos *gråtfärdig*. Det är emellertid värt att notera att alla informanter, oavsett språklig bakgrund, gör bruk av analogiska mönster. I alla informantgrupper använder informanterna samma relation mellan leden för de tre instanserna *gråtfärdig*, *skrattfärdig* och *skrikfärdig*, liksom samma mönster används för att tolka sammansättningarna *manspräst* och *kvinnopräst*. Skillnaderna i tolkning mellan första- och andraspråkstalarna beror sålunda inte på skillnader i tolkningsstrategi, utan på skillnader i befästning av en etablerad sammansättning som kan tjäna som analogibas (*gråtfärdig* i tolkningen av *skrattfärdig* och *skrikfärdig*; *kvinnopräst* i tolkningen av *manspräst*). Tillfälliga bildningar kan också tolkas med hjälp av mer schematiska analogibaser. Skillnader i befästningen av schemat $[NN_{\text{huvud}}]_N$ förklarar exempelvis skillnader mellan första- och andraspråkstalarnas tolkningar av den tillfälliga bildningen *vagnbarn*.

Frekvens är en viktig faktor när det gäller befästning, men naturligtvis inte den enda (se avsnitt 3.1.9). Andra aspekter som är viktiga med avseende på befästning och som kan påverka informanternas tolkningar diskuteras i det följande, avsnitt 7.2–7.5. Analysen avslutas genom en konklusion i avsnitt 7.6.

7.2. Jämförelse

Oavsett språklig bakgrund är informanterna i studien benägna att använda sig av den semantiska relationen jämförelse när de tolkar tillfälliga eller mindre bekanta sammansättningar. Detta gäller i hög grad NA-sammansättningar av vissa slag (såsom *snökall* och *honungssöt*). Det gäller också i stor utsträckning sammansättningar där båda leden betecknar ett djur (*zebralejon*, *korpörn*, *apörn*), där jämförande egenskapsprojicering dominerar över tematisk relationslänkning. Jämförande egenskapstolkningar ligger också nära till hands vid tolkningen av tillfälliga, personbetecknande sammansättningar, såsom *flickpojke* och *kvinnoman*.

Informanternas benägenhet att välja egenskapsprojicering framför relationslänkning vid djursammansättningarna ligger delvis i linje med tidigare forskningsrön. En stor andel jämförande egenskapsprojicering har noterats i språkbrukares tolkningar av tillfälliga djursammansättningar där förledet betecknar ett djur med ett karakteristiskt eller salient drag (t.ex. *tiger-lion*; se avsnitt 2.3; se vidare avsnitt 7.3 om saliens). När informanterna i min studie gör en jämförande egenskapstolkning av *korpörn*, *apörn* och *zebralejon*, följer de således ett tolkningsmönster som visat sig vara lättillgängligt vid nya sammansättningar. Detta mönster är dessutom frekvent förekommande i etablerade djursammansättningar (se avsnitt 2.3).

Även mindre kända NA-sammansättningar tolkas jämförande. Denna benägenhet sammanhänger med förekomsten av det som Ryder (1994) kallar för språkliga förebilder eller mallar, som kan utgöra analogibaser (se avsnitt 3.1.5). Som diskuterats är många NA-sammansättningar jämförande, dvs. instantieringar av jämförande scheman på olika abstraktionsnivåer (se avsnitt 2.1).

Genomslaget för ett jämförande mönster i min studie kan relateras till liknande fynd i en annan studie som behandlar gymnasieelevers språk. Prentice & Sköldberg (2010) undersöker gymnasieelevers bruk av figurativa uttryck i uppsatser skrivna inom ramen för det nationella provet i svenska. De noterar att av de nybildade figurativa uttrycken är liknelser den i särklass vanligaste typen (2010:29–30).

Om man vidgar perspektivet är jämförelse överhuvudtaget en grundläggande kognitiv princip hos människan. All kategorisering utgår från jämförelse och baseras på tillräcklig likhet (se avsnitt 3.1.1 och 3.1.7). Rikedomen av jämförande uttryck i språket är synliga tecken på denna princip. Det allmänmänniskliga behovet av att uttrycka jämförelse framkommer i en mängd sammansättningar, som inte sällan hänvisar till det som Langacker (2008) kallar basala domäner (se avsnitt 3.1.2). I adjektiviska sammansättningar utgörs jämförelseledet ofta av olika naturfenomen. Detta gäller

i hög grad sammansättningar som betecknar färg (*snövit, nattsvart, blodröd, solgul, åskgrå* och *himmelsblå*), men också sammansättningar som betecknar andra egenskaper (*stenhård, iskall, skyhög*). Sammansättningar som betecknar djur jämförs, som redan konstaterats, ofta med andra djur. Benägenheten att jämföra avspeglas också i liknelser och metaforer i språket. Inom kognitiv semantik ses språket som metaforiskt till sin natur, och metaforik bygger på jämförelse som princip. Ryder (Ryder 1994:124) sammanfattar betydelsen av jämförelse i mänsklig (språklig) aktivitet:

computing similarities between different objects is a fairly common human activity. It is involved, for example in interpretation of metaphors, which is a universal activity that even fairly young children will engage in.

Jämförelse har sammanfattningsvis identifierats som en grundläggande kognitiv princip (se avsnitt 3.1.1). Detta förklarar varför informanterna i min studie – oavsett språklig bakgrund – har en preferens att välja en jämförande relation mellan leden i de mindre etablerade NA-sammansättningarna, liksom i de djur- och personbetecknande sammansättningarna i enkäten.

7.3. Saliens

Saliens är ytterligare en aspekt som inverkar på informanternas tolkningar av enkätens sammansättningar.

Som antytts i föregående avsnitt kan man anta att saliens är en bidragande orsak till den stora andelen egenskapsprojicerande tolkningar av *zebralejon* och *corpörn* i enkäten hos både första- och andraspråkstalare av svenska. De aktuella sammansättningarna kännetecknas av att förledens referenter har ontologiskt salienta egenskaper, vilket ordboksdefinitionerna tar fasta på, där zebrans randighet och korpens svärta behandlas som konstanter, dvs. är de drag som anges som särskiljande i relation till hyperonymen *hästdjur* respektive *kråkfågel* (se exempelvis *Svensk ordbok* 2009):

zebra	typ av svart- och vitrandigt hästdjur (...)
corp	en stor kråkfågel som har kolsvart, metallglänsande fjäderdräkt (...)

Eftersom zebrans randighet och korpens svärta är så utmärkande är det rimligt att tänka sig att dessa egenskaper lätt aktualiseras när språkbrukaren möter sammansatta ord med *zebra* och *corp* som sammansättnings-

led i enkäten, dvs. att associationen *zebra* > *randig*, respektive *corp* > *svart* är befäst. De aktualiserade egenskaperna projiceras i nästa steg på sammansättningen som helhet. Annorlunda uttryckt tillämpas ett lagrat schema [*zebra-N*] 'randig N' respektive [*corp-N*] 'svart N', som är befäst hos språkbrukaren p.g.a. *saliens* (snarare än frekvens). En sådan process förklarar den stora andelen randiga lejon, randiga väskor och svarta korpar i informanternas tolkningar av *zebralejon*, *zebraväska* och *corpörn*.²⁷ I fritextdelen av enkäten tolkas *zebraväska* i nästan tre fjärdedelar av fallen egenskapsprojicering, t.ex. 'väska som är randig', trots att det rör sig om en tillfällig sammansättning med många potentiella betydelser. Liknande resultat beträffande djursammansättningar har presenterats i tidigare forskning (se avsnitt 2.3). Sambandet mellan *saliens* och jämförelse sammanfattar Ryder (1994:124) på följande sätt: "the characteristic upon which the similarity relation is based is likely to be a rather salient feature of a central schema for the modifying element, and usually also of the compound's referent". Med Ryders analysmodell kan informanternas tolkningar av djursammansättningarna i enkäten antas utgå från ett semantiskt särdragsschema (se avsnitt 3.1.5).

Saliens kan också ha betydelse för tolkningarna av den etablerade sammansättningen *isvatten*. Som väntat har L1-grupperna en större andel konventionella, relationslänkande tolkningar än L2-grupperna. Om sammansättningen är obekant som lexikal enhet är egenskapsprojicering däremot en fullt rimlig tolkningsstrategi, eftersom *kyla* är en salient egenskap hos *is*, vilket även utnyttjas i den jämförande sammansättningen *iskall* 'kall som is'. Det är därför inte förvånande att andelen egenskapsprojicering tolkningar är större hos andraspråkstalarna, särskilt hos dem som har en högre startålder för svenskan.

Salienta egenskaper hos modifierarnas referenter frammanar en hög grad av jämförande tolkningar också vid mindre etablerade NA-sammansättningar som *honungssöt*, *snömjuk* och *glasskall*.

Olika typer av fenomen i världen tilldrar sig uppmärksamhet i varierande grad, dvs. de har olika uppmärksamhetsvärde, eller är olika salienta. Enligt Langackers salienhierarki gäller följande: "a whole is more salient than its parts; a physical object is more salient than an abstract entity; and a person has maximal cognitive salience" (Langacker 1991:171; se avsnitt 3.1.6). Det finns indikationer i mitt material om att dessa styrkeförhållanden påverkar informanternas tolkningar.

27 Det är dock något överraskande att även L2-talarna i stor utsträckning betraktar *corpörn* som en svart örn. Medan *zebra* är ett internationellt ord, är *corp* ett språkspecifikt och mindre frekvent ord. Om man inte känner till korpen och dess egenskaper är en egenskapsprojicering tolkning av *corpörn* inte resultatet av *saliens*. Det kan snarare vara det jämförande mönstrets allmänna tillgänglighet som slår igenom.

Det första ledet i citatet från Langacker rör förhållandet mellan salien- sen för å ena sidan helheten, å andra sidan delarna, vilket har betydelse i en analys av sammansättningar. Att de frekventa sammansättningarna i enkäten tolkas med större samstämmighet än de mindre frekventa är en indikation på att de förra i högre grad än de senare uppnått det stadium då helheten är mer salient än delarna. Ett sammansatt ord som språk- brukaren hört ofta övergår till att uppfattas som en helhet, dvs. som en enhet med en befäst betydelse. När språkbrukaren inte behöver pussla ihop delarna framträder betydelsen hos den befästa enheten som det själv- klara alternativet och andra potentiella alternativ i sammansättningens meningspotential överskuggas. Detta förhållande ger också L1-talaren en relativ fördel framför L2-talaren, som inte kan antas ha lika många ord befästa i sitt andraspråk. L2-talarna förlitar sig i högre grad än L1-talarna på kompositionalitetsprincipen (Wray 2002, 2008; se avsnitt 3.2.3).

Enligt det andra ledet i citatet från Langacker är konkreta företeelser mer salienta än abstrakta. Det är också känt sedan tidigare att ord som betecknar det konkreta är mer lätthanterliga för inlärare än ord som be- tecknar abstrakta företeelser, förutsatt att ordens frekvens är jämförbar (se avsnitten 2.5 och 3.2.3). Detta förhållande kan appliceras på de sig- nifikanta L1–L2-skillnaderna vid sammansättningarna *båtsommar* och *boktext*. L2-talarna tolkar dessa omvänt med *båt* respektive *bok* som huvudled i betydligt högre grad än L1-talarna. Resultatet kan relateras till strukturen med en konkret modifierare och ett abstrakt huvudled. Om man inte känner till sammansättningarna *båtsommar* och *boktext* som enheter är det kognitivt rimligare – utifrån antagandet om relativ saliens – att tänka sig att man kan prata om olika sorters konkreta båtar och böcker, än om olika sorters abstrakta somrar och texter.

Likaså är det förväntat att inlärare lär sig konkreta betydelser hos po- lysema ord före abstrakta, om betydelserna är jämförbara med avseen- de på frekvens. En sådan effekt kan ligga bakom (den icke signifikanta) tendensen till mer bokstavliga tolkningar av *skräp* i sammansättningen *skräpkruka* bland L2-talarna med en hög startålder för svenskan än bland L1-talarna i SVE. L2-talarna är benägna att välja relationslänkningen 'kruka med skräp i', medan informanterna i SVE har en viss övervikt av egenskapsprojicerande tolkning 'kruka som ser skräpig ut'. För att kun- na göra en egenskapsprojicerande tolkning av *skräpkruka* krävs att man gör en abstrakt tolkning av *skräp* som något dåligt eller fult, medan den relationslänkande definitionen bygger på en konkret förståelse av ordet. Den abstrakta tolkningen av *skräp* förefaller med andra ord vara mindre tillgänglig för informanter med en högre startålder för svenskan.

Den lägre graden av saliens för abstrakta företeelser, jämfört med konkreta, har betydelse för tolkningarna av AA-sammansättningar som

betecknar bl.a. färg och smak. Dessa sammansättningar tolkas påfallande divergerande i enkäten. Det kan bero på avsaknad av ett konkret och tydligt betydelseinnehåll hos sammansättningarna. Färgbeteckningarna är av naturen diffusa och låter sig svårligen definieras – är *gulgrön* mest grön eller mest gul? – eftersom färger rör individuell perception. Följaktligen blir skillnaden mellan två språkbrukares tolkningar av en AA-sammansättning sällan salient. Eftersom sådana skillnader i tolkning inte uppmärksammas avstannar den konventionaliseringsprocess, som annars skulle kunnat leda till en ökad samstämmighet om AA-sammansättningens semantiska innehåll.

Det tredje ledet i Langackers citat om saliensen hos företeelser i världen förutsäger att personer har maximal kognitiv saliens. Detta förhållande har betydelse för informanternas tolkningar av de tillfälliga bildningarna *vagnbarn* och *stövelbarn*. Dessa sammansättningar tolkas i stor utsträckning med *barn* som sammansättningens huvudled, trots att det torde röra sig om förhållandevis okända begrepp. Av betydelse för resultatet är det faktum att sammansättningarna har ett agentivt huvudled. I en tidigare studie av Onysko (2016) uppmärksammas det att nybildningar där huvudledets referent är agentiv förefaller ha större motståndskraft mot att tolkas omvänt med förledet som det semantiska huvudledet än andra typer av nybildningar (se avsnitt 2.4.1). Resultaten från Onyskos studie och resultaten från föreliggande studie ligger således i linje med Langackers konstaterande att en person har maximal kognitiv saliens.

7.4. Blockering

Blockering är en annan kognitiv process med potential att påverka språkliga tolkningar. Processen innebär att en etablerad beteckning för ett fenomen blockerar eller hämmar uppkomsten av en annan beteckning för samma sak. Särskilt gäller det morfologiska former, såsom när förekomsten av *thief* blockerar **stealer*, trots att det personbetecknande suffixet *-er* är produktivt (Aronoff 1976:43). En sådan blockeringsprocess kan förklara informanternas tolkning när de i enkätens fritextdel möter sammansättningen *husbil*. Alla utom två informanter tolkar ordet konventionellt som en 'bil som man kan bo i'. Om man inte känner till sammansättningen sedan tidigare, och inte har strukturen [efterled=huvud] helt internaliserad, är det rimligt att utvärdera den potentiella betydelsen 'hus för bil'. I enkäten är det bara två informanter som ger en sådan tolkning av *husbil*. Eftersom 'hus för bil' redan har en beteckning – *garage* – är det möjligt att tolkningsalternativet 'hus för bil' på grund av blockering inte aktualiseras för informanterna.

Å andra sidan förekommer flera svar i enkäten där blockering inte verkar slå igenom, när informanter tolkar en ny beteckning (t.ex. *vagnbarn*) på samma sätt som en gammal, känd beteckning (*barnvagn*), dvs. där de tillämpar en ny beteckning på ett känt semantiskt innehåll, 'vagn för barn' (se vidare det följande avsnittet 7.5).

7.5. Omvärldskunskap

Kognitiva perspektiv utgår från en encyklopedisk semantik, där språklig betydelse ses som nära sammanflätad med kunskap om och erfarenhet av världen (se avsnitt 3.1.2). Omvärldskunskapen ingår i meningspotentialen hos språkliga uttryck, såsom framgår av följande citat: "[m]eaning potentials contain both kinds of information – information deriving from use of language and information deriving from other experience with the world" (Allwood 2003:43).

Omvärldskunskapen har givetvis stor betydelse när vi tolkar obekanta sammansättningar, där flera alternativ i meningspotentialen är tillgängliga. Med Ryders (1994) terminologi måste sådana checkas av mot semantiska scheman (se avsnitt 3.1.5). I det följande ger jag några exempel från enkäten, där pragmatik inverkar på informanternas tolkningar.

Djursammansättningar får åter tjäna som ett illustrativt exempel. I sin studie av nybildade, engelska sammansättningar visar Ryder (1994) att det egenskapsprojicerande mönstret slår igenom starkt vid vissa djursammansättningar, t.ex. *elephant-fish*. Djur med karakteristiska drag frammanar särdragsschemat. Vid andra djursammansättningar, såsom *rabbit-dog* fördelar sig tolkningarna jämnt mellan egenskapsprojicering och ett relationslänkande jaktschema. Detta indikerar att valet av tolkningsmönster förstås är avhängigt vår kunskap om specifika djur och deras inbördes relationer. Det är pragmatiskt rimligt med en hund som jagar kaniner. Med Ryders analysmodell kan en sådan tolkning stödjas av händelsescheman. Men det är knappast rimligt med en fisk som jagar elefanter. Eftersom en sådan tolkning inte kan baseras på något händelseschema, aktualiseras i stället ett särdragsschema (se avsnitt 3.1.5). Samma förhållande förklarar varför *tiger-lion* i Ryders (1994) studie tolkas egenskapsprojicerande (såsom ett 'randigt lejon'), snarare än med en jaktrelation. I det här fallet rör det sig om en sammansättning där både ledens referenter är rovdjur.

I föreliggande studie testas tolkningen av en liknande sammansättning, *zebralejon*. *Zebralejon* skiljer sig på ett avgörande sätt från *tiger-lion*, eftersom det i zebralejonets fall rör sig om en sammansättning där huvudledets referent (lejonet) är ett rovdjur och modifierarens referent (zebran) en gräsätare som utgör ett bytesdjur för lejonet. Trots att det relationslän-

kande jaktschemat därmed är relevant är det en överväldigande majoritet som väljer bort detta till förmån för en egenskapsprojicerande tolkning 'lejon som är randigt som en zebra'. I det här fallet kan det paradoxalt nog vara rimligheten i ett zebrajagande lejon som utesluter jaktschemat. Som Ryder påtalar är det enligt Grice' maxim om kvantitet förväntat att förledet i en specifik sammansättning har funktionen att modifiera eller specificera huvudledet, eftersom huvudbegreppet annars hade kunnat användas ensamt (Ryder 1994:87). I Downings studie (1977:832) visade det sig exempelvis svårt för informanterna att tolka nybildningen *egg-bird*. Eftersom alla fåglar lägger ägg ger förledet i detta fall redundant information, vilket skapar en osäkerhet om vad förledet egentligen tillför sammansättningen. En sådan pragmatisk aspekt skulle kunna inverka på informanternas tolkning av *zebralejon*. Lejon jagar normalt zebror, och en jaktrelationstolkning skulle därmed bygga på redundant information. Endast vid en egenskapsprojicerande tolkning får förledet den förväntade funktionen att modifiera eller specificera huvudledet. Informanterna överser därmed med en annan aspekt som har med omvärldskunskap att göra, nämligen frånvaron av randiga lejon i världen.

Ett annat exempel från enkäten där omvärldskunskapen spelar in, rör tolkningarna av *husbåt* respektive *båthus*. Medan *husbåt* nästan uteslutande tolkas konventionellt som en sorts båt, visar sig *båthus* vara ett svårtolkat ord för informanterna, som i stor utsträckning väljer vänster-tolkningen 'en sorts båt' även vid *båthus*. Tendensen är särskilt tydlig hos de flerspråkiga informanterna (i synnerhet hos informanterna i SVE+). Samma tendens – att vända på mindre frekventa sammansättningar – förekommer i tolkningarna av *vagnbarn*, *stövelbarn*, *boktext* och *skalpotatis*. Resultatet kan relateras till omvärldskunskap. Om de enskilda sammansättningarna i enkäten (t.ex. *båthus*, *vagnbarn* och *skalpotatis*) är okända, och om det strukturella schemat för sammansättningar inte är helt befäst, är det rimligt att språkbrukaren för att skapa en mening i sammansättningarna orienterar sig mot ett känt semantiskt innehåll, såsom 'husliknande båt', 'vagn för barn' och 'skal från potatis'. En sådan strategi hos informanterna uppmärksammas av Ryder (1994:140). I hennes studie förklaras exempelvis de nya beteckningarna *mountain-chair* och *sea-lamp* med de kända beteckningarna *ski lift* respektive *lighthouse* (Ryder 1994:137). Liksom i Ryders studie indikerar resultatet i min studie att omvärldskunskapen i detta fall väger tyngre än principen om blockering: informanterna orienterar sig mot ett mer känt begrepp för att skapa en mening i en okänd sammansättning, trots att en sådan strategi innebär att de accepterar en ny beteckning för ett specifikt semantiskt innehåll som redan har en känd beteckning, vilket strider mot principen om blockering (Aronoff 1976:43; se resonemanget i avsnitt 7.4).

Tolkningarna av de ambigüosa sammansättningarna med efterledet *-misshandel* kan också kopplas till omvärldskunskap och i detta fall erfarenhet av maktrelationer. I fritextdelen av enkäten tolkar alla informanter utom en (dvs. 99,5 % av informantgruppen) barnet som drabbad part (dvs. PATIENT) i sammansättningen *barnmisshandel*. Vid *kvinnomisshandel* (i flervalssdelen av enkäten) är motsvarande siffra något lägre, 93 % patienttolkning, medan 7 % väljer tolkningen där kvinnan är AGENT och utför misshandel (enligt exempelmeningen riktad mot en man). I fallet med *mansmisshandel* är det 85 % som tolkar mannen som PATIENT, medan resterande 15 % identifierar mannen som den AGENT som utövar misshandel (enligt exempelmeningen riktad mot en kvinna). Patienttolkningen är tydligt prefererad för dessa tre ord. Siffrorna tyder dock på att informanterna är något mer benägna att tolka barn än vuxna respektive kvinnor än män som drabbad part i sammansättningarna med mönstret [N-misshandel]. Detta sammanhänger med kunskap om maktrelationer och styrkeförhållanden.

Som framgått förekommer i många fall skillnader mellan L1- och L2-talarnas tolkningar av sammansättningarna i enkäten. I vissa fall kan det bero på att L2-talaren ännu inte internaliserat den nya beteckningen (i andraspråket) för ett begrepp som är känt från modersmålet. Men det kan också bero på skillnader på den konceptuella nivån (se avsnitt 3.1.8). Begrepp som 'båthus', 'båtsommar', 'kvinnopräst', 'småkakor', 'grönyta' och 'skalpotatis' har begränsad kulturell räckvidd och kan därmed vara okända för L2-talaren. I sådana fall är det inte bara ett nytt ord som ska införlivas, utan också ett nytt begrepp. Enskilda sammansättningar, t.ex. *solvarm*, *isvatten* och *fågelrädd*, skulle dessutom kunna ha ekvivalenter i andra språk med andra konventionaliserade betydelserelationer mellan leden än i svenskan, t.ex. 'varm som solen', 'kallt vatten', 'rädd som en fågel'.

7.6. Konklusion

Resultaten av min studie indikerar att språkbrukaren i första hand hanterar en sammansättning på den specifika instansens nivå. Om sammansättningen finns befäst som en enhet eller ett exemplar behöver den inte analyseras, utan betydelsen är direkt tillgänglig att hämta fram ur det mentala lexikonet (se avsnitt 3.1.4). Att en befäst sammansättning inte behöver analyseras utifrån delarna kan också relateras till Langackers antagande om saliensen hos fenomen i världen: "a whole is more salient than its parts" (Langacker 1991:171). Befästning är ofta resultatet av hög frekvens, men även andra aspekter, såsom saliens, kan leda till befästning

av instanser och scheman. Instansen *zebra* och schemana [zebra-A] och [zebra-N] skulle exempelvis ha potential att nå befästning utifrån ett fåtal instantieringar, eftersom zebrans randighet är en så framträdande egen-skap.

Om en sammansättning inte är befäst som en enhet söker sig språkbrukaren vidare på mönsternivå. I detta steg analyseras sammansättningen utifrån de ingående leden. Det är tydligt i mitt material att informanterna jämför och tolkar obekanta sammansättningar utifrån liknande sammansättningar och mönster, såsom när den tillfälliga sammansättningen *manspräst* tolkas med samma relation mellan leden som den etablerade sammansättningen *kvinnopräst*, eller när tillfälliga NA-sammansättningar tolkas jämförande i enlighet med ett frekvent jämförande mönster vid etablerade NA-sammansättningar. Liksom flera kognitivt inriktade forskare ser jag ingen avgörande, principiell skillnad mellan tolkning utifrån ett schema och analogisk tolkning utifrån en specifik instans (dvs. en enhet eller ett exemplar; se avsnitt 3.1.5). Båda processerna handlar om att identifiera och använda mönster i tolkningen. Samtidigt är det en teoretiskt intressant fråga vilken abstraktionsnivå som tillämpas i denna mönstertolkning.

Det är principiellt svårt att avgöra om en tolkning går via ett produktivt schema – abstrakt eller mer ytnära – eller om tolkningen görs med en specifik instans som analogibas. Detta diskuteras ingående av Ryder (1994; se avsnitt 3.1.5). Som Langacker (2008) framhåller är det ett utmärkande drag i vår kognition att vi kan konceptualisera situationer på olika schematicitetsnivåer (se avsnitt 3.1.3). I mitt material finner jag emellertid empiriska belägg för att ytnära scheman och analogi spelar en viktig roll.

När en stor andel av informanterna tolkar *vagnbarn* och *stövelbarn* som en sorts barn, trots att bildningarna inte är etablerade, är det svårt att avgöra om de använder sig av det abstrakta schemat $[NN_{\text{huvud}}]_N$ eller om tolkningen utgår från ett lågnivåschema [N-barn], eller om den görs i analogi med ett befäst exemplar, såsom *skolbarn*. Produktiviteten för mönstret skulle kunna tala för det övergripande abstrakta schemat. Upprepningen av instanser med huvudet till höger, och frånvaron av instanser med huvudet till vänster, kan antas leda till en stark befästning av mönstret, möjligen starkare än motsvarande befästning av underordnade, delvis fyllda scheman, såsom [N-barn] och [N-sommar], vilka instantieras av få kandidater. Schemat för huvudledets placering i sammansättningar är etablerat och typprekvent. I många fall ger sådana övergripande scheman en för abstrakt information, men i detta fall ger schemat tillräcklig information för att språkbrukaren ska lyckas identifiera huvudledet. Däremot för det endast med sig begränsad semantisk information, dvs. schemat vägleder språkbrukaren fram till en mycket schematisk tolkning, 'en sorts

barn', men det framgår inte av detta abstrakta schema vilken semantisk relation som föreligger mellan barnet och vagnen. Med Ryders (1994) terminologi leder schemat fram till ett semantiskt skelett (se avsnitt 3.1.5). I flervaldsdelen är samstämmigheten förhållandevis stor om att ett *vagnbarn* är en sorts barn, men de fria förklaringarna av *vagnbarn* visar en stor spridning när det gäller informanternas tolkning av den exakta relationen mellan barnet och vagnen.

I de andra kategorierna av sammansättningar finns inte så entydiga scheman att följa. I stället rör det sig om konkurrerande, semantiska mönster. NA-sammansättningar som *snökall* och *glasskall* kan tolkas som instanser av ett jämförande mönster [A som N] eller ett orsaksrelaterat mönster [A av N]. Eftersom dessa mönster är abstrakta kan språkbrukaren behöva mer konkreta förebilder, t.ex. använda sig av ett mer ytnära schema [kall som N] (vilket är mer frekvent än [kall av N] och därmed sannolikt mer befäst) eller tolka *snökall* och *glasskall* i analogi med en befäst instantiering såsom *iskall*. I en sådan analogisk tolkning aktualiseras endast det jämförande mönstret i tolkningen av *glasskall* och *berghård*. Ryder (1994) för ett liknande resonemang, om behovet av konkreta analogibaser, utifrån sina studier om tolkning av nya sammansättningar (se avsnitt 3.1.5).

Tolkningarna i mitt material indikerar att specifika instanser är lättillgängliga analogibaser. Informanter i alla grupperna använder sig i stor utsträckning av analogiska mönster, såsom när *gråtfärdig*, *skrattfärdig* och *skrikfärdig* tolkas med samma relation mellan leden, liksom *manspräst* och *kvinnopräst* tolkas parallellt. När den tillfälliga bildningen *skrattfärdig* tolkas med betydelsen 'färdig att börja skratta' är det troligt att tolkningen görs i analogi med den specifika, etablerade instansen *gråtfärdig*, även om man inte kan utesluta att tolkningen baseras på en något mer abstrakt språklig mall [färdig att börja VERBA]. I fallet med *manspräst* ger schemat [N-präst] begränsad vägledning, eftersom det kan instantieras av antingen en identitetsrelation (som i *kvinnopräst*) eller en tematisk relation (som i *studentpräst*). Därför är det sannolikt att *manspräst* tolkas med den specifika sammansättningen *kvinnopräst* som analogibas. Det är tydligt i materialet att både första- och andraspråkstalare använder sig av analogiska mönster när de tolkar dessa sammansättningar. Skillnaderna i samstämmighet mellan första- och andraspråkstalarnas tolkningar beror således inte på skillnader avseende tolkningsprincip, utan på skillnader i befästning av de instanser som kan utgöra analogibaser vid tolkningen (i det här fallet *gråtfärdig* och *kvinnopräst*). L1-talarna är mer bekanta än L2-talarna med de konventionaliserade betydelserna hos *gråtfärdig* och *kvinnopräst*, dvs. de stöder sig i högre grad på ytnära semantiska mönster. Sammanfattningsvis bekräftar min empiri den bild som presenterats

i tidigare forskning om att analogi är en grundläggande, kognitiv princip (se avsnitt 3.1.5 och där refererad litteratur).

Den analogiska tolkningen står sig stark även i fall där den krockar med andra principer. Detta framgår av flera exempel i enkäten. *Issaft* i fritextdelen av enkäten tolkas i större utsträckning som 'saft med is i' – dvs. i analogi med *isvatten* – än som 'kall saft', trots att kyla är en så framträdande egenskap hos is att den skulle kunna frammana en jämförande tolkning. Likaledes tolkar förstaspråkstalarna i stor utsträckning *lökstark* med en orsaksrelation mellan leden, parallellt med *kryddstark*, medan andraspråkstalarna är mer benägna att använda ett övergripande jämförande mönster, som noterats fungera som default i tolkningen av nya NA-sammansättningar. Detta tyder på skillnader i befastning av en potentiell analogibas, *kryddstark*. Samma sak gäller tolkningarna av *musrädd*. L1-talarna tolkar i stor utsträckning (och i högre grad än L2-talarna) *musrädd* enligt det etablerade, språkliga lågnivåmönstret [rädd för N], snarare än ett allmänt jämförande mönster. *Glaslåda* i fritextdelen av enkäten tolkas i mycket stor utsträckning som en låda av materialet glas, i analogi med *trälåda* och *plastlåda* (eller möjligen i enlighet med ett mer abstrakt schema [N-låda]), trots att en sådan referent, dvs. en låda av glas, är mindre vanligt förekommande. Sammantaget indikerar detta att språkbrukare i första hand stöder sig på ytnära språkliga mallar i tolkningen av nya sammansättningar.

Min empiri visar att analogisk tolkning väger tyngre än blockering. Vid tolkningen av *glasskall* och *snökall* leder förekomsten av *iskall* inte till blockering av tolkningen 'mycket kall'. Tvärtom tolkar informanterna i stor utsträckning *glasskall* och *snökall* i analogi med betydelsen hos den etablerade sammansättningen *iskall*. Likaledes förekommer en hel del tolkningar där *kvinnoman* bedöms ha samma semantiska innehåll som *kvinnokarl*, dvs. den analogiska tolkningen vinner över principen om blockering. Även dessa resultat ligger i linje med resultaten hos Ryder (1994:137, 140) och tyder på att ytnära analogi ligger nära till hands när språkbrukare möter okända sammansättningar.

Avslutningsvis ger denna studie ett empiriskt bidrag till teoretiska antaganden inom lexikal och kognitiv semantik och i viss mån även till flerspråkighetsforskning. När det gäller de antaganden som rör sammansättningsstruktur och betydelser (se avsnitt 4.2) ger empirin stöd för hypotes 1a: schemat för huvudledets placering sist i nominala sammansättningar är etablerat i svenskan, men samtidigt framgår det av denna undersökning att det inte fungerar tvingande. Hypotes 1b infrias: huvudledets placering i AA-sammansättningar är mindre etablerat än huvudledets placering i NN-sammansättningar. Generellt föreligger en korrelation mellan frekvens och konventionaliserad betydelse, såsom antas i hypotes

1c. I materialet finns dock vissa undantag från detta samband, undantag som diskuterats ovan (se avsnitt 7.1.13).

Dessa slutsatser avseende etablering i språkgemenskapen (hypotes 1) följer av det tydliga resultatet när det gäller tolkningarna i informantgruppen som helhet. Huvudledets placering till höger i NN-sammansättningar får stort genomslag i informantgruppen (i linje med hypotes 2a), medan spridningen i tolkningarna är stor när det gäller identifieringen av AA-sammansättningarnas huvudled (såsom antas i hypotes 2b). Den generella bilden är att frekventa sammansättningar ges mer samstämmiga tolkningar än mindre frekventa (i enlighet med hypotes 2c). Undantag från denna generella bild förekommer dock och har diskuterats ovan (avsnitt 7.1.13).

När det gäller de antaganden som utgör hypotes 3 och avser skillnader mellan grupperna, bekräftas att L1-talarna följer det etablerade schemat för nominala sammansättningsstruktur i betydligt högre grad än L2-talarna (hypotes 3a). L1-talarna har en signifikant större andel höger-tolkningar än L2-talarna, och signifikanta skillnader föreligger vid alla nio NN-sammansättningar. Däremot falsifieras hypotes 3b: L1-talarna har inte en mindre andel vänstertolkningar av AA-sammansättningar än L2-talarna, utan tvärtom en signifikant större andel. Samstämmigheten är större bland L1- än L2-talarna vid tolkningen av etablerade sammansättningar i de kategorier som testar mer semantiska mönster. Det finns en tendens till att skillnaderna mellan L1- och L2-talarna är större vid etablerade men mindre frekventa sammansättningar än i etablerade och frekventa sammansättningar (såsom antas i hypotes 3c–d). Även hypotes 3e bekräftas: tillfälliga bildningar tolkas i analogi med betydelsen hos etablerade sammansättningar i högre grad av L1- än av L2-talarna.



8

Sammanfattande diskussion

I detta kapitel diskuteras hur de empiriskt grundade resultaten från min studie förhåller sig och kan bidra till sammansättningsforskningen, som växt fram som ett eget forskningsfält. Främst fokuseras det teoretiska bidraget till kognitiv lingvistik (i vid mening) och lexikal semantik, men även aspekter som kan bidra till flerspråkighetsforskningen uppmärksammas och integreras i diskussionen.

Det ger en förenklad bild av sammansättningskomplexitet att utgå från att deras betydelser enkelt kan förklaras utifrån ett fåtal potentiella semantiska relationer mellan leden (såsom antogs i generativa språkmodeller). Polysemi är grundläggande för sammansättningar, och därmed aktualiseras ett begrepp som meningspotential (se avsnitt 3.2.1). Konventionaliseringsprocesser bidrar emellertid till att skilja ut de betydelser som är mest brukliga i språkgemenskapen, varvid andra möjliga betydelser hos sammansättningarna hämmas.

Föreliggande studie, där en informantgrupp om 190 gymnasieelever fått i uppgift att tolka sammansatta ord av olika etableringsgrad, visar att språkbrukare utvärderar sammansättningskomplexiteter endast vid behov. När en sammansättning blivit befäst hos individen överskuggar den befästa betydelsen andra potentiella betydelser. När det gäller de tillfälliga sammansättningarna, som saknar en konventionaliserad betydelse, bekräftar mina resultat tidigare resultat, nämligen att språkbrukares

tolkningar påverkas av språkliga mallar på olika abstraktionsnivåer, dvs. etablerade och befästa sammansättningar, liksom ytnära och mer abstrakta scheman.

8.1. Befästning, analogi och mönster vid tolkning av sammansättningar

Den här studien bekräftar även ett grundläggande antagande inom bruksbaserad, kognitiv forskning, nämligen det att frekvens har betydelse för våra språkliga tolkningar, i det här fallet tolkning av sammansatta ord. Informanterna är generellt mer samstämmiga i sina tolkningar av de sammansättningar som är frekventa. Detta förklaras av att en frekvent sammansättning som språkbrukaren mött många gånger har större potential än en mindre frekvent att bli befäst som en *enhet* eller ett *exemplar* i språkbrukarens mentala lexikon. När en sammansättning blivit befäst hos den enskilde språkbrukaren behöver den således inte analyseras utifrån sina delar, utan den hanteras som en helhet. Det uppstår med andra ord en automatiserad koppling mellan uttrycket och dess innehåll, varvid helhetens kända betydelse överskuggar andra potentiella betydelser (se avsnitt 3.1.4).

Studien visar emellertid att även tillfälliga sammansättningar tolkas relativt samstämmigt, förutsatt att tolkningarna kan göras i analogi med liknande instanser eller baseras på ett etablerat mönster på ytnära eller mer schematisk nivå. Resultaten tyder på att språkbrukaren i första hand tolkar en ny sammansättning i analogi med en liknande, konventionaliserad sammansättning (såsom när *manspräst* tolkas i analogi med *kvinnopräst*). Detta överensstämmer med tidigare forskningsrön (se avsnitt 3.1.5). Om ingen sådan ytnära analogibas är tillgänglig för informanten kan ett något mer abstrakt mönster, som bara är delvis lexikalt fyllt, fungera vägledande, t.ex. [N-präst] i tolkningen av *manspräst* eller [N-rädd] i tolkningen av *musrädd*.

I mötet med en ny sammansättning kan språkbrukaren alternativt ta stöd i ett mer schematiskt mönster, t.ex. $[NN_{\text{huvud}}]_N$ som anger att efterledet är sammansättningens semantiska och grammatiska huvudled. Som diskuterats ovan (avsnitt 7.6) har ett sådant schematiskt mönster å ena sidan stor potential att bli befäst, p.g.a. hög typfrekvens. Å andra sidan ger det ingen detaljerad semantisk information, utan utgör snarare ett semantiskt skelett. Det ger därför begränsad vägledning när det gäller den specifika relationen mellan sammansättningens led (se avsnitt 3.1.5). Detta framgår av svaren vid *vagnbarn* i fritextdelen av enkäten. Med hjälp

av det övergripande schemat identifierar en stor del av informantgruppen den övergripande strukturen och förklarar att det rör sig om ett barn. Men hur barnet förhåller sig till förledets referent, dvs. vagnen, är inte självklart, och många olika semantiska relationer aktualiseras i informanternas tolkningar.

När språkbrukaren möter en ny sammansättning kan hon (också) utnyttja lagrade och befästa relationspreferenser, såsom antas inom CARIN-modellen. Företrädare för CARIN-modellen lyfter fram modifierarnas relationspreferenser som mer avgörande än efterledens, medan andra argumenterar för en interaktionell effekt där de semantiska kategorierna för sammansättningens båda delar är involverade (se avsnitt 2.2). Min empiri visar att båda konstituenternas relationspreferenser är viktiga. Förledets betydelse syns exempelvis i tolkningarna av *zebralejon* som tar sin utgångspunkt i ett yt nära schema [zebra-N], där en salient egenskap hos förledets referent, zebran, överförs till huvudledets referent så att sammansättningen kommer att beteckna ett 'lejon som är randigt som en zebra'. Efterledets betydelse framgår av tolkningarna av *fågelrädd*, där mönstret [N-rädd] 'rädd för N' fungerar vägledande. En stor andel MATERIAL-tolkningar av *glaslåda* skulle kunna baseras på relationspreferenserna för *glas-* (jfr CARIN-modellen), eller utgå från semantiska kategorier och baseras på mönstret [material-objekt] som involverar båda led (se avsnitt 2.2).

Befästning på individnivå motsvaras av konventionalisering inom språkgemenskapen, dvs. förekomsten av en socialt etablerad koppling mellan uttryck och innehåll. Vid en tillfällig eller lågfrekvent sammansättning kan många alternativ i meningspotentialen aktualiseras, och betydelsen kan då i en enskild användning bli föremål för förhandling. Ju mer frekvent, etablerad och spridd en sammansättning är, desto tydligare framstår *en* betydelse som den självklara, medan andra potentiella betydelser hämmas. Sammansättningens betydelse har blivit konventionaliserad i språkgemenskapen (se avsnitt 3.1.7. och 3.2.2).

8.2. Konventionalitet – en fråga om tid och input

Som framgår av avsnitt 3.2.3 har andraspråksforskare noterat en lägre grad av konventionalitet och idiomaticitet och en högre grad av variation i L2-talares än i L1-talares språkbruk. Min empiri visar samma sak: andraspråkstalarna är inte lika samstämmiga i sina tolkningar som förstaspråkstalarna, och de väljer i lägre grad de konventionaliserade betydelserna, eller

de alternativ som följer konventionaliserade mönster. När det gäller andraspråkstalarna är samstämmigheten och konventionaliteten i tolkningarna betydligt större i den grupp som börjat lära sig svenska tidigt (före sex års ålder) än i den grupp som börjat lära sig svenska vid sex års ålder eller senare. Skillnader i tolkningarna mellan informanter med olika språklig bakgrund förklaras av skillnader i exponeringstid för svenska språket. De informanter som haft svenskan med sig från början (L1-talarna) har större tillgång till befästa enheter på olika språkliga nivåer än de L2-informanter som börjat med svenskan under tidig barndom eller något senare.

Startåldern i sig (inte bara exponeringstiden) skulle kunna ha betydelse, på så sätt att biologiska förhållanden gör det lättare att tillägna sig ett språk före en viss ålder. Många andraspråksforskare har undersökt och diskuterat vilken betydelse startåldern för andraspråket har, när det gäller uppnådd slutnivå i andraspråket. Abrahamsson & Hyldenstam (2013:245–246) har iakttagit förekomsten av en skenbart kritisk inlärningsperiod före 6–7 års ålder. Detta innebär att en andraspråkstalare som börjat med andraspråket före 6–7 års ålder i vuxen ålder skulle kunna passera som infödd talare (även om man genom specifika test skulle kunna påvisa icke inföddlika drag hos talaren). Denna brytpunkt har i mitt material fått markera gränsen mellan tidig och sen inlärn timer av andraspråket. Det är emellertid sannolikt att den högre graden av samstämmighet och konventionalitet i tolkningarna från informantgruppen med startålder före sex år än i gruppen med startålder sex år eller äldre, är ett resultat av skillnader i exponeringstid och antal år i svensk skola, snarare än direkt avhängigt av startåldern i sig. Mina informanter är ungdomar vars språk är under utveckling, och några av dem har haft kontakt med svenska språket under så kort tid som fyra år. Skillnader i exponeringstid överskuggar därmed skillnader som skulle kunna ha med mognadsbegränsningar att göra. Inverkan från startålder och mognadsbegränsningar kan undersökas först när språkbrukare nått det som Abrahamsson & Hyldenstam (2013:245–246) kallar för ”slutlig behärskningsnivå”.

Signifikanta skillnader mellan L1- och L2-talarnas tolkningar av de konventionaliserade orden i enkäten indikerar skillnader i befästning av enskilda instanser, och tendensen i materialet är att dessa skillnader tilltar vid lägre frekvens. Mina resultat visar att informanterna i alla grupper använder sig av analogiska mönster i tolkningen av tillfälliga sammansättningar. Signifikanta skillnader mellan första- och andraspråkstalarnas tolkningar av dessa sammansättningar indikerar skillnader i befästning av språkliga mallar, dvs. specifika instanser, liksom ytnära och mer schematiska mönster som kan fungera som analogibaser i tolkningen.

Skillnader mellan enspråkiga L1-talare och flerspråkiga talare med svenska som L1 eller L2 sammanhänger med skillnader i input på ytter-

ligare ett sätt, eftersom flerspråkiga talare har tillgång till en ytterligare repertoar av mönster befästa genom input som de fått på sitt (andra) L1 (se avsnitt 3.1.8 och 3.2.3).

Om man betraktar informanternas förklaringar av orden i fritextdelen av enkäten i ljuset av resultatet från flervalsdelen, framstår skillnaderna i samstämmighet/konventionalitet mellan första- och andraspråkstalarnas tolkningar som större än vad som framkommer i flervalsdelen av enkäten. Detta illustreras exempelvis av de ord som förekommer i enkätens båda delar. De informanter som har svenska som andraspråk har en betydligt större andel konventionella tolkningar i flervalsdelen än i sina svar i fritextdelen. Skillnaderna mellan fritext- och flervalsdelen är däremot små för de informanter som har svenska som förstaspråk. Detta tyder på att L2-talarna är mer beroende av tillgång till kontext än L1-talarna. Resultatet tyder på att de skillnader i konventionalitet som iakttagits mellan första- och andraspråkstalarnas tolkningar i flervalsdelen snarare än att vara överskattade är något underskattade (se vidare avsnitt 6.5).

Den informantgrupp som består av talare som är flerspråkiga med svenska och ytterligare ett (eller flera) L1 (dvs. SVE+) bekräftar två viktiga antaganden. Å ena sidan illustrerar variationen i denna grupp, med avseende på de språkliga bakgrundsförhållandena, att L1/L2-distinktionen är metodologiskt svårhanterlig. Å andra sidan visar de skillnader i resultat som framkommer mellan undergrupperna i denna grupp, dvs. mellan de som har en förälder som har svenska som L1 (SVE+/S) och de som har föräldrar med andra modersmål (SVE+/A), att det är motiverat att göra en närmare analys av språkbrukares språkliga bakgrund. Med hjälp av en sådan analys har jag kunnat påvisa en skillnad avseende språklig konventionalitet mellan flerspråkiga ungdomar som skiljer sig åt när det gäller variabeln förälders modersmål. Därmed indikeras att en lägre grad av konventionalitet således inte primärt är ett resultat av flerspråkigheten i sig, utan ett resultat av skillnader i inputens omfattning.

8.3. Att identifiera det semantiska huvudledet

När det gäller schematiska mönster har frågan om det semantiska och grammatiska huvudledets placering uppmärksammats i tidigare forskning. För svenskans del har man fastslagit att sammansättningarna följer en universellt prefererad struktur, där det grammatiska och semantiska huvudledet är placerat till höger i sammansättningen. Hur talare förhåller sig till ett sådant schema i produktion såväl som reception har emellertid i begränsad omfatt-

ning undersökts empiriskt. Det är symptomatiskt att man, i den mån man alls intresserat sig empiriskt för sammansättningsstruktur, har studerat hur barn förhåller sig till strukturen hos sammansättningar. Eftersom huvudledets placering sist i sammansättningen tagits för given, har det endast varit aktuellt att studera vid vilken ålder språkbrukare internaliserar denna struktur, inte att undersöka i vilken mån de alls gör det (se avsnitt 2.4.1). Min empiri visar att det finns skäl att modifiera antagandet om ett till fullo etablerat schema. När det gäller nominala sammansättningar är huvudledets placering sist visserligen den prefererade tolkningen i materialet, men andelen omvända tolkningar där förledet betraktas som huvudled är förhållandevis stor i informantgruppen som helhet. Vänstertolkningar hos informanterna med svenska som L2 kan förklaras som antingen transfereffekter från deras L1 (om det språket har en vänsterstruktur) eller som resultatet av begränsad befästning av ett konventionaliserat mönster i L2. Andelen vänstertolkningar är dock inte obetydlig heller om man avgränsar sig till den grupp som har svenska som modersmål, trots att [efterled = huvudled] är ett mycket tyfrequent schema i svenskan. Resultatet kan jämföras med liknande resultat i informantstudier för engelskans del, där man också identifierat en icke obetydlig andel vänstertolkningar hos modersmåls-talare av engelska (se avsnitt 2.4.1). Som Ryder (1994) påpekar kan dessa återkommande fynd av vänstertolkningar i språk som antas ha en struktur med huvudledet till höger inte avfärdas som performansfel. Fynden falsifierar snarare antagandet om en absolut syntaktisk regel som förhindrar vänstertolkning (se avsnitt 2.4.1).

När man närmar sig de adjektiviska AA-sammansättningarna blir bilden ännu mer komplex. Utifrån lexikologisk tradition finns det två förväntade mönster: ett determinativt där efterledet utgör det semantiska huvudledet (såsom i *gröngul*), och ett kopulativt där sammansättningen innehåller två semantiska huvudled (såsom i *bitterljuv*). I enkäten förekommer ett tredje alternativ som distraktor, närmare bestämt en vänsterstruktur där förledet utgör det semantiska huvudledet och efterledet fungerar som sammansättningens modifierare. Anmärkningsvärt nog får detta mönster stort genomslag i informanternas tolkningar – i synnerhet hos L1-talarna i SVE. Informanter med svenska som andraspråk har en låg grad av determinativa tolkningar och använder kopulativ tolkning som default, vilket kan betraktas som en form av underspecificering. De informanter som har svenska som modersmål skiljer tydligare mellan traditionellt determinativa och traditionellt kopulativa sammansättningar. När de väljer en determinativ tolkning är emellertid det alternativ som speglar en struktur med huvudet till vänster det prefererade alternativet. L1-talarna väljer med andra ord i stor utsträckning en struktur som inte anses vara etablerad i svenska språket. Det sammantagna resultatet tyder på att förekomsten

av två etablerade och konkurrerande mönster (ett determinativt och ett kopulativt) luckrar upp gränserna för vilka tolkningar som är möjliga vid adjektiviska sammansättningar. Det har förutsatts att schemat [efterled = huvud] gäller alla determinativa sammansättningar oavsett ordklassstruktur. Denna empiriska studie visar att schemat inte aktualiseras i större utsträckning vid tolkningen av AA-sammansättningar.

Många av enkätens AA-sammansättningar är mycket frekventa, men samstämmigheten i informanternas tolkningar är trots det mycket begränsad. Detta indikerar att dessa sammansättnings betydelse är konventionaliserad i låg grad i den svenska språkgemenskapen. Resultatet synliggör en diskrepans mellan sammansättnings uttrycks- och innehållssida. Sammansättningarnas uttryckssida är etablerad (p.g.a. frekvens), medan innehållssidan, dvs. begreppen eller betydelsena, inte blivit konventionaliserade (se Svanlund 2009:227–228, Hannesdóttir 2012:54–55; se även avsnitt 7.1.8). Detta i sin tur sammanhänger med att adjektiviska sammansättningar, till skillnad från nominala, refererar till egenskaper som svårigen låter sig beskrivas på ett objektivt sätt, utan grundas i subjektiva perceptionsintryck av exempelvis färg och smak. Skillnaden mellan två språkbrukares tolkningar av en AA-sammansättning (är *gulgrön* 'mest grön' eller 'mest gul'?) blir sällan uppenbar. Därmed är förutsättningarna begränsade för en konventionaliseringsprocess, som skulle kunnat leda till en konsolidering av relationen mellan AA-sammansättnings form och semantiska innehåll. Detta resonemang kan kopplas till Schmidts (2015) argument om en avgörande skillnad mellan konventionalisering och befestning: ett ord kan spridas och etableras i en språkgemenskap, men en sådan konventionaliseringsprocess är "blind" för detaljerna i den mentala representationen hos de enskilda språkbrukarna. Att ordet är etablerat i språkgemenskapen innebär inte att det representeras på samma sätt i de enskilda språkbrukarnas mentala lexikon (se avsnitt 3.1.7).

Resultaten från min studie illustrerar att det inte är trivialt – och inte heller fruktbart – att entydigt fastställa strukturen i en AA-sammansättning. Detta illustreras tydligt i fritextdelen av enkäten, där informanterna själva formulerar förklaringar till AA-sammansättningarna *fulsnygg* och *dumsnäll*. Spännvidden i förklaringarna visar att det för det första är problematiskt att utgå från ett dikotomiskt förhållande mellan kopulativa och determinativa AA-sammansättningar. För det andra visar de att det i många fall är svårt att avgöra om informantens förklaring speglar struktur där efterledet utgör semantiskt huvudled, eller en struktur där förledet fungerar som semantiskt huvudled, såsom i förklaringen "en som inte ser bra ut men ändå gör det" eller "i vissa sammanhang ful o vissa snygg". Förklaringar som dessa låter sig inte inordnas efter dikotomin

kopulativ – determinativ, och de låter sig inte heller beskrivas i termer av höger- eller vänstertolkning.

8.4. Mångfald av meningspotentialer eller konventionella mönster?

I sammansättningsforskningen föreligger ett spänningsförhållande mellan å ena sidan forskare som (utifrån olika teoretiska perspektiv) framhåller sammansättningsregelsbundenhet och å andra sidan forskare som framhåller sammansättningsregelsvida meningspotentialer. Genom designen av föreliggande studie belyser jag i flervalssdelen av enkäten styrkeförhållanden mellan utvalda alternativ i sammansättningsregelspotentialer, dvs. där studeras generella drag och regelsbundenheter. Informanternas fritt formulerade förklaringar i fritextdelen av enkäten illustrerar snarare sammansättningsregelsrika meningspotentialer, och nyanserar därmed resultaten från flervalssdelen. I stora drag styrker förklaringarna i fritextdelen resultatet från flervalssdelen. Jämförbara sammansättningsregels tolkas i stor utsträckning med samma relation mellan leden i de båda delarna av enkäten (se dock diskussionen om de skillnader som noterats mellan L2-talarnas tolkningar i de båda delarna av enkäten, avsnitt 8.2), och vissa mönster får stort genomslag.

Detta förhållande illustrerar att möjligt inte är lika med brukligt. Att nya sammansättningsregels tenderar att bli alltmer konventionaliserade beror på att det har ett kommunikativt värde att vi förstår uttryck på samma sätt. Det har med andra ord ett kommunikativt värde att använda ett konventionaliserat språk. Något som däremot inte uppmärksammas i större utsträckning är att konventionalisering inte bara kan öka utan också minska (se dock Svanlund 2002:27–28). När en företeelse blivit mindre bruklig eller bekant blir ordets betydelse mindre konventionaliserad. Sannolikt är det en sådan effekt som syns i enkättolkningarna av *herrfrisör* och *båthus* som tolkas mycket divergerande i informantgruppen. Andra sammansättningsregels kan etableras som uttryck utan att betydelsen konventionaliseras, som framgår i diskussionen om AA-sammansättningsregels (se avsnitt 7.1.8 och strax ovan i avsnitt 8.3).

Ett mönster som får stort genomslag i informanternas tolkningar är jämförelse. Detta visar sig i egenskapsbetecknande NA-sammansättningsregels (*snökall*), såväl som person- eller djurbetecknande nominala sammansättningsregels (*flickpojke*; *zebralejon*). Sådana tolkningar av nya eller tillfälliga sammansättningsregels baseras inte på befästa lexikala instanser utan är resultatet av analogisk tolkning och tillgång till mönster. Det jämförande

mönstret förefaller ligga nära till hands. Särskilt när förledets referent har en salient egenskap, såsom i fallet med den randiga zebran, frammanas egenskapsprojicerande och jämförande tolkningar. Detta resultat bekräftar liknande fynd i tidigare forskning som dels identifierat jämförelse som en frekvent förekommande relation mellan leden i sammansättningar som betecknar växter och djur, dels identifierat jämförelse som en prominent tolkningsstrategi (se avsnitt 2.3). Som Wisniewski & Love (1998) framhåller framstår jämförelse därmed som grundläggande i tolkningar, snarare än som en ”last resort strategy” (se avsnitt 2.2). Också rikedom av metaforik i språket illustrerar att jämförelse är en stark kognitiv princip, som tar sig uttryck i språket. Egenskapsprojicerande tolkningar av *zebralejon* som ett ’randigt lejon’ skulle kunna ses som en sådan metaforisk tolkning. Det är anmärkningsvärt att ett jämförande mönster slår igenom i så hög grad i tolkningarna i alla informantgrupper. Resultatet kan vara en effekt av att jämförelse är en grundläggande kognitiv princip. All kategorisering utgår från att vi jämför och sorterar utifrån kriteriet likhet (se avsnitt 3.1.1). Detta förhållande kan också tänkas bidra till en benägenhet hos språkbrukare att tolka nya ambigüosa sammansättningar jämförande.

8.5. Genomskinlighet, kompositionalitet och befästning

Hur kompositionalitetsprincipen ska förstås i relation till begrepp som förutsägbarhet och genomskinlighet har diskuterats i tidigare forskning. Flera forskare har ifrågasatt en strikt tolkad kompositionalitetsprincip där helhetens betydelse är förutsägbar utifrån de ingående delarna (se avsnitten 3.1.2 och 3.2.1).

I föreliggande studie kan sammansättningarna i enkäten betraktas som kompositionala i betydelsen att de kan förstås utan metaforisk tolkning. Men detta betyder inte att de är genomskinliga i bemärkelsen att de har en självklar betydelse utifrån de ingående leden. Förståelsen av dem kan alltså inte härledas enbart ur kompositionalitetsprincipen.

Definitionen av genomskinlighet utgår från att tolkningen av en sammansättning går via en analys av de ingående leden, men min empiri indikerar snarare att förståelsen av sammansättningar i många fall kräver att de är befästa som enheter. Man kan anta att merparten av informanterna i min studie känner till orden *sol* och *varm*, men som framgått råder det stor spridning i tolkningarna av det sammansatta ordet *solvarm*. Detta gäller framför allt andraspråkstalarna, men även förstaspråkstalarna av svenska väljer i anmärkningsvärt hög grad det mindre konventionella alternativet

'varm som solen'. Exemplet illustrerar att språkbrukare inte kan förlita sig på kompositionalitetsprincipen. På grund av relationspolysemi kan språkbrukare sällan a priori pussla sig fram till den vedertagna betydelsen. Eftersom enbart kännedom om de ingående orden inte leder fram till en specifik betydelse blir det missvisande att karakterisera *solvarm* som en genomskinlig sammansättning. Som Langacker framhåller innebär polysemin hos sammansättningar att betydelsen i regel är *motiverad* av men inte *förutsägbar* utifrån de ingående leden. Sammansättningen *solvarm* skulle med Langackers terminologi kunna betraktas som kompositionell men inte analyserbar, motiverad men inte förutsägbar (se vidare avsnitt 3.1.2 och 3.2.1). Överlag synliggör L2-tolkningarna av de etablerade sammansättningarna i mitt material att det vanligen i sammansättnings meningspotentialer finns rimliga alternativ till den konventionaliserade betydelsen, dvs. de flesta sammansättningar är inte genomskinliga utifrån de ingående leden. L2-talarna förlitar sig i högre grad än L1-talarna på kompositionalitetsprincipen. När de analyserar sammansättningarna kompositionellt leder det i vissa fall till en underspecificering, såsom när *grönyta* tolkas som en 'yta som är grön'. I andra fall leder den kompositionella analysen fram till en mindre vedertagen betydelse, såsom i tolkningen där *solvarm* ges betydelsen 'varm som solen' (se avsnitt 3.2.3).

Min slutsats är att genomskinlighet eller analyserbarhet först och främst är en fråga om en subjektiv upplevelse (se också avsnitten 3.1.2 och 3.2.1), och upplevd genomskinlighet kan vid konventionaliserade sammansättningar vara en fråga om befästning. En sammansättning som finns befäst hos individen har en självklar betydelse för henne, och det har den utan att hon behöver göra en analys utifrån de ingående leden. I så måtto är den genomskinlig. Däremot kan man ifrågasätta hur lyckad genomskinlighetsmetaforen är i detta fall, eftersom man inte behöver "se igenom" befästa sammansättningar.

Ett ord som har en självklar betydelse för en språkbrukare som mött det så många gånger att det blivit befäst kan upplevas som ogenomskinligt för den som möter det första gången och har att utvärdera olika alternativ i sammansättningens meningspotential, såsom visades ovan i exemplet med *solvarm*. Det är också detta förhållande, svårigheten att bedöma genomskinligheten för den som möter ett ord för första gången, som gör det till en grannlaga uppgift att bedöma vilka sammansättningar som bör förklaras i ordböcker och läromedel, liksom i undervisningssituationer. Huruvida en sammansättning som är ny för den enskilde språkbrukaren framstår som genomskinlig sammanhänger med vilken tillgång hon har till analogibaser – dvs. scheman och mönster på olika abstraktionsnivåer – som kan utgöra tolkningsram och hjälpa språkbrukaren att "se igenom" sammansättningen.

9

Slutsatser

Denna avhandling bygger på en studie där 190 gymnasieelever tolkar sammansatta ord. Här beskrivs vilka mönster – strukturella, semantiska och lexikala – som aktualiseras i informanternas tolkningar. Vidare diskuteras vilka kognitiva principer som vägleder dem i deras tolkningar. I det följande presenteras några övergripande slutsatser från undersökningen.

Av min empiriska studie framgår att det finns skäl att modifiera axiomet att svenska sammansättningar följer en universellt prefererad struktur, där sammansättningens efterled utgör det grammatiska och semantiska huvudledet.

När det gäller NN-sammansättningar får strukturen där efterledet utgör huvudled visserligen stort genomslag i mitt material. Men språkbrukarna är inte heller främmande för att betrakta sammansättningens förled som det semantiska huvudledet vid mindre kända sammansättningar. Detta gäller inte bara de deltagare i studien som har annat modersmål än svenska. Även bland modersmålstalarna av svenska förekommer en inte obetydlig andel sådana tolkningar med huvudledet till vänster.

När det gäller AA-sammansättningar avviker tolkningarna av dessa i hög grad från förväntade mönster. Informanterna vacklar inte bara i valet mellan de etablerade strukturerna kopulativ (XY är både X och Y) och determinativ (XY är ett slags Y) tolkning. Vid determinativ tolkning är

alternativet där förledet utgör det semantiska huvudledet det prefererade valet, dvs. *gulgrön* uppfattas som 'mer gul än grön'. I motsats till vad man kunde förvänta sig är vänstertolkningar av detta slag mer frekvent förekommande bland de informanter som har svenska som sitt enda modersmål än bland de informanter som är flerspråkiga med svenska som L1 eller L2. De flerspråkiga informanterna föredrar kopulativ tolkning framför determinativ.

Min studie bekräftar följaktligen tidigare studier om ett förhållandevis etablerat schema $[NN_{\text{huvud}}]_N$, där efterledet utgör huvud. Det finns dock anledning att uppmärksamma att schemat inte är så etablerat att det fungerar tvingande för språkbrukare. När det gäller AA-sammansättningar visar studien att schemat $[AA_{\text{huvud}}]_A$ med efterledet som huvudled får begränsat genomslag i informanternas tolkningar, dvs. AA-sammansättningarna hanteras inte på samma sätt som NN-sammansättningarna. Detta resultat illustrerar ett generellt problem inom sammansättningsforskningen: den har till stor del koncentrerats till NN-sammansättningar, samtidigt som man uttalat sig om sammansättningar per se. Min empiri visar att det inte finns något etablerat schema på en högre abstraktionsnivå $[XX_{\text{huvud}}]_X$, där X utgörs av vilken ordklass som helst. Sammantaget aktualiserar behovet av vidare forskning om sammansättningar av olika ordklassstruktur.

En annan sanning som det finns skäl att ifrågasätta är *s*-fogens betydelse för det semantiska huvudledets räckvidd i treledade nominala sammansättningar. I min empiriska undersökning förefaller förekomsten eller avsaknaden av fog ha begränsad betydelse för den segmentering som informanterna gör. Det finns en svag preferens för en segmentering mellan led 2 och 3, dvs. en segmentering som innebär att ledet längst till höger utgör det semantiska huvudledet, och denna gäller oavsett förekomst eller avsaknad av fog.

Resultaten av min studie belyser också svårigheterna med att bedöma sammansättningsgenomskinlighet och indikerar att tillgången till en sammansättningskonventionaliserad betydelse ofta förutsätter befästning.

Föreliggande studie bekräftar däremot det grundläggande antagandet inom kognitiva perspektiv om att frekvens har betydelse för graden av befästning (på individnivå) och konventionalisering (på språkgemenskapens nivå). De sammansättningar som är frekventa tolkas mer samstämmigt än de sammansättningar som är mindre frekventa. Vid mindre frekventa eller tillfälliga sammansättningar har tillgången till språkliga mallar på olika schemacitetsnivåer, dvs. specifika instanser, liksom ytnära eller mer schematiska mönster som kan utgöra analogibaser, betydelse för språkbrukarnas sammansättningsstolkningar.

De av informanterna som har svenska som modersmål är generellt mer samstämmiga och konventionella i sina tolkningar av enkätens sammansättningar än de informanter som har svenska som andraspråk. Dessa skillnader kan vid de etablerade sammansättningarna förklaras som skillnader avseende befästning av de enskilda instanserna. När det gäller de tillfälliga sammansättningarna indikerar skillnaderna däremot varierande tillgång till sådana språkliga mallar som kan fungera som analogibaser i tolkningen. De skillnader som föreligger mellan L1- och L2-talarna illustrerar skillnader i befästning på två sätt. För det första är mönstren i svenskan mindre befästa hos L2- än hos L1-talarna, p.g.a. mindre svensk-språkig input. För det andra har de flerspråkiga talarna tillgång till en ytterligare uppsättning av mönster befästa genom input från sitt L1, som kan aktualisera andra alternativ i meningspotentialen.

När det gäller språkliga mönster som kan vägleda tolkningen av en obekant sammansättning, ger denna studie empiriskt stöd för slutsatsen att ytnära analogibaser spelar en särskilt viktig roll, både för L1- och L2-talare. Men det skulle krävas vidare forskning – och även forskning med andra metoder – för att komma närmare ett svar på frågan om vilket inflytande språkliga mönster av olika schematicitet har på språkbrukares tolkningar.

Som konstaterades i inledningen är sammansättningsforskningen ett område som rymmer många invecklade, svårlösta problem. Med hjälp av nyare paradig har jag i denna studie kastat nytt ljus över en del gamla sanningar om sammansättningar. Eftersom de i dubbel bemärkelse är ett komplext fenomen finns det behov av fortsatt forskning som ur olika perspektiv belyser sammansättningarna och deras betydelse och roll i vår språkliga kognition.



Summary

Introduction

In this thesis a cognitive questionnaire-based study on the interpretation of compounds is presented. The analysis is based on 190 questionnaires, where Swedish-speaking informants give their interpretations of 85 compounds. The informants are students in upper secondary schools. Included in this study are not only L1 speakers of Swedish, but also multilingual speakers with Swedish as their second language (L2). The reason for this is twofold. Firstly, I have done so to reflect the Swedish language as it is spoken in the multilingual Sweden of today. Secondly, this approach enables a contrastive analysis where I – from a usage-based perspective – can compare the interpretations from speakers with varying exposure time to Swedish.

The overall aim of the study is to test the assumption of usage-based approaches that linguistic phenomena become entrenched in the minds of speakers as a result of repetition. Through comparison between the interpretations of one group of speakers who have been exposed to Swedish from birth (or even earlier) and one group of speakers who have acquired

the Swedish language later (L2 speakers), I can analyse the impact of exposure time on the entrenchment of different compounds and compounding patterns. To be able to analyse such entrenchment effects, I have included not only conventionalised but also non-established compounds.

Compounds are fundamentally polysemic. They can be characterised as having wide meaning potentials (e.g. Hanks 2000, Allwood 2003). The aim of this study is to analyse on the one hand the meaning potentials and the activated meanings of the compounds. On the other hand, the aim is to analyse the interpretation strategies of the speakers, to see to what extent they rely on entrenched words, and to what extent they take advantage of schemas of various granularity. Frequent compounds are hypothesised to have more conventionalised meanings than less frequent ones, reflected in concordant interpretations. L1 speakers are hypothesised to have more concordant interpretations of established compounds and compounding patterns than L2 speakers.

Method and data

The questionnaire includes two parts with test items and one part with background questions, see appendix 1. The multiple choice part includes 69 compounds with two or three alternative paraphrases the informants can choose between. Each alternative is illustrated with a contextual sentence. This section gives information on the strength of different potential meanings. The results of the multiple choice part are analysed with statistical methods. Since the data is not normally distributed, non-parametric tests are used. In the free text part, the informants formulate their own paraphrases to each one of 20 compounds, which are similar (or in some cases identical) to the ones in the multiple choice part. The informants filled out the free text part before they filled out the multiple choice part. The free text part functions as a “control part”, to test the reliability of the results from the multiple choice part, where the informants had to choose from fixed alternatives. Since frequency is hypothesised to influence the interpretations, the concordance in the interpretation of each word is related to the frequency of the word, measured in terms of its representation in a reference corpus that is compiled of texts from daily newspapers, novels and blogs (Korp/Språkbanken) and teaching aids (Ribeck 2015) (see Table 4:13).

The informants are divided into four main groups, according to their answers to the background questions: speakers who report Swedish as their only first language (abbreviated SVE, $n=94$), speakers who report Swedish as one of their first languages (SVE+, $n=30$), speakers who have Swedish as

their second language and report an early age of onset of Swedish, before the age of six ($START < 6$, $n=34$), and finally, speakers who have Swedish as their second language and report an onset of Swedish at the age of six or later ($START \geq 6$, $n=32$) (see Table 4:15). The informants in the SVE+ group are further divided into two subgroups depending on the status of the Swedish language as an L1 in their families, since that variable turned out to be influential, regarding the degree of concordance and conventionality of the interpretations. The informants in SVE+/S report that one of their parents has Swedish as his/her L1, whereas the informants in SVE+/A report that their parents have other L1s than Swedish.

Theoretical perspectives

This study is conducted within a cognitive framework. In the analysis of the material, tools from cognitive grammar as well as from construction grammar are used.

Central in a cognitive and usage-based approach is the assumption that linguistic structure is emergent. Experience with language creates and impacts the cognitive representations for language. Frequency and entrenchment are core aspects in this process. What is frequent in language is more easily accessed, and repetition leads to conventionalisation of associations and categories (e.g. Langacker 1987, 2000, 2017, Kemmer and Barlow 2000, Bybee 2013, Schmid 2015). According to Langacker, the repetition of linguistic items leads to increasing entrenchment resulting in the establishment of *units*, i.e. pre-packaged items whose meanings are easily and automatically accessed by the speaker:

A **unit** is a structure that a speaker has mastered quite thoroughly, to the extent that he can employ it in largely automatic fashion, without having to focus his attention specifically on its individual parts or their arrangement. Despite its internal complexity, a unit constitutes for the speaker a “pre-packaged” assembly; because he has no need to reflect on how to put it together, he can manipulate it with ease as a unitary entity. It is effectively simple, since it does not demand the **constructive effort** required for the creation of novel structures. Psychologists would speak of a “habit”, or say that “automatization” has occurred. (Langacker 1987:57; emphasis in original)

Not only words and idioms/expressions, but also schemas, are established as units (Langacker 2017:44, 53). *Schema* is a central notion in cognitive linguistics, which can be defined as “a superordinate concept, one which specifies the basic outline common to several, or many, more

specific concepts” (Tuggy 2007:83). Patterns and schemas are established through a categorisation of similar linguistic items in a schematisation process, i.e. they are abstractions over concrete expressions: “[p]atterns and regularities of any sort, at any level of specificity, reside in schematic units abstracted from occurring expressions” (Langacker 2008:219). The notion of schema is similar to the notion of *constructions* in construction grammar, which is defined as form–meaning pairings “at varying levels of complexity and abstraction” (Goldberg 2013:17).

The repetition of a specific linguistic item or token, i.e. token frequency, leads to entrenchment of this specific item. The representation of different, unique items or types in a slot of a schema or a construction (i.e. type frequency) leads to entrenchment of the schema or the construction. In the context of compounds, high token frequency for the string *moonless* leads to entrenchment of this specific item as a unit, whereas high type frequency for N-*less*, e.g. *moonless*, *hopeless*, *fruitless*, *treeless* and *cordless*, leads to entrenchment of a superordinate schema [N-*less*] as a unit (Langacker 2008:24).

Schemas exist at different levels of abstraction. There is an ongoing discussion whether language users are more influenced by higher-level or lower-level schemas, or if they interpret new items using specific, established items as analogy bases. Cognitive researchers tend to emphasise the importance of lower-level schemas (e.g. Langacker 2000:29) and analogy (Bybee 2013:57–58, 61, Ryder 1994:89, 197). In the context of compounds, and based on informant studies, Ryder (1994) draws the conclusion that speakers tend to use low-level linguistic templates, such as established compounds or lower-level schemas, as analogy bases rather than more abstract schemas when they interpret novel compounds (1994:193, 197). This question is also addressed in the present study.

As compounds are fundamentally polysemic, they are interesting in relation to the notions of compositionality and meaning potential. From the perspective of conceptual combination, Costello & Keane (2000) question a strict definition of compositionality, where the meaning of the whole is considered predictable from the parts. In contrast, they propose a gradable compositionality concept. Likewise, Langacker (2008) states that the meaning of compounds is motivated, but not predictable, from their parts. From a lexicographic perspective, Hanks (2000) argues that words do not have meanings. Rather, they have meaning potentials. Meanings emerge from the context in specific situations. The notion of meaning potential is also discussed by Allwood (2003) and Norén & Linell (2006).

Since multilingual persons have rich linguistic experience from more than one language, their L2 is affected by prior associations from their L1. Therefore, they have access to wider meaning potentials than L1 speakers,

resulting in a less conventionalised language usage. From a cognitive perspective Ellis (2006) states that L2 acquisition in contrast to the “*tabula rasa*” situation of L1 acquisition is a situation of “*tabula repleta*”:

In the first few years of life, the human learning mechanism optimizes its representations of first language from the cumulative sample of first language input. One result of this process is that the initial state for SLA is a *tabula repleta*; it is no longer a plastic system, it is one that is already tuned and committed to the L1. Our later experience is shaded by prior associations; it is perceived through the memories of what has gone before. Since the optimal solution for L2 is not that for L1, SLA suffers from various types of L1 interference. Transfer phenomena pervade SLA. (Ellis 2006:109; emphasis in original)

To conclude, compounds have rich meaning potentials, rather than compositional meanings. From a cognitive perspective the meaning of a compound is motivated, but not predictable, from the components. When speakers interpret compounds, they rely on entrenched units of variable granularity. If they are unfamiliar with a compound, their interpretations are affected by entrenched linguistic and encyclopedic patterns. As shown by Ellis (2006), L2 speakers have a wider range of patterns to draw upon than L1 speakers (see also Pavlenko 2009).

Previous research

In the 1960s and 1970s, research on compounding was influenced by the generative paradigm. From such a perspective compounds were considered to be transformations from underlying relative and prepositional clauses (e.g. Lees 1960, Levi 1978). The aim of these generative approaches was to present semantic classifications, where all established compounds were sorted according to a finite set of semantic relations such as PART-WHOLE, PURPOSE and LOCATION.

Later, many researchers have objected to this “list approach”. Rather than establishing a finite list of semantic relations between the components, researchers have highlighted that compounds have a wide range of potential meanings (e.g. Downing 1977:840–841, Murphy 1988:555–557, Kay & Zimmer 1990:239–240, Costello & Keane 2000:325, Svanlund 2002:22–24, Plag 2003:148). However, what is possible is not necessarily equivalent to what is customary. While everything is theoretically possible, several empirical studies show that meanings follow conventionalised patterns (e.g. Eiesland 2016) and that interpretations are influenced by established units at different levels of abstraction. When a compound

is conventionalised in a specific meaning, this meaning overrides other potential meanings and novel compounds tend to be interpreted using linguistic templates, such as established compounds and schemas, as analogy bases (e.g. Ryder 1994:87–89, 193, Gagné & Shoben 1997, Maguire et al. 2010a, 2010b). Within the CARIN (= Competition Among Relations In Nominals) framework, Gagné & Shoben (1997) highlight the importance of the patterns for individual modifiers. For example, *chocolate-* as a modifier is associated with the <made of> relation. In contrast, Maguire et al. (2010a, 2010b) argue for a model where conceptual combination is based on semantic categories, rather than individual items. They propose an interactional model where speakers take advantage of frequent semantic patterns such as [substance-object], involving in the interpretation process semantic categories from both slots of the compound.

It has been observed that language users' interpretations of novel NN compounds can be property interpretations or thematic relation interpretations. In thematic relation interpretations the referents of the constituents play different functional roles, cf. a *robin hawk* is 'a hawk that preys on robins'. In property interpretations, the referents of the constituents are only indirectly related: the referent of the head is compared to one property of the modifier's referent (rather than the referent itself), cf. a *robin hawk* is 'a hawk with a red breast' (Wisniewski & Love 1998). The prominence of these different strategies has been discussed. According to Downing (1977) and Gagné & Shoben (1997) thematic relation is the preferred choice. A comparing property interpretation is the strategy chosen only if the language user can find no reasonable thematic relation, cf. "the last resort hypothesis". This hypothesis is dismissed by Wisniewski & Love (1998), who find empirical evidence that property interpretations are as likely to be used as thematic relation interpretations. Property interpretations are likely to happen when the referent of the modifier has salient features, cf. the stripes of the tiger or zebra (Ryder 1994:124, 361–362). Conventionalised compounds, denoting animals and plants often instantiate property relations, e.g. *zebra mussel* (Wisniewski & Love 1998:195, 198–199, Eiesland 2016:130–131).

The Swedish language is considered to follow a universal preference for right-headed compounds (SAG 2:42, 186, 526; cf. Williams 1981). This rule of right-headedness has been treated as an axiom, rather than being empirically analysed, with the exception of child language studies, which show a preference of right-headed structures over left-headed structures from an early age (Mellenius 1997:98–99; cf. Clark, Gelman & Lane 1985:86 regarding right-headedness in children's language in English).

Headedness has been studied from a multilingual perspective. Several contrastive studies show transfer effects from L1 to L2, concerning the

order of the head and the modifier (e.g. Nicoladis 2002, Foroodi-Nejad & Paradis 2009, Onysko 2016), regardless of the existence or not of a structural overlap between the languages (Müller 1998, Döpke 1998).

Results

The compounds of the questionnaire are chosen in order to represent different structural and semantic patterns. The compounds instantiate 12 different categories, of which six are discussed further here, focussing aspects such as headedness (in nominal and adjectival compounds), compositionality, preference for property or thematic relation interpretations and preference for a comparing versus a causal relation between the components (see Table 4:1–4:12). The distribution of the interpretations in the multiple choice part is presented category by category, in tables and boxplots (chapter 5). The test items are presented in a frequency order in the tables, with the most frequent compounds at the top of the tables. The frequency is measured in terms of representation in a reference corpus, (see chapter 4), and the figures are supplemented in appendix 3.

Frequency turned out to be a factor of great impact on the interpretations. Frequent compounds result in concordant interpretations, while less frequent ones result in discordant interpretations.

Headedness is tested in one category of NN compounds, i.e. compounds with a nominal modifier and a nominal head. The influence on the informants' interpretations of the headedness rule or schema is tested with nine NN compounds of different frequency, ranging from frequent to novel compounds. Definitions that follow the right-headedness rule are supplemented with left-headed distractors. The interpretations in this category show that the right-headed alternative is the preferred choice. But a considerable share of the interpretations does not follow this pattern (see Table 5:1, where the right-headed alternative is marked in bold). There are significant differences between the interpretations from different informant groups (this is marked with significance symbols in the table). The L1 speakers follow the right-headedness structure to a significantly greater degree than do the L2 speakers. The L2 speakers with an early age of onset (<6 years) follow this structure to a greater degree than the L2 speaker with a later age of onset (≥6 years). These differences are summarised in a boxplot (see Figure 5:1).

Headedness in adjectival AA compounds is an intricate issue in Swedish. Swedish AA compounds, for example colour terms such as *gulvit* (literally 'yellow white') and *svartvit* (lit. 'black white'), are traditionally considered to follow two different patterns. In determinative compounds such

as *gulvit*, the last constituent of the compound is considered the head of the compound: something that is *gulvit* is ‘white with a nuance of yellow’. This holds true for compounds where the colours are close to each other. Colour compounds with distant colours, such as *svartvit* are considered copulative, ‘black and white’. The headedness of AA compounds is tested with eight compounds, see Table 5:10. In this category of compounds, the informants could choose between three alternative definitions. The definitions instantiate either a copulative relation or a determinative relation (right-headed or left-headed). The most conventionalised alternative is represented in line 1 and the least conventionalised, left-headed structure is represented in line 3 in the table. The informants’ interpretations turned out to be discordant with the headedness rule. Not only are traditionally copulative compounds confused with determinative right-headed compounds, but a considerable proportion of the interpretations follow a left-headed structure. The multilingual speakers with Swedish as one of their first languages (SVE+) and the multilingual speakers with Swedish as their second language (START<6 and START≥6) have a tendency to choose the copulative alternative, also for compounds that are traditionally analysed as determinative. While the L1 speakers who report Swedish as their only first language (SVE) follow a more traditional distinction between copulative and determinative compounds, they have a tendency to choose the determinative but left-headed alternative, instead of the right-headed one.

The choice between thematic relation interpretations and property interpretations is tested in two categories of NN compounds; see Table 5:4 (with six compounds denoting artefacts and animals) and Table 5:5 (with six compounds denoting persons). The interpretations are affected by aspects such as frequency and degree of conventionalisation. Established and frequent compounds result in concordant interpretations. The interpretations from L1 speakers are more concordant than the interpretations from L2 speakers, especially L2 speakers with a later age of onset (cf. Table 5:4 and Table 5:5). However, the concordance in the interpretations is surprisingly high when the informants meet novel compounds, e.g. *zebralejon* ‘zebra lion’ and *korpörn* ‘raven eagle’, where the property interpretation (i.e. a comparing pattern) is the preferred choice. *Zebralejon* is interpreted as a ‘striped lion’ rather than a ‘lion that hunts zebras’ (see Table 5:4). The novel compound *flickpojke* (‘girl boy’) generates a large proportion of property interpretations. This suggests that it is interpreted using as an analogy base the established compound *pojkflicka* (lit. ‘boy girl’), i.e. ‘girl considered to be “boyish”’.

The preference for a comparing relationship between the components is observed in another category of compounds as well, namely in compounds with a nominal modifier and an adjectival head (i.e. NA compounds) such

as *iskall* (lit. ‘ice cold’) (see Table 5:11 with 14 NA compounds). Many Swedish compounds of this sort are conventionalised with a comparing pattern. But this comparing pattern is preferred over a causal pattern also for less conventionalised or novel compounds, e.g. *snömjuk* ‘snow soft’. Established compounds that differ from this frequent pattern generate discordant interpretations, especially in the L2 groups. This is true for *solvarm* (lit. ‘sun warm’), which is conventionalised with a causal relation ‘warmed by the sun’. In the L2 groups the distribution is fifty-fifty between the conventionalised definition and the definition that instantiates a comparing pattern.

The issue of compositionality is tested in one category of AN compounds, i.e. four compounds with an adjectival modifier and a nominal head, such as *finkläder* (lit. ‘nice clothes’) i.e. ‘formalwear’, see Table 5:9. This type of compounds are conventionalised in a more specific meaning than apparent from the components – *finkläder* is not identical to ‘nice clothes’ – i.e. they are non-compositional. The L2 speakers (especially the informants in START \geq 6) have a larger proportion of compositional, non-conventionalised interpretations than the L1 speakers (see Table 5:9 and the boxplot in Figure 5:4).

The interpretations in the multiple choice part are compared to the interpretations in the free text part. The overall result indicates that the results from the free text part follow the results of the multiple choice part. However, the L2 speakers follow conventionalised patterns in the multiple choice part (where they have access to fixed alternatives and a contextual sentence) to a larger extent than in the contextless free text part, while the differences between the parts are insignificant for the L1 speakers.

Discussion

The relational polysemy of compounds entails that the meaning of a compound is motivated, rather than predictable, from its components. Since compounds have wide meaning potentials, compositionality itself does not lead to analysability or transparency. For example, the conventionalised meaning of *solvarm* (lit. ‘sun warm’) ‘warmed by the sun’ is not apparent from its parts, i.e. this compound is not transparent. The L2 interpretations in the questionnaire illustrate that conventionalised compounds usually have alternative, compositional and reasonable meaning potentials, in the case of *solvarm* a comparing relation ‘warm as the sun’. Furthermore, AN compounds like *finkläder* (lit. ‘nice clothes’) are conventionalised in a more specific meaning than indicated by the parts, resulting in non-compositionality. Therefore, access to the conventionalised meaning of a

compound often requires that the compound is entrenched as a unit in the mind of the speaker (rather than analysed from its components). If the language user does not have access to a conventionalised meaning, he/she takes advantage of entrenched schemas. Whether or not a compound that is novel to the language user is perceived as transparent or analysable depends on which access he/she has to relevant analogy bases of different granularity that can help him/her to “see through” the compound.

The comparing pattern proved influential in the interpretation process. This pattern is preferred over a causal pattern in the interpretations of NA compounds. Likewise, a comparing property interpretation is favoured over a thematic relation in the interpretation of novel NN compounds denoting animals, artifacts or persons, especially if the referent of the modifier has a salient feature. This finding indicates that property interpretation is a productive pattern for interpretation (cf. Wisniewski & Love 1998, Ryder 1994) rather than a last-resort strategy (cf. Downing 1977, Gagné & Shoben 1997). The tendency to choose a comparing relation between the components of compounds is in line with cognitive research stating comparison as a fundamental, cognitive activity that is involved in all types of categorization (Langacker 1999:94, Schmid 2015:11).

This study nuances the picture of a straight-forward right-headedness rule or schema. The results indicate that the right-headedness schema is partly established for NN compounds. The L1 speakers follow this schema to a greater degree than the L2 informants, but the left-headedness alternatives are not blocked for any group of informants. Similar findings have been reported for English. Left-headed interpretations of novel English NN compounds have been reported in several informant studies (see Ryder 1994:137–138, Costello & Keane 2000:325, Onysko 2016:487). As Ryder (1994:138) points out, these recurrent findings of left-headed interpretations in languages which are characterised as “right-headedness languages” must not be dismissed as performance errors. Rather, they falsify the assumption of an absolute syntactical rule that blocks left-headed interpretations (cf. Williams 1981:248, Selkirk 1982).

The interpretations of AA compounds indicate that the right-headedness rule or schema is not as established for this type of compounds as previously assumed. In this study, the determinative AA compounds tend to be interpreted as left-headed (by the monolingual L1 speakers of Swedish) or copulative (by the multilingual L1 and L2 speakers of Swedish), rather than as instantiations of a right-headed pattern. These findings can be related to the very nature of AA compounds. They typically denote subjective and perceptual qualities. A language user seldom can check if his/her perception of *gulgrön* (lit. ‘yellow green’) is equal to another person’s, and if he/she interprets this colour term in the same way as the

interlocutor. The results falsify the assumption for Swedish that adjectival compounds with few exceptions follow a right-headedness pattern (SAG 2:186). Furthermore, the results from the free text part indicate that AA compounds are not easily inserted in a dichotomy separating copulative meanings from determinative meanings. Rather than being a matter of equal properties (i.e. a copulative meaning) or different-weighting properties (i.e. a determinative meaning), the free text interpretations illustrate that the properties of the modifier and head in AA compounds can be valid for different domains. For example *fulsnygg* (lit. ‘ugly good-looking’) is in many cases interpreted as a ‘good-looking person who has an ugly personality’, or vice versa, i.e. it is not obvious how to analyse the word in relation to headedness.

The overall result indicates that frequency is an influential factor in the interpretations. Frequent compounds result in more concordant interpretations than less frequent ones. From a cognitive perspective, this is the result of entrenchment of specific items as units. If the compound is entrenched in the mind, as a unit, the language user needs not analyse it, cf. “a whole is more salient than its parts” (Langacker 1991:171). Less frequent or novel compounds result in more concordant interpretations if they can be interpreted using a familiar compound as an analogy base, or if they follow an established schema. This result confirms findings from Ryder (1994:193), that informants are guided by linguistic templates, preferably lower-level ones.

The L1 speakers have more concordant interpretations than the L2 speakers. Regarding conventionalised compounds, this result can be explained as a difference in entrenchment of these specific compounds. Regarding novel compounds this result reflects L1–L2 differences in entrenchment of linguistic templates at different levels of abstraction (i.e. specific compounds that can serve as analogy bases as well as lower-level and higher-level schemas). L2 speakers with an early age of onset are more concordant in their interpretations than L2 speakers with a later age of onset. In the multilingual SVE+ group, which consists of informants reporting Swedish as one of their first languages, the informants who have had support for the Swedish language from one parent who has Swedish as their first language (i.e. SVE+/S) act more like the SVE group (Swedish as their only L1) than do the informants whose parents have other first languages (SVE+/A). In conclusion, these differences between informant groups are the result of differences in input and exposure time to the Swedish language, i.e. a result of entrenchment differences.

The differences in interpretation (and conventionality) between the L1 and L2 speakers are more significant in the free text part than in the multiple choice part. This indicates that L2 speakers – more than L1 speakers

Summary

– take advantage of the test design of the multiple choice part. The access to fixed alternatives and a contextual sentence means that they do not have to evaluate all meaning potentials that are activated in the free text part. They also get help to find an alternative that seems reasonable, one they would not come up with themselves. Summing up, this indicates that the differences in conventionality, observed in the multiple choice part, between the L1 and L2 speakers, rather than being overestimated, are underestimated.

It is noteworthy that the results of the study indicate that all informants, irrespective of background, tend to interpret novel compounds using established compounds or schemas as analogy bases. This is reasonable from a cognitive perspective: if a particular item is unfamiliar, the language user compares this item with entrenched compounds seeking an established pattern to follow. The differences between the informant groups discussed above are not the result of different strategies for interpretation between the groups, but rather the result of a varying degree of access to analogy bases and their semantic content.

Conclusion

The result of this study indicates that the meaning potential of a compound is evaluated only when necessary. If the compound is entrenched as a unit, no analysis is needed. If the compound is unfamiliar to the language user, he/she makes use of entrenched linguistic templates, preferably lower-level ones, as analogy bases in the interpretation.

This study further indicates that the right-headedness pattern is not as established for Swedish compounds as previously assumed. This is especially true for adjectival compounds. The structural and semantic aspects of adjectival compounds is an understudied area: an area that calls for further research.

Referenser

- Abrahamsson, Niclas 2009. *Andraspråksinläring*. Lund: Studentlitteratur.
- Abrahamsson, Niclas & Kenneth Hyltenstam 2013. Mognadsbegränsningar och den kritiska perioden för andraspråksinläring. I: Hyltenstam, Kenneth & Inger Lindberg (red.), 221–257.
- Ackesjö, Helena & Sven Persson 2010. Skolförberedelse i förskoleklass. Att vara *lärare-i-relation* i gränslandet. *Pedagogisk forskning i Sverige* 15:2–3, 142–163.
- Allwood, Jens 2003. Meaning potentials and context. Some consequences for the analysis of variation in meaning. I: Cuyckens, Hubert, René Dirven & John R. Taylor (red.), *Cognitive Approaches to Lexical Semantics*. (Cognitive Linguistics Research 23.) Berlin/New York: Mouton de Gruyter, 29–65.
- Aronoff, Mark 1976. *Word Formation in Generative Grammar*. Cambridge: MIT Press.
- Bakken, Kristin 1998. *Leksikalisering av sammensetninger. En studie av leksikaliseringsprosessen belyst ved et gammelnorsk diplommateriale fra 1300-tallet*. (Acta Humaniora nr 38.) Oslo: Scandinavian University Press.
- Barlow, Michael & Suzanne Kemmer (red.) 2000. *Usage-Based Models of Language*. Stanford: Stanford University.

- Bauer, Laurie 1983. *English word-formation*. Cambridge: Cambridge University Press.
- Bauer, Laurie 2009. Typology of compounds. I: Lieber, Rochelle & Pavol Štekauer (red.), 343–356.
- Benczes, Réka 2006. *Creative Compounding in English. The Semantics of Metaphorical and Metonymical Noun-Noun Combinations*. (Human Cognitive Processing 19.) Amsterdam/Philadelphia: John Benjamins.
- Bergman, Gösta 1962. *Rätt och fel i språket*. (Skrifter utg. av Nämnden för svensk språkvård 26.) Stockholm: Svenska Bokförlaget.
- Booij, Geert 2009. Compounding and construction morphology. I: Lieber, Rochelle & Pavol Štekauer 2009 (red.), 201–216.
- Booij, Geert 2010. Compound construction: Schemas or analogy? A construction morphology perspective. I: Scalise, Sergio & Irene Vogel (red.), 93–107.
- Booij, Geert 2013. Morphology in Construction Grammar. I: Hoffmann, Thomas & Graeme Trousdale (red.), 255–273.
- Borin, Lars, Markus Forsberg & Johan Roxendal 2012. Korp – the corpus infrastructure of Språkbanken. I: Calzolari, Nicoletta, Khalid Choukri, Thierry Declerck, Mehmet Uğur Doğan, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odijk & Stelios Piperidis (red.), *Proceedings of LREC 2012, Istanbul, Turkey, 23–25 May*. Istanbul: ELRA, 474–478.
- Bybee, Joan L. 2010. *Language, Usage and Cognition*. Cambridge: Cambridge University Press.
- Bybee, Joan L. 2013. Usage-Based Theory and Exemplar Representations of Constructions. I: Hoffmann, Thomas & Graeme Trousdale (red.), 49–69.
- Clark, Eve V., Susan Gelman & Nancy M. Lane 1985. Compound Nouns and Category Structure in Young Children. *Child Development* 56:1, 84–94.
- Costello, Fintan J. & Mark T. Keane 2000. Efficient Creativity: Constraint-Guided Conceptual Combination. *Cognitive Science* 24:2, 299–349.
- Croft, William 2001. *Radical Construction Grammar. Syntactic Theory in Typological Perspective*. Oxford/New York: Oxford University Press.
- Cummins, Jim 2008. BICS and CALP. Empirical and Theoretical Status of the Distinction. I: Street, Brian & Nancy H. Hornberger (red.), *Encyclopedia of Language and Education 2: Literacy*. 2 uppl. New York: Springer, 71–83.

- Dabrowska, Ewa 2010. The mean lean grammar machine meets the human mind. Empirical investigations of the mental status of linguistic rules. I: Schmid, Hans-Jörg & Susanne Handl (red.), *Cognitive foundations of linguistic usage patterns*. New York: Mouton de Gruyter, 151–170.
- Downing, Pamela A. 1977. On the Creation and Use of English Compound Nouns. *Language* 53:4, 810–842.
- Döpke, Susanne 1998. Competing Language Structures: The Acquisition of Verb Placement by Bilingual German-English Children. *Journal of Child Language* 25:3, 555–584.
- Eiesland, Eli Anne 2016. *The Semantics of Norwegian Noun-Noun Compounds. A corpus-based study*. Oslo: University of Oslo.
- Ekberg, Lena 2013. Grammatik och lexikon i svenska som andraspråk på nästan infödd nivå. I: Hyltenstam, Kenneth & Inger Lindberg (red.), 259–279.
- Ellis, Nick C. 2006. Cognitive Perspective on SLA. The Associative-Cognitive CREED. *AILA Review* 19, 100–121.
- Enström, Ingegerd 2013. Ordförråd och ordinläring – med särskilt fokus på avancerade inlärare. I: Hyltenstam, Kenneth & Inger Lindberg (red.), 169–195.
- Enström, Ingegerd 2016. *Ordens värld. Svenska ord – struktur och inläring*. 2 uppl. Uppsala: Hallgren & Fallgren.
- Fillmore, Charles J. 1982. Frame Semantics. I: *Linguistics in the Morning Calm* (ed. by the Linguistic Society of Korea). Seoul: Hanshin, 111–137.
- Fillmore, Charles J., Paul Kay & Mary Catherine O'Connor 1988. Regularity and Idiomaticity in Grammatical Constructions. The Case of *Let Alone*. *Language* 64:3, 501–538.
- Foroodi-Nejad, Farzaneh & Johanne Paradis 2009. Crosslinguistic Transfer in the Acquisition of Compound Words in Persian-English Bilinguals. *Bilingualism: Language and Cognition* 12:4, 411–427.
- Forza, Francesca, Emiliano Guevara & Sergio Scalise 2009. Compound-ing Adjectives. *Lingue E Linguaggio* 8:1, 137–162.
- Fraurud, Kari & Sally Boyd 2011. The native–non-native speaker distinction and the diversity of linguistic profiles of young people in multilingual urban contexts in Sweden. I: Källström, Roger & Inger Lindberg (red.), *Young Urban Swedish. Variations and change in multilingual settings*. Göteborg: Göteborgs universitet, 67–87.
- Gagné, Christina L. 2002. Lexical and Relational Influences on the Processing of Novel Compounds. *Brain and Language* 81:1–3, 723–735.

- Gagné, Christina L. & Edward J. Shoben 1997. Influence of Thematic Relations on the Comprehension of Modifier-Noun Combinations. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 23:1, 71–87.
- Gagné, Christina L. & Thomas L. Spalding 2007. Semantic property activation during the interpretation of combined concepts. *The Mental Lexicon* 2:1, 25–47.
- Gagné, Christina L. & Thomas L. Spalding 2010. Relational competition during compound interpretation. I: Scalise, Sergio & Irene Vogel (red.), 287–300.
- García, Ofelia & Wei Li 2018. *Translanguaging: Flerspråkighet som resurs i lärandet*. Stockholm: Natur & Kultur.
- Geeraerts, Dirk & Hubert Cuyckens (red.) 2007. *The Oxford Handbook of Cognitive Linguistics*. Oxford/New York: Oxford University Press.
- Gleitman, Lila R. & Henry Gleitman 1970. *Phrase and Paraphrase: Some Innovative Uses of Language*. New York: W.W. Norton & Company, Inc.
- Goldberg, Adele E. 2006. *Constructions at Work: The Nature of Generalization in Language*. Oxford/New York: Oxford University Press.
- Goldberg, Adele E. 2013. Constructionist Approaches. I: Hoffmann, Thomas & Graeme Trousdale (red.), 15–31.
- Golden, Anne 2006. Minoritetslever og ordforrådet i läreböcker. I: Lindberg, Inger & Karin Sandwall (red.), *Språket och kunskapen – att lära på sitt andraspråk* (ROSA. Rapporter om svenska som andraspråk 7.) Göteborg: Göteborgs universitet, 115–127.
- Golden, Anne & Anne Hvenekilde 1983. *Rapport fra Prosjektet Lærebokspråk*. Oslo: Universitetet i Oslo.
- Grosjean, Francois 2001. The bilingual's language modes. I: Nicol, Janet (red.), *One Mind, Two Languages. Bilingual Language Processing*. Oxford: Blackwell, 1–23.
- Hanks, Patrick 2000. Do Word Meanings Exist? *Computers and the Humanities* 34:1–2, 205–215.
- Hannessdóttir, Anna Helga 2012. Ekvivalensrelationer i tvåspråkig lexikografi. *LexicoNordica* 19, 39–58.
- Hene, Birgitta 1993. *Utlandsadopterade barns och svenska barns ordförståelse. En jämförelse mellan barn i åldern 10–12 år.* (SPRINS-projektet 41.) Göteborg: Göteborgs universitet.
- Henriksen, Birgit 1999. Three Dimensions of Vocabulary Development. *Studies in Second Language Acquisition* 21:2, 303–317.
- Heyvaert, Lisbeth 2009. Compounding in Cognitive Linguistics. I: Lieber, Rochelle & Pavol Štekauer 2009 (red.), 233–254.

- Hoffmann, Thomas & Graeme Trousdale (red.) 2013. *The Oxford Handbook of Construction Grammar*. Oxford/New York: Oxford University Press.
- Holmegaard, Margareta 2007. Långa ord – en svårighet för flerspråkiga studerande? I: Lindberg, Inger & Sofie Johansson Kokkinakis (red.), *OrdiL: En korpusbaserad kartläggning av ordförrådet i läromedel för grundskolans senare år* (ROSA. Rapporter om svenska som andraspråk 8.) Göteborg: Göteborgs universitet.
- Hyltenstam, Kenneth & Inger Lindberg (red.) 2013. *Svenska som andraspråk: i forskning, undervisning och samhälle*. 2 uppl. Lund: Studentlitteratur.
- Håkansson, Gisela 2003. *Tvåspråkighet hos barn i Sverige*. Lund: Studentlitteratur.
- Josefsson, Gunlög 1997. *On the Principles of Word Formation in Swedish*. (Lundastudier i nordisk språkvetenskap 51.) Lund: Lund Univ. Press/Chartwell-Bratt.
- Josefsson, Gunlög 2005. *Ord*. Lund: Studentlitteratur.
- Karlsson, Ola (red.) 2017. *Svenska skrivregler*. 4 uppl. (Språkrådets skrifter 22.) Stockholm: Liber AB.
- Kay, Paul, & Karl Zimmer 1990. On the Semantics of Compounds and Genitives in English. I: Tsohatzidis, Savas L. (red.), *Meanings and Prototypes: Studies in linguistic categorization*. London: Routledge, 239–246.
- Kemmer, Suzanne & Michael Barlow 2000. Introduction: A Usage-Based Conception of Language. I: Barlow, Michael & Suzanne Kemmer (red.), vii–xxii.
- Krott, Andrea 2009. The role of analogy for compound words. I: Blevins, James P. & Juliette Blevins (red.), *Analogy in Grammar: Form and Acquisition*. Oxford: Oxford University Press.
- Krott, Andrea, Christina L. Gagné & Elena Nicoladis 2009. How the Parts Relate to the Whole: Frequency Effects on Children's Interpretations of Novel Compounds. *Journal of Child Language* 36:1, 85–112.
- Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar. Vol. I: Theoretical Prerequisites*. Stanford: Stanford University Press.
- Langacker, Ronald W. 1991. *Foundations of Cognitive Grammar. Vol II: Descriptive application*. Stanford: Stanford University Press.
- Langacker, Ronald W. 1999. *Grammar and Conceptualization*. (Cognitive Linguistics Research 14.) Berlin/New York: Mouton de Gruyter.
- Langacker, Ronald W. 2000. A Dynamic Usage-Based Model. I: Barlow, Michael & Suzanne Kemmer (red.), 1–63.
- Langacker, Ronald W. 2008. *Cognitive grammar: A Basic Introduction*. Oxford/New York: Oxford University Press.

- Langacker, Ronald W. 2009. A dynamic view of usage and language acquisition. *Cognitive Linguistics* 20:3, 627–640.
- Langacker, Ronald W. 2017. Entrenchment in Cognitive Grammar. I: Schmid, Hans-Jörg (red.), 39–56.
- Lees, Robert 1960. *The Grammar of English Nominalizations*. (Indiana University Research Center in Anthropology, Folklore and Linguistics 12.) Bloomington.
- Levi, Judith 1978. *The Syntax and Semantics of Complex Nominals*. New York: Academic Press.
- Li, Charles 1971. *Semantics and the structure of compounds in Chinese*. Berkeley: University of California.
- Lieber, Rochelle & Pavol Štekauer (red.) 2009. *The Oxford Handbook of Compounding*. Oxford: Oxford University Press.
- Liljestrand, Birger 1993. *Så bildas orden. Handbok i ordbildning. 2.* [omarb.] uppl. Lund: Studentlitteratur.
- Lindberg, Inger & Sofie Johansson Kokkinakis (red.) 2007. *OrdiL. En korpusbaserad kartläggning av ordförrådet i läromedel för grundskolans senare år* (ROSA. Rapporter om svenska som andraspråk 8.) Göteborg: Göteborgs universitet.
- Loenheim, Lisa 2016. Korpusundersökning om NA-sammansättningar. Institutionen för svenska språket, Göteborgs universitet. [Opublicerad kursuppsats].
- Lyngfelt, Benjamin, Lars Borin, Linnéa Bäckström, Markus Forsberg, Leif-Jöran Olsson, Julia Prentice, Rudolf Rydstedt, Emma Sköldberg, Sofia Tingsell & Jonatan Uppström 2014. Ett svenskt konstruktikon. Grammatik möter lexikon. *Svenskans beskrivning* 33. Helsingfors, 268–279.
- Lyngfelt, Benjamin 2018. Introduction. Constructions and constructicography. I: Lyngfelt, Benjamin, Lars Borin, Kyoko Ohara, Tiago Timponi Torrent (red.), *Constructicography: Construction Development across Languages*. Amsterdam: John Benjamins, 1–18.
- Maguire, Phil, Rebecca Maguire, Arthur W. S. Cater & Randi C. Martin 2010a. The Influence of Interactional Semantic Patterns on the Interpretation of Noun–Noun Compounds. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 36:2, 288–297.
- Maguire, Phil, Edward J. Wisniewski & Gert Storms 2010b. A Corpus Study of Semantic Patterns in Compounding. *Corpus Linguistics and Linguistic Theory* 6:1, 49–73.
- Malmgren, Sven-Göran 1994. *Svensk lexikologi. Ord, ordbildning, ordböcker och orddatabaser*. Lund: Studentlitteratur.
- McLaughlin, Barry 1978. *Second-Language Acquisition in Childhood*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

- Mellenius, Ingmarie 1997. *The Acquisition of Nominal Compounding in Swedish*. (Travaux de L'institut de linguistique de Lund 31.) Lund/Bromley: Lund University Press/Chartwell-Bratt.
- Murphy, Gregory L. 1988. Comprehending Complex Concepts. *Cognitive Science* 12, 529–562.
- Müller, Natascha 1998. Transfer in bilingual first language acquisition. *Bilingualism: Language and Cognition* 1, 151–171.
- Nesset, Tore 2016. Spøkelsesfiske, makrellfotball og traktoregg: norske sammensetninger og konseptuell integrasjon. *Maal og minne* 2, 85–110.
- Nicoladis, Elena. 2002. What's the difference between 'toilet paper' and 'paper toilet'? French-English bilingual children's crosslinguistic transfer in compound nouns. *Journal of Child Language* 29:4, 843–863.
- Nicoladis, Elena. 2003. What compound nouns mean to preschool children. *Brain and Language* 84:1, 38–49.
- Niday, Sara Marie 2017. *Shit-cool and Giant-interesting Words. A study of noun+adjective compounds in Norwegian*. [Master thesis.] Universitetet i Oslo: Department of Linguistics and Scandinavian Studies.
- Noreen, Adolf 1906. *Vårt språk. Nysvensk grammatik i utförlig framställning*. Fjärde delen: Formlära (Morfologi). Sjunde bandet. Lund: Gleerups förlag.
- Norén, Kerstin & Per Linell 2006. *Meningspotentialer i den språkliga praktiken*. (MISS: Meddelanden från Institutionen för svenska språket 54.) Göteborg: Göteborgs universitet.
- Onysko, Alexander 2016. Crosslinguistic influence on headedness of novel English compounds: Evidence from bilingual speakers of te reo Māori and English. *Language Typology and Universals = Sprachtypologie und Universalienforschung*, 69:4, 471–494.
- Ortega, Lourdes 2013. *Understanding Second Language Acquisition*. London/New York: Routledge.
- Parault, Susan J., Paula J. Schwanenflugel & Heather R. Haverback 2005. The Development of Interpretations for Novel Noun-Noun Conceptual Combinations During the Early School Years. *Journal of Experimental Child Psychology* 91:1, 67–87.
- Pavlenko, Aneta 2009. *The Bilingual Mental Lexicon: Interdisciplinary approaches* (Bilingual education and bilingualism 70.) Bristol: Multilingual Matters.
- Plag, Ingo 2003. *Word-Formation in English*. Cambridge: Cambridge University Press.
- Prentice, Julia, & Emma Sköldberg 2010. Klättra på väggarna eller bara vara ett med soffan? Om figurativa ordförbindelser bland ungdomar i flerspråkiga miljöer. *Språk och stil* 20, 5–35.

- Ribeck, Judy 2015. *Steg för steg. Naturvetenskapligt ämnesspråk som räknas.* (Data linguistica 28.) Göteborg: Göteborgs universitet.
- Rogström, Lena 1998. *Jacob Serenius lexikografiska insats.* (Meijerbergs arkiv för svensk ordforskning 22.) Göteborg: Göteborgs universitet.
- Rudebeck, Lisa 2006. Ordförklaringar som betydelserepresentationer eller yttranden. *Språk och stil* 16, 112–147.
- Ryder, Mary Ellen 1994. *Ordered Chaos: The Interpretation of English Noun-Noun Compounds.* (University of California Publications in Linguistics 123.) Berkeley: University of California Press.
- SAG = Teleman, Ulf, Staffan Hellberg & Erik Andersson 1999. *Svenska Akademiens grammatik.* Stockholm: Norstedts ordbok.
- Scalise, Sergio & Irene Vogel (red.) 2010. *Cross-Disciplinary Issues in Compounding.* (Current Issues in Linguistic Theory 311.) Amsterdam: John Benjamins.
- Schmid, Hans-Jörg 2007. Entrenchment, Saliency, and Basic Levels. I: Geeraerts, Dirk & Hubert Cuyckens (red.), 117–138.
- Schmid, Hans-Jörg 2015. A blueprint of the Entrenchment-and-Conventionationalization Model. *Yearbook of the German Cognitive Linguistics Association* 3:1, 3–25.
- Schmid, Hans-Jörg 2017. A Framework for Understanding Linguistic Entrenchment and its Psychological Foundations. I: Schmid, Hans-Jörg (red.), 9–35.
- Schmid, Hans-Jörg (red.) 2017. *Entrenchment and the Psychology of Language Learning.* Washington D.C.: De Gruyter Mouton.
- Selkirk, Elisabeth O. 1982. *The Syntax of Words.* Cambridge: MIT Press.
- Simonsen, Hanne Gram, Rolf Theil Endresen & Even Hovdhaugen 1988. (1 utg. 1978). *Språkvitenskap. En elementær innføring.* Oslo: Universitetsforlaget.
- Skolverket 2018. <<https://www.skolverket.se/regler-och-ansvar/ansvar-i-skolfragar/obligatorisk-forskoleklass>> Hämtat november 2018.
- Skolverket 2016. Siris elevstatistik. *Gymnasieskolan: avgångselever, nationellt program.* Läsåret 2015/16. Tillgänglig via: <https://siris.skolverket.se/reports/rwservlet?cmdkey=common&geo=1&report=gy_betyg&p_flik=G&p_programkod=&p_ar=2016&p_lan-kod=&p_kommunkod=&p_hmantyp=&p_skolkod=&p_hman-kod=>>
- Skolverket 2018. <<https://www.skolverket.se/regler-och-ansvar/ansvar-i-skolfragar/obligatorisk-forskoleklass>> Hämtat november 2018.
- Skutnabb-Kangas, Tove 1981. *Tvåspråkighet.* Lund: LiberLäromedel.
- Språkbanken/Korp, <spraakbanken.gu.se/korp>.
- Svanlund, Jan 2002. Lexikalisering. *Språk och stil* 12, 7–45.

- 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.
- Svanlund, Jan 2013. Har sammansättningar en grammatik? [Föredrag på grammatikdagen 2013-03-25]. < http://grammatikdagen.se/wp-content/uploads/2013/2012/11/svanlund_grammatikdagen_2013.pdf>.
- Svensén, Bo 2004. *Handbok i lexikografi. Ordböcker och ordboksarbete i teori och praktik.* 2 uppl. Stockholm: Norstedts Akademiska Förlag.
- Svensk ordbok 2009 = *Svensk ordbok utgiven av Svenska Akademien.* Stockholm: Norstedts [distributör].
- Svensson, Gudrun 2017. *Transspråkande i praktik och teori.* Stockholm: Natur & Kultur.
- Søgaard, Anders 2005. Compounding theories and linguistic diversity. I: Frajzyngier, Zygmunt, Adam Hodges & David S. Rood (red.), *Linguistic Diversity and Language Theories.* (Studies in Language Companion Series 72.) Amsterdam: John Benjamins, 319–337.
- Teleman, Ulf 1970. *Om svenska ord.* Lund: Gleerups.
- Teleman, Ulf, Staffan Hellberg & Erik Andersson 1999. *Svenska Akademiens grammatik.* Stockholm: Norstedts Ordbok.
- Thomas, Wayne & Virginia P. Collier 2002. *A National Study of School Effectiveness for Language Minority Students' Long-Term Academic Achievement.* UC Berkely: Center for Research on Education, Diversity & Excellence (CREDE). Tillgänglig via: <<https://escholarship.org/uc/item/65j213pt>>
- Thorell, Olof 1981. *Svensk ordbildningslära.* Solna: Esselte Studium.
- Tomasello, Michael 2003. *Constructing a language: A usage-based theory of language acquisition.* Cambridge, Mass.: Harvard Univ. Press.
- Tuggy, David 2007. Schematicity. I: Geeraerts, Dirk & Hubert Cuyckens. (red.), 82–116.
- Wray, Alison 2002. *Formulaic Language and the Lexicon.* New York: Cambridge University Press.
- Wray, Alison 2008. *Formulaic Language: Pushing the Boundaries.* Oxford: Oxford University Press.
- Yoxsimer Paulsrud, BethAnne, Jenny Rosén, Boglárka Straszer & Åsa Wedin 2018. *Transspråkande i svenska utbildningssammanhang.* Lund: Studentlitteratur.
- Warren, Beatrice 1978. *Semantic patterns of noun-noun compounds.* (Gothenburg studies in English 41.) Göteborg: Acta Universitatis Gothoburgensis.
- Williams, Edvin 1981. On the Notions "Lexically Related" and "Head of a Word". *Linguistic Inquiry* 12, 245–274.

Referenser

Wisniewski, Edward J. & Bradley Love 1998. Relations versus Properties in Conceptual Combination. *Journal of Memory and Language* 38:2, 177–202.

Bilaga 1: Elevenkät

I denna bilaga återfinns enkätversion BC, en av de fyra versioner som distribuerades i klasserna. De fyra versionerna av enkäten innehåller samma ord, men sammansättningarna förekommer i olika, framslumpad ordning (se vidare avsnitt 4.8).

BC

Samtycke till deltagande: Denna enkät ingår som en del i en studie om vilka strategier vi använder när vi tolkar ord. Det är **frivilligt** att delta i studien. Genom att lämna in enkäten godkänner du, och samtycker till, att ditt enkätsvar används i forskningssyfte. Du samtycker till att forskningsresultaten används för publicering (i t.ex. min avhandling och i tidskrifter) och för presentation i olika sammanhang (konferenser och presentationer riktade till allmänheten). Ditt enkätsvar är **anonymt**. Du är välkommen att ställa frågor om studien innan, under och efter det att du svarat på enkäten. Kontaktuppgifter lämnas i samband med introduktionen. / Lisa Loenheim (lisa.loenheim@svenska.gu.se)

DEL 1: Skriv förklaringar till orden

Förklara följande ord i en eller ett par meningar. Skriv vad du tror att ordet betyder även om du inte hört det förut, eller om du är osäker på betydelsen.

En barnbokhylla är _____

En person som är fulsnugg är _____

Något som är honungssött är _____

Ett grönområde är _____

Issaft är _____

En glaslåda är _____

Ett vagnbarn är _____

En husbil är _____

Något som är dammgrått är _____

En kvinnoläkare är _____

En kvinnoman är _____

En person som är dumsnäll är _____

En blomkruka är _____

Politikerförakt är när _____

Något som är sandmjukt är _____

En person som är musrädd är _____

Barnmisshandel är när _____

Skalpotatis är _____

Något som är blodrött är _____

En zebraväska är _____

BC

Samtycke till deltagande: Denna enkät ingår som en del i en studie om vilka strategier vi använder när vi tolkar ord. Det är **frivilligt** att delta i studien. Genom att lämna in enkäten godkänner du, och samtycker till, att ditt enkätsvar används i forskningssyfte. Du samtycker till att forskningsresultaten används för publicering (i t.ex. min avhandling och i tidskrifter) och för presentation i olika sammanhang (konferenser och presentationer riktade till allmänheten). Ditt enkätsvar är **anonymt**. Du är välkommen att ställa frågor om studien innan, under och efter det att du svarat på enkäten. Kontaktuppgifter lämnas i samband med introduktionen. / Lisa Loenheim (lisa.loenheim@svenska.gu.se)

DEL 2: Välj ut och kryssa för en förklaring till orden

Läs ordet och förklaringarna till varje ord. Kryssa sedan i rutan för den förklaring som **du** tycker **stämmer bäst**. Obs! Kryssa bara i **en** ruta för varje ord. Svara på orden i tur och ordning och hoppa inte över något ord.

Citronvatten	<input type="checkbox"/> Vatten med citron i T.ex. "Jag beställde citronvatten till maten på restaurangen." <input type="checkbox"/> Vatten som smakar surt som citron T.ex. "Usch, det kom citronvatten ur kranen."
Snövit	<input type="checkbox"/> Vit av snö T.ex. "Jackan blev snövit när hon gick ut i snöfallet." <input type="checkbox"/> Vit som snö T.ex. "Hon ville ha en snövit jacka."
Småkakor	<input type="checkbox"/> Kakor som är små T.ex. "Han bakade två småkakor: en liten sockerkaka och en liten kladdkaka." <input type="checkbox"/> Kakor som är små och hårda T.ex. "Han bakade småkakor nästan varje vecka, t.ex. havrekakor, pepparkakor och kolakakor."
Herrfrisör	<input type="checkbox"/> Frisör som bara klipper män T.ex. "Hon kunde inte gå till frisören mittemot, för det var en herrfrisör." <input type="checkbox"/> Frisör som är man T.ex. "Det arbetade både män och kvinnor på frisersalongen. Hon valde att bli klippt av en herrfrisör."
Polismisshandel	<input type="checkbox"/> När polisen slår eller misshandlar en person T.ex. "När poliserna gripit mannen slog och sparkade de honom. De blev senare åtalade för polismisshandel." <input type="checkbox"/> När polisen blir misshandlad T.ex. "Killarna slog och sparkade polismannen. De blev senare åtalade för polismisshandel."

<p>Kvinnomisshandel</p>	<ul style="list-style-type: none"> <input type="checkbox"/> När kvinnor blir misshandlade T.ex. "Mannen hade slagit kvinnan flera gånger och blev dömd för kvinnomisshandel." <input type="checkbox"/> När kvinnor misshandlar T.ex. "Kvinnan hade slagit mannen flera gånger och blev dömd för kvinnomisshandel."
<p>Stövelbarn</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Stövel som passar ett barn T.ex. "Den här stöveln finns inte i vuxenstorlek utan bara som stövelbarn." <input type="checkbox"/> Barn som gillar att ha stövlar T.ex. "Hon ville alltid ha stövlar på sig. Hon var ett riktigt stövelbarn."
<p>Flodstrand</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Strand vid floden T.ex. "Vi stod på flodstranden och tittade ut över vattnet." <input type="checkbox"/> Flod vid stranden T.ex. "Akta dig så att du inte ramlar i flodstranden."
<p>Husbåt</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Hus där man kan lägga båten T.ex. "Innan vintern tar vi upp båten och lägger den i husbåten." <input type="checkbox"/> Båt som man kan bo i T.ex. "På semestern ska vi göra en resa med husbåten."
<p>Kryddstark</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Mycket stark (som kryddor) T.ex. "Rummet var målat i en kryddstark färg." <input type="checkbox"/> Stark av kryddor T.ex. "Till köttet serverades en kryddstark sås."
<p>Flickpojke</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Pojke som har intressen som anses vara "flickiga" T.ex. "Han tyckte om att leka med dockor och blev ibland kallad för flickpojke." <input type="checkbox"/> Pojke som gillar flickor T.ex. "Han lekte helst med flickor och blev ibland kallad för flickpojke."
<p>Skrattfärdig</p>	<ul style="list-style-type: none"> <input type="checkbox"/> När man tycker något är så roligt att man nästan börjar skratta T.ex. "Hon såg så rolig ut att han blev skrattfärdig." <input type="checkbox"/> När man skrattat färdigt T.ex. "Hon hade skrattat så mycket att hon fått ont i magen, men äntligen var hon skrattfärdig."
<p>Burköl</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Öl i burk T.ex. "Jag handlar oftare burköl än flasköl." <input type="checkbox"/> Burk för öl T.ex. "Släng inte burkölen i naturen när du druckit upp!"

Bilaga 1: Elevenkät

<p>Spisvarm</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Mycket varm (som spisen) T.ex. "Bränn dig inte – det är spisvarmt!" <input type="checkbox"/> Varm (eller uppvärmd) av spisen T.ex. "Maten är spisvarm och klar att ätas."
<p>Skrikfärdig</p>	<ul style="list-style-type: none"> <input type="checkbox"/> När man är så arg och upprörd att man nästan börjar skrika T.ex. "Hon blev så arg att hon kände sig skrikfärdig." <input type="checkbox"/> När man skrikit färdigt T.ex. "Hon hade blivit så arg att hon skrek rakt ut, men nu var hon skrikfärdig."
<p>Berghård</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Hård av berg T.ex. "Det är bara berg som kan göra marken berghård." <input type="checkbox"/> Väldigt hård (som berg) T.ex. "Jag sov dåligt för madrassen som jag låg på var berghård."
<p>Snömjuk</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Mycket mjuk (som snö) T.ex. "Kudden var härligt snömjuk." <input type="checkbox"/> Mjuk av snö T.ex. "Det hade snöat och hon lade sig ner på den härligt snömjuka gräsmattan."
<p>Gråtfärdig</p>	<ul style="list-style-type: none"> <input type="checkbox"/> När man gråtit färdigt T.ex. "Efter flera timmars gråt var han nu gråtfärdig." <input type="checkbox"/> När man är så ledsen att man nästan börjar gråta T.ex. "Barnens elaka ord gjorde honom alldeles gråtfärdig."
<p>Hundrädd</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Rädd för hundar T.ex. "Han var hundrädd och tyckte inte om när hundar kom för nära." <input type="checkbox"/> Rädd som en hund T.ex. "Han kände sig hundrädd när någon skällde på honom."
<p>Sötsur</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Både söt och sur, men mest söt T.ex. "Han ville ha en sås som var mer söt än sur – alltså sötsur." <input type="checkbox"/> Både söt och sur, men mest sur T.ex. "Han ville ha en sås som var mer sur än söt – alltså sötsur." <input type="checkbox"/> Lika söt som sur (eller lika sur som söt) T.ex. "Han ville ha en sås som var lika delar söt och sur – alltså sötsur."

Apörn	<ul style="list-style-type: none"> <input type="checkbox"/> Örn som låter eller tjartrar som en apa T.ex. "Det är lätt att känna igen apörnens märkliga låte." <input type="checkbox"/> Örn som jagar apor T.ex. "När apörnen visar sig flyr alla apor undan."
Olivkorg	<ul style="list-style-type: none"> <input type="checkbox"/> Korg som är flätad av kvistar från olivträd T.ex. "Han hittade en flätad korg av olivkvist på marknaden – alltså en olivkorg." <input type="checkbox"/> Korg som det ligger oliver i T.ex. "Det stod en korg på bordet som det låg oliver i – alltså en olivkorg." <input type="checkbox"/> Speciell korg som man använder när man plockar oliver T.ex. "Hämta olivkorgen, så går vi ut och plockar oliver."
Svartvit	<ul style="list-style-type: none"> <input type="checkbox"/> Både svart och vit T.ex. "Katten var så söt med sin långa svartvita päls." <input type="checkbox"/> Vit färg som samtidigt har en lite svart nyans T.ex. "Han ville ha en helt och hållet vit färg – den fick inte vara svartvit." <input type="checkbox"/> Svart färg som samtidigt har en lite vit nyans T.ex. "Han ville ha en helt och hållet svart färg – den fick inte vara svartvit."
Skräpkruka	<ul style="list-style-type: none"> <input type="checkbox"/> Kruka med skräp i T.ex. "Jag slängde bananskalet i skräpkrukan." <input type="checkbox"/> (Blom)kruka som inte är så fin utan ser lite skräpig ut T.ex. "Jag tror att jag ska slänga den här gamla skräpkrukan."
Gulgrön	<ul style="list-style-type: none"> <input type="checkbox"/> Grön färg som samtidigt har en lite gul nyans eller färgton T.ex. "Limefrukten var rejält mogen. Den gröna färgen hade fått en svag nyans av gult." <input type="checkbox"/> Både grön och gul T.ex. "Mössan hade gröna och gula ränder – den var alltså gröngul." <input type="checkbox"/> Gul färg som samtidigt har en lite grön nyans eller färgton T.ex. "Bananen var inte riktigt mogen. Den gula färgen hade fortfarande en svag nyans av grönt."
Solgul	<ul style="list-style-type: none"> <input type="checkbox"/> Gul som solen T.ex. "Citronerna lyste alldeles solgula i träden." <input type="checkbox"/> Gul av solen T.ex. "Papperen hade legat i solen och blivit solgula."

Bilaga 1: Elevenkät

<p>Stenhård</p>	<p><input type="checkbox"/> Väldigt hård (som sten) T.ex. "Jag sov dåligt för madrassen som jag låg på var stenhård."</p> <p><input type="checkbox"/> Hård av sten T.ex. "Det är bara sten som kan göra marken stenhård."</p>
<p>Helgfri</p>	<p><input type="checkbox"/> När man är fri på helgen T.ex. "Äntligen fredag! Vad härligt att snart vara helgfri!"</p> <p><input type="checkbox"/> När det inte är helg (= fri från helg) T.ex. "Affären är öppen helgfria dagar mellan 10 och 18."</p>
<p>Barnbokshylla</p>	<p><input type="checkbox"/> Hylla med barnböcker, t.ex. på ett bibliotek T.ex. "När man vill låna en barnbok på biblioteket får man gå till barnbokshyllan."</p> <p><input type="checkbox"/> Liten hylla som passar i ett barnrum T.ex. "Hon åkte till en möbelaaffär och köpte en barnbokshylla till dotterns rum."</p>
<p>Gulvit</p>	<p><input type="checkbox"/> Vit färg med en lite gul nyans eller färgton T.ex. "Hans tänder hade blivit gulvita – de var alltså fortfarande mest vita men hade en lite gul färgton."</p> <p><input type="checkbox"/> Både gul och vit T.ex. "Mössan hade gula och vita ränder – den var alltså gulvit."</p> <p><input type="checkbox"/> Gul färg med en lite vit nyans eller färgton T.ex. "Hans tänder såg smutsigt gulvita ut – de var alltså mer gula än vita."</p>
<p>Vagnbarn</p>	<p><input type="checkbox"/> Barn som sitter i vagn T.ex. "Barnen är fortfarande små. De är vagnbarn båda två."</p> <p><input type="checkbox"/> Vagn som man har det lilla barnet i T.ex. "Det är dyrt att köpa vagnbarn."</p>
<p>Lökstark</p>	<p><input type="checkbox"/> Stark av lök T.ex. "Han lagade en härligt lökstark soppa."</p> <p><input type="checkbox"/> Mycket stark (som lök) T.ex. "Ljuset är så starkt – alldeles lökstarkt!"</p>
<p>Manspräst</p>	<p><input type="checkbox"/> Präst som är man T.ex. "En del personer tycker att vi har för många manspräster i församlingen."</p> <p><input type="checkbox"/> Präst som tar emot och hjälper män T.ex. "Män som har det svårt kan få hjälp hos en manspräst."</p>

Solvarm	<input type="checkbox"/> Mycket varm (som solen) <i>T.ex. "Akta dig för spisen! Den är solvarm!"</i> <input type="checkbox"/> Varm (eller uppvärmd) av solen <i>T.ex. "Solen hade lyst hela dagen och gjort klipporna solvarma."</i>
Studentförakt	<input type="checkbox"/> Förakt som studenter visar <i>T.ex. "Studenter som ser ner på människor med lägre utbildning visar studentförakt."</i> <input type="checkbox"/> Förakt som är riktat mot studenter <i>T.ex. "Människor som ser ner på studenter i allmänhet visar studentförakt."</i>
Grönyta	<input type="checkbox"/> Yta som är grön <i>T.ex. "Hon målade väggarna gröna, för hon älskade grönytor."</i> <input type="checkbox"/> Naturområde med stor gräsplan <i>T.ex. "Hon uppsökte ofta stadens grönytor, för hon älskade att promenera."</i>
Kvinnokarl	<input type="checkbox"/> Kvinnlig man <i>T.ex. "Han hade traditionellt kvinnliga intressen, som att sy och sticka. En del tyckte därför att han var en kvinnokarl."</i> <input type="checkbox"/> Man som är förtjust i och har många kärleksrelationer med kvinnor <i>T.ex. "Han flörtade hela tiden med nya kvinnor. Han var en riktig kvinnokarl!"</i>
Blågul	<input type="checkbox"/> Blå färg som samtidigt har en lite gul nyans <i>T.ex. "Han ville ha en helt och hållet blå färg – den fick inte vara blågul."</i> <input type="checkbox"/> Gul färg som samtidigt har en lite blå nyans <i>T.ex. "Han ville ha en helt och hållet gul färg – den fick inte vara blågul."</i> <input type="checkbox"/> Både blå och gul <i>T.ex. "Sveriges blågula flagga vajade i vinden."</i>
Barnbokhylla	<input type="checkbox"/> Liten hylla som passar i ett barnrum <i>T.ex. "Hon åkte till en möbelaffär och köpte en barnbokhylla till dotterns rum."</i> <input type="checkbox"/> Hylla med barnböcker, t.ex. på ett bibliotek <i>T.ex. "När man vill låna en barnbok på biblioteket får man gå till barnbokhyllan."</i>

Bilaga 1: Elevenkät

<p>Sockersöt</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Söt av socker <i>T.ex. "Hon längtade efter en kopp sockersött te."</i> <input type="checkbox"/> Mycket söt (som socker) <i>T.ex. "Hon hade en sockersöt klänning i rosa med många pärlor."</i>
<p>Iskall</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Kall (eller kyld) av is <i>T.ex. "Vattnet i sjön kan bara vara iskallt på vintern när det finns is."</i> <input type="checkbox"/> Mycket kall (som is) <i>T.ex. "Det var iskallt på mitt rum igår."</i>
<p>Snökall</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Kall av snö <i>T.ex. "Marken var snökall och vit."</i> <input type="checkbox"/> Mycket kall (som snö) <i>T.ex. "Det var snökallt på mitt rum igår."</i>
<p>Kvinnopräst</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Präst som är kvinna <i>T.ex. "En del personer tycker att det är fel med kvinnopräster i församlingen."</i> <input type="checkbox"/> Präst som tar emot och hjälper kvinnor <i>T.ex. "Kvinnor som har det svårt kan få hjälp hos en kvinnopräst."</i>
<p>Fågelrädd</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Rädd som en fågel <i>T.ex. "Han kände sig fågelrädd där han stod ensam på den stora skolgården."</i> <input type="checkbox"/> Rädd för fåglar <i>T.ex. "Han var fågelrädd och tyckte inte om när fåglarna kom för nära."</i>
<p>Fruktkorg</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Speciell korg som man använder när man plockar frukt <i>T.ex. "Hämta fruktkorgen, så plockar vi ned frukten från äppelträdet."</i> <input type="checkbox"/> Korg som det ligger frukt i <i>T.ex. "Det stod en korg på bordet som det låg apelsiner, äpplen och bananer i – alltså en fruktkorg."</i> <input type="checkbox"/> Korg som är dekorerad med frukter på utsidan <i>T.ex. "Han hittade en liten korg med frukter på – alltså en fruktkorg."</i>

<p>Gröngul</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Både grön och gul T.ex. "Mössan hade gröna och gula ränder – den var alltså gröngul." <input type="checkbox"/> Grön färg som samtidigt har en lite gul nyans eller färgton T.ex. "Limefrukten var rejält mogen. Den gröna färgen hade fått en svag nyans av gult." <input type="checkbox"/> Gul färg som samtidigt har en lite grön nyans eller färgton T.ex. "Bananen var inte riktigt mogen. Den gula färgen hade fortfarande en svag nyans av grönt."
<p>Sommardagdröm</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Dröm om en sommardag T.ex. "Jag hade en dröm i natt om en vacker sommardag – en underbar sommardagdröm." <input type="checkbox"/> Dagdröm som man har på sommaren T.ex. "På sommaren fantiserar jag mycket – och jag älskar mina sommardagdrömmar."
<p>Sötsaker</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Söta saker som man kan äta T.ex. "Hon borstade tänderna noga, eftersom hon ätit så mycket sötsaker." <input type="checkbox"/> Saker som är söta T.ex. "Hon köpte två sötsaker till sin treåring: en nalle och en liten docka."
<p>Zebralejon</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Lejon som äter zebror T.ex. "När zebralejonet kommer flyr alla zebror undan." <input type="checkbox"/> Lejon som är randigt som en zebra T.ex. "Det är lätt att känna igen det randiga zebralejonet."
<p>Boktext</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Text ur/i en bok T.ex. "Läraren hade kopierat upp några boktexter om hinduismen." <input type="checkbox"/> Bok med texter T.ex. "I franskan hade de en boktext som de läste ur varje vecka."
<p>Båthus</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Båt som man kan bo i T.ex. "På semestern ska vi göra en resa med båthuset." <input type="checkbox"/> Hus där man kan lägga båten T.ex. "Innan vintern tar vi upp båten och lägger den i båthuset."

Bilaga 1: Elevenkät

<p>Skolboks-försäljning</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Försäljning av olika typer av böcker, som äger rum i skolan T.ex. "För att öka intresset för att läsa ordnade lärarna en bokförsäljning i skolan – alltså en skolboks-försäljning." <input type="checkbox"/> Försäljning av skolböcker T.ex. "Skolan ville göra sig av med gamla skolböcker och anordnade en skolboks-försäljning."
<p>Finkläder</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Kläder som är fina T.ex. "Hon hade köpt funkläder: ett par mjukisbyxor och tre par sockar." <input type="checkbox"/> Kläder som passar vid finare, högtidligare tillfällen T.ex. "Hon hade köpt funkläder: en långklänning och en kavaj."
<p>Spelkort</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Spel som spelas med speciella kort T.ex. "Jag spelar hellre schack än spelkort." <input type="checkbox"/> Speciella kort (i en kortlek) som man använder för att spela med T.ex. "Hon hade fått bra spelkort: flera kungar och ess."
<p>Fulsnygg</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Mittemellan ful och snygg T.ex. "Han var inte ful, men inte heller snygg – alltså fulsnygg" <input type="checkbox"/> Mer ful än snygg T.ex. "Han hade ett fint leende men som helhet tyckte hon att han ändå var mer ful än snygg – alltså fulsnygg." <input type="checkbox"/> Mer snygg än ful T.ex. "Han var inte snygg på det perfekta sättet men han var ändå attraktiv och snygg på något sätt – alltså fulsnygg."
<p>Plastlåda</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Låda som det ligger plast i T.ex. "Det stod en låda i hyllan som det låg plast i – alltså en plastlåda." <input type="checkbox"/> Speciell låda som man använder för att samla plast i T.ex. "Vi måste sortera återvinningen och lägga all plast i plastlådan!" <input type="checkbox"/> Låda som är gjord av plast T.ex. "Maten som blev över lade han i en liten plastlåda som han ställde in i kylan."
<p>Mansläkare</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Läkare som arbetar med sjukdomar som bara män kan få T.ex. "Han hade ont i pungen och sökte hjälp hos en mansläkare." <input type="checkbox"/> Läkare som är man T.ex. "Det arbetar mest mansläkare på den här avdelningen. Vi skulle behöva fler kvinnor."

Svampkorg	<ul style="list-style-type: none"> <input type="checkbox"/> Korg som är dekorerad med svampar på utsidan T.ex. "Han hittade en liten korg med svampar på – alltså en svampkorg." <input type="checkbox"/> Korg som det ligger svamp i T.ex. "Det stod en korg på bordet som det låg svamp i – alltså en svampkorg." <input type="checkbox"/> Speciell korg som man använder när man plockar svamp T.ex. "Hämta svampkorgen, så går vi ut i skogen och plockar svamp."
Mansmisshandel	<ul style="list-style-type: none"> <input type="checkbox"/> När män misshandlar T.ex. "Mannen hade slagit kvinnan flera gånger och blev dömd för mansmisshandel." <input type="checkbox"/> När män blir misshandlade T.ex. "Kvinnan hade slagit mannen flera gånger och blev dömd för mansmisshandel."
Honungssöt	<ul style="list-style-type: none"> <input type="checkbox"/> Mycket söt (som honung) T.ex. "Hon hade en honungssöt klänning i rosa med många pärlor." <input type="checkbox"/> Söt av honung T.ex. "Hon längtade efter en kopp honungssött te."
Kvinnovård	<ul style="list-style-type: none"> <input type="checkbox"/> Vård som kvinnor ger T.ex. "Hon vände sig till kvinnovården eftersom det bara arbetade kvinnor där." <input type="checkbox"/> Vård som kvinnor får T.ex. "På mottagningen för kvinnovård arbetade män och kvinnor med att hjälpa kvinnliga patienter."
Isvatten	<ul style="list-style-type: none"> <input type="checkbox"/> Vatten som är kallt som is T.ex. "När jag tar vatten ur kranen spolar jag länge tills det kommer isvatten." <input type="checkbox"/> Vatten med is i T.ex. "Jag beställde isvatten till maten på restaurangen."
Glasskall	<ul style="list-style-type: none"> <input type="checkbox"/> Mycket kall (som glass) T.ex. "Maten från frysen hade inte riktigt tinat. Den var fortfarande glasskall." <input type="checkbox"/> Kall av glass T.ex. "Hon var glasskall efter att ha ätit en stor glasspinne."

Bilaga 1: Elevenkät

<p>Blomlåda</p>	<ul style="list-style-type: none"> ❑ Låda som är blommig T.ex. "Hon ville inte ha en randig låda till sitt rum utan en blommig – alltså en blomlåda." ❑ Speciell låda för blommor som man kan hänga upp på balkongen T.ex. "Hon köpte blåa och lila blommor att plantera i blomlådan." ❑ Låda/kartong som det är blommor i T.ex. "Hon köpte så många blommor att hon fick släpa ut en hel blomlåda till bilen."
<p>Båtsommar</p>	<ul style="list-style-type: none"> ❑ Sommar i båt T.ex. "Vi har varit ute i båten nästan hela sommaren. Det har varit en härlig båtsommar." ❑ Båt som man bara kan använda på sommaren T.ex. "Vår båt är en typisk båtsommar som bara funkar när det är varmt."
<p>Korpörn</p>	<ul style="list-style-type: none"> ❑ Örn som jagar korpfåglar T.ex. "När korpörnen dyker upp flyr alla korpar undan." ❑ Örn som är svart som en korp T.ex. "Det är lätt att känna igen den svarta korpörnen."
<p>Bilfri</p>	<ul style="list-style-type: none"> ❑ Fri på grund av bil T.ex. "Den här veckan är det jag som har bilen. Känner mig härligt bilfri." ❑ Fri från bilar = utan bilar T.ex. "Man får inte köra bil här. Det här är en bilfri gata!"
<p>Vitgul</p>	<ul style="list-style-type: none"> ❑ Gul färg med en lite vit nyans eller färgton T.ex. "Hans tänder såg smutsigt vitgula ut – de var alltså mer gula än vita." ❑ Vit färg med en lite gul nyans eller färgton T.ex. "Hans tänder hade blivit vitgula – de var alltså fortfarande mest vita men hade en lite gul färgton." ❑ Både vit och gul T.ex. "Mössan hade vita och gula ränder – den var alltså vitgul."
<p>Valfri</p>	<ul style="list-style-type: none"> ❑ Fri från val = utan val T.ex. "De önskade en valfri skola som var lika för alla." ❑ Som kan väljas fritt T.ex. "Till lunchen ingick valfri efterrätt."

DEL 3: Språk och språkanvändning

Till sist följer här några frågor om dig och din språkanvändning. Kryssa för de alternativ som beskriver dig. Finns det frågor där inget av alternativen stämmer, så skriv en kommentar vid frågan.

1. Kön

- Tjej
- Kille
- Transperson

2. Vilket år är du född?

- 1993
- 1994
- 1995
- 1996
- 1997
- 1998
- 1999
- Annat: _____

3. Är du född i Sverige?

- Ja
- Nej

4. Om nej, hur gammal var du när du kom till Sverige?

- | | |
|-------------------------------|--------------------------------|
| <input type="checkbox"/> 0 år | <input type="checkbox"/> 10 år |
| <input type="checkbox"/> 1 år | <input type="checkbox"/> 11 år |
| <input type="checkbox"/> 2 år | <input type="checkbox"/> 12 år |
| <input type="checkbox"/> 3 år | <input type="checkbox"/> 13 år |
| <input type="checkbox"/> 4 år | <input type="checkbox"/> 14 år |
| <input type="checkbox"/> 5 år | <input type="checkbox"/> 15 år |
| <input type="checkbox"/> 6 år | <input type="checkbox"/> 16 år |
| <input type="checkbox"/> 7 år | <input type="checkbox"/> 17 år |
| <input type="checkbox"/> 8 år | <input type="checkbox"/> 18 år |
| <input type="checkbox"/> 9 år | <input type="checkbox"/> 19 år |
| | <input type="checkbox"/> 20 år |

5. Vilket/vilka språk kan du tala? _____

6. Vilket/vilka språk lärde du dig tala först (ditt/dina modersmål)? _____

Bilaga 1: Elevenkät

7. Hur gammal var du när du började lära dig svenska?

- | | |
|-------------------------------|--------------------------------|
| <input type="checkbox"/> 0 år | <input type="checkbox"/> 10 år |
| <input type="checkbox"/> 1 år | <input type="checkbox"/> 11 år |
| <input type="checkbox"/> 2 år | <input type="checkbox"/> 12 år |
| <input type="checkbox"/> 3 år | <input type="checkbox"/> 13 år |
| <input type="checkbox"/> 4 år | <input type="checkbox"/> 14 år |
| <input type="checkbox"/> 5 år | <input type="checkbox"/> 15 år |
| <input type="checkbox"/> 6 år | <input type="checkbox"/> 16 år |
| <input type="checkbox"/> 7 år | <input type="checkbox"/> 17 år |
| <input type="checkbox"/> 8 år | <input type="checkbox"/> 18 år |
| <input type="checkbox"/> 9 år | <input type="checkbox"/> 19 år |
| | <input type="checkbox"/> 20 år |

8. Var började du först lära dig svenska?

- Hemma
- På dagis/förskola
- Hos dagmamma
- I skolan
- Med kompisar
- Annat, nämligen _____

9. Har du under din skoltid läst svenska som andraspråk?

- Ja
- Nej

Om du svarar nej på fråga 9 kan du gå direkt till fråga 14

10. Har du läst svenska som andraspråk under lågstadiet?

- Ja
- Nej

11. Har du läst svenska som andraspråk under mellanstadiet?

- Ja
- Nej

12. Har du läst svenska som andraspråk under högstadiet?

- Ja
- Nej

13. Har du läst svenska som andraspråk under gymnasiet?

- Ja
- Nej

14. Har du under din skoltid läst ämnet modersmål/hemspråk?

- Ja
- Nej

Om du svarar nej på fråga 14 kan du gå direkt till fråga 19

15. Har du läst ämnet modersmål/hemspråk under lågstadiet?

- Ja
- Nej

16. Har du läst ämnet modersmål/hemspråk under mellanstadiet?

- Ja
- Nej

17. Har du läst ämnet modersmål/hemspråk under högstadiet?

- Ja
- Nej

18. Har du läst ämnet modersmål/hemspråk under gymnasiet?

- Ja
- Nej

19. Vilken typ av svenskundervisning deltar du i nu?

- Svenska
- Svenska som andraspråk
- Ingen

20. Vilket gymnasieprogram läser du?

- Ekonomiprogrammet
- Estetiska programmet
- Humanistiska programmet
- Naturvetenskapsprogrammet
- Samhällsvetenskapsprogrammet
- Teknikprogrammet
- Annat: _____

21. Vilken årskurs på gymnasiet läser du?

- År 1
- År 2
- År 3

Bilaga 1: Elevenkät

22. Är din mamma född i Sverige?

- Ja
- Nej

Kommentar _____

23. Är din pappa född i Sverige?

- Ja
- Nej

Kommentar _____

24. Föräldrars modersmål

- Mina föräldrar har svenska som modersmål
- En av mina föräldrar har svenska som modersmål.
- Mina föräldrar har inte svenska som modersmål

Kommentar _____

25. Vilket/vilka är din mammas modersmål? _____

26. Vilket/vilka är din pappas modersmål? _____

27. Vilket/vilka språk talar ni hemma i familjen? _____

28. Vilket/vilka språk talar du mest med dina vänner? _____

29. Om du har en pojkvän/flickvän, vilket språk talar du oftast med honom/henne?

30. Övriga kommentarer

Stort tack för din medverkan! /Lisa

*Lisa Loenheim (lisa.loenheim@svenska.gu.se)
Doktorand i nordiska språk, Göteborgs universitet*

Bilaga 2: Korpusstudie om NA-sammansättningar (ur Loenheim 2016)

I denna bilaga återfinns tabeller som sammanfattar resultatet av en mindre pilotstudie om NA-sammansättnings semantiska relationer (Loenheim 2016; se vidare avsnitt 2.1). Till denna studie användes fem korpusar med tidningstext, skönlitteratur och bloggtext ur Språkbankens korpusar (se Språkbanken/Korp; se även Borin et al. 2012). Korpusarnas och deras storlek redovisas i tabell 1.

I tabell 2 redovisas frekvenser för de semantiska (och grammatiska) relationer som identifierades för studiens 293 sammansättningar.

Därefter presenteras i tabell 3 en fördjupad studie av de semantiska relationer som förekommer mellan leden i NA-sammansättningar med efterledet *-kall* i samma korpusmaterial.

I tabell 4 återges resultaten av en motsvarande undersökning av sammansättningar med efterledet *-varm* i samma material.

TABELL 1. *Fördelning mellan studerade korpusar.*²⁵

Korpus	Token	Andel
GP 2013	16 871 215	28,4 %
Två Dagar	1 035 966	1,7 %
Bloggmix 2014	34 298 071	57,8 %
Norstedtsromaner	2 534 307	4,3 %
SUC-romaner	4 653 784	7,8 %
Totalt	59 393 343	

²⁵ Sökningarna utfördes i Språkbankens korpusar med sökverktyget Korp feb.-mars 2016.

TABELL 2. NA-sammansättningar fördelade efter relation (typ- och tokenfrekvenser).

Relation/kategori	Exempel på ord i kategorin	Typfrekvens Antal lemman	Procent	Total tokenfrekvens, per miljon ord
Förstärkande relation [mycket led 2]	<i>jättebra</i>	51	17,4	253,61
Bundet adverbial med -lös [lös från = utan led 1]	<i>arbetslös</i>	39	13,3	102,62
Bundet adverbial innehåll [av/med led 1]	<i>kraftfull</i>	38	13,0	141,89
Jämförande relation [som led 1]	<i>stenhård</i>	26	8,9	84,74
Bundet adverbial [från led 1]	<i>felfri</i>	24	8,2	71,33
Led 1 = objekt	<i>fruktansvärd</i>	21	7,2	99,47
Omständighetsadverbial [i fråga om led 1]	<i>färgglad</i>	15	5,1	51,38
Omständighetsadverbial [med avseende på led 1]	<i>konkurrens- kraftig</i>	14	4,8	38,70
Bundet adverbial [på led 1]	<i>framgångsrik</i>	11	3,8	57,53
Bundet adverbial [för led 1]	<i>livsavgörande</i>	10	3,4	22,04
Bundet adverbial orsak [av led 1]	<i>avundsjuk</i>	7	2,4	45,10
Fritt adverbial plats	<i>världsberömd</i>	7	2,4	11,57
Bundet adverbial [mot led 1]	<i>miljövänlig</i>	5	1,7	19,64
Bundet adverbial [i led 1]	<i>sakkunnig</i>	4	1,4	12,52
Led 1 = geografisk modifiering	<i>nordöstra</i>	4	1,4	12,73
Omständighetsadverbial [enligt led 1]	<i>sedvanlig</i>	3	1,0	8,31

Bilaga 2: Korpusstudie om NA-sammansättningar (ur Loenheim 2016)

Relation/kategori	Exempel på ord i kategorin	Typfrekvens Antal lemman	Procent	Total tokenfrekvens, per miljon ord
[i relation till led 1]	<i>jordnära</i>	3	1,0	8,06
Bundet adverbial [med led 1]	<i>skräckblandad</i>	2	0,7	2,72
Bundet adverbial [till led 1]	<i>stadiebehörig</i>	2	0,7	6,67
Fritt adverbial orsak [p.g.a. led 1]	<i>pappaledig</i>	2	0,7	6,63
Fritt adverbial tid	<i>morgonpigga</i>	2	0,7	3,97
Bundet adverbial [om led 1]	<i>målmedveten</i>	1	0,3	3,79
Led 1 = mått	<i>livslång</i>	1	0,3	3,07
Bundet adverbial agent [av led 1]	<i>kvinnodominerad</i>	1	0,3	1,05
Total		293	100	

TABELL 3. NA-sammansättningar med -kall fördelade efter relation (typ- och tokenfrekvenser).

Relation/kategori	Exempel	Typfrekvens		Total tokenfrekvens, per miljon ord
		Antal lemman	Procent	
Förstärkande relation [mycket led 2]	<i>svinkall</i>	9	34,6	4,14
Jämförande relation [som led 1]	<i>iskall</i>	7	26,9	12,87
Bundet adverbial orsak [av led 1]	<i>kylskåpskall</i>	6	23,1	0,78
Grad	<i>knasterkall</i>	1	3,8	0,02
Additiv relation	<i>solkall</i>	1	3,8	0,02
Omständighetsadverbial [i fråga om led 1]	<i>känslökall</i>	1	3,8	0,71
Fritt adverbial tid [på led 1]	<i>nattkall</i>	1	3,8	0,02
Totalt		26	100,0	

TABELL 4. Fördelning av NA-sammansättningar med -varm som efterled.

Relation/kategori	Exempel	Typfrekvens		Total tokenfrekvens, per miljon ord
		Antal lemman	Procent	
Förstärkande relation [mycket led 2]	<i>tokvarm</i>	13	37,1	1,75
Jämförande relation [som led 1]	<i>fingervarm</i>	9	25,7	1,58
Bundet adverbial orsak [av led 1]	<i>solvarm</i>	8	22,9	3,55
Grad	<i>t-shirtvarm</i>	2	5,7	0,03
Fritt adverbial tid [på/i led 1]	<i>höstvarm</i>	2	5,7	0,03
Additiv relation	<i>fuktvarm</i>	1	2,9	0,03
Total		35	100,0	

Bilaga 3: Enkät sammansättningarnas frekvenser i avhandlingens referenskorpus

I denna bilaga redovisas korpusfrekvenser för de sammansättningar som ingår i avhandlingens enkätstudie. Sammansättningarnas frekvenser i referenskorpusen (se avsnitt 4.4) redovisas per sammansättningskategori i tabell 1–12. I tabell 13 ges en alfabetiskt ordnad sammanställning över enkätens samtliga sammansättningar och deras frekvenser i referenskorpusen.

Tabellförteckning

- TABELL 1. Huvudled i nominala sammansättningar
- TABELL 2. Huvudled i treledade sammansättningar
- TABELL 3. Semantisk relation i sammansättningar som betecknar behållare
- TABELL 4. Jämförande egenskapsprojicering eller tematisk relationslänkning
- TABELL 5. Identitet/jämförande egenskapstolkning eller tematisk relationslänkning i personbetecknande sammansättningar
- TABELL 6. Agens vid verbavledning
- TABELL 7. Betydelsesammansmälta AN-sammansättningar
- TABELL 8. Huvudled i AA-sammansättningar
- TABELL 9. Jämförelse eller orsak i NA-sammansättningar
- TABELL 10. Sammansättningar med *-radd*
- TABELL 11. Sammansättningar med *-fri*
- TABELL 12. Sammansättningar med *-färdig*
- TABELL 13. Alla sammansättningar i bokstavsordning

Tabeller, per sammansättningskategori

TABELL 1. *Huvudled i nominala sammansättningar.*

	tokenfrekvens	
	absolut	relativ per miljon token
Husbil	272	3,8611
Husbåt	59	0,8375
Båthus	54	0,7665
Spelkort	28	0,3975
Flodstrand	18	0,2555
Skalpotatis	9	0,1278
Burköl	4	0,0568
Båtsommar	1	0,0142
Boktext	0	0,0000
Stövelbarn	0	0,0000
Vagnbarn	0	0,0000

TABELL 2. *Huvudled i treledade sammansättningar.*

	tokenfrekvens	
	absolut	relativ per miljon token
Barnbokhylla	1	0,0142
Barnbokshylla	0	0,0000
Skolboksförsäljning	0	0,0000
Sommardagdröm	0	0,0000

TABELL 3. *Semantisk relation i sammansättningar som betecknar behållare.*

	tokenfrekvens	
	absolut	relativ per miljon token
Blomkruka	105	1,4905
Blomlåda	58	0,8233
Plastlåda	56	0,7949
Fruktkorg	19	0,2697
Glaslåda	12	0,1703
Svampkorg	7	0,0994
Olivkorg	0	0,0000

TABELL 4. *Jämförande egenskapsprojicering eller tematisk relationslänkning.*

	tokenfrekvens	
	absolut	relativ per miljon token
Citronvatten	72	1,0221
Isvatten	48	0,6814
Apörn	0	0,0000
Issaft	0	0,0000
Korpörn	0	0,0000
Skräpkruka	0	0,0000
Zebralejon	0	0,0000
Zebraväska	0	0,0000

TABELL 5. *Identitet/jämförande egenskapstolkning eller tematisk relationslänkning i personbetecknande sammansättningar.*

	tokenfrekvens	
	absolut	relativ per miljon token
Kvinnopräst	11	0,1561
Kvinnokarl	9	0,1278
Flickpojke	2	0,0284
Herrfrisör	2	0,0284
Kvinnoläkare	1	0,0142
Kvinnoman	0	0,0000
Mansläkare	0	0,0000
Manspräst	0	0,0000

TABELL 6. *Agens vid verbavledning.*

	tokenfrekvens	
	absolut	relativ per miljon token
Barnmisshandel	51	0,7240
Politikerförakt	42	0,5962
Kvinnomisshandel	38	0,5394
Polismisshandel	1	0,0142
Mansmisshandel	0	0,0000
Kvinnovård	0	0,0000
Studentförakt	0	0,0000

TABELL 7. *Betydelsesammansmälta AN-sammansättningar.*

	tokenfrekvens	
	absolut	relativ per miljon token
Sötsaker	394	5,5929
Grönområde	151	2,1435
Småkakor	135	1,9164
Finkläder	60	0,8517
Grönyta	24	0,3407

TABELL 8. *Huvudled i AA-sammansättningar.*

	tokenfrekvens	
	absolut	relativ per miljon token
Svartvit	664	9,4257
Blågul	368	5,2239
Gulgrön	68	0,9653
Gulvit	60	0,8517
Sötsur	54	0,7665
Fulsnygg	35	0,4968
Gröngul	15	0,2129
Dumsnäll	13	0,1845
Vitgul	11	0,1561

TABELL 9. *Jämförelse eller orsak i NA-sammansättningar.*

	tokenfrekvens	
	absolut	relativ per miljon token
Stenhård	1229	17,4461
Iskall	853	12,1086
Sockersöt	86	1,2208
Blodrött	81	1,1498
Solvarm	64	0,9085
Kryddstark	49	0,6956
Solgul	39	0,5536
Snövit	36	0,5110
Honungssöt	3	0,0426
Dammgrått	2	0,0284
Berghård	0	0,0000
Glasskall	0	0,0000
Lökstark	0	0,0000
Sandmjuk	0	0,0000
Snökall	0	0,0000
Snömjuk	0	0,0000
Spisvarm	0	0,0000

TABELL 10. *Sammansättningar med -rädd.*

	tokenfrekvens	
	absolut	relativ per miljon token
Fågelrädd	0	0,0000
Hundrädd	0	0,0000
Musrädd	0	0,0000

TABELL 11. *Sammansättningar med -fri.*

	tokenfrekvens	
	absolut	relativ per miljon token
Valfri	1578	22,4003
Bilfri	41	0,5820
Helgfri	6	0,0852

TABELL 12. *Sammansättningar med -färdig.*

	tokenfrekvens	
	absolut	relativ per miljon token
Gråtfärdig	104	1,4763
Skrattfärdig	1	0,0142
Skrikfärdig	0	0,0000

Tabell, alla sammansättningar i bokstavsordning

TABELL 13. *Alla sammansättningar.*

tabell	tokenfrekvens	
	absolut	relativ per miljon token
Apörn	4	0,0000
Barnbokhylla	2	0,0142
Barnbokshylla	2	0,0000
Barnmisshandel	6	0,7240
Berghård	9	0,0000
Bilfri	11	0,5820
Blodrött	9	1,1498
Blomkruka	3	1,4905

Bilaga 3: Enkät sammansättningarnas frekvenser i avhandlingens referenskorpus

	tabell	tokenfrekvens	
		absolut	relativ per miljon token
Blomlåda	3	58	0,8233
Blågul	8	368	5,2239
Boktext	1	0	0,0000
Burköl	1	4	0,0568
Båthus	1	54	0,7665
Båtsommar	1	1	0,0142
Citronvatten	4	72	1,0221
Dammgrått	9	2	0,0284
Dumsnäll	8	13	0,1845
Finkläder	7	60	0,8517
Flickpojke	5	2	0,0284
Flodstrand	1	18	0,2555
Fruktkorg	3	19	0,2697
Fulsnygg	8	35	0,4968
Fågelrädd	10	0	0,0000
Glaslåda	3	12	0,1703
Glasskall	9	0	0,0000
Gråtfärdig	12	104	1,4763
Gröngul	8	15	0,2129
Grönområde	7	151	2,1435
Grönyta	7	24	0,3407
Gulgrön	8	68	0,9653
Gulvit	8	60	0,8517
Helgfri	11	6	0,0852
Herrfrisör	5	2	0,0284
Honungssöt	9	3	0,0426
Hundrädd	10	0	0,0000
Husbil	1	272	3,8611
Husbåt	1	59	0,8375
Iskall	9	853	12,1086
Issaft	4	0	0,0000

Bilaga 3: Enkät sammansättningarnas frekvenser i avhandlingens referenskorpus

	tabell	tokenfrekvens	
		absolut	relativ per miljon token
Isvatten	4	48	0,6814
Korpörn	4	0	0,0000
Kryddstark	9	49	0,6956
Kvinnokarl	5	9	0,1278
Kvinnoläkare	5	1	0,0142
Kvinnoman	5	0	0,0000
Kvinnomisshandel	6	38	0,5394
Kvinnopräst	5	11	0,1561
Kvinnovård	6	0	0,0000
Lökstark	9	0	0,0000
Mansläkare	5	0	0,0000
Mansmisshandel	6	0	0,0000
Manspräst	5	0	0,0000
Musrädd	10	0	0,0000
Olivkorg	3	0	0,0000
Plastlåda	3	56	0,7949
Polismisshandel	6	1	0,0142
Politikerförakt	6	42	0,5962
Sandmjuk	9	0	0,0000
Skalpotatis	1	9	0,1278
Skolboks försäljning	2	0	0,0000
Skrattfärdig	12	1	0,0142
Skrikfärdig	12	0	0,0000
Skräpkruka	4	0	0,0000
Småkakor	7	135	1,9164
Snökall	9	0	0,0000
Snömjuk	9	0	0,0000
Snövit	9	36	0,5110
Sockersöt	9	86	1,2208
Solgul	9	39	0,5536
Solvarm	9	64	0,9085

Bilaga 3: Enkät sammansättningarnas frekvenser i avhandlingens referenskorpus

	tabell	tokenfrekvens	
		absolut	relativ per miljon token
Sommardagdröm	2	0	0,0000
Spelkort	1	28	0,3975
Spisvarm	9	0	0,0000
Stenhård	9	1229	17,4461
Studentförakt	6	0	0,0000
Stövelbarn	1	0	0,0000
Svampkorg	3	7	0,0994
Svartvit	8	664	9,4257
Sötsaker	7	394	5,5929
Sötsur	8	54	0,7665
Vagnbarn	1	0	0,0000
Valfri	11	1578	22,4003
Vitgul	8	11	0,1561
Zebralejon	4	0	0,0000
Zebraväska	4	0	0,0000