

På rak sak

GÖTEBORGSSTUDIER I NORDISK SPRÅKVETENSKAP 13

REDAKTÖRER FÖR SERIEN: Elisabet Engdahl, Roger Källström,
Sven-Göran Malmgren, Lena Rogström, Barbro Wallgren Hemlin

Det finns många sätt att uttrycka ett visst innehåll på ett grammatiskt korrekt sätt i ett språk, men inte alla accepteras som naturliga uttryckssätt av infödda talare. Ett idiomatiskt, dvs. inföddligt, språkbruk kännetecknas bl.a. av användningen av konventionaliserade flerordsenheter. Dessa uttryck är ofta språk-specifika och måste alltså läras in av andraspråkstalare. Inom andraspråksforskning har man visat att svårigheter att använda konventionaliserade uttryck på ett idiomatiskt sätt kvarstår även för inlärare på avancerad nivå. På ett semantiskt plan utmärks dessutom konventionaliserade uttryck som *på rak arm*, *hålla någon i schack* eller *lyckligt ovetande* genom att deras betydelse inte är helt förutsägbar utifrån de ingående komponenternas betydelse.

Denna doktorsavhandling i svenska som andraspråk består av tre studier om användning och behärskning av idiom, kollokationer och andra typer av konventionaliserade uttryck hos ungdomar i flerspråkiga storstadsmiljöer. Julia Prentice har undersökt på vilka sätt dessa unga språkbrukare med varierande språklig bakgrund avviker från standardenligt bruk av fraseologiska enheter, hur de använder figurativa ordförbindelser i skrift och i vilken mån de behärskar ett femtiotal svenska konventionaliserade uttryck. Med hjälp av såväl kvantitativa som kvalitativa analyser och utifrån ett huvudsakligen kognitivt perspektiv beskriver författaren på vilka sätt dessa uttryck kan sätta färg på språket, men även sätta käppar i hjulen för idiomaticiteten i språkbruket bland ungdomar med varierande språklig bakgrund.



JULIA PRENTICE

På rak sak

Om ordförbindelser och konventionaliserade uttryck bland unga språkbrukare i flerspråkiga miljöer



GÖTEBORGSSTUDIER I NORDISK SPRÅKVETENSKAP

REDAKTÖRER: Elisabet Engdahl, Roger Källström,
Sven-Göran Malmgren, Lena Rogström, Barbro Wallgren Hemlin

1. Magnus Fernberg: *Käseristil*. 2004.
2. Tore Otterup: "Jag känner mej begåvad bara". *Om flerspråkighet och identitetskonstruktion bland ungdomar i ett multietniskt förortsområde*. 2005.
3. Per Fröjd: *Att läsa och förstå svenska. Läsförmågan hos elever i årskurs 9 i Borås 2000–2002*. 2005.
4. Jenny Nilsson: *Adverb i interaktion*. 2005.
5. Susanna Karlsson: *Positioneringsfraser i interaktion. Syntax, prosodi och funktion*. 2006.
6. Thorwald Lorentzon: *Mellan frihet och jämlikhet. Jämförande studier av lexikala förändringar i moderaternas och vänsterpartiets valmanifest 1948–2002*. 2006.
7. Maia Andréasson: *Satsadverbial, ledföljd och informationsdynamik i svenskan*. 2007.
8. Sofia Tingsell: *Reflexivt och personligt pronomen. Anaforisk syftning hos ungdomar i flerspråkiga stadsmiljöer*. 2007.
9. Claes Ohlsson: *Folkets fonder? En textvetenskaplig studie av det svenska pensionssparandets domesticering*. 2007.
10. Peter Andersson: *Modalitet och förändring. En studie av må och kunna i fornsvenska*. 2007.
11. Gudrun Rawoens: *Kausativa verbkonstruktioner i svenskan och nederländskan. En korpusbaserad syntaktisk-semantisk undersökning*. 2008.
12. Sven Lövfors: *Historiska angivelser i allmänna ordböcker med särskild hänsyn till Svensk ordbok utgiven av Svenska Akademien*. 2010.
13. Julia Prentice: *På rak sak. Om ordförbindelser och konventionaliserade uttryck bland unga språkbrukare i flerspråkiga miljöer*. 2010

Göteborgsstudier i nordisk språkvetenskap 13

På rak sak

Om ordförbindelser och konventionaliserade uttryck
bland unga språkbrukare i flerspråkiga miljöer

(English Summary)

av Julia Prentice



GÖTEBORGS UNIVERSITET 2010

TITLE: Word combinations and conventionalized expressions among adolescent language users in multilingual environments.

SWEDISH TITLE: På rak sak. Om ordförbindelser och konventionaliserade uttryck bland unga språkbrukare i flerspråkiga miljöer.

LANGUAGE: Swedish (English summary).

AUTHOR: Julia Prentice

Abstract

The overall aim of the thesis is to investigate and describe the use and mastery of different types of phraseological units and conventionalized expressions in the language of adolescents in multilingual environments. The thesis comprises three studies focusing on different aspects of phraseology in relation to first and second language acquisition, development and use. The studies have been carried out mainly from a cognitive perspective.

The first study focuses on deviation from standard Swedish in adolescents' use of phraseological units. Data from the *SUF*-project (*Language and language use among adolescents in multilingual urban settings*) have been analyzed in relation to standard Swedish reference data and compared to a constructed variety, rendered in a novel placed in a Swedish multiethnic setting. In the novel deviations from standard Swedish were found to be mostly lexico-grammatical, whereas the authentic data appears to be more complex and deviate both on a lexico-grammatical and a semantic-pragmatic level.

The second paper reports on a study on the use of figurative word combinations and their degree of conventionalization in written data from the *SUF*-project. The results support the initial hypothesis that the frequency and types of figurative expressions in the participants' texts can, at least to some degree, be linked to the participants' linguistic background. L1 students use more conventionalized figurative word combinations than L2 students, whereas modifications of conventionalized figurative expressions are more frequent in the L2 students' texts. The study also shows that the students create novel figurative word combinations based on a great variety of source domains.

The third study focuses on the mastery of Swedish conventionalized expressions among adolescents with various linguistic backgrounds. Complementary data have been collected by means of a survey that contains four tests focusing on various aspects of the mastery of conventionalized expressions. The quantitative results show that L1 students reach significantly higher test results than L2 students. Furthermore, L2 students with an early age of onset reach significantly higher results than the later L2 learners. Thus, age of onset appears to be an important factor for the participants' mastery of conventionalized expressions in Swedish, a conclusion that is also confirmed by the qualitative analysis of part of the data.

KEYWORDS: Second language acquisition and use, phraseology, conventionalization, multilingualism, metaphor, cognitive linguistics, lexical competence, youth language, multiethnic varieties, Swedish

© Julia Prentice

DISTRIBUTION: Institutionen för svenska språket
Box 200
405 30 Göteborg

OMSLAGSBILD: Philipp Große

ISSN: 1652-3105

ISBN: 978-91-87850-40-0

SÄTTNING: Janne Saaristo

TRYCKNING: Intellecta Infolog, Göteborg 2010

Für Papa
Du fehlst

Förord

Att skriva en avhandling kan ibland kännas som det mest ensamma jobbet i världen, men med jämna mellanrum, och särskilt mot slutet, inser man att så inte är fallet. Det behövs många människor som delar med sig av sin tid och energi för att ett projekt som detta ska ros i land. Dessa människor vill jag tacka här.

Jag börjar med informanterna, alla elever som har deltagit i projektet *Språk och språkbruk bland ungdomar i flerspråkiga storstadsmiljöer* och de elever som har deltagit i enkätundersökningen, ett stort tack! Utan er hade denna avhandling inte blivit till! Ett speciellt tack till alla lärare på de berörda skolorna som har delat med sig av lektionstid och engagerat sig för forskningens skull!

Näst på tur är min Huvudhandledare Inger Lindberg, som med sin otroligt breda kunskap har varit en aldrig sinande källa till inspiration och nya idéer. Jag är tacksam för många goda samtal, klok vägledning, mycket omtanke och allt roligt som jag har fått vara med om som Ingers doktorand! Kenneth Hyltenstam har varit bihandledare och i olika skeden drivit arbetet framåt med sina kloka och sakliga synpunkter, ett stort tack för det! Emma Sköldberg bordade skeppet under andra halvan av doktorandtiden, först som samarbetspartner och sedan som bihandledare. Tack Emma, för att du drog mig tillbaka in i idiomens värld under en tuff period i mitt liv och sedan dess har stöttat mig i mitt arbete, noggrant och rakt på sak! Trots ibland heta diskussioner har det i första hand varit roligt att jobba ihop.

Elisabet Engdahl har på olika sätt varit ett viktigt stöd under doktorandtiden och med klarsynthet och omtanke visat vägen genom forskarutbildningsdjungeln. Hon har dessutom läst och kommenterat delar av texten som redaktör för avhandlingen, ett hjärtligt tack för båda delarna!

Att få ingå i ett projektsamarbete har varit både roligt och lärorikt, vilket har känts som en stor förmån särskilt i början av doktorandtiden. Därför vill jag tacka alla medlemmar i projektet *Språk och språkbruk bland ungdomar i fler-*

språkiga storstadsmiljöer. Tack Elin Almér, Ellen Bilvoet, Petra Bodén, Sally Boyd, Lena Ekberg, Kari Fraurud, Natalia Ganuza, Sofie Johansson Kokkinakis, Roger Källström, Inger Lindberg, Tore Otterup, Gudrun Svensson, Sofia Tingsell, Alexandra Utrén och Marie Werdin, för samarbetet och för att ni delat med er av material, kunskap och viktiga synpunkter! Jag vill särskilt tacka Lena Ekberg har läst delar av mitt arbete i olika skeden och varit så generös med sin kunskap!

Roger Källström har också läst och kommenterat olika delar av texten, både av ren vänlighet och som redaktör för min avhandling, tack Roger!

Tack även till Maria Wiktorsson som har varit en utmärkt slutseminarieopponent! Hennes frågor och kommentarer har betytt mycket i arbetets slutskede.

Jag vill också tacka Sofie Johansson Kokkinakis, Sofia Tingsell och Hans Landkvist som har läst och granskat delar av texten.

För hjälp med de statistiska analyserna vill jag tacka Tommy Johnsson och Sofie Johansson Kokkinakis. Tack även till Jenny Mattson, som har granskat de engelska sammanfattningar och Eva Lundin som har hjälpt till med korrekturläsningen i slutskedet!

Margit Nelson Wareborn och Carina Carlund har öppnat viktiga dörrar för mig i jakten på informanter, tack så hemskt mycket!

Ett stort tack till min gamle studiekollega Janne Saaristo som har gjort denna fina bok av alla mina texter! Det var roligt att jobba ihop igen. Min bror Philipp Große har skapat omslagsbilden, vielen Dank dafür!

Institutionen för svenska språket och Institutet för svenska som andraspråk har på olika sätt stöttat mig finansiellt under doktorandtiden, bl.a. genom resebidrag till olika konferensresor. Även för detta vill jag tacka här!

Ulrika Magnusson har varit min rumskompis under en stor del av doktorandtiden och blivit en mycket uppskattad kollega och vän. Vi har under vissa perioder varit märkligt synkade både vad gäller doktorerandet och livet i övrigt. Jag saknar redan nu våra dagliga samtal och diskussioner om allt från språkvetenskap till kultur, världspolitik, föräldraskap och världen i övrigt. Tack också för värdefulla synpunkter på mitt arbete i olika skeden, speciellt för noggrann läsning och ”gnetiga kommentarer” på delar av slutmanuset!

Jag vill även tacka alla andra kollegor, bland doktorander, medarbetare på ISA och institutionen i övrigt, som öppet och vänligt tog emot mig som ny doktorand, har hjälpt mig på olika sätt under åren och helt enkelt varit goda kollegor. Här vill särskilt nämna Tore, Sofia, Karin H, Susanna, Jenny, Claes, Peter, Annika, Ida, Ann-Marie, Eva, Carina, Sofie, Qarin, Karin S, Anette, Margareta, Sören, Anders-Börje, Britt-Marie, Pirjo och Malin (som har lovat att bära mig och mitt manus till tryckeriet ifall det skulle behövas).

Man kan ha mer eller mindre gynnsamma förutsättningar för att ge sig in i ett avhandlingsprojekt, även på det personliga planet. Jag känner mig utan tvekan lyckligt lottad på den punkten.

Utan all fysisk aktivitet och annat kul som jag ägnat mig åt tack vare mina tränare och lagkamrater i landhockeyn hade den mentala hälsan lidit betydligt mer än den har gjort under avhandlingsskrivandet, därför även ett tack till alla er! Ett särskilt tack till Sara, för blåbär och jordgubbar (bortsett från vänskap, mentalt stöd och barnpassning).

Von meinen Eltern und Geschwistern habe ich vieles, aber vor allem immer eines erfahren: vollkommene bedingungslose Unterstützung. Danke!

Mein kleiner Junge Joshua har varit ovanligt förstående och förlåtande för en treåring. Jag ska verkligen försöka vara lite mindre av en Rabenmutter, nu när detta projekt är avslutat!

And, last but not at all least, Jim, the Excell wizard, who – despite my somewhat confusing instructions – managed to turn my results into some not so confusing diagrams, but who most of all has been a very supportive husband and made it possible for me to put the time and effort required into this project, while at the same time still have a functioning family and social life. THANK YOU!

Göteborg i juni 2010
Julia Prentice

Innehållsförteckning

1. Inledning.....	1
1.1. Avhandlingens övergripande syfte.....	2
2. Tidigare forskning, teoretiska utgångspunkter och begrepp	5
2.1. Andraspråksinläring och flerspråkighet	5
2.1.1. Kritisk period och mognadsbegränsningar	5
2.1.2. Ordförråd i andraspråket.....	7
2.1.3. Vem är flerspråkig?	9
2.1.4. Teoretiska modeller för flerspråkighet	11
2.1.5. Svenska på mångspråkig grund	12
2.2. Kognitiva teorier om språk	13
2.2.1. Kognitiv grammatik	14
2.2.2. Det mentala lexikonet i kognitiva modeller.....	17
2.2.2.1. Det flerspråkiga lexikonet	20
2.2.3. Metaforteori.....	22
2.3. Fraseologi	26
2.3.1. Fraseologiska enheter och konventionaliserat språk	27
2.3.2. Idiomaticitet	33
2.3.3. Hur fasta är fasta fraser?.....	34
2.3.4. Fraseologiska enheter och andraspråksinläring.....	35
3. Material och metod	37
3.1. Material	37
3.2. Kvantitativa och kvalitativa metoder.....	38

4. Sammanfattning av artiklarna	41
I. Fraseologiska enheter på flerspråkig grund. En jämförelse av verklighet och fiktion	41
II. <i>Klättra på väggarna</i> eller <i>bara vara ett med soffan</i> . Om figurativa ordförbindelser bland ungdomar i flerspråkiga skolmiljöer	42
III. <i>Käppen i hjulen</i> . Behärskning av svenska konventionaliserade uttryck bland gymnasieelever med varierande språklig bakgrund	43
5. Återstående frågor och fortsatt forskning	47
Summary.....	49
Referenser.....	59

Ingående publikationer

I. Fraseologiska enheter på flerspråkig grund. En jämförelse av verklighet och fiktion	103–132
II. <i>Klättra på väggarna</i> eller <i>bara vara ett med soffan</i> . Om figurativa ordförbindelser bland ungdomar i flerspråkiga skolmiljöer	5–35
III. <i>Käppen i hjulen</i> . Behärskning av svenska konventionaliserade uttryck bland gymnasieelever med varierande språklig bakgrund	1–145

1. Inledning

Denna avhandling är en sammanläggning av tre arbeten som ur olika perspektiv behandlar användningen av fraseologiska enheter och konventionaliserade uttryck bland ungdomar i flerspråkiga miljöer.

En gemensam utgångspunkt för delstudierna är att användningen av fraseologiska enheter, och i synnerhet konventionaliserade flerordsenheter, är ett av de språkliga drag som i tidigare studier har visats vara av stor betydelse för det som man brukar kalla idiomatiskt eller inföddligt språkbruk (se bl.a. Pawley & Syder 1983, Ekberg 1997, 1998, Wray 2002). Språkliga konventioner är också specifika för en viss språkgemenskap, vilket gör flerspråkiga miljöer särskilt intressanta att studera i detta avseende. Ett återkommande inslag i samtliga delstudier är också jämförelsen mellan ungdomar med enspråkig svensk och flerspråkig bakgrund. Ur ett andraspråksperspektiv är startålderns (eng. *age of onset*) betydelse också en viktig och återkommande fråga för användningen och behärskningen av fraseologiska enheter och konventionaliserat språk i andraspråket (jfr Abrahamsson & Hyltenstam 2009).

Den forskning som presenteras i avhandlingen har genomförts inom ramen för projektet *Språk och språkbruk bland ungdomar i flerspråkiga storstadsmiljöer* (SUF).¹ Projektet har ett tvådelat syfte; man avser att avgränsa och beskriva en språklig varietet som talas bland ungdomar i multietniska storstadsområden i Sverige samt att belysa språkanvändningen i dessa områden med avseende på bl.a. sociolingvistiska och inlärningsrelaterade aspekter (Ganuza 2008, Svensson 2009, se även Lindberg, Boyd, Ekberg, Fraurud & Källström 2001).

Det avhandlingsarbete som presenteras i denna volym består av tre olika studier som fokuserar olika aspekter av användning och behärskning av olika typer av ordförbindelser och konventionaliserade uttryck bland ungdomar i

¹ Projektet finansierades av Riksbankens jubileumsfond 2002–2006.

flerspråkiga miljöer. Studie I utgörs av artikeln ”Fraseologiska enheter på mångspråkigt grund. En jämförelse mellan verklighet och fiktion”. Undersökningen baseras på fraseologiska enheter i både muntligt och skriftligt SUF-material som jämförs med litterärt konstruerade avvikelser från svenska konventionaliserade uttryck (Khemiri 2003). I artikeln tar jag upp kognitiva, inlärningsrelaterade, fraseologiska och i viss mån sociolingvistiska aspekter i samband med användningen av fraseologiska enheter bland ungdomar i flerspråkiga miljöer och ger en kvalitativ beskrivning av avvikelser i förhållande till svenskt standardspråk på både ett lexiko-grammatiskt och ett semantiskt-pragmatiskt plan.

Studie II betår av artikeln ”*Klättra på väggarna eller bara vara ett med soffan? Om figurativa ordförbindelser bland ungdomar i flerspråkiga skolmiljöer*”, skriven tillsammans med Emma Sköldberg. Den fokuserar figurativitet, som kan sägas vara ett utmärkande drag hos många typer av konventionaliserade ordförbindelser, men tar samtidigt upp figurativa nybildningar i gymnasieelevers texter. Undersökningen baseras på det skriftliga SUF-materialet. Centrala frågor i denna studie är bl.a. i vilken mån eleverna använder figurativa ordförbindelser i sitt skrivande samt graden av konventionalisering hos de figurativa ordförbindelser som används. Även i studie II spelar både kognitiva, inlärningsrelaterade och fraseologiska aspekter en central roll i analysen.

Studie III består av rapporten ”*Käppen i hjulen*. Behärskning av svenska konventionaliserade uttryck bland gymnasieelever med varierande språklig bakgrund” och baseras på en enkätstudie som genomförts som komplement till SUF-materialet i syfte att bedöma gymnasieelevers behärskning av ett antal konventionaliserade svenska uttryck. Enkäten innehåller ett förståelsetest, två lucktest där eleverna ombes att komplettera konventionaliserade uttryck i och utanför en kontext samt ett matchningstest som går ut på att para ihop givna ord till de mest idiomatiska kollokationerna. Studien skiljer sig från de första två bl.a. genom en mer omfattande kvantitativ analys. Dessutom har enkäterna med de lägsta respektive högsta resultaten analyserats kvalitativt. Även i denna tredje, mest omfattande studie analyseras materialet ur de tre ovannämnda teoretiska perspektiven.

1.1. Avhandlingens övergripande syfte

Avhandlingens övergripande syfte är att utforska, beskriva och analysera användningen och behärsningen av olika typer av fraseologiska enheter och konventionaliserade uttryck i svenskan bland ungdomar i flerspråkiga miljöer. Inom ramen för avhandlingens tre studier görs detta ur olika perspektiv och med fokus på

- i vilken utsträckning och på vilka sätt ungdomar i en flerspråkig miljö
 - använder fraseologiska enheter i vissa språksituationer och
 - avviker från standardspråket i sin användning av fraseologiska enheter och konventionaliserade uttryck,
- figurativitet och grad av konventionalisering i ungdomars användning av flerordskombinationer i skrift,
- ungdomars behärskning av olika typer av konventionaliserade uttryck i förhållande till faktorer som språklig bakgrund och startålder för andraspråksinläring.

Mot bakgrund av SUF-projektets syfte, som inledningsvis refererades, består denna avhandlings bidrag huvudsakligen i att belysa frågor kring språkanvändning och språklig variation i flerspråkiga storstadsmiljöer utifrån kognitiva och inlärningsrelaterade faktorer. Även om analyserna alltså i första hand präglas av dessa perspektiv berörs i viss mån också sociolingvistiska aspekter i samband med diskussionen kring olika ungdomsvarieteter.

2. Tidigare forskning, teoretiska utgångspunkter och begrepp

I detta avsnitt redogörs för tidigare forskning inom för avhandlingen relevanta områden samt för de teoretiska utgångspunkter som är centrala för de ingående arbeten. Inledningsvis berörs frågor om andraspråksinlärning och tvåspråkighet. Relevanta aspekter är startålderns betydelse för andraspråksinlärning, ordförrådet i ett andraspråk, olika syn på flerspråkighet samt ungdomsspråk i flerspråkiga sammanhang. I anslutning till detta tas några grundläggande utgångspunkter och antaganden inom kognitiva språkliga teorier upp. Här berörs också frågor om det flerspråkiga mentala lexikonet och metafor-teori. Vidare introduceras den grundläggande terminologin och viktiga teoretiska utgångspunkter på fraseologiområdet. *Idiom* och *kollokationer* presenteras som centrala typer av konventionaliserade uttryck och avslutningsvis tas teoretiska modeller och forskning om inlärning och användning av fraseologiska enheter i ett andraspråk upp.

2.1. Andraspråksinlärning och flerspråkighet

2.1.1. Kritisk period och mognadsbegränsningar

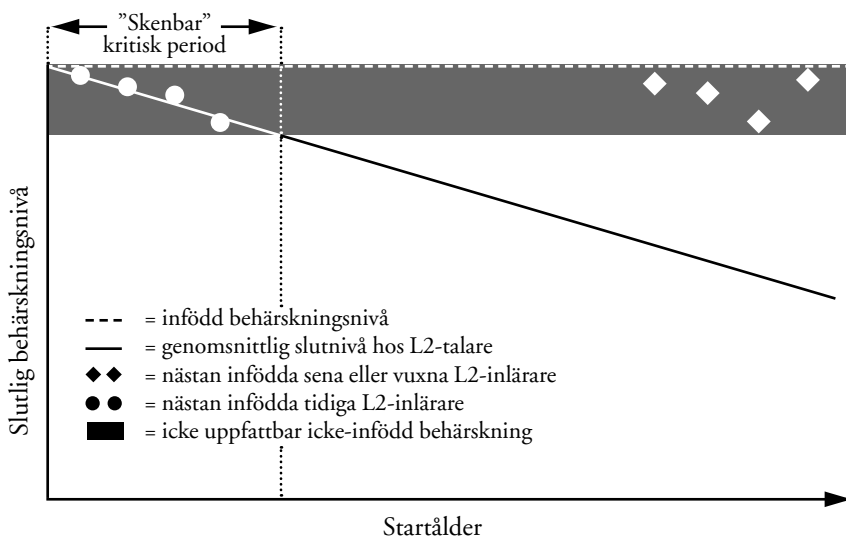
En av de mest omdiskuterade frågorna inom andraspråksforskningen är sambandet mellan startålder och slutnivå för inlärning (eng. *ultimate attainment*). Antagandet att barn lättare lär in ett andraspråk än vuxna inlärare formulerades i ett vetenskapligt sammanhang redan på sent 1950-tal av Penfield & Roberts (1959). De menar att barns förmåga att lära in ett språk enbart genom in-

flödet upphör, eller åtminstone avtar, vid nioårsåldern på grund av biologiska och neurologiska förändringar (se Abrahamsson & Hyltenstam 2004). Lenneberg (1967) har också formulerat en hypotes som senare har använts i varierande formuleringar som går under benämningen *hypotesen om en kritisk period* (eng. *the critical period hypothesis*, CPH). Lennebergs ursprungliga formulering innebär att möjligheten att tillägna sig ett första- eller andraspråk upp till infödd nivå genom icke-formell inläring i princip försvinner efter puberteten. Detta sker på grund av biologiska mognadsprocesser i hjärnan. Senare inläring måste, enligt Lenneberg, ske genom formell undervisning och kräver analytisk förmåga och språklig medvetenhet (1967:176). CPH kan sägas vara en av de mest omdebatterade frågorna inom andraspråksforskningen. Debatten har dock förts på mycket varierande grunder och ofta utan hänsyn till en eller flera viktiga aspekter av Lennebergs ursprungliga formulering, t.ex. att han avser en infödd slutnivå på inläringen, och inte i vilken mån det generellt är möjligt att tillägna sig ett andraspråk efter en viss ålder (se Abrahamsson & Hyltenstam 2004:223f).

En gemensam nämnare för studierna kring slutnivå för andraspråksinläring är, som Abrahamsson & Hyltenstam (2004:234) påpekar, att de visar på ett signifikant samband mellan startålder för inläring och slutnivå på inläringen. Dessutom kommer samtliga studier fram till att de flesta inlärare som påbörjar inläringen av ett språk tidigt och före en viss ålder verkar nå en slutnivå som liknar infödda talares, även om åldersgränserna och vilka faktorer som framhävs som negativt påverkande i samband med äldre inlärares förmåga till informell inläring varierar. Huvudorsakerna anses ligga i ovannämnda neurologiska mognadsprocesserna (se t.ex. Johnson & Newport 1989, DeKeyser 2000), och olika sociala och psykologiska faktorer som anses gynnsamma för barns språkinläring, bl.a. ”högre motivation, mer och ’bättre’ språkligt inflöde, skolundervisning, färre psykologiska spärrar, lösare identitetsmässiga band till den ursprungliga kulturen och modersmålet (se t.ex. Bialystock & Hakuta 1999, Bialystock & Miller 1999, Flege 1999 [...])” (Abrahamsson & Hyltenstam 2004:236).

Även de slutsatser som har dragits utifrån olika studier på området varierar. Vissa forskare menar att endast de yngsta andraspråksinlärarna har möjlighet att uppnå en slutnivå i sin språkfärdighet som är jämförbar med förstaspråkstalarars (se t.ex. Bialystock & Miller 1999, Johnson & Newport 1989). Andra drar slutsatsen att åtminstone enstaka sena inlärare kan komma upp i en nivå på andraspråket där de inte kan skiljas från infödda talare i olika typer av språkfärdighetstest, t.ex. grammatikalitetsbedömningstest eller perceptionstest där infödda lyssnare uppmanas att identifiera talare som första- respektive andraspråkstalare av språket ifråga (jfr bl.a. Birdsong 1992, Bongaerts 1999, Moyer 1999). Under det senaste decenniet har det emellertid genomförts ett antal stu-

dier som tyder på att varken tidiga eller sena andraspråksinlärare uppnår en slutnivå på andraspråket som motsvarar infödda talares språkbehärskningsnivå; istället avtar förmågan att tillägna sig ett språk upp till denna nivå kontinuerligt från födseln (bl.a. Hyltenstam & Abrahamsson 2002, 2003, Butler 2000, se även Abrahamsson & Hyltenstam 2004, 2009). Abrahamsson & Hyltenstam (2004) har efter en utvärdering av forskningsläget samt en egen undersökning om startålder för andraspråksinlärning och uppfattad inföddhet föreslagit en modell som innebär att möjligheten att uppnå inföddnivå på ett språk avtar linjärt från födseln. Detta innebär i sin tur att andraspråksinlärare aldrig uppnår en nivå som helt motsvarar infödd nivå. Modellen innehåller dessutom en gräns för när det är möjligt att identifiera icke-infödda talare som sådana. Både tidiga och enstaka sena andraspråksinlärare med exceptionellt gynnsamma förutsättningar för språkutveckling antas kunna passera denna gräns. Denna "icke uppfattbara ickeinföddhet" (Abrahamsson & Hyltenstam 2004:247) kan alltså vara en förklaring till de motstridiga forskningsresultat som tidigare refererats; tidiga andraspråksinlärare och även vissa individer med sen startålder för sin andraspråksinlärning kan alltså helt enkelt nå en nivå i sin språkbehärskningsnivå där man endast med väldigt specifika testmetoder kan skilja dem från infödda talare. Att andraspråksinlärare som påbörjar sin inlärning före en viss ålder passerar gränsen för uppfattbar inföddhet innebär att det uppstår en *skenbar kritisk period*, under vilken starten på andraspråksinlärning kan leda till en slutbehärskningsnivå på en skenbar inföddnivå. Abrahamsson & Hyltenstam nämner 6–7 årsåldern som gräns för en sådan skenperiod (2004:246ff).



FIGUR 1: Startålder för språkinlärning och slutlig behärskningsnivå (Abrahamsson & Hyltenstam 2004:247)

2.1.2. Ordförråd i andraspråket

Utvecklingen av ordförrådet pågår under hela livet, vilket gäller såväl för förstaspråket som för eventuella andraspråk (Viberg 2004, se även 1993). Skolåren har inom ordinlärningsforskningen identifierats som en period där ordförrådet utvecklas speciellt snabbt, vilket naturligtvis inte är särskilt förvånande. Man utgår allmänt ifrån att ett barn vid skolstarten har ett ordförråd på mellan 8 000 och 10 000 ord på sitt modersmål och att de sedan kan utöka lexikonet med upp till 3 000 ord per år (Nagy & Herman 1987). I detta sammanhang bör det dock påpekas att dessa uppskattningar i första hand bygger på läsförståelsetest och att den individuella variationen är stor (Viberg 1993:35f). Även om forskarna inte är helt överens och det heller inte går att ange några exakta siffror, innebär skolstarten alltså en dramatisk expansion av ordförrådet för förstaspråkselever och samtidigt en stor utmaning för elever som har undervisningsspråket som andraspråk. Andraspråkseleverna behöver ofta komma i kapp de jämnåriga förstaspråkseleverna när det gäller det basordförråd som dessa redan behärskar vid skolstarten, samtidigt som de därutöver ska bygga upp ett skolrelaterat ordförråd på andraspråket och utöka det med flera tusen ord om året (Viberg 1993:36, se även Studie III). Mot denna bakgrund är det inte särskilt förvånande att Thomas & Collier (2002) har funnit att det kan ta mellan fem och tio år för andraspråkselever att uppnå en nivå som motsvarar jämnåriga förstaspråkselevens i sin behärskning av undervisningsspråket. Samtidigt är förmågan att möta de lexikala krav som andraspråkstalare ställs inför i olika utbildningssammanhang naturligtvis inte enbart en fråga om ordförrådets storlek. Den lexikala kompetensen är inte endimensionell utan komplex och sammansatt som Nation påpekar i följande citat:

Words are not isolated units of language, but fit into many interlocking systems and levels. Because of this, there are many things to know about any particular word and there are many degrees of knowing (Nation 2001:23).

Även Henriksen (1999) betonar att full behärskning av ett ord har många dimensioner. Hon föreslår en tredimensionell modell för lexikal utveckling, enligt vilken lexikal kompetens består av ”(a) *partial to precise knowledge*, (b) *depth of knowledge*, and (c) *receptive to productive use ability*” (1999:303). Enkelt uttryckt betyder det att en grammatiskt felaktig eller icke-idiomatisk användning av ett ord eller flerordsuttryck inte ska tolkas som att språkbrukaren inte har någon kunskap om ordet eller uttrycket i fråga. Att ordet eller uttrycket överhuvudtaget används är i stället ett bevis på partiell kunskap om den lexikala enheten ifråga, även om den inte behärskas till fullo (jfr Nation 2001:27, se vidare Studie III)

En viktig del av lexikonkunskapen består i kunskaper om olika slags relationer mellan de ord som lärts in (Viberg 1993:37). Det handlar bl.a. om att lära sig i vilka semantiska kontexter och tillsammans med vilka andra ord en lexikal enhet kan förekomma (jfr Nation 2001:27). Sådan kunskap har bl.a. undersökts med hjälp av associationstest (se bl.a. Håkansson & Norrby 2005, Meara 1983, Namei 2002). I studier där dessa typer av test har använts på andraspråksinlärare har man bl.a. kommit fram till att dessa, i likhet med barn som lär sig sitt förstaspråk, i högre utsträckning gör fonologiska associationer än infödda talare, som snarare verkar associera på det semantiska planet utifrån ett givet ord. Dessa resultat kan tolkas som att fonologiska associationer utvecklas tidigare än semantiska (Viberg 1993:37). De kognitiva aspekterna av ordförrådets utveckling och uppbyggnad diskuteras med tonvikt på inlärningen av fraseologiska enheter och konventionaliserade flerordsuttryck vidare i avsnitt 2.2.2 nedan.

Viberg (1993, 2004) har i en longitudinell studie studerat utvecklingen av verblexikonet hos 20 flerspråkiga barn som han har följt från förskoleklassen till årskurs fyra. Han har funnit en överanvändning av det han kallar för *nukleära verb* i andraspråksbarnens muntliga produktion, jämfört med språkbruket hos jämnåriga barn med svenska som förstaspråk. Med nukleära verb avses verb som är ”typologiskt omarkerade och har samma grundläggande betydelse i ett stort antal språk” (Viberg 2004:203). Det är högfrekventa verb som ofta har betydelse som ’göra’, ’ge’, ’ta’, ’säga’, ’se’ eller ’veta’ och som kan sägas utgöra kärnan i semantiska fält som FÖRFLYTTNING, PRODUKTION, ÄGANDE, VERBAL KOMMUNIKATION eller PERCEPTION (Viberg 1994, 2004: 210). De nukleära verben kan sägas spela en speciell roll i ordförrådet, både för första- och andraspråkstalare (Viberg 2004:218); de lärs in tidigt under ordförrådsutveckling och överanvänds alltså ofta av andraspråkanvändare på bekostnad av semantiskt mer differentierade eller polysema verb. Dessa mönster förändras i takt med att barnen utvecklar sin lexikonkunskap, både i fråga om ordförrådets bredd och djup. Enligt Viberg (2004:217) utvecklas språkspecifik semantisk differentiering i ordförrådet på ett andraspråk successivt, vilket innebär att vissa nukleära verb överanvänds längre än andra. Detta kan sägas spegla ett allmänt mönster för lexikal utveckling i både första och andraspråk; lexikala enheter med mer generella betydelser, som finns representerade i många olika språk, utvecklas till ett mer och mer differentierat, specialiserat och språkspecifikt ordförråd.

2.1.3. Vem är flerspråkig?

Vilka omständigheter som krävs för att en individ ska betraktas som två- eller flerspråkig är, som Bialystok (2001:1) poängterar, inte på något sätt självklart. De flesta människor har vid något tillfälle kommit i kontakt med ett annat språk än det egna modersmålet och många har kunskaper på olika nivåer i fler än ett språk, t.ex. genom språkliga studier, utan att varken se sig själva eller bli betraktade som flerspråkiga (Bialystok 2001:2). Även bland de individer som allmänt anses vara flerspråkiga är förutsättningarna och omständigheterna för utvecklingen av deras flerspråkighet högst varierande.

Picture a bilingual child. What languages does this child speak? What kind of neighbourhood does she live in? What are the educational arrangements that either support or demand bilingualism? Are any of the child's languages spoken in the community outside the home? What were the circumstances that led to her bilingualism? How long has the child been in the present country? By changing even a single answer to this small sample of questions, the child being described is importantly different from one who would have elicited a different answer (Bialystok 2001:1).

Definitionsproblemet återspeglas i forskningen om flerspråkig utveckling där kriterierna för flerspråkighet varierar från full behärskning av två (eller flera) språk (jfr Bloomfield 1933) till funktionell behärskning utifrån den enskilda individens språkliga behov (Grosjean 1982). Därutöver finns en rad definitioner som ligger emellan dessa två ytterligheter (se Otterup 2005:11ff). Mackey (1970) menar, som framgår av citatet nedan, att flerspråkighet helt och hållet är att betrakta som ett relativt begrepp.

It seems obvious that, if we are to study the phenomenon of bilingualism we are forced to consider it as something entirely relative. We must moreover include the use not only of two languages, but any number of languages. We shall therefore consider bilingualism as the alternative use of two or more languages by the same individual (Mackey 1970:55).

Citatet antyder även ett terminologiskt problem när det gäller begreppen tvåspråkighet respektive flerspråkighet. Begreppen används i dagens läge alltmer synonymt, från att flerspråkighet ursprungligen huvudsakligen använts i samband med flerspråkiga samhällen som t.ex. Schweiz eller Belgien (Otterup 2005:12). Skutnabb-Kanges (1981) anser att definitionerna av begreppet tvåspråkighet kan delas in i fyra kategorier, nämligen utgåenden ifrån *ursprung*, *kompetens*, *funktion* eller *attityder* (se Otterup 2005:12). Flerspråkighetsbegreppet, som det används i denna avhandling, kan sägas utgå från såväl informanternas ursprung, i termer av föräldrarnas språkliga bakgrund, som attityder, dvs. i vilken mån eleverna själv anser sig vara flerspråkiga.

Fraurud & Boyd (2006:53ff) diskuterar samma typ av komplexitet vad gäller begreppen första- och andraspråksinlärare i samband med den språkligt och kulturellt heterogena informantgruppen i SUF-projektet. I en studie baserad på de bakgrundsintervjuer som genomfördes med informanterna beskriver de upp till 88 olika språkliga profiler (eng. *linguistic profiles*) för deltagarna i projektet, baserade på olika kombinationer av bakgrundsvariabler. Med andra ord visar Fraurud & Boyd att informanterna inte kan kategoriseras utifrån en enkel, binär distinktion mellan första- och andraspråkstalare (2006:54). De visar i sin analys att variabler som vanligtvis inte beaktas vid indelningen av olika grupper, t.ex. bästa vänners modersmål, i hög grad kan vara relevanta när det gäller vissa informanter (Fraurud & Boyd 2006:66). För att få ett operationellt antal kategorier av informanter måste man naturligtvis förenkla bilden genom att generalisera utifrån ett urval av bakgrundsfaktorer. Det är dock viktigt att komma ihåg att informantgrupperna i de här aktuella undersökningarna inte bör förstås som helt homogena ifråga om språklig bakgrund. Indelningen av informanterna, både i SUF-projektet (studierna I och II) och enkätundersökningen (studie III), har alltså inte varit helt oproblematisk. Då startålder för inlärningen av svenska har varit en central utgångspunkt, har indelningen i första hand baserats på föräldrarnas språkliga bakgrund och vilket språk eleverna själva anger som förstaspråk.

2.1.4. Teoretiska modeller för flerspråkighet

Traditionellt har man inom forskningen utgått från två olika typer av flerspråkighet – nämligen *subtraktiv* och *additiv flerspråkighet*. I korta ordalag innebär subtraktiv flerspråkighet att andraspråksinläraren i en huvudsakligen enspråkig miljö successivt överger sitt modersmål till förmån för andraspråket och på så sätt blir en enspråkig medlem av majoritetsspråksgemenskapen (se García 2009:52f). Additiv flerspråkighet siktar däremot mot en balanserad två- eller flerspråkighet. I denna modell, som enligt García i första hand speglar verkligheten för ”prestigious groups and the elite” (2009:52), adderas andraspråket till inlärarens modersmål och båda språken används och utvecklas.

García påpekar dock att dessa traditionella modeller för två- eller flerspråkighet inte räcker till för att beskriva den komplexa flerspråkiga kompetens som i dagens läge krävs i många samhällen (2009:53). Ett problem med de tidigare nämnda modellerna är, enligt García, att de utgår från en enspråkig norm, vilket inte speglar verkligheten i dagens alltmer globaliserade samhällen. Hon introducerar istället en modell för *dynamisk flerspråkighet*.

A more heteroglossic conception of bilingualism recognizes its adjustments as it shifts and bounces. Bilingualism is not simply linear but *dynamic*, drawing from the

different contexts in which it develops and functions. [...] the world's globalization is increasingly calling on people to interact with others in ways that defy traditional categories. [...] With language interaction taking place on different planes including multimodalities, that is, different modes of language (visuals as well as print, sound as well as text, and so on) as well as multilingualism, it is possible for individuals to engage in multiple complex communicative acts [...] (Garcia 2009:53f).

Modellen tar hänsyn till att flerspråkighet i vår tid innebär annat än dubbel enspråkighet (García 2009:52ff). Flerspråkiga individer i den dynamiska modellen antas alltså inte använda sina olika språk på samma sätt som enspråkiga modersmålstalare av dessa språk. Flerspråkighet anses istället inkludera alla olika språkliga praktiker och funktioner där individen använder de olika språk som han eller hon behärska på olika nivåer. Modellen inkluderar alltså, utöver behärsknigen av två eller flera standardspråk, också det som Garcia (2009:55) kallar *hybrid language practices*, dvs. att man på olika sätt använder flera språk i en och samma språksituation, vilket inte minst är relevant i samband med ungdomars språkbruk i flerspråkiga miljöer (jfr även Lindberg 2009).

2.1.5. Svenska på mångspråkig grund

Svenska på mångspråkig grund (SMG) är den term som används inom SUF-projektet för att beteckna de (talade) varieteter som används av ungdomarna som deltagit i projektet. Begreppet är tänkt som en övergripande beteckning som innefattar olika varieteter bland ungdomar i alla de tre storstäderna, nämligen Stockholm, Göteborg och Malmö.

Behovet av en neutral term motiveras bl.a. utifrån det faktum att det förekommer en rad olika beteckningar för de muntliga varieteter som talas bland ungdomar i de aktuella områdena, t.ex. *rinkebysvenska*, *rosengårdssvenska*, *gårdstenska* eller *blattesvenska*. Vissa av dessa beteckningar implicerar att varieteten endast används lokalt i en viss stadsdel. Andra, som *blattesvenska*, är klart stigmatiserande. Dessutom är de varieteter som betecknas på ovannämnda sätt inte de enda som används av ungdomar i de aktuella områden.

Ett av syftena med SUF-projektet har även varit att beskriva ungdomarnas språkbruk mer generellt, dvs. i olika språksituationer i tal och skrift. Som framgår av redogörelsen ovan, berör forskningen inom SUF-projektet olika aspekter av språkbruket bland ungdomar i flerspråkiga miljöer. Vissa av studierna uppfyller i första hand syftet att undersöka och avgränsa språkliga drag utmärkande för SMG som ungdomsvariant (se Bodén 2007, Bodén & Große [Prentice] 2006, Ekberg 2006, 2007, Ganuza 2008, Hansson & Svensson 2004, Svensson 2009). Andra arbeten fokuserar i större utsträckning på beskrivningar av olika sociolingvistiska och inlärningsrelaterade aspekter ifråga om språk-

användning och språklig variation bland de berörda ungdomarna (se Tingsell 2007). Man kan dock säga att båda de ovannämnda aspekterna, om än i olika utsträckning, berörs i de studier som genomförts inom ramen för projektet.

Den forskning som presenteras i detta arbete syftar inte till att beskriva användningen av konventionaliserade uttryck utmärkande för SMG som tal-språksvarietet. Det skulle kräva en djupare analys av talspråksmaterialet och en jämförelse av sådana projektdeltagare som har identifierats som talare av varietetet med en kontrollgrupp av SUF-informanter som inte anses vara SMG-användare (se bl.a. Hansson & Svensson 2004, Bodén & Große 2006). Som antytts tidigare, berör denna forskning i första hand kognitiva och inlärningsrelaterade aspekter av språkanvändningen bland ungdomar i flerspråkiga miljöer. Frågan om sociolingvistisk variation, och därmed också om SMG som varietet, berörs i viss utsträckning, bl.a. i samband med delstudie I, men framför allt indirekt i den meningen att studierna visar att språkbruket bland ungdomarna ifråga är varierat. SMG är på intet vis den enda språkliga varietet som används bland informanterna och det är, som påpekats i tidigare forskning (se bl.a. Fraurud & Bijvoet 2004), viktigt att skilja mellan variation som är sociolingvistisk betingad och variation och avvikelser som är inlärningsrelaterade och t.ex. kan kopplas till informanternas startålder för andraspråksinläring. Både sociolingvistiska och inlärningsrelaterade avvikelser och variation i förhållande till standardenligt språkbruk måste beskrivas i sin egen rätt och sättas i relation till varandra om man vill ge en fullständigare bild av språkbruket bland ungdomar i flerspråkiga miljöer (jfr Bijvoet & Fraurud 2008).

2.2. Kognitiva teorier om språk

Till skillnad från formella teorier om språk som utgår från språket som system med syntaxen i centrum, har en del funktionellt inriktade teorier ett gemensamt mål i försöket att förstå och beskriva språkets struktur utifrån realiseringar av kognitiva strukturer i olika språksituationer i kontext. Utifrån ett funktionellt synsätt sker språkutveckling i samband med att språket fyller specifika kognitiva, sociala och kommunikativa funktioner (Bialystok 2001:40f). Ett annat grundläggande drag inom funktionella grammatiska teorier, som är relevant för föreliggande arbete, är att syntaxen inte innehar en särställning på samma sätt som i vissa formellt inriktade teorier, utan att språkliga strukturer ses som sammansatta av grammatiska, semantiska och pragmatiska mönster (jfr Fillmore 1988, se även Bialystok 2001:41).

Samtliga studier i detta arbete är, liksom en stor del av fraseologiforskningen, kognitivt inriktad. Det kan därför vara på sin plats att i korthet redogöra

för några av de kognitiva teorier som tillämpats inom detta område. Tanken är inte att ge en heltäckande bild utan snarare att teckna en bakgrund till de teoretiska resonemang som förs i samband med de tre undersökningarna och sätta dem i relation till andra begrepp på området samt att redogöra för var de termer som används ursprungligen hör hemma.

2.2.1. Kognitiv grammatik

Den kognitiva grammatiken bygger på den uppsättning av betydelser eller den *begreppsbildning* (eng. *conceptualization*, se vidare nedan) som är resultatet av vad vi upplever och kognitivt processar i vår verklighet i olika situationer (Bialystok 2001:43). Utgångspunkten för teorin är att vår kunskap om världen består av konceptuella strukturer lagrade i minnet som vid behov kan överföras och bidra till förståelsen av fenomen i verkligheten (Langlotz 2006:61). Så kan en människa t.ex. föreställa sig en katt utan att se, höra eller röra vid en katt i den reella världen. Detta tyder på att människan har ett begrepp lagrat i minnet, som består i information om hur t.ex. en katt ser ut, låter och känns (jfr även Lamb 2000). Dessa begrepp fungerar, enligt Langlotz, som kategorier som hjälper oss att hantera och organisera informationen som vi hela tiden utsätts för, gällande vår omgivning:

...meaningful experience is not only related to our immediate perception of the physical world, but draws upon conceptual routines [...]. Concepts allow us to make sense of our environments [...], conceptual units work as [...] categorisation-standards that can be mapped onto the flow of experience to group and organize it (Langlotz 2006:61f).

De betydelser och den begreppsbildning som språket, enligt kognitiv grammatik, anses baserat på, har sitt ursprung i interaktionen med hela den sociala och pragmatiska kontext i vilken språk uppstår. Att inkludera hela kontexten innebär bl.a. att en aktiv konstruktion som *Anna kysste Kalle*, enligt ett kognitivt-grammatiskt synsätt, inte har samma betydelse som den passiva konstruktionen *Kalle kysstes av Anna*, då samtliga nyanser och bilder som framkallas av de olika formuleringarna alltså är en del av deras respektive betydelse. En följd av denna breda definition av semantikens roll i språkssystemet är att semantiken, som Bialystok påpekar, blir en faktor i grammatikalitetsbedömningen:

In this broadened conception of meaning, semantics becomes part of decisions about grammatical well-formedness, making such judgments relative rather than absolute. Judging grammaticality is a problem-solving task that incorporates information from any source that may be relevant to the nature or purpose of the utterance and evaluates it against similarity to linguistic convention (Bialystok 2001:43).

Detta antagande är i hög grad relevant för de undersökningar som presenteras i föreliggande arbete, inte minst med tanke på analyserna av enkätaterialet i studie III. Här bör dock påpekas, att definitionen av begreppet *språklig konvention* och därmed också definitionen av konventionaliserade flerordsuttryck inom kognitiv grammatik skiljer sig något från den som tillämpas i detta arbete. Som framgår av citatet nedan rymmer begreppet alla konventionaliserade strukturer i ett språk, oavsett i vilken mån de är *kompositionella*, *genomskinliga* (se vidare om dessa termer i avsnitt 2.3) eller, med Langackers terminologi, *motiverade*.

The units comprising a grammar represent the speaker's grasp of linguistic convention. I do not limit the term "conventional" to structures that are arbitrary, unmotivated, or unpredictable: linguistic structures form a gradation with respect to their degree of motivation, and no coherence or special status attaches to the subclass of structures that are fully motivated or those that are fully arbitrary; speakers operate with the entire spectrum of structures as an integrated system (Langacker 1988:11).

I de arbeten som presenteras i denna avhandling tillämpas begreppet konventionalisering däremot i första hand på flerordsenheter som Langacker benämner arbiträra eller oförutsägbara, dvs. de som inte är helt kompositionella i den meningen att deras betydelse bara till viss del eller inte alls motsvarar summan av de ingående lexikala komponenternas betydelser.

Langacker beskriver ett språks grammatik som

..."a structured inventory of conventional linguistic units". By "unit" I mean a thoroughly mastered structure (i.e. a cognitive routine). A unit may be highly complex internally, yet it is simple in the sense of constituting a prepackaged assembly that speakers can employ in essentially automatic fashion, without attending to the details of its composition. Examples of units include a lexical item, an established concept, or the ability to articulate a particular sound or sound sequence (1988:11).

Det Langacker kallar för enheter kan alltså vara allt från enstaka morfem till längre fraser, som betraktas som enheter, då de fungerar som färdigbildade helheter (se även Bialystok 2001).

En aspekt som den kognitiva grammatiken kan sägas uppmärksamma i relativt liten utsträckning är den regelbundenhet som trots allt råder inom grammatiken och de olika typerna av enheter emellan. Detta problematiseras av Bialystok (2001:44), som påpekar att den kognitiva grammatiken i Langackers mening, genom att avvisa idén om en inre grammatik, bl.a. inte kan förklara strukturella likheter mellan olika språk. Ur språkutvecklingssynpunkt kan man dessutom tycka att det är svårt att förklara barns förmåga att lära in grammatiska strukturer utan större svårigheter, om man inte tar hänsyn till regelbundenheterna i språksystemet. Att förstaspråksinlärare skulle lära sig behärska sys-

tematiken som råder mellan de språkliga enheterna enbart på basis av empiriska belägg i interaktionen verkar, som Bialystok anmärker, vara en förenklad syn på språkutvecklingsprocessen (ibid.).

I samband med de undersökningar som presenteras i denna avhandling kan en del teoretiska utgångspunkter föras tillbaka på ett kognitivt grammatiskt förhållningssätt. Detta gäller t.ex. semantikens roll som central och rätt sammanlänkad med både lexikon och grammatik.

Många studier om flerordsenheters natur och användning har genomförts inom det teoretiska ramverk för konstruktionsgrammatik, Bialystok (2001:44) beskriver som en kusin till den kognitiva grammatiken. Den centrala teoretiska utgångspunkten inom denna teoribildning är att

”ALL LEVELS OF GRAMMATICAL ANALYSIS INVOLVE CONSTRUCTIONS: LEARNED PAIRINGS OF FORM WITH SEMANTIC OR DISCOURSE FUNCTION, including morphemes or words, idioms, partially lexically filled and fully general phrasal patterns (Goldberg 2006:5).

Här finns alltså tydliga beröringspunkter med föremålen för de undersökningar som presenteras i denna avhandling. Goldberg betonar vidare:

Any linguistic pattern is recognized as a construction as long as some aspect of its form or function is not strictly predictable from its component parts or from other constructions recognized to exist. In addition, patterns are stored as constructions even if they are fully predictable as long as they occur with sufficient frequency [...] (Goldberg 2006:5).

Vidare kan det ses som utmärkande för konstruktionsgrammatiken att studier inom denna forskningsansats ofta fokuserar detaljerade semantiska aspekter och distributionen av specifika ord, morfem och språkspecifika fraseologiska mönster. Man utgår från hypotesen att en beskrivning av de mångfacetterade semantiska, pragmatiska och formella restriktionerna i samband med dessa mönster även kan tillämpas på mer generella, enkla och regelbundna språkliga mönster (Goldberg 2006:5)

Flera av de teoretiska resonemang som nedan presenteras som relevanta för detta avhandlingsarbete kan sägas bygga på ett konstruktionsgrammatiskt synsätt (se Bybee 1998, 2006, Wray 2008). Av detta följer att även mina egna teoretiska resonemang i samband med vissa delar av analyserna har inslag av denna teoribildning, utan att det för den delen kan sägas att avhandlingen är skriven utifrån ett konstruktionsgrammatiskt förhållningssätt. En av de centrala utgångspunkterna för föreliggande arbete som uppvisar tydliga likheter med konstruktionsgrammatikens begreppsbildning är synen på konventionaliserade uttryck som *form-function* eller *form-meaning pairings* som delvis definieras genom att någon aspekt av deras form, betydelse eller funktion inte är förut-

sägbar utifrån de ingående konstituenterna (jfr Goldberg 2006:5). En tydlig skillnad mellan studierna som genomförts inom ramen för detta avhandlingsarbete och studier inom konstruktionsgrammatik är att man diskuterar språklig form på ett betydligt mer formellt sätt inom den senare forskningstraditionen än vad som är fallet i samband med mina analyser (jfr Bialystok 2001:45).

2.2.2. Det mentala lexikonet i kognitiva modeller

En fråga som diskuterats livligt i samband med kognitivt inriktad forskning om fraseologiska enheter och konventionaliserade uttryck är hur dessa är representerade och på vilket sätt de är tillgängliga i språkanvändarens mentala lexikon. Antaganden om hur minnet och det mentala ordförrådet är organiserade är naturligtvis svåra att föra i bevis och förblir i viss mån alltid något spekulativa. Det finns en rad olika teoretiska modeller som skiljer sig i vissa grundläggande antaganden, bl.a. om hur olika kognitiva mekanismer samverkar, i vilken mån det finns ett integrerat eller flera olika (i viss mån interagerande) system för exemplarrepresentation och regel användning och hur systemet är organiserat på detaljnivå. Samtidigt överlappar dessa modeller också på en del punkter. Jag kommer i det följande att presentera några modeller som är relevanta för det perspektiv som intas i denna avhandling.

Wray (2008) presenterar flera nyckelantaganden i samband med sin senaste forskning om det som hon kallar för *morpheme equivalent units* (MEUs). En MEU är enligt Wray

a word or word string, whether incomplete or including gaps for inserted variable items, that is processed like a morpheme, that is, without recourse to any form-meaning matching of any sub-parts it may have (Wray 2008:12).

Som Wray betonar, speglar denna definition i första hand vissa teoretiska antaganden om det som hon kallar för *formulaic language* och om dess roll i språk-systemet och representation i det mentala lexikonet. Detta innebär att definitionen inte är lämpad för att identifiera fraseologiska enheter i naturligt språkbruk eller för att hitta väl lämpade exempel för mer experimentella studier. Det handlar alltså snarare om ett teoretiskt antagande än om en operationaliserbar definition. Huruvida fraseologiska enheter processas helt automatiskt, utan någon som helst medvetenhet om enskilda komponenters betydelse, kan diskuteras (jfr Wray 2008:12). Det relevanta i samband med undersökningarna i denna avhandling är emellertid ett annat teoretiskt antagande som återspeglas i definitionen av MEUs, nämligen att lexikonet betraktas som heteromorfiskt (ibid.). Med andra ord består ordförrådet av såväl enstaka ord och morfem som holistiska flerordsenheter.

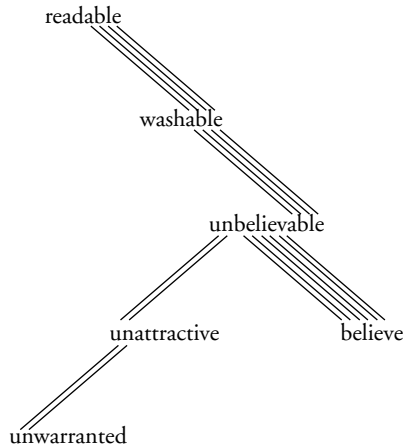
The theoretical position taken in this book [...] is that linguistic material is stored in bundles of different sizes [...] That is, the mental lexicon contains not only morphemes and words but also many multiword strings, including some that are partly lexicalized frames with slots for variable material, treated as if they were single morphemes [...] (Wray 2008:12).

Som antyds ovan, är det naturligtvis inte helt oproblematiskt att göra antaganden om hur det mentala lexikonet exakt är organiserat, bl.a. annat med tanke på relationen mellan det som Wray kallar för morfemekvivalenta enheter (MEUs) och deras komponenter i det heteromorfska lexikonet. Av förklarliga skäl förblir även Wrays modell något vag i det avseendet, vilket samtidigt gör den flexibel med tanke på den variation av fraseologiska enheter som förekommer i språkbruket.

The heteromorphic lexicon model (Wray 2002b) permits multiple partmappings of the same information in the lexicon. Alongside an entire phrase, that is stored whole with its associated meaning and phonological form, subparts [...] may also be stored. Not all possible subcomponents are necessarily so stored, nor are subcomponents necessarily recognisable grammatical constituents, though they may be (Wray 2008:15). The determination of which components are separately stored and, thus, how much flexibility there is for varying the form of the complete string, depends on input evidence and user need [...] (Wray 2008:15).

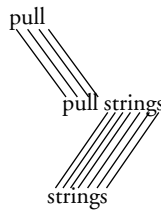
Wray föreslår dessutom att många flerordsenheter inte är lagrade i helt fixerad form utan som ”partly-fixed frames”, vilket, enligt henne, är en förutsättning för en effektiv integration av det hon kallar för *formulaicity* och kreativ regel användning (2008:16, se även Studie III).

Även Bybee (1998, 2006) utgår ifrån en modell som integrerar exemplarrepresentation och regel användning i ett heteromorft lexikon. Hon presenterar relationen mellan holistiska flerordsenheter, enskilda ord och olika typer av mindre enheter som ett nätverk där större enheter är sammanlänkade med de mindre enheter som utgör deras beståndsdelar. Detta gör det möjligt, enligt Bybee, att ta fram större enheter (t.ex. konventionaliserade flerordsuttryck) i sin helhet från minnet samtidigt som de är tillgängliga för analys genom länkar till olika konstituenten. Hon beskriver nätverket som i figurerna 2 och 3 nedan.



FIGUR 2: *Exempel relationen mellan lexikala enheter och mindre enheter som utgör deras komponenter (Bybee 1998:3).*

Figur 2 visar alltså relationerna mellan fonem, morfem och större lexikala enheter i Bybees modell. Bybee tillämpar samma nätverksmodell när det gäller relationen mellan flerordsenheter, t.ex. idiom, och deras lexikala komponenter, vilket illustreras i figur 3.



FIGUR 3: *Exempel på relationen mellan ett idiom och dess lexikala komponenter (Bybee 2006:713)*

Huruvida det finns en länk mellan en större lexikal enhet och en viss komponent beror, enligt Bybee, på om komponenten är representerad som en självständig enhet i nätverket, vilket i sin tur beror på enhetens frekvens i inflödet. Med andra ord har inte alla enheter som är representerade i nätverket samma status: ”those that are used frequently have very strong representations, or possibly even multiple representations, but those that are used infrequently, or are potential but have not been used, may have no representation” (Bybee 1998: 423). Hur nätverket ser ut, och i vilken mån ett uttryck som lagras som helhet

kan länkas till olika enheter och därmed bli analyserbart, beror alltså på inlära-rens tidigare språkliga erfarenheter, dvs. på i vilken mån dessa enheter i sin tur finns lagrade i minnet (se vidare Studie III). Viktiga frågor i sammanhanget är alltså vad språkbrukaren antas lagra i lexikonet och i vilken mån och när fler-ordsenheter processas som holistiska enheter respektive görs till föremål för analys och regelanvändning. Wray (2002, 2008) föreslår att lexikonets innehåll bestäms av det hon kallar för *Needs Only Analysis* (NOA), vilket – enkelt uttryckt – innebär att barnet vid förstaspråksinläring inte automatiskt i ett visst utvecklingsstadium börjar analysera allt språkligt inflöde. Istället analyseras endast strukturer där den sociala kontexten visar sig kräva förmågan att variera och därmed även förmågan att använda språkssystemets regler (Wray 2002:131). Språkliga strukturer som aldrig varieras när det gäller form eller pragmatik – som t.ex. sånger och ramsor men även kortare ordförbindelser, t.ex. konventionaliserade uttryck som *thank you for having me* – är, enligt Wray, alltså inte heller föremål för analys utan lagras och används som färdiga enheter. Detta innebär också att eventuella grammatiska oregelbundenheter eller atypiska användningar av enskilda komponenter inte uppmärksammas. Den atypiska användningen av verbet *have* i *thank you for having me* lagras alltså, enligt detta resonemang, inte som en separat enhet i förstaspråksinlärares lexikon (ibid.). Som Wray påpekar är språkbrukarnas behov dynamiska och olika för olika grupper av språkbrukare och en högst relevant fråga i sammanhanget är naturligtvis hur motsvarande utveckling av det mentala lexikonet ser ut för sena andraspråksinlärare. Det senare diskuteras vidare i avsnitt 2.3.4 i samband med flerordsenheterens roll i andraspråksinläringen.

2.2.2.1. Det flerspråkiga lexikonet

Särskilda aspekter som rör det flerspråkiga mentala lexikonet är, med tanke på informantgrupperna i detta arbete, naturligtvis högst relevanta här. Det ska dock också påpekas att ingen av studierna behandlar användningen och behärskningsen av konventionaliserade uttryck från ett kontrastivt perspektiv i någon djupare mening.

Tidiga studier i flerspråkighet har, som Pavlenko påpekar, undersökt frågan om flerspråkighet innebär två separata ordförråd som är länkade till en uppsättning av begrepp eller om även begreppsbildningen sker separat för varje språk. De flesta forskningsresultat tyder hittills på att flerspråkiga individer har ett lexikon för varje språk länkade till samma uppsättning av begrepp (Pavlenko 2009:125). Bialystok (2001), som ger en liknande bild av forskningsläget, nämner dock också en, enligt henne, trovärdig modell av Dijkstra & van Heuven (1998) som förespråkar ett integrerat flerspråkigt lexikon. Bialystok

skriver vidare att forskarna i stort sett är eniga om att båda ordförråden är aktiva när tvåspråkiga individer använder vart och ett av sina språk (2001:102).

En annan relevant aspekt är i detta sammanhang relationen mellan lexikala enheter på de olika språken och deras betydelser, begrepp och referenter i den reella världen. Bialystok menar att enkla dikotomier som mellan oberoende och integrerade ordförråd inte styrks av empirin. Det flerspråkiga mentala lexikonet kan i stället antas vara mer komplext.

Adequate descriptions of the organization of mental representations include the effect of such factors as level of proficiency and the circumstances of second-language learning. Moreover it appears likely that multiple arrangements can even coexist in the mind of an individual speaker [...] (Bialystok 2001:103).

Bialystok antar att flerspråkiga barns lexikala utveckling är lika komplex i förhållande till mental representation som hos vuxna andraspråksinlärare. Språk och begrepp antas vara sammanlänkande på ett dynamiskt sätt och strukturen ses som öppen för förändring för att kunna anpassas till ändrade språkliga behov och ökad språkfärdighet (Bialystok 2001:103, jfr Wray 2008). Även Pavlenko (2009) beskriver komplexa relationer mellan lexikala enheter och begrepp i det flerspråkiga mentala lexikonet. Hon refererar till forskning under det senaste decenniet (för en översikt se Jarvis & Pavlenko 2008 med anförda referenser) som inte i första hand undersöker uppsättningen av begrepp är gemensam eller separat för olika språk hos den flerspråkiga individen utan snarare vad hos de specifika begrepp som är gemensamt för de olika språken och vad som inte är det. Inom denna forskning använder man termen *linguistic category* som beteckning för konceptuella mentala representationer sammanlänkade med bl.a. lexikala enheter. De forskningsfrågor som har ställts är bl.a. relaterade till jämförelser av dessa språkliga kategoriers struktur mellan flerspråkiga individer och enspråkiga modersmålsalare av språken ifråga (Pavlenko 2009:126). Ett problem i sammanhanget är att lexikala översättningsekvivalenter inte nödvändigtvis behöver vara konceptuella ekvivalenter. Med andra ord kan en kategori på ett språk motsvara flera på det andra språket och därmed också flera olika begrepp. I det senare språket finns det då alltså två lexikala enheter som är länkade till var sitt begrepp. Som exempel kan relationerna mellan (eng.) *jealousy* och (sv.) *avund* och *svartsjuka* anföras (jfr Pavlenko 2009:134ff), där *jealousy* kan sägas vara länkade både till begreppet 'avund med anledning av en annan persons lycka, framgång, materiella förmögenheter etc.' och 'svartsjuka i en intim relation'. Ordet *avund* är däremot endast länkade till det förra och *svartsjuka* endast till det senare begreppet. Dessutom kan det finnas språkliga kategorier som helt saknar motsvarighet i det andraspråket. Pavlenko nämner som exempel bl.a. eng. *privacy* och *personal space* som, enligt hennes forskningsresultat, verkar sakna konceptuell ekvivalens i ryska.

En slutsats som kan dras utifrån den forskning som refereras ovan – och som i så fall har stora konsekvenser för synen på andraspråksinläring – är enligt Pavlenko att de centrala processerna vid senare andraspråksinläring är konceptuell omstrukturering och utveckling, snarare än bildandet av nätverk mellan redan existerande begrepp och de nya lexikala enheterna på andraspråket.

...the studies reviewed here reveal that the center of L2 vocabulary learning are the processes of *conceptual restructuring*, that is readjustment of the category structure and boundaries in accordance with the constraints of the target linguistic category, and *conceptual development*, that is development of new multimodal representations that allow speakers to map new words onto real-world referents similar to native speakers of the target language (Pavlenko 2009:141).

Dessa slutsatser leder alltså till starka teoretiska antaganden om själva andraspråksinläringens natur och det kan diskuteras om de undersökningar som Pavlenko (2009:126ff) refererar verkligen rättfärdigar en sådan drastisk förändring av synen på ordinläring. Utifrån den forskning som genomförts inom ramen för detta avhandlingsarbete kan inga slutsatser dras angående dessa frågor. Problemet är dock relevant i den meningen att många konventionaliserade flerordsenheter är relativt språkspecifika, dvs. de kan sakna fullständig konceptuell ekvivalens i andra språk. Ett sådant exempel skulle kunna vara det ryska konventionaliserade uttrycket *kak ob stenu goroch* som ordagrant kan översättas med 'som ärtor mot väggen' och betyder 'effektlös'. Uttrycket har konceptuella ekvivalenter och översättningsekvivalenter i andra slaviska språk (t.ex. polska) men saknar helt konceptuell ekvivalens i t.ex. de germanska språken. Pavlenkos antaganden skulle alltså i så fall kunna vara en del av förklaringen till de svårigheter som inlärare även på mer avancerad nivå upplever vid både receptiv och produktiv användning av konventionaliserade uttryck på andraspråket.

2.2.3. Metafor-teori

Som antytts tidigare, är kognitiv semantik och närmare bestämt konceptuell metafor-teori relevant för samtliga tre studier som presenteras inom ramen för detta arbete. I delstudie I används termen *tankemönster*, som kan sägas ligga nära termen begrepp och idén om domänmappning inom konceptuell metafor-teori som går tillbaka på Lakoff & Johnson (bl.a. 1980). Med tankemönster menas själva bilden, begreppet eller den metaforiska grundtanken som ligger till grund för ett visst figurativt uttryck. Lakoff & Johnson skriver

Primarily on the basis of linguistic evidence, we have found that most of our ordinary conceptual system is metaphorical in nature. [...]. To give some idea of what it could mean for a concept to be metaphorical [...], let us start with the concept ARGUMENT and the conceptual metaphor ARGUMENT IS WAR. This metaphor is reflected in our everyday language by a wide variety of expressions:

ARGUMENT IS WAR

Your claims are *indefensible*.

He *attacked every weak point* in my argument.

His criticisms were *right on target*.

[...] (Lakoff & Johnson 1980:4)

Som ett annat ofta citerat exempel på konceptuell metafor kan THE DEVELOPMENT OF A LOVE RELATIONSHIP IS A JOURNEY anföras (Langlotz 2006:67, jfr Lakoff 1993), där kunskap om resors eller resandets väsen överförs till kunskap om kärleksförhållanden (Lakoff 1993:207). Denna konceptuella metafor kan språkligt realiseras i yttranden som *our relationship has hit the buffers, look how far we've come* eller *it's been a long and winding road* (Langlotz 2006:66f, jfr Lakoff 1993:206). Enligt Langlotz fyller konceptuella metaforer bland annat funktionen att göra en abstrakt måldomän, som KÄRLEKSRELATION i det ovan nämnda fallet, tillgängligare genom referens till en mer konkret domän med en tydligare struktur som RESA.

Instead of manipulating the empirically-abstract domain LOVE RELATIONSHIP, the cognitive agent can draw inferences about it by referring to the more clearly structured and more directly comprehensible domain [...] JOURNEY (Langlotz 2006:67).

Att förstå en domän innebär alltså att kognitivt manipulera och strukturera den. Mer konkreta domäner kan vi, som tidigare nämnts, *begripa*, i ordets rätta bemärkelse, genom att använda våra sinnen (se avsnitt 2.2.1). När det gäller abstrakta domäner som EMOTIONER, KÄRLEKSRELATIONER, TID, LIVET eller KOGNITION så kan dessa inte "testas" eller manipuleras och därmed begripas på samma sensomotoriska sätt: "they are not open to the direct formation of fine-grained cognitive models, but can only be accessed indirectly" (Langlotz 2006:68). Den kognitiva manipulation som ger en abstrakt domän en mer konkret, begriplig struktur kan alltså ske med hjälp av metaforisk överföring från ett mer direkt tillgängligt begrepp från en konkret källdomän (Langlotz 2006:68).

En senare utveckling inom metafor-teorin är Fauconnier & Turners idé om *conceptual blending* eller *conceptual integration* (Fauconnier 1997, Fauconnier

& Turner 1998, Turner & Fauconnier 2000).² Dessa forskare anser att synen på metaforisk överföring som den beskrivits ovan är för enkelriktad och restriktiv för att beskriva vissa mer komplexa metaforiska överföringar (se Langlotz 2006:70). Inom blendteorin ses domänerna vid en metaforisk överföring i större utsträckning som interagerande än i traditionell konceptuell metafor-teori. Överföringen ses alltså inte bara som koppling mellan käll- och måldomän. Snarare handlar det om att fler domäner eller, med Fauconniers terminologi, *mentala rymder* är inblandade i den kognitiva processen (Svanlund 2001:16ff). Figur 4 beskriver strukturen i en *blend*, som alltså är den domän som innehåller element både från källdomänen (här *Input 1*), från måldomänen (här *Input 2*) och utöver detta vissa element som varken ingår i käll- eller måldomän. Det som Fauconnier kallar för *generic space* är domänen med den abstrakta struktur som är gemensam för input 1 och input 2 och som därmed möjliggör en metaforisk överföring emellan dem (Fauconnier 1997:149, se även Svanlund 2001:17). Som exempel används det konventionaliserade uttrycket *gräva sin egen grav* som enligt Svenskt språkbruk betyder att 'förstöra för sig själv genom omdömeslösa handlingar' (2003:425). Källdomänen är alltså GRAVGRÄVANDE och måldomänen kan beskrivas som SJÄLVFÖRVÅLLAT PERSONLIGT MISSLYCKANDE. Som illustreras i figur 4, är den metaforiska överföringen mellan dessa två domäner dock inte tillräcklig för att göra metaforen begriplig:

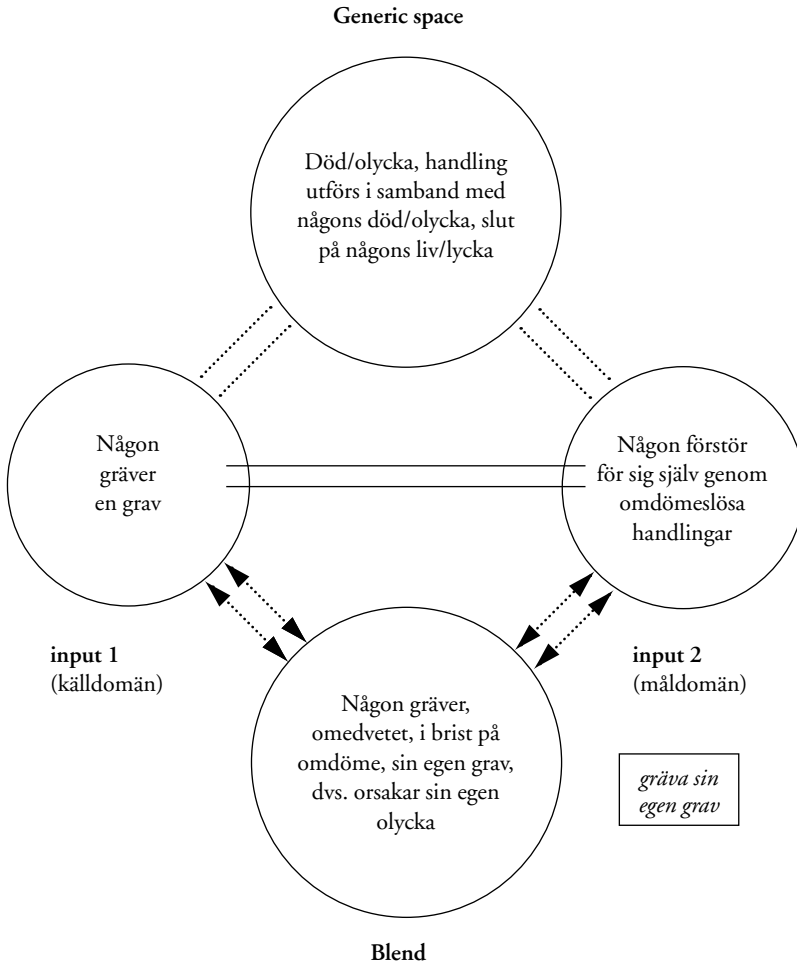
...gravgrävandet (källdomänen) är inte direkt och utan vidare projicerat till självförvållade personliga misslyckanden (måldomänen). I stället integreras vissa element från domänerna i en *blend*, som också uppvisar emergent struktur, dvs. drag som inte kommer från någondera av domänerna (Svanlund 2001:18).

I fallet *gräva sin egen grav* kan det konstateras att det uppstår en ny logisk struktur i blenden. Den kausala strukturen kan inte överföras från input 1 till input 2; "[f]oolish actions cause failure but grave digging does not cause death. It is typically someone's death that prompts others to dig a grave".³ Här kan man alltså prata om en inverterad kausal struktur i den konceptuella blandningen, blenden. Vidare överförs den intentionella aspekten inte till blenden; dödgravaren i källdomänen utför handlingen med avsikt medan det bildliga grävandet av den egna graven i blenden innebär en oavsiktlig, missriktad handling. De semantiska rollerna från käll- och måldomän bevaras heller inte i blenden;

² Idén lanserades på sent 1980-tal och utvecklades under 1990-talet, från början inom ramen för Fauconniers *mental space theory* (se vidare bl.a. Fauconnier 1994, 1997).

³ Fauconnier fortsätter: "And if the grave is atypically prepared in advance, to secure a plot, to keep workers busy, or because the person is expected to die, there is still not the slightest causal connection from the digging to the dying" (1997:169).

grävaren och den döde är vanligtvis inte samma person i den verklighet som källdomänen bygger på, men blir ett i den konceptuella blandningen. Som antyds ovan är händelsernas ordning dessutom omkastad; i källdomänen grävs graven vanligtvis efter att personen ifråga har dött (Fauconnier 1997:169, se även Svanlund 2001:18).



FIGUR 4: *Conceptual blending* (efter Fauconnier 1997:150, exempel från Fauconnier 1997:168f)

Denna konceptuella blandning fungerar alltså som metafor trots att det alltså finns ”extraordinary mismatches” mellan input 1 och input 2 (Fauconnier

1997:169). Att den gör det bygger på att inferenserna som beskrivs ovan uppstår i blenden och sedan kan föras tillbaka till käll- respektive måldomän. Det är alltså med hjälp dessa inferenser – och inte enbart med hjälp av källdomänens konkreta struktur – som måldomänens struktur blir tillgänglig och den konceptuella metaforen SJÄLVFÖRVÅLLAD PERSONLIGT MISSLYCKANDE ÄR ATT GRÄVA SIN EGEN GRAV därmed begriplig (se Svanlund 2001:18).

Blendteorin har ingen direkt relevans för de analyser av figurativt språk som har genomförts i detta avhandlingsarbete. Resonemangen baseras här i stället på en mer traditionell syn på domänmappning (se Studie I, Studie II). Modellen har dock fått allt större betydelse för forskningen om metaforer och bör därför inte förbli onämnd i samband med studier om figurativa ordförbindelser. Konceptuella blandningar är heller inte att förväxla med det som i denna avhandling klassas som *kontaminationer*. Det senare begreppet innebär att element från två eller flera konventionaliserade uttryck blandas i en kontaminerad form (jfr bl.a. Hyltenstam 1992), t.ex. i *skriva rakt på pucken* (Khemiri 2003, jfr Studie I). Här blandas alltså komponenter från de konventionaliserade uttrycken *rakt på sak* och *raka puckar*, vilket kan antas bero på vissa likheter i uttryckens form och betydelse. Här är det alltså inte fråga om en konceptuell blandning. Att något av de kontaminerade uttrycken i vissa fall kan analyseras med hjälp av blendmodellen är inte relevant för själva kontaminationsbegreppet.

2.3. Fraseologi

Terminologin inom fraseologin – och i synnerhet i samband med det som i detta arbete kallas för *fraseologisk enhet* – är tämligen svårhanterlig. Det finns en lång rad olika benämningar för det fenomen som beskrivs här (och olika begrepp som hänger ihop med det), vilket illustreras av figur 5.

amalgams – automatic – chunks – clichés – co-ordinate constructions – collocations – complex lexemes – composites – conventionalized forms – F[ixed] E[xpressions] including I[dioms] – fixed expressions – formulaic language – formulaic speech – formulas/formulae – fossilized forms – frozen metaphors – frozen phrases – gambits – gestalt – holistic – holophrases – idiomatic – idioms – irregular – lexical simplex – lexical(ized) phrases – lexicalized sentence stems – listemes – multiword items/units – multiword lexical phenomena – noncompositional – nonproductive – nonpropositional – petrifications – phrasemes – praxons – preassembled speech – precoded conventionalized routines – prefabricated routines and patterns – ready-made expressions – ready-made utterance – recurring utterance – rote – routine formulae – schemata – semi-preconstructed phrases that constitute single choices – sentence builders – set phrases – stable and familiar expressions with specialized substances – stereotyped phrases – stereotypes – stock utterances – synthetic – unanalyzed chunks of speech – unanalyzed multiword chunks - units

FIGUR 5: *Ett urval av termer som används i samband med fraseologiska enheter (efter Figur 1.2. i Wray 2002:9)*

Som Burger (1998:33) nämner är fraseologerna idag dock i stort sett överens om den övergripande terminologin på forskningsområdet och huvudtyperna av fraseologiska enheter och konventionaliserade flerordsuttryck definieras på liknande sätt. Två av dessa huvudtyper kan sägas vara de som kallas för *kollokationer* och *idiom*, som är centrala begrepp i samband med de föreliggande arbetena (se vidare avsnitt 2.3.1).

2.3.1. Fraseologiska enheter och konventionaliserat språk

Det som ibland kallas för *prototypiska idiom* (jfr Fernando & Flavell 1981, Sköldberg 2004) eller *pure idioms* (Howarth 1998, se vidare nedan) exemplifierar Sköldberg med uttrycken *väcka ont blod* och *i ett nötskal* (2004:26). I engelskspråkig litteratur anförs gärna bl.a. *kick the bucket* och *pull someone's leg* som prototypiska exempel för denna typ av konventionaliserade uttryck (se även Moon 1998:4, Wray 2002:4). De för idiomerna utmärkande egenskaper som Sköldberg framhåller i sin definition är bl.a. en relativt fixerad form, figurativ betydelse, *icke-kompositionalitet* och *konventionalisering* (som hon kallar för *institutionalisering*, jfr Moon 1998:7). *Kompositionalitet* definieras av Partee (1984) på följande sätt:

The meaning of an expression is a function of the meaning of its parts and of the way they are syntactically combined (Partee 1984:281).

Kompositionalitetsbegreppet ligger nära det som i svensk litteratur ofta beskrivs som grad av semantisk *genomskeinlighet*, dvs. i vilken mån det är möjligt

att härleda frasens betydelse utifrån de enskilda konstituenternas betydelse. Svanlund (2009:36f) diskuterar begreppet genomskinlighet i förhållande till en sådan strikt syn på kompositionalitet. Genomskinlighet anges ofta som en av de centrala egenskaper hos ett uttryck som är nära kopplat till adjektiv av typen *förutsägbar* eller *härledbar*. Svanlund själv använder begreppet genomskinlighet på ett, som han anser, mer preciserat sätt i sin undersökning av hur nya sammansättningar etableras i språket:

Om vi uppfattar en sammansättning som genomskinlig så innebär det att vi lätt kan generera en föredragen defaultbetydelse. Detta innebär inte att andra betydelser är otänkbara, men en viss betydelse uppfattas som mer naturlig än andra kandidater och aktiveras därmed först – om ingenting i sammanhanget indikerar annorlunda (2009:36f).

Om ett uttryck är genomskinligt eller inte är alltså, enligt detta synsätt, en fråga om språkbrukarens upplevelse. Vet man vad ett uttryck betyder har man alltså bara sin egen intuition att tillgå för att avgöra hur pass genomskinligt uttrycket är, dvs. i vilken mån en språkbrukare som läser eller hör det för första gången kan förväntas tolka uttrycket på det konventionaliserade sättet. I samband med de arbeten som presenteras i denna avhandling är det även viktigt att poängtera att genomskinlighet – liksom kompositionalitet – i enlighet med Svanlund, bör ses som ett graduellt fenomen. Med andra ord kan den konventionaliserade betydelsen hos ett uttryck vara mer eller mindre uppenbar (Svanlund 2009:37).

En annat återkommande term i denna avhandling är *konventionalisering*. Termen är, som Svanlund påpekar, inte helt oproblematiskt att tillämpa med tanke på att konventionalisering måste ses som en ständigt pågående, dynamisk process, vilket innebär att uppsättningen av etablerade mönster kontinuerligt förändras (Svanlund 2001:33). Språklig konvention definieras i detta avhandlingsarbete, i enlighet med Svanlund (2001:11), som etablerade mönster i språkbruket gentemot nyskapade eller tillfälliga sätt att uttrycka ett visst innehåll. En fördel med begreppet är att det framhäver den diakrona aspekten som har betydelse för användningen av konventionaliserade uttryck och som uppstår genom att språkbrukarna inom en språkgemenskap upprepade gånger gör samma språkliga val som därmed så småningom etableras som ett naturligt sätt att uttrycka ett visst innehåll. För att avgöra graden av konventionalisering hos ett uttryck har olika referensmaterial i form av en frasordbok (Svenskt språkbruk 2003) samt olika korpusar i Språkbanken (<<http://spraakbanken.gu.se/>>) och i viss mån även Google⁴ (<www.google.se>) konsulterats.

⁴ För en diskussion kring Googles lämplighet som referensmaterial se bl.a. Moon (2008:28), Lorentzen (2008:260) och Studie III.

Som diskuteras vidare nedan, är idiom, liksom andra typer av konventionaliserade uttryck, ingen homogen kategori som kan ses som helt isolerad från andra typer av konventionaliserade uttryck. Vissa forskare har t.ex. använt idiombegreppet i så vid bemärkelse att det innefattar olika typer av konstruktioner som inte är fritt formade som kollokationer (se vidare nedan), men även icke-kompositionella sammansättningar som *blackbird* (se Makkai 1972, jfr Moon 1998:4). O’Keefe, McCarthy & Carter (2007:62) refererar till ”the ordinary idioms of every day” och avser bl.a. öppningsfraser i samtal (eng. *gambits*) av typen *guess what* och diskurspartiklar (jfr Coulmas 1979).

En definition som motsvarar tillämpningen av idiombegreppet i studierna i denna avhandling är den som Langlotz (2006) har presenterat:

An idiom is an institutionalised construction that is composed of two or more lexical items and has the composite structure of a phrase or semi-clause, which may feature constructional idiosyncrasy. An idiom primarily has an ideational discourse-function and features figuration, i.e. its semantic structure is derivationally non-compositional [or not fully compositional]. Moreover, it is considerably fixed and collocationally restricted (Langlotz 2006:5).

Även Langlotz är dock noggrann med att poängtera att gränsdragningen mellan idiom och intelligande fraseologiska enheter på intet vis är skarp (jfr O’Keefe et al. 2007:81f). Langlotz (2006:6) menar vidare att en sådan definition kan inkludera så pass olika konstruktioner som *red herring* (’villospår’), *take the bull by the horns* (’ta tjuren vid hornen’) och *fall flat* (’misslyckas’) (se även Studie II).

När det gäller kollokationsforskning skiljer man i huvudsak mellan två olika inriktningar, nämligen *kontextualistisk* (frekvensbaserad) och *systemorienterad* kollokationsforskning. I litteraturen framställs ofta Firth (bl.a. 1957) som den kontextualistiska kollokationsforskningens grundare även om Palmer (1933) använde begreppet långt tidigare (se Nation 2001:317). Den kontextualistiska kollokationsforskningen definierar kollokation som frekvent samförekomst (eng. *co-occurrence*) av två eller flera ord i en textkorpus (Malmgren 2003:124, jfr Stubbs 2001:29). Sinclair (1966, 1987) kan sägas ha vidareutvecklat kollokationsbegreppet efter Firth (1957) och Halliday (bl.a. 1966). Sinclairs teoretiska utgångspunkt är att man behöver tillämpa två olika principer för att kunna förklara en texts uppkomst. Här talar han om (a) *the open choice principle* och (b) *the idiom principle*. Medan man med hjälp av (a) kan förklara ett yttrandes uppkomst i den mån som det produceras genom kreativa val i alla dessa delar (ord) utifrån ett antal olika möjliga val, är (b) viktigt med tanke på valet av större – mer eller mindre fixerade – enheter, dvs. idiomatiska flerordskombinationer, bl.a. kollokationer. Både Halliday och Sinclair använder sig av begreppen *nod* (eng. *node*) och *kollokator* (eng. *collocate*). Med det

förra avses det lexikala element som studeras med avseende på dess "kollokationsbeteende" medan det senare begreppet betecknar de lexikala element som kollokerar med en viss nod. Halliday, McIntosh & Strevens (1964:33) exemplifierar denna relation utifrån ordet *chair* som nod: "Given the item 'chair', we are more likely to find in the same utterance the items 'sit' or 'comfortable' [...] than, say, 'haddock'..." . Enligt detta citat skulle alltså ordkombinationen *comfortable chair* klassas som kollokation, vilket gör tydligt att man ser kollokation som en frekvensbaserad företeelse där den statistiska sannolikheten av ordens samförekomst är ett avgörande kriterium. Ett problem i samband med denna syn på kollokationsbegreppet är att frekvens är en relevant aspekt, men att den i sig inte är tillräcklig för att definiera kollokationer som en "meningsfull kategori för en analys av lexikala mönster i språkssystemet" (Lindberg 2004:68, jfr Malmgren 2003:125). Kollokationsforskningen har dock avancerat betydligt, inte minst på 1970- och 80-talet tack vare stora empiriska studier i svenska (Allén 1975, jfr även Kjellmer 1994 för engelska) som har lett till en åtskillnad mellan "grammatisk välformade och rent tillfälliga kollokationer" (Malmgren 2003:125). Inom den datoriserade korpuslingvistikens har man utvecklat statistiska metoder för att kunna skilja mellan tillfälliga (om än frekventa) samförekomster av ord, t.ex. *och det är...* (jfr Biber & Conrad 1999 om *lexical bundles*), och ordförbindelser där det finns starka associationer konstituenterna emellan, som i kollokationen *fatta beslut* (jfr Ellis, Simpson-Vlach & Maynard 2008). Det är alltså inte bara möjligt att beräkna en ordförbindelses frekvens i en korpus, utan man kan även mäta det som kallas för *mutual information* (MI). MI är ett statistiskt mått som anger i vilken mån konstituenterna i en ordförbindelse förekommer tillsammans oftare än slumpen kan förutsäga; ju högre frasens MI desto starkare associeras de ingående orden med varandra (jfr Church & Hanks 1990).

Inom systemorienterad kollokationsforskning är det relationen konstituenterna emellan som fokuseras, snarare än ordförbindelsernas frekvens. Inom denna skola tillämpar man bl.a. begreppet *direktionalitet*, en aspekt som helt saknas i kontextualistisk kollokationsforskning. Direktionalitetsbegreppet utvecklades av bl.a. Hausmann (1985) inom ramen för den systemorienterade kollokationsforskningen i Tyskland (se Malmgren 2003:127). Hausmann, som gäller som hård kritiker av den kontextualistiska skolan, ser inte konstruktioner som *köpa en bok* som kollokationer, eftersom de är semantiskt triviala. Enligt ett systemorienterat synsätt kännetecknas en kollokation av att en av konstituenterna i viss mån är oförutsägbar och begränsad i sin kombinerbarhet (Malmgren 2003:127), som t.ex. verbet *dra* i *dra en slutsats*. Hausmann använder begreppen *bas* och *kollokator*, som alltså skiljer sig från kontextualisternas nod och kollokater i att det råder direktionalitet, dvs. ett slags hierarkiskt förhållande, emellan dem. I den ovannämnda kollokationen är *slutsats* basen och *dra*

kollokator. Denna direktionalitet kan, enligt Hausmann, motiveras av det faktum att man som språkbrukare kan tänkas leta efter det rätta verbet att använda tillsammans med substantivet *slutsats*, om man vill uttrycka innehållet i *dra en slutsats* på ett idiomatiskt sätt. Det skulle däremot vara föga meningsfullt att utgå ifrån verbet *dra* och leta efter rätt substantiv att använda i samma syfte (jfr Hausmann 1985:119ff, se även Malmgren 2003:127).

Kollokationsbegreppet som det tillämpats i detta avhandlingsarbete baseras på ett systemorienterat synsätt. De kollokationer som har undersökts består av en bas som starkt associeras med en eller ett fåtal kollokatorer. Kollokationerna, liksom övriga de fraseologiska enheter som undersökts i denna avhandling, är dessutom i viss mån *icke-kompositionella*. Med det menas att det utifrån basens betydelse inte är helt förutsägbart vilka som är de, enligt konventionen, prefererade kollokatorerna. Av denna definition följer att *fatta ett beslut* är en relevant kollokation, med substantivet *beslut* som bas och verbet *fatta* kollokator. *Läsa en bok* ses däremot inte som kollokation, dels eftersom det inte finns några starka restriktioner vad gäller konstituenternas kombinerbarhet med andra ord, dels för att frasen är helt kompositionell (jfr Malmgren 2003:127, Hausmann 1985, se vidare Studie III).

En viktig teoretisk utgångspunkt i denna avhandling är att tidigare nämnda egenskaper, som fixering i form, kompositionalitet och genomskinlighet, hos idiom och andra typer av fraseologiska enheter ses som graduella. Detta innebär i sin tur att gränserna emellan olika typer av konventionaliserade uttryck är flytande och kollokationer, idiom och andra typer av fraseologiska enheter således betraktas som del av ett kontinuum, snarare än från varandra isolerade fenomen (Wulff 2008:11, jfr bl.a. Moon 1998, Nattinger & De Carrico 1992). Detta kontinuum har beskrivits på något olika sätt av olika forskare på det fraseologiska området. Fernandos kontinuum för flerordsenheter utgår ifrån idiom och kollokationer som två separata typer av lexikala enheter som dock endast skiljer sig ifråga om varierbarhet (1996:32). Fernando beskriver vissa kollokationer som fritt varierbara (t.ex. *catch a bus/plane* etc. eller *beautiful/lovely/sweet woman*), medan idiomen aldrig ses som mer än begränsat varierbara (se även Wulff 2008:12). Detta ligger i linje med de forskningsresultat som Wulff (2008:1f) refererar, nämligen att man genom perceptionstest har visat att informanternas bedömningar av hur prototypiskt ett idiom är korrelerar med i vilken mån modifieringar av uttrycket ifråga bedöms som acceptabla enligt ett inföddlikt (dvs. idiomatiskt) språkbruk. Kompositionalitet är, enligt Fernandos modell, inget kriterium som diskriminerar mellan idiom och kollokationer. Hon beskriver uttryck som *arm in arm* 'arm i arm' och *in the meantime* 'under tiden' som *literal idioms*, vilket innebär att det förekommer kompositionella uttryck både bland de uttryck som hon klassar som idiom och de som klassas som kollokationer. Det kan dessutom noteras att Fernando även räknar

sammansättningar som *backlash* ('bakslag, häftig motreaktion') som idiom (Fernando 1996:32). Howarth (1998:28) presenterar en något enklare modell som skiljer mellan *free combinations*, *restricted collocations*, *figurative idioms* och *pure idioms*. Den första kategorin innehåller ordförbindelser som är sammansatta av komponenter som används i sin bokstavliga betydelse och är fritt utbytbara. *Restricted collocations* innerhåller en komponent som används i en speciell konventionaliserad, ofta figurativ betydelse som endast samförekommer i ett begränsat antal kollokationer. *Figurative idioms* har metaforisk betydelse såväl som en aktuell bokstavlig betydelse, medan *pure idioms* däremot är icke-kompositionella, ogenomskinliga och den mest fixerade kategorin i Howarths modell (1998:28). Figur 6 ger svenska exempel på Howarth *collocational continuum*).

	Fria kombinationer	Kollokationer	Figurativa idiom	Rena idiom
Lexikala kollokationer verb + substantiv	<i>blåsa i en trumpet</i>	<i>blåsa håret</i>	<i>blåsa i den stora trumpetten</i>	<i>blåsa ngn på konfekten</i>
Grammatiska kollokationer preposition + substantiv	<i>på bordet</i>	<i>på håll</i>	<i>på krigstigen</i>	<i>på håret</i>

FIGUR 6: *Kontinuum av fraseologiska enheter i svenska.*

Denna modell korresponderar väl med de typer av fraseologiska enheter som har studerats inom ramen för detta avhandlingsarbete. Som framgår av figur 5 är det flera av de ovannämnda faktorerna som ligger till grund för denna typologi; medan Howarth i den mindre prototypiska änden av detta kontinuum i första hand åberopar fixeringsgrad som kriterium, handlar kategoriseringen i den andra änden i högre grad om icke-kompositionalitet (se Wulff 2008:12). Modellen är väl lämpad för att illustrera de graduella skillnaderna mellan olika typer av fraseologiska enheter. Det finns däremot många grader mellan kategorierna i Howarths kontinuum, vilket är en anledning till att jag avstått ifrån en detaljerad kategorisering i de undersökningar som presenteras här. Kategorierna sammanfattas i stället som fraseologiska enheter, konventionaliserade figurativa ordförbindelser respektive konventionaliserade uttryck (se studie I, II respektive III) och beskrivs vid behov i detalj i termer av konventionalisering, fixering, kompositionalitet och genomskinlighet (jfr även Sköldberg 2004:25f).

2.3.2. Idiomaticitet

En grundläggande teoretisk utgångspunkt i samband med studier av fraseologiska enheter är det som ofta kallas för *idiomaticitet* (eng. *idiomaticity*). Ett problem med denna term är att den används med olika innebörd i litteratur på det fraseologiska området. I de arbeten som presenteras här används termen i enlighet med Yorio (1989), som försöker komma åt idiomaticitetsbegreppet genom att hänvisa till hur infödda talare uppfattar bristande idiomaticitet:

Idiomaticity is a non-phonological “accent”, not always attributable to surface language errors, but to a certain undefined quality which many frustrated L2 composition teachers define as “I don’t know what’s wrong with this, but we just don’t say that in English” (Yorio 1989:64).

Pawley & Syder (1983) använder termerna *nativelike selection* och *nativelike fluency* med en liknande innebörd. Med andra ord skulle man alltså kunna definiera idiomaticitet som ’det som gör att språkanvändning uppfattas som inföddlik’ (se även Studie III). I detta sammanhang är det, som Yorio (1989:64) antyder, viktigt att poängtera att idiomaticitet inte enbart är en fråga om grammatisk korrekthet. Som bl.a. Pawley & Syder har påpekat är andelen idiomatiska yttranden i ett språk, dvs. sådana som accepteras som naturliga av infödda talare, relativt liten jämfört med andelen grammatisk korrekta yttranden (1983:196).

Många forskare på fraseologiområdet använder termen idiomaticitet för att ange graden av bl.a. fixering, konventionalisering, kompositionalitet och genomskinlighet hos ett visst uttryck. Med denna terminologi är ett uttryck mer idiomatiskt ju mer fixerat och konventionaliserat och ju mindre kompositionalt och genomskinligt det anses vara (jfr bl.a. Coulmas 1979, Fleischer 1997, Gläser 1986, 1988, och Wulff 2008).

Begreppsförvirringen förvärras av att de olika användningarna av idiomaticitetsbegreppet delvis går i varandra; ett yttrande är inte idiomatiskt, dvs. inte inföddligt, eftersom det inte är konventionaliserat. Detsamma gäller delvis för det som här kallas mer eller mindre *prototypiskt* med avseende på idiom eller konventionaliserade uttryck. (Det sistnämnda är alltså den innebörd som bl.a. Fleischer (1997) och Wulff (2008) lägger i begreppet idiomaticitet.) Ett idiom är mindre prototypiskt delvis därför att det är mindre konventionaliserat än andra, mer prototypiska uttryck. Det kan alltså vara svårt att hålla isär de olika användningarna av idiomaticitetsbegreppet. I detta arbete avses alltså den innebörd av *idiomatiskt* som framgår av Yorios definitionsförsök ovan, dvs. det som accepteras av infödda talare som naturligt, inföddligt språkbruk. De flersidigheterna som är aktuella här beskrivs som mer eller mindre prototypiska el-

ler, mer detaljerat, som mer eller mindre konventionaliserade, fixerade, kompositionella eller genomskinliga (se avsnitt 2.3.1 ovan).

2.3.3. Hur fasta är fasta fraser?

Variationen, även inom de kategorier av flerordsenheter som bl.a. Howarth (1998) beskriver som de mest fixerade i form, har diskuterats av Moon (1998). Hon använder begreppet *idiomschema* i samband med variationen av sådana uttryck som hon kallar *fixed expressions and idioms* (FEIs). Ett idiomschema grundar sig på en metafor eller ett tankemönster som är gemensamt för olika realiseringar av en fraseologisk enhet. Såväl strukturen som de lexikala konstituenterna kan alltså i vissa fall variera rätt avsevärt inom ett sådant schema, så länge tankemönstret förbli intakt, det vill säga så länge den fraseologiska enhetens olika realiseringar kan föras tillbaka till samma metaforiska grundtanke. En sådan modell för variationen hos fraseologiska enheter förklarar, enligt Moon [1998:163], också varför det oftast inte är något problem att avkoda även sådana realiseringar som ligger relativt långt ifrån ett uttrycks kanoniska form (Studie I:109). Begreppet *tankemönster* kan sägas ligga nära källdomänsbegreppet inom konceptuell metafor-teori (se avsnitt 2.2.3). Även Sköldberg (2004:74ff) har visat att de fraseologiska enheter som enligt Howarths modell hamnar i den mest fixerade kategorin, t.ex. *sätta käppar i hjulet/hjulen*, varierar i förhållande till den mest konventionaliserade formen, både ifråga om form och innehåll i en del av pressmaterialet i Språkbanken.

Avvikelser från konventionaliserade mönster är alltså inget rent andraspråksfenomen. Det är heller inte enkelt att veta, speciellt i naturlig produktion, i vilken mån avvikelser från ett uttrycks idiomatiska och konventionaliserade form är ett avsiktligt, språkligt kreativt grepp eller i vilken mån det beror på bristande kunskap om uttrycket ifråga, vilket naturligtvis gäller för både första- och andraspråksanvändare. Det finns däremot skäl att anta att avvikelser i förhållande till konventionaliserad användning av flerordsuttryck som produceras av första- och andraspråkstalare skiljer sig både i fråga om kvantitet och kvalitet (se t.ex. Hyltenstam 1992). Det sistnämnda är en central fråga i denna avhandling.

2.3.4. Fraseologiska enheter och andraspråksinläring

Till skillnad från de teoretiska modellerna för mental representation av lexikala enheter som utgår från integrerade nätverkssystem för både analys och automatiserad användning av flerordsenheter (se Bybee 1998, 2006, Wray 2008),

utgår vissa kognitivt inriktade andraspråksforskare som Skehan (1998) från att den mentala representationen av flerordsenheter fungerar med hjälp av ett s.k. *dual-mode system* (1998:88ff). Utgångspunkten för ett sådant system är antagandet att det finns två med varandra integrerade system som har olika roller i en individs språkutveckling. Med hjälp av ett *exemplarbaserat system* sker hel-frasinlärningen. De inlärdade flerordsenheterna är, på samma sätt som enkla ord, representerade som oanalyserade helheter i det mentala lexikonet. Genom att använda flerordsenheter sparar språkanvändaren processkapacitet vid pågående språkanvändning (Skehan 1998:89; jfr Pawley & Syder 1983). Det *regelbaserade systemet* behövs däremot för att inläraren ska kunna analysera de lexikala flerordsenheterna, vilket i sin tur är en förutsättning för att språkutveckling ska ske (jfr Wray & Perkins 2000).

Skehans (1998) modell skiljer sig alltså från de antaganden som görs av Wray (2008) och Bybee (1998, 2006) i det avseendet att den utgår ifrån två olika, om än interagerande kognitiva mekanismer för processande av holistiska flerordsenheter. Wray (2008) och Bybee (1998, 2006) antar, som tidigare nämnts, att det mentala lexikonet är heteromorft. Vad gäller Wrays antagande om Needs Only Analysis konstaterar hon att vuxna andraspråksinlärares språkliga behov naturligtvis skiljer sig från förstaspråksinlärares.

Adult foreign language learners have different needs (and often substantially different input) from native speaker children. [...]. [T]he adult mind seems better able than the child's to rationalize patterns explicitly [...]. [I]t is hypothesized that adult learners tend overall to analyse input more and to store smaller lexical units (cf. Kjellmer 1991). These then require additional rules to combine. This proposal offers an explanation for the difficulties reported for adult learners in relation to restricting their output to only the idiomatic patterns that they have encountered in input – that is, there is a tendency to express messages in alternative ways, which are meaningful but not nativelike (Wray 2008:19f).

Wrays hypotes innebär alltså att andraspråkstalare har ett större behov av regelanvändning än infödda talare, då det, enligt henne, finns skäl att anta att de i större utsträckning lagrar mindre lexikala enheter (dvs. fler enskilda ord) i sitt lexikon. De behöver alltså analysera detta inflöde i större utsträckning för att sedan kunna kombinera de mindre enheterna i längre yttranden (jfr Ekberg 1998).

Den ovannämnda hypotesen styrks bl.a. av en del korpuslingvistiska studier kring användningen av fraseologiska enheter i inlärartexter i engelska som främmandespråk. Gemensamt för dessa studier är att resultaten visar att inlärarna, jämfört med infödda talare av engelska, har en tendens att underanvända konventionaliserade uttryck i sitt skrivande (se t.ex. Granger 1998, jfr Wiktorsson 2003).

Även Ekberg (1997, 1998, 1999, 2004) har i sina studier av användningen av konventionaliserade uttryck bland flerspråkiga barn i Rosengård kommit fram till att de enspråkiga barnen i hennes kontrollgrupp använder ”fler konventionaliserade strukturer på alla språkliga nivåer än de tvåspråkiga barnen” (2004:272, jfr Wray 2008:19f).

Går vi ännu ett steg längre i resonemanget kan man tänka sig att detta förhållande tyder på att de tvåspråkiga barnen har relativt sett färre ”färdiga val” av lingvistiska strukturer, och därmed i större utsträckning får förlita sig på grammatikens produktiva regler i kommunikationen (Ekberg 2004:272).

Av de forskningsresultat som refereras ovan följer att andraspråksanvändare tenderar att använda färre konventionaliserade strukturer och strategier i vissa givna språksituationer än infödda brukare av språket ifråga. På så sätt kan andraspråkstalare sägas använda språkssystemet mer kreativt än förstaspråkstalare som är mer stereotypa i sina språkliga val. ”Och intressant nog är det kanske just stereotypin som i vissa fall utmärker ett idiomatiskt språkbruk” (Ekberg 2004: 274).

3. Material och metod

3.1. Material

Materialet för studierna I och II har samlats in inom ramen för det ovannämnda projektet *Språk och språkbruk bland ungdomar i flerspråkiga storstadsmiljöer*. Sammanlagt har åtta gymnasieklasser och 222 elever från åtta olika skolor (tre i Stockholm, tre i Malmö och två i Göteborg) deltagit i SUF-projektet. Samtliga skolor ligger, i något varierande grad, i flerspråkiga och flerkulturella områden (se Ganuza 2008:34ff). Allt material har avidentifierats och skolorna har, liksom informanterna, fått kodnamn. Skolorna som valts ut för projektet har, som tidigare nämnts, det gemensamt att andelen elever med utländsk bakgrund, som enligt definitionen i Skolverkets databaser innebär att eleven och/eller båda föräldrarna är födda utomlands (se <<http://salsa.artisan.se/definitioner.htm>>), är relativt hög. Skolorna skiljer sig dock sinsemellan både vad det gäller just andel elever med utländsk bakgrund och en del andra faktorer. Tanken har varit att välja två gymnasieklasser i årskurs 2 från varje stad, en med hög andel elever med utländsk bakgrund och en där fördelningen mellan eleverna med svensk respektive utländsk bakgrund är något jämnare (se Ganuza 2008:34f). Materialet består av drygt 300 timmar inspelat tal och fler än 200 uppsatser från elevernas nationella prov i svenska B (se Studie I). Det muntliga materialet spelades in i olika situationer i skolan, t.ex. under lektioner med hela klassen, grupparbeten, diskussioner i mindre grupper och muntliga redovisningar, samt intervjuer om elevernas språkliga bakgrund (se vidare nedan). Dessutom har en del av informanterna fått med sig inspelningsutrustning hem för att spela in sig själva med vänner eller familj på fritiden. Vissa medarbetare i projektet har också spelat in material under mer kontrollerade förhållanden, t.ex. i form av styrda fokusgruppsdiskussioner, återberättelser av film eller serieteckningar (se t.ex. Tingsell 2007 och Ganuza 2008) och genomfört ett grammatikalitetsbedömningstest. Bakgrundsintervjuerna har genomförts utifrån en enkät med frågor om bl.a. eleverna födelseland, modersmål, startålder

för inlärnigen av svenska, språkanvändning i hemmet och på fritiden, föräldrarnas bakgrund, födelse-land och modersmål, elevernas medievanor och framtidsplaner angående utbildning och arbete. Intervjuaren har antecknat elevens svar i enkäten under samtalet, som samtidigt spelades in. Bakgrundsdata för informanterna i SUF-projektet är alltså, liksom i enkätundersökningen, självrapporterade, då det är svårt att komma åt den aktuella typen av information på annat sätt. Samtidigt är metoden inte helt oproblematisk, bl.a. då eleverna själva kan vara osäkra och ha svårt att minnas i vissa fall, t.ex. när det gäller deras startålder för inläring av svenska.

I studierna I och II som presenteras i detta arbete har enbart delar av det material som samlats in i samband med SUF-projektet använts. Studie I är enbart baserat på materialet som samlats in i Göteborg. För att kunna undersöka flerspråkiga ungdomars behärskning av olika typer av konventionaliserade uttryck, har jag inför Studie III samlat in ett kompletterande material i form av en enkätundersökning. På så sätt har varit möjligt att studera konventionaliserade uttryck i svenskan bland ungdomar i flerspråkiga miljöer både i muntligt material, i text samt utifrån ett speciellt konstruerat test.

Från början har 220 informanter i sammanlagt nio gymnasieklasser från två olika skolor i Göteborgsområdet deltagit i enkätundersökningen i studie III. 186 av dessa bildar en informantgrupp från flerspråkiga skolmiljöer som sedan vidare delats in i olika grupper efter språklig bakgrund och startålder för andraspråksinläring. 26 informanter tillhör en enspråkig jämförelseklass. Resterande informanter har av olika anledningar tagits bort ur analysen. De skolor som har deltagit i enkätundersökningen har tilldelats kodnamnen *Kampusgymnasiet* och *Prismagymnasiet*. Skillnaden mellan skolorna är relativt stor, både vad gäller andel elever med flerspråkig bakgrund och deras upptagningsområden. Kampusgymnasiet är en innerstadsskola, som tar in elever från många olika områden i Göteborg. Andelen elever med utländsk bakgrund på Kampusgymnasiet är 40 %. Prismagymnasiet ligger i en förortskommun. Man kan alltså anta att skolan, bl.a. med tanke på det geografiska läget, tar in relativt många elever som bor i närliggande områden, som kan sägas vara relativt homogent svenskspråkiga. Andelen elever med utländsk bakgrund på Prismagymnasiet är 17 %.

3.2. Kvantitativa och kvalitativa metoder

En kombination av olika metoder har använts i de tre undersökningarna och i samtliga fall har analyserna haft både kvalitativa och kvantitativa inslag.

Ett kvantitativt inslag i undersökningarna har varit att redovisa förekomster av olika typer av ordförbindelser och konventionaliserade uttryck i de olika materialen och att jämföra frekvenser av förekomst mellan olika informantgrupper.

Korpuslingvistiska metoder, som alltid är av huvudsakligen kvantitativ karaktär (se Stubbs 2001), har också använts i alla tre studier. I första hand har olika korpusar i Språkbanken (<<http://spraakbanken.gu.se>>) använts som standardspråklig referens. För att få en uppfattning om ordförbindelsen ifråga förekommer i det allmänna språkbruket och i så fall hur frekvent den verkar vara har den sökts i det ovannämnda korpusmaterialen. I viss mån har även sökmaskinen Google (<www.google.se>) använts i detta syfte (för en diskussion om användningen av webben som korpus se Moon 2008:28, jfr Lorentzen). Korpuslingvistiska metoder har dessutom till viss del använts i konstruktionen av enkäten, med hjälp av vilken materialet för studie III har samlats in.

Enkätundersökningen som metod kan i de alla flesta fall sägas ha både kvantitativa och kvalitativa inslag (Trost 2007:20), vilket också gäller studie III. Självva utgångspunkten är dock i detta fall en kvantitativ ansats, då forskningsfrågorna till stor del syftar på att synliggöra vissa mönster i svaren hos ett större antal informanter. Här är jag således intresserad av hur många (men även vilka) informanter som svarar på det förväntade, konventionaliserade sättet. Enkätsvaren tolkas alltså i första hand på ett kvantitativt sätt. Skillnader och samband mellan dessa kvantitativa resultat för olika informantgrupper och olika komponenter i enkäten (t.ex. enskilda konventionaliserade uttryck) har signifikansprövats med hjälp av statistiska analyser (se Studie III, jfr Trost 2007, Körner & Wahlgren 2005). En del av materialet har också tolkats kvalitativt, i den meningen att jag har undersökt hur de svar som inte motsvarar den efterlysta konventionaliserade formen ser ut.

Den kvalitativa beskrivningen av materialet har varit en viktig del i alla tre undersökningarna. En fråga har i detta sammanhang varit hur olika avvikelser från konventionaliserade ordförbindelser ser ut i materialet. I studie II har den kvalitativa beskrivningen av vilka typer av figurativa ordförbindelser som förekommer varit en viktig del av analysen, bl.a. när det gäller exakt hur dessa ordförbindelser kan se ut och vilka källomäner nybildningar bygger på.

4. Sammanfattning av artiklarna

I de följande avsnitten beskrivs i korta ordalag forskningsfrågorna samt de viktigaste teoretiska utgångspunkterna och resultaten av de tre studierna som ingår i denna avhandling.

I. Fraseologiska enheter på flerspråkig grund. En jämförelse av verklighet och fiktion

Det har på senare år förts en stundtals livlig debatt kring multietniska ungdomsvarieteter och deras roll i flerspråkiga ungdomars utveckling av andraspråket svenska. Det är bl.a. mot bakgrund av denna debatt som projektet *Språk och språkbruk i flerspråkiga storstadsmiljöer* har initierats. Dessutom har romanen *Ett öga rött* (Khemiri 2003) spelat en viss roll i debatten. Författaren använder bl.a. förvrängningar av fraseologiska enheter för att skapa huvudpersonens idiolekt, som används i syfte att konstruera en multietnisk identitet. I detta arbete används romanen som referenspunkt i beskrivningen av fraseologiska enheter i ungdomsspråk i flerspråkiga miljöer. Data från SUF-projektet jämförs med avseende på användningen av fraseologiska enheter dels med utdrag ur Khemiris roman, dels med ett standardspråkligt referensmaterial. Undersökningen har genomförts utifrån följande frågeställningar:

- Hur kan man beskriva olika avvikelser i förhållande till standardspråket i bruket av fraseologiska enheter hos ungdomarna i SUF-materialet?
- Hur kan man beskriva olika avvikelser i förhållande till standardspråket i bruket av fraseologiska enheter i Khemiris roman *Ett öga rött*?
- Vilka likheter och skillnader kan observeras mellan det avvikande bruket av fraseologiska enheter i ungdomarnas och romanfigurens språkbruk?

Moons (1998) modell för beskrivning av variationen hos fraseologiska enheter med olika grad av fixering i termer av *idiom schemas*, har tillämpats för att beskriva variationen i de två olika materialen i förhållande till ett standardspråkligt jämförelsematerial. Modellen har utvecklats i viss mån, bl.a. för att, utöver lexiko-grammatiska avvikelser, även kunna beskriva avvikelser på ett semantiskt-pragmatiskt plan. Andra relevanta begrepp för analysen är *kontamination*, dvs. en blandning av komponenter från två eller flera olika fraseologiska enheter, och *approximation*, dvs. ordförbindelser som skiljer sig endast marginellt från den standardenliga formen av den fraseologiska enheten i fråga (se Hylténstam 1992).

I den analyserade delen av romanen visar sig avvikelserna från standardenligt bruk av de fraseologiska enheterna i fråga i första hand ligga på lexiko-grammatisk nivå. När det gäller SUF-materialet är variationen däremot delvis mer komplex, då avvikelserna i vissa fall ligger på såväl lexiko-grammatisk som semantisk-pragmatisk nivå.

II. *Klättra på väggarna eller bara vara ett med soffan.* Om figurativa ordförbindelser bland ungdomar i flerspråkiga skolmiljöer

I denna artikel, som är samförfattad med Emma Sköldberg, beskrivs en studie av figurativa ordförbindelser i texter skrivna av gymnasieelever som lever och studerar i flerspråkiga urbana miljöer i Sverige. Materialet består av sammanlagt 175 uppsatser som skrivits inom ramen för det nationella provet i Svenska B och har samlats in i samband med SUF-projektet. Texterna har delats in i fem olika kategorier, beroende på skribenternas språkliga bakgrund. (Detta arbete har genomförts av Julia Prentice.)

Författarna har utgått ifrån följande forskningsfrågor:

- I vilken omfattning använder eleverna figurativa ordförbindelser i de analyserade texterna?
- Vilka typer av figurativa ordförbindelser finns det i materialet och hur konventionaliserade är de?
- Finns det skillnader mellan olika informanternas användning av figurativa ordförbindelser, och kan de eventuella skillnaderna kopplas till informanternas språkliga och kulturella bakgrund?

Författarna har gått igenom textkorporus och manuellt excerperat figurativa ordförbindelser var för sig och sedan gemensamt tagit ut det slutgiltiga materialet på sammanlagt 839 figurativa ordförbindelser. Dessa har sedan analyserats i samarbete mellan båda författarna.

Analysen visar att de figurativa ordförbindelserna i materialet kan klassificeras i tre huvudkategorier efter grad av konventionalisering: (1) *konventionaliserade ordförbindelser*, (2) *delvis modifierade konventionaliserade ordförbindelser* och (3) *nybildade ordförbindelser*. Konventionaliserade ordförbindelser visar sig vara den mest frekventa typen av ordförbindelser i materialet, följd av de delvis modifierade ordförbindelserna. Nybildade figurativa ordförbindelser är därmed den minst frekventa typen i elevernas texter. Resultaten bekräftar också den ursprungliga hypotesen att förekomst och typer av figurativa ordförbindelserna i texterna i viss mån kan kopplas till informanternas språkliga bakgrund. Förstaspråkseleverna använder fler konventionaliserade figurativa ordförbindelser än andraspråkseleverna. Modifieringar av konventionaliserade figurativa ordförbindelser är däremot mer frekventa i andraspråkselevernas texter (jfr Ekberg 1997, 98, 99). Studien visar dessutom att det finns stor variation i vilka källdomäner eleverna väljer att utgå från vid skapande av nya figurativa ordförbindelser.

III. *Käppen i hjulen*. Behärskning av svenska konventionaliserade uttryck bland gymnasieelever med varierande språklig bakgrund

Tidigare forskningen tyder på att språklig bakgrund, och inte minst startåldern för inlärningen, är en viktig faktor när det gäller behärsningen av konventionaliserade uttryck i ett andraspråk (se bl.a. Abrahamsson & Hyltenstam 2009). Denna studie syftar till att närmare undersöka i vilken mån gymnasieelever med varierande språklig bakgrund behärskar, dvs. känner igen, förstår och kan producera, några olika typer av konventionaliserade svenska uttryck som de

kan förväntas möta och använda. Studien är baserad på en kognitiv teoriram och ett sekundärt syfte är att ge en bättre bild av vilka typer av konventionaliserade uttryck som vållar mest problem för informanterna samt vad svårigheterna i samband med dessa konventionaliserade uttryck kan tänkas bestå i. Frågeställningarna för analysen är följande:

- Finns det skillnader i behärskning av konventionaliserade uttryck mellan olika elevgrupper
 - som kan relateras till elevernas språkliga bakgrund, dvs. modersmål och/eller startålder för inläring?
 - vad gäller olika aspekter av lexikal kompetens, t.ex. förståelse och produktion?
- Finns det skillnader i behärskning av enskilda konventionaliserade uttryck mellan elevgrupper med olika startålder för inläring?
- På vilket sätt avviker eleverna från konventionaliserade mönster i sin produktion av konventionaliserade uttryck?

Som framgår av forskningsfrågorna har det krävts såväl kvantitativa som kvalitativa metoder för att besvara dem. Materialet har samlats in med hjälp av av en enkät som delats ut till sammanlagt tio gymnasieklasser på två olika gymnasieskolor. Enkäten består av fyra olika deltest, ett förståelsetest i flervalsformat, två lucktest och ett matchningstest, som alltså syftar till att studera elevernas receptiva och produktiva behärskning av olika typer av konventionaliserade uttryck i svenskan. Sammanlagt 186 enkäter har inkluderats i analysen. Resultaten för gymnasieklasserna har jämförts, varefter informanterna delats in i fyra startåldersgrupper, vars olika delresultat har jämförts. En del av materialet har dessutom analyserats kvalitativt, för att undersöka vilka typer av avvikelser från de aktuella konventionaliserade strukturerna som producerats av eleverna med lägst resultat i varje grupp respektive eleverna med högst resultat i undersökningen överlag.

Analysen visar att det finns signifikanta skillnader i elevernas testresultat som kan kopplas till informanternas startålder för inläringen av svenska; förstaspråkseleverna (för vilka inlärningsstarten sammanfaller med födseln) når signifikant högre resultat än andraspråkseleverna i samtliga deltest i enkäten. De signifikanta skillnaderna mellan förstaspråkseleverna och andraspråkseleverna med den tidigaste inläringen för svenska ligger i linje med en del tidigare forskning som visat att även avancerade andraspråksinlärare ställs inför betydande svårigheter när det gäller ett idiomatiskt bruk av konventionaliserade uttryck i andraspråket (se bl.a. Wray 2002, Hyltenstam 1992, Abrahamsson & Hyltenstam 2009). Analysen visar dessutom på en rad signifikanta skillnader mellan andraspråksgrupperna med olika startålder, vilket bekräftar hypotesen

att startålder för inlärnigen av svenska är en viktig faktor när det gäller behärsknigen av den typ av konventionaliserade uttryck som har studerats i föreliggande undersökning. Tittar man på resultaten för de enstaka uttrycken tyder dessa på att graden av genomskinlighet och kompositionalitet är faktorer som i viss mån kan påverka elevernas resultat.

De kvalitativa resultaten visar vissa (inte helt oväntade) samband mellan typ av avvikelser från de väntade konventionaliserade formerna och informanternas språkliga bakgrund; andraspråkseleverna väljer i större uträkning än eleverna med svenska som modersmål att avstå från att lösa uppgifterna medan förstaspråkselever verkar vara mer benägna att pröva olika typer av lösningar med hjälp av olika strategier när de inte behärskar ett uttryck. Det förekommer bl.a. en del kontaminerade former som tyder på att enskilda komponenter i ett uttrycksfragment kan väcka associationer till andra konventionaliserade uttryck som innehåller delvis samma komponenter. Detta är intressanta resultat, då de stödjer vissa teoretiska antaganden om mental representation av lexikala enheter (jfr Bybee 1998, 2006).

5. Återstående frågor och fortsatt forskning

Flera av de aspekter som har tagits upp i denna avhandling behöver undersökas vidare och såväl SUF-materialet som enkätundersökningen kan tjäna som underlag för ytterligare studier på området. Något som är värt att undersöka djupare med utgångspunkt i resultaten som presenteras här är användningen av konventionaliserade uttryck i talade ungdomsvarieteter i flerspråkiga miljöer. En sådan studie kan med fördel baseras på systematiskt utvalda delar av det muntliga SUF-materialet. En relevant forskningsfråga skulle i det sammanhanget vara vilka typer av konventionaliserade uttryck som används i talspråket bland ungdomar i flerspråkiga miljöer. En annan intressant fråga är om det finns konventionaliserade uttryck som är kännetecknande för svenska på mångspråkig grund (SMG) generellt. Dessutom skulle man utifrån SUF-materialet kunna studera skillnader i ungdomarnas användning av konventionaliserade uttryck mellan de regionala varieteterna som talas i storstadsområdena. Med hjälp av samtalsanalytiska metoder kan också utforskas vilken funktion de konventionaliserade uttryck som används i talade ungdomsvarieteter i flerspråkiga miljöer kan ha (jfr Svensson 2009).

Det vore också intressant att relatera resultaten av en sådan undersökning till de resultat som presenteras i Tingsell (2007), Ganuza (2008) och Svensson (2009), som har undersökt valda delar av det muntliga SUF-materialet. En sådan jämförelse kan tjäna till att närmare utreda i vilken mån de olika språkliga drag som studerats, som fonetiska särdrag (Bodén, bl.a. 2007), användning av reflexivt och personligt pronomen (Tingsell 2007), syntaktiskt variation (Ganuza 2008) och användningen av diskurspartiklar (Svensson 2009) samt användningen av olika typer av konventionaliserade uttryck kan beskrivas som olika drag i en gemensam talad varietet (eller eventuell flera liknande talspraksvarieteter) bland ungdomar i flerspråkiga storstadsmiljöer.

Från ett inlärningsperspektiv vore en omfattande undersökning av sambanden mellan behärskningen av konventionaliserade uttryck och ordinlärning i ett andraspråk mer generellt intressant. Ett möjligt förfarande skulle i detta sammanhang vara att korrelera kvantitativa resultat från ett test, liknande det som använts i studie III (med en något mer utvecklat metodik och i större skala), med resultat av test som mäter olika aspekter av informanternas ordkunskap, som t.ex. ordförrådets storlek (eng. *vocabulary breadth*), produktiv ordkunskap och ordkunskapens djup (eng. *vocabulary depth*) (se Milton 2009, jfr Henriksen 1999, Nation 2001). Resultaten av en sådan studie skulle dessutom kunna sättas i relation till statistiska mått, som frekvens och mutual information (MI). Dessa mått kan tas fram för de testade konventionaliserade uttrycken med hjälp av korpuslingvistiska metoder. Resultaten kan dessutom jämföras med de test och korpuslingvistiska undersökningar som genomförts inom ramen för OrdiL-projektet (se Lindberg & Johansson Kokkinakis 2007, jfr Studie III). En sådan storskalig studie kan, i kombination med kvalitativa data, bl.a. bidra till en bättre förståelse av flerordsenheterens roll i ordinlärningen.

Summary

The overall aim of the present thesis is to investigate and describe the use and mastery of different types of phraseological units and conventionalized expressions in the language of adolescents in multilingual environments. The subject is investigated from different angles in three studies, which are summarized individually after an outline of the essential terminology and the most important theoretical issues in the present thesis. The issue of phraseological units and conventionalized structures in second language use and youth language in multilingual environments is interesting from several different perspectives. The relevant research areas are phraseology and first and second language acquisition, development and use. The present studies are carried out from a mainly cognitive perspective and the research questions focus on frequency and deviation from standard Swedish in the participants' use of phraseological units, on figurativeness and degree of conventionalization in the use of phraseological units and on the participants' mastery of Swedish conventionalized expressions.

Data and methodology

The present studies have been performed within the research project *Language and language use among adolescents in multilingual urban settings (SUF)*. The project is carried out in cooperation between the universities of Stockholm, Gothenburg and Lund. The data collected for this project consists of more than 300 hours oral recordings and more than 200 essays written for the national exam in Swedish by students in upper secondary school. The students were recorded in different situations in their school environment and were al-

so interviewed about their linguistic background, language choices and language use in general. For the first and second study of this thesis, both written and oral data from the project have been used, whereas for the third study, complementary data were collected by means of a survey administered to a participant group, similar to the one in the larger research project as for age and language background. The survey consists of four different tests focusing on different aspects of the participants' mastery of conventionalized Swedish expressions as well as a short questionnaire aimed at mapping information about the participants' linguistic background (see Appendix 1). A total of 186 surveys have been included in the analysis.

Both quantitative and qualitative methods have been applied in all three studies. Deviations from conventionalized forms have been described in terms of cognitive, SLA related as well as phraseological aspects. Quantitative and statistical methods have been applied for the comparison of frequencies of different types of expressions and deviations from standard Swedish use of conventionalized expressions in different data sets and between different participant groups.

Conventionalized expressions and idiomaticity

There are many types of phraseological units and many different terms have been used to describe them. There are, however, certain types that play a central role in this thesis and which are therefore described in some detail.

Idioms are often described as the most fixed *conventionalized, non-compositional* and semantically *non-transparent* phraseological units in the lexicon. However, there are many different kinds of idioms, collocations and other types of phraseological units, and fixedness in form, *compositionality* and semantic *transparency* have to be seen as gradual properties. This implies that also the differences between types of phraseological units should be seen as gradual and that different types can be viewed along a continuum (cf. Wulff 2008:12). The definition of idiom applied in the present thesis corresponds to the one presented by Langlotz (2006):

An idiom is an institutionalised construction that is composed of two or more lexical items and has the composite structure of a phrase or semi-clause, which may feature constructional idiosyncrasy. An idiom primarily has an ideational discourse-function and features figuration, i.e. its semantic structure is derivationally non-compositional [or not fully compositional -JP]. Moreover, it is considerably fixed and collocationally restricted (Langlotz 2006:5).

This definition includes, according to Langlotz (2006:6), different types of constructions, e.g. *take the bull by the horns*, *red herring* and *fall flat* (see study II).

The definition of *collocation* as applied in this thesis is based on a system oriented view (cf. Malmgren 2003:126ff). A collocation has a *base* that is strongly associated with one *collocate* or a restricted number of collocates. Furthermore collocations are to some degree non-compositional. From this definition follows that *draw a conclusion* is a collocation with the noun base *conclusion* and the verb collocate *draw*. *Read a book*, on the other hand, is not seen as a relevant collocation (cf. Malmgren 2003:127, see study III).

A related issue is the notion of idiomaticity. The term has been used in somewhat different ways both within the fields of second language acquisition (SLA) and phraseological research. In this thesis, the term corresponds to *native-like*, used for example by Pawley & Syder (1983).

Idiomaticity is a non-phonological “accent”, not always attributable to surface language errors, but to a certain undefined quality which many frustrated L2 composition teachers define as “I don’t know what’s wrong with this, but we just don’t say that in English” (Yorio 1989:64).

As Yorio points out above, the notion of idiomaticity, in the present sense, is not easy to define because it is based on a consensus within a language community on which expressions are natural or native-like, i.e. a consensus on how to express certain contents in the language in question.

Conventionalized expressions in second language acquisition

Several studies in the field of SLA have found that *age of onset* is an important factor in relation to the use of conventionalized language patterns and expressions (e.g. Abrahamsson & Hyltenstam 2009). Earlier research has also shown that native speakers seem to have a larger set of conventionalized patterns in their lexicon to choose from than L2 learners (e.g. Ekberg 1997, 1998, 1999). This often makes the learners’ language less stereotypical but at the same time less idiomatic (Ekberg 2004). L2 learners, on the other hand, seem to rely more on creative and less conventionalized language choices.

Different models and theories have been developed to describe the acquisition of conventionalized multiword expressions in both first and second language (e.g. Wray & Perkins 2000, Skehan 1998, Bybee 2006, Wray 2008).

What these models have in common is the need to deal with the fact that multiword expressions have to be stored as holistic units and analysable word combinations at the same time, for the learner to achieve both full mastery of the expressions and reasonable fluency in the language. Crucial differences between some of the models concern theoretical assumptions about the cognitive mechanisms that are active in the acquisition process and the organisation of the mental representation for the learned lexical units. Some researchers propose two different systems or modes for holistic storage of multiword units and grammatical analysis, which initially operate separately but successively start to interact (Skehan 1998) or even become integrated to some degree (Wray & Perkins 2000) as language development progresses. Other researchers have proposed a heteromorphic system where both single words and multiword units are stored together and are available for automatic retrieval from memory and grammatical analysis at the same time (Bybee 2006, Wray 2008). Keeping in mind that all theoretical models for the organisation of mental lexical representation are based on assumptions that are difficult to prove, the latter view is adopted in this thesis on the ground that the assumption of a heteromorphic lexical system is more compatible with other theoretical assumptions made here, e.g. that properties like formal fixedness, compositionality and semantic transparency are viewed as gradual.

Like Ekberg (1997, 1998, 1999), Wray (2008) proposes that L2 learners have access to a smaller set of multiword units than native speakers, which implies a bigger need for analysis and application of grammatical rules:

[I]t is hypothesized that adult learners tend overall to analyse input more and to store smaller lexical units (cf. Kjellmer 1991). These then require additional rules to combine. This proposal offers an explanation for the difficulties reported for adult learners in relation to restricting their output to only the idiomatic patterns that they have encountered in input – that is, there is a tendency to express messages in alternative ways, which are meaningful but not nativelike (Wray 2008:19f).

Concepts and metaphor

As evident from the previous sections, all three studies have been carried out from a cognitive perspective. Thus one of the central issues is the relationship between conceptualization and (more or less) conventionalized word combinations in the mental lexicon, which is especially interesting in the case of bilingual individuals. According to cognitive theories, the conceptual structures we use to organize the information about the real world are in one way or another linked to the lexicon. An issue that has been investigated in early studies

of bilingualism is whether the individuals' different languages are linked to the same set of concepts, or if conceptualization goes on separately for every language (Pavlenko 2009:125). Currently, most researchers have come to the conclusion that the bilingual individual's lexical systems (whether separate or integrated) are linked to the same set of conceptual structures (Pavlenko 2009:125, cf. Bialystok 2001:102). Recent studies have, however, given some indication that L2 vocabulary learning not only involves the building of new networks between the existing concepts and the new lexical items, but also considerable conceptual reconstruction. The reason for this is, according to Pavlenko, that not every concept that is linked to a lexical item in the L2 necessarily has an equivalent in the learners L1 (Pavlenko 2009:134ff, cf. Jarvis & Pavlenko 2008 with references).

Conceptual structures also form the basis for what is called *source domains* in conceptual metaphor theory. Metaphor is an important aspect when discussing conventionalized expressions, which often contain at least some figurative elements. A metaphor arises through conceptual transfer from a source domain to a target domain, i.e. when a semantic structure from a certain domain (often concrete) of experience is being used to express a meaning that belongs within another (often more abstract) domain of experience (Ekberg 1993:107, cf. Lakoff & Johnson 1980:4). An abstract concept can thus be made more comprehensible by mapping the structure of a more concrete concept on to it (cf. Langlotz 2006:65ff). Examples of conceptual metaphor are ARGUMENT IS WAR (Lakoff & Johnson 1980:4) and THE DEVELOPMENT OF A LOVE RELATIONSHIP IS A JOURNEY (Langlotz 2006, cf. Lakoff 1993). Because perception of the world and consequently conceptualization obviously is influenced by an individual's cultural background, the issue of metaphor and figurative expressions is interesting to investigate in a multilingual and multicultural environment. This is done in the second study.

I. Phraseological units on multilingual ground – a comparison of reality and fiction

Multiethnic youth varieties and their influence on second language development among adolescents in multicultural urban areas have been in focus for public debate in Sweden over the last decade. In this context, the novel *One Eye Red* (Sw. *Ett öga rött*) (Khemiri 2003), in which the leading character creates his own multiethnic idiolect, has played an important role. Since the modification of phraseological units and conventionalized expressions in Swedish is one of the stylistic features used to create this literary version of a multiethnic variety,

these modifications have been used as a point of reference in this study of deviations from conventionalized Swedish forms of phraseological units in authentic data collected in a multiethnic environment. Data from the SUF-project have been compared to standard language reference data on one hand, and part of the novel on the other hand. The research questions for the study have been the following:

- How can different types of deviations from standard Swedish use of phraseological units in the language of the participants in the SUF-project be described?
- How can different types of deviations from standard Swedish use of phraseological units in Khemiri's novel *One Eye Red* be described?
- What similarities and differences concerning deviating use of phraseological units can be observed in the comparison between the two data sets?

Moon's (1998) model, which describes variation of phraseological units with different degrees of formal fixedness in terms of *idiom schemas*, has been applied in the description of variation in the two data sets in relation to standard Swedish reference data. The model has been developed further to some degree, in order to be able to describe not only lexico-grammatical deviations, but also deviations on a semantic-pragmatic level. Other relevant terms for the analysis are *contamination*, i.e. a blend of components from two or several different phraseological units, and *approximation*, i.e. word combinations that deviate only marginally from the conventionalized form of the phraseological unit in question (cf. Hyltenstam 1992).

The deviations from standard Swedish use of the phraseological units found in the novel are predominantly lexico-grammatical whereas deviations in the SUF-data appear to be more complex, in the sense that in some cases the modified expressions deviate both lexico-grammatically and semantic-pragmatically from the standard Swedish reference material.

II. *Climb the walls or become one with the sofa?*

On the use of figurative word combinations by adolescents in multilingual school environments

This article, co-authored with Emma Sköldberg, reports on a study of figurative word combinations in essays written by upper secondary students who live and study in multilingual urban environments in Sweden. The data consists of

175 essays written as part of the participants' national Swedish exam. The texts have been divided into five different categories according to the authors' linguistic background.

The analysis has been based on the following research questions:

- To what degree do the students use figurative word combinations in the analyzed texts?
- Which types of figurative word combinations are found in the data and how conventionalized are those word combinations?
- Are there any differences between the participants' use of figurative word combinations which can be linked to the participants' linguistic and cultural background?

The analysis shows that the figurative word combinations in the data can be divided into three main categories, according to their degree of conventionalization: (1) *conventionalized word combinations*, (2) *partially modified conventionalized word combinations* and (3) *novel word combinations*. Conventionalized word combinations appear to be the most frequent type of figurative word combinations in the data, followed by partially modified word combinations. Thus, novel figurative word combinations are the least frequent type in the students' texts. The results also confirm the initial hypothesis that frequency and type of figurative word combinations to some degree can be linked to the participants' linguistic background. The L1 students use more conventionalized figurative word combinations than the L2 students. Modifications of conventionalized figurative expressions are, on the other hand, more frequent in the L2 students' texts (cf. Ekberg 1997, 1998, 1999). The study also shows that the students create novel figurative word combinations based on a rather large variety of source domains.

III. *The spanner into the works*. Mastery of Swedish conventionalized expressions among upper secondary school students with varying linguistic backgrounds

Several studies in the field of conventionalized language in second language acquisition point to age of onset as an important factor for the acquisition and idiomatic use of conventionalized expressions in a second language (e.g. Abrahamsson & Hyltenstam 2009). The aim of this study is to further investigate

to what degree upper secondary school students with varying linguistic backgrounds master, i.e. recognize, understand and are able to produce, different types of Swedish conventionalized expressions. The study has been carried out from a cognitive perspective and a secondary aim is to examine what types of expressions are shown to present most problems for the participants and in what way. The aim of the study is to answer the following research questions:

- Are there differences to be found between different groups of students
 - that can be related to the students' linguistic background, i.e. mother tongue and age of onset for Swedish?
 - concerning different aspects of lexical competence, e.g. receptive and productive knowledge?
- Are there differences in the mastery of individual expressions between participant groups with different ages of onset?
- In what ways do the students deviate from conventionalized patterns in their production of conventionalized expressions?

As indicated in the research questions above, both quantitative and qualitative methods have been applied in the analysis. The data has been collected in the form of a survey that has been administered in two different upper secondary schools in the Gothenburg area. The survey consist of four different tests, which aim at testing both perceptive and productive mastery of different kinds of Swedish conventionalized expressions. A total of 186 surveys were included in the analysis and the test results for the different classes were compared. Furthermore, the participants were divided into four groups according their age of onset for Swedish to enable comparisons between the groups' results for the whole survey, for the different tests and for individual expressions. Part of the data has also been analysed qualitatively in order to investigate which types of deviations from the conventionalized expressions in question have been produced by the participants with the lowest and the highest test results overall.

The statistical analysis shows significant differences between the students' test results that can be linked to their age of onset for Swedish. The L1 students (for which age of onset coincides with birth) reach significantly higher results than the L2 students in all of the tests in the survey. The significant differences between the L1 and the L2 students with the lowest age of onset are consistent with some earlier studies that have shown that even advanced L2 learners are experiencing considerable difficulties concerning the native-like use of conventionalized expressions in the L2 (e.g. Wray 2002, Hyltenstam 1992, Abrahamsson & Hyltenstam 2009). Furthermore, the analysis shows a number of significant differences between the L2 groups with different ages of onset, which confirms the initial hypothesis of age of onset as an important factor in

relation to mastery of the types of expressions investigated in the present study. The test results for the individual expressions indicate that degree of semantic transparency and compositionality are factors with a possible impact on the students' test results in the present survey.

The qualitative results indicate some (not entirely unexpected) connections between the types of deviations from the conventionalized forms in question and the participants' linguistic background; L2 students refrained from solving the tasks to a higher degree than L1 students, whereas L1 students who did not fully master a form seemed more ready to try different solutions and strategies in order to solve the problem. Some of the contaminated forms in the data indicate that participants make associations between individual components of two or several different expressions that partially contain the same components, a result which is in line with theoretical assumptions concerning the mental representation of lexical units (cf. Bybee 1998, 2006).

Conclusions

All three studies presented in this thesis investigate deviations from conventionalized linguistic constructions in the language of adolescents in Swedish multilingual settings from a cognitive perspective, focusing on different aspects of phraseology and conventionalized multiword expressions. The combined results provide insights into the kind of difficulties that especially L2 learners are confronted with in terms of idiomatic comprehension and production of phraseological units in the L2.

The comparison with literary constructed data in the first study reveals the complexity of the analysis of deviations and modifications of conventionalized expressions in language use in multilingual contexts. This complexity is further investigated in the second and third study, which confirm the conclusion that the variation in relation to conventionalized structures is not only a lexicogrammatical issue. The qualitative analysis in the third study shows for example that contaminations of two or several expressions can be caused by both formal and semantic associations between lexical components.

The studies also show that some types of deviations from the conventionalized forms, e.g. contaminations, are equally frequent among L1 students and L2 students (cf. Hyltenstam 1992). The results in studies II and III are, however, consistent in the sense that they confirm that deviations from conventionalized forms of the expressions in question are considerably more frequent in the data produced by L2 students and that the L1 students seem to have a larger set of conventionalized patterns to choose from (cf. Ekberg 1997, 1998,

1999, Hyltenstam 1992). Studies II and III also clearly highlight age of onset as a crucial factor in relation to an idiomatic, i.e. native-like, use of conventionalized expressions (cf. Abrahamsson & Hyltenstam 2009).

The sociolinguistic aspects of the use of conventionalized expressions in multiethnic youth language are briefly discussed in the first study, but need to be further investigated by means of a larger scale study of oral data, e.g. from the SUF-project.

One benefit of the kind of studies presented here is the detailed qualitative and quantitative analysis in relation to a number of individual conventionalized Swedish expressions given certain background variables. The results can shed light on ways in which participants with different linguistic backgrounds make associations between different lexical components and on the role of metaphor and figurative expressions in a multilingual context. Although there is still a need for further investigation and deeper analysis of the issues raised in this thesis, I hope that the combined results of the three studies can contribute to a better understanding of cognitive aspects concerning acquisition and use of multiword expressions in multilingual contexts and benefit further research within the field.

Referenser

- Abrahamsson, Niclas & Hyltenstam, Kenneth 2004. Mognadsbegränsningar och den kritiska perioden för andraspråksinläring. I: Hyltenstam, Kenneth & Lindberg, Inger, *Svenska som andraspråk – i forskning, undervisning och samhälle*. Lund: Studentlitteratur, s. 221–258.
- Abrahamsson, Niclas & Hyltenstam, Kenneth 2009. Age of Onset and Native-likeness in a Second Language: Listener Perception Versus Linguistic Scrutiny. *Language Learning* 59, s. 249–306.
- Allén, Sture 1975. *Nusvensk frekvensordbok 3. Ordförbindelser* [Frequency Dictionary of Present-Day Swedish 3. Collocations]. (Data Linguistica 10.) Stockholm: Almqvist & Wiksell.
- Bialystok, Ellen 2001. *Bilingualism in Development. Language, Literacy, & Cognition*. Cambridge: Cambridge University Press.
- Bialystok, Ellen & Hakuta, Kenji 1999. Confounded age: Linguistic and cognitive factors in age differences for second language acquisition. I: Birdsong, David (ed.), *Second Language Acquisition and the Critical Period Hypothesis*. Mahwah, NJ: Lawrence Erlbaum, s. 161–181.
- Bialystok, Ellen & Miller, Barry 1999. The problem of age in second-language acquisition: Influences from language structure and task. *Bilingualism: Language and Cognition* 2, s. 127–145.
- Biber, Douglas & Conrad, Susan 1999. Lexical bundles in conversation and academic prose. I: Hasselgard, Hilde & Oksefjell, Signe (eds.), *Out of Corpora: Studies in Honor of Stig Johansson*. Amsterdam: Rodopi, s. 181–189.
- Bijvoet, Ellen & Fraurud, Kari 2008. Svenskan i dagens flerspråkiga stors stadsmiljöer: en explorativ studie av unga stockholmars perceptioner av variation och varieteter. *Nordand* 3, s. 7–38.
- Birdsong, David 1992. Ultimate attainment in second language acquisition. *Language* 68, s. 706–755.

- Bloomfield, Leonard 1933. *Language*. New York: Holt, Rinehart & Winston.
- Bodén, Petra 2007. "Rosengårdssvensk" fonetik och fonologi. I: Ekberg, Lena (red.), *Språket hos ungdomar i en flerspråkig miljö i Malmö*. (Nordlund 27.) Lund: Språk- och Litteraturcentrum, Lunds universitet, s. 1–47.
- Bodén, Petra & Große, Julia 2006. Youth language in multilingual Gothenburg. I: *Proceedings FONETIK 2006, The Swedish Phonetics Conference. Lund, June 7–9, 2006*. (Working Papers 52.) Lund: Språk- och Litteraturcentrum, Lunds universitet, s. 17–20.
- Bongaerts, Theo 1999. Ultimate attainment in L2 pronunciation: The case of very advanced late L2 learners. I: Birdsong, David (ed.), *Second Language Acquisition and the Critical Period Hypothesis*. Mahwah, NJ: Lawrence Erlbaum, s. 133–159.
- Boyd, Sally och Fraurud, Kari, und. utg.: Challenging the homogeneity assumption in language variation analysis. Findings from a study of multilingual urban spaces. I: Auer, Peter & Schmidt, Jürgen Erich (eds.), *Language and Space – an International Handbook of Language Variation*. Berlin & New York: Mouton de Gruyter.
- Burger, Harald, 1998: *Phraseologie. Eine Einführung am Beispiel des Deutschen*. Berlin: Schmidt.
- Butler, Yuko G. 2000. The age effect in second language acquisition: Is it too late to acquire native-level competence in a second language after the age of seven? I: Oshima-Takane, Y., Shirai, Y. & Sirai, H. (eds.), *Studies in Language Science 1*. Tokyo: The Japanese Society for Language Science, s. 159–169.
- Bybee, Joan 1998. The Emergent Lexicon. I: Gruber, M. Catherine, Higgins, Derrick, Olson, Kenneth S. & Wysocki, Tamra (eds.), *CLS 34: The Panels*. Chicago: Chicago Linguistic Society, s. 421–439.
- Bybee, Joan 2006. From Usage to Grammar: The Mind's Response to Repetition. *Language* 82, s. 711–733.
- Church, Kenneth W. & Hanks, Patrick 1990. Word association norms, mutual information, and lexicography. *Computational Linguistics* 161, s. 22–29.
- Coulmas, Florian 1979. On the sociolinguistic relevance of routine formulae. *Journal of Pragmatics* 3, s. 239–266.
- DeKeyser, Robert M. 2000. The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition* 22, s. 499–533.
- Dijkstra, Ton & Van Heuven Walter J. B. 1998. The BIA model and bilingual word recognition. I: Grainger, Jonathan & Jacobs, Arthur M. (eds.), *Localist connectionist approaches to human cognition*. Mahwah, NJ: Erlbaum, s. 189–225.

- Ekberg, Lena, 1993. Verbet ta i metaforisk och grammatikaliserad användning. I: *Språk och stil* NF 3. S. 105–139.
- Ekberg, Lena 1997. Diskursiva och syntaktiska mönster i återberättelser hos invandrarbarn i Rosengård. I: Håkansson, Gisela et al. (red.), *Svenskans beskrivning 22. Förhandlingar vid Tjugoandra sammankomsten för svenskans beskrivning*. Lund: Lund University Press, s. 97–110.
- Ekberg, Lena 1998. Regeltillämpning kontra lexikonkunskap i svenskan hos invandrarbarn i Rosengård. I: Møller, Janus, Quist, Pia, Holmen, Anne & Norrman Jørgensen, Jens (red.), *Nordiske sprog som andetsprog. Københavnerstudier i tosprogethed 30*. København: Københavns universitet, s. 247–263.
- Ekberg, Lena 1999. Användningen av komplexa predikat hos invandrarbarn i Rosengård. I: Andersson, Lars-Gunnar, Lundqvist, Aina, Norén, Kerstin & Rogström, Lena (red.), *Svenskans beskrivning. 23. Förhandlingar vid Tjugotredje sammankomsten för svenskans beskrivning*. Lund: Lund University Press, s. 86–95.
- Ekberg, Lena 2004. Grammatik och lexikon i svenska som andraspråk på nästan infödd nivå. I: Hyltenstam, Kenneth & Lindberg, Inger (red.), *Svenska som andraspråk – i forskning, undervisning och samhälle*. Lund: Studentlitteratur, s. 259–276.
- Ekberg, Lena 2006. Bruket av sån i ungdomsspråk på mångspråkig grund. I: Ask, Sofia, Byrman, Gunilla, Hammarbäck, Solveig, Lindgren, Maria & Stille, Per (red.), *Lekt och lärt. Vänskrift till Jan Einarsson 2006*. Växjö: Växjö University Press, s. 65–73
- Ekberg, Lena 2007. ”Sån svensk å blond å sånt du vet” Lexiko-grammatiska drag i Malmöungdomars talspråk. I: Ekberg, Lena (red.) *Språket hos ungdomar i en flerspråkig miljö i Malmö*. (Nordlund 27.) Lund: Lund University, s. 48–77
- Ellis, Nick. C., Simpson-Vlach, Rita & Maynard, Carson 2008. Formulaic Language in Native and Second Language Speakers: Psycholinguistics, Corpuslinguistics and TESOL. *TESOL Quarterly* 42, s. 375–396.
- Fauconnier, Gilles 1994. *Mental spaces. Aspects of Meaning Construction in Natural Language*. Cambridge: Cambridge University Press.
- Fauconnier, Gilles 1997. *Mappings in Thought and Language*. Cambridge: Cambridge University Press.
- Fauconnier, Gilles & Turner, Mark 1998. Conceptual integration networks. *Cognitive Science* 22, s. 133–187.
- Fernando, Chitra 1996. *Idioms and Idiomaticity*. Oxford. Oxford University Press.
- Fernando, Chitra & Flavell, Roger 1981. *On idiom: Critical views and perspectives*. (Exeter Linguistic Studies 5.) Exeter: University of Exeter.

- Fillmore, Charles J. 1988. The mechanisms of construction grammar. I: Axmaker, Shelley, Jaisser, Annie & Singmaster, Helen (eds.), *Proceedings of the fourteenth annual meeting of the Berkeley Linguistic Society, February 13–15, 1988: General session and parasession on grammaticalization*. Berkeley: BLS, s. 35–55.
- Firth, John R. 1957. *Papers in Linguistics 1934–1951*. London: Oxford University Press.
- Flege, James E. 1999. Age of learning and second language speech. I: Birdsong, David (ed.), *Second Language Acquisition and the Critical Period Hypothesis*. Mahwah, NJ: Lawrence Erlbaum, s. 101–131.
- Fleischer, Wolfgang 1997. *Phraseologie der deutschen Gegenwartssprache*. (2., durchgesehene und ergänzte Auflage.) Tübingen: Niemeyer.
- Fraurud, Kari och Bijvoet, Ellen 2004. Multietniskt ungdomsspråk och andra varieteter av svenska i flerspråkiga miljöer. I: Hyltenstam, Kenneth och Lindberg, Inger (red.), *Svenska som andraspråk – i forskning, undervisning och samhälle*. Lund: Studentlitteratur, s. 389–418.
- Fraurud, Kari & Boyd, Sally 2006. The native/non-native speaker distinction and the diversity of linguistic profiles of young people in Swedish multilingual urban contexts. I: Hinskens, Frans (ed.), *Language variation – European perspectives: Selected papers from the Third International Conference on Language Variation in Europe (ICLaVE 3), Amsterdam, June 2005*. Amsterdam: John Benjamins, s. 53–70.
- Ganuza, Natalia 2008. *Syntactic Variation in the Swedish of Adolescents in Multilingual Urban Settings. Subject-verb order in Declaratives, Questions and Subordinate Clauses*. (Dissertation in Bilingualism 15.) Stockholm: Center for Research on Bilingualism, Stockholm University.
- García, Orfelia 2009. *Bilingual Education in the 21st Century. A Global Perspective*. Malden, MA: Wiley-Blackwell.
- Gläser, Rosemarie 1986. *Phraseologie der englischen Sprache*. Tübingen: Niemeyer.
- Gläser, Rosemarie 1988. The grading of idiomaticity as a presupposition for a taxonomy of idioms. I: Hüllen, Werner & Schulze, Rainer (eds.), *Understanding the lexicon. Meaning, sense and world knowledge in lexical semantics*. Tübingen: Niemeyer, s. 264–279.
- Goldberg, Adele 2006. *Constructions at Work. The Nature of Generalization in Language*. Oxford: Oxford University Press.
- Granger, Sylviane 1998. Prefabricated Patterns in Advanced EFL Writing: Collocations and Formulae. I: Cowie, Anthony P. (ed.), *Phraseology, Theory, Analysis and Applications*. Oxford: Oxford University Press, s. 145–160.

- Große [Prentice], Julia 2008. Fraseologiska enheter på flerspråkig grund. En jämförelse mellan verklighet och fiktion. *Nordand. Nordisk tidsskrift för andrespråksforskning* 3, s. 103–132.
- Grosjean, Francois 1982. *Life with Two Languages*. Cambridge: Harvard University Press.
- Halliday, M. A. K. 1966. Lexis as a linguistic level. I: Bazell, Charles. E., Catford, John C., Halliday, M. A. K. & Robins, Robert H. (eds.), *In Memory of J.R. Firth*. London: Longmans, s. 148–162.
- Halliday, M. A. K., McIntosh, Angus & Stevens, Peter D. 1964. *The Linguistic Sciences and Language Teaching*. London: Longmans.
- Hansson, Petra & Svensson, Gudrun 2004. *Listening for "Rosengårds Swedish"*. I: Branderud, Peter & Traunmüller, Hartmut (eds.), *Proceedings FONETIK 2004, The Swedish Phonetics Conference, May 26–28 2004*. Stockholm: Dept. of Linguistics, Stockholm University, s. 24–27.
- Hausmann, Franz Josef 1985. Kollokationen im deutschen Wörterbuch. Ein Beitrag zur Theorie des lexikographischen Beispiels. I: Bergenholtz, Henning & Mugdan, Joachim (Hrsg.), *Lexikographie und Grammatik. Akten des Essener Kolloquiums zur Grammatik im Wörterbuch 28.–30.6. 1984*. Tübingen: Niemeyer, s. 118–129.
- Henriksen, Birgit 1999. Three Dimensions of Vocabulary Development. *Studies in Second Language Acquisition* 21, s. 303–317.
- Howarth, Peter 1998. Phraseology and Second Language Proficiency. *Applied Linguistics* 19, s. 24–44.
- Hyltenstam, Kenneth 1992. Non-native features of Near-native speakers: On the Ultimate Attainment of Childhood L2 Learners. I: Harris, Richard (ed.), *Cognitive Processing in Bilinguals*. Amsterdam, London, New York & Tokyo: North Holland, s. 351–368.
- Hyltenstam, Kenneth & Abrahamsson, Niclas 2002. Startålder och slutlig behärskning hos nästan infödda talare av svenska som andraspråk. I: Hauksdóttir, A, Arbjörnsdóttir, B, Gardarsdóttir, M. & Porvaldsdóttir, S (red.), *Forskning i nordiske sprog som andet- og fremedsprog. Rapport fra konference i Reykjavík 23–25 maj 2001*. Reykjavík: Háskóli Íslands, Háskólaútgáfan.
- Hyltenstam, Kenneth & Abrahamsson, Niclas 2003. Age of onset and ultimate attainment in near-native speakers of Swedish. I: Fraurud, Kari & Hyltenstam, Kenneth (eds.), *Multilingualism in Global and Local Perspectives. Selected Papers from the 8th Nordic Conference on Bilingualism, November 1–3, 2001, Stockholm Rinkeby*. Stockholm: Rinkeby Institute of Multilingual Research, City of Stockholm and Stockholm University.
- Håkansson, Gisela & Norrby, Catrin 2005. Flicka – dricka eller tjej? Ordassociationer hos andraspråksinlärare. I: Byrman, Gunilla et al. (red.), *Svenskans beskrivning 27*. Växjö: Växjö University Press, s. 125–134.

- Jarvis, Scott & Pavlenko, Aneta 2008. *Crosslinguistic Influence in Language and Cognition*. New York: Routledge.
- Johnson, Jaqueline S. & Newport, Elissa L. 1989. Critical period effects in second language learning: The influence of maturational state on acquisition of English as a second language. *Cognitive Psychology* 21, s. 60–99.
- Khemiri, Jonas Hassen 2003. *Ett öga rött*. Stockholm: Norstedts.
- Kjellmer, Göran 1991. A mind of phrases. I: Aijmer, Karin & Altenberg, Bengt (eds.). *English Corpus Linguistics: Studies in Honour of Jan Svartvik*. London & New York: Longman, s. 111–127.
- Kjellmer, Göran 1994. *A Dictionary of English Collocations. Based on the Brown Corpus. I–III*. Oxford: Clarendon Press.
- Körner, Svante & Wahlgren, Lars 2005. *Statistiska metoder* (2. uppl.). Lund: Studentlitteratur.
- Lakoff, George 1993. The contemporary theory of metaphor. I: Ortony, Andrew (ed.), *Metaphor and Thought*. Cambridge: Cambridge University Press, s. 202–251.
- Lakoff, George & Johnson, Mark 1980. *Metaphors We Live By*. Chicago & London: The University of Chicago Press.
- Lamb, Sidney M. 2000. Neuro-cognitive structure in the interplay of language and thought. I: Pütz, Martin & Verspoor, Marjolijn, *Explorations in linguistic relativity*. Amsterdam: John Benjamins, s. 173–196.
- Langacker, Ronald 1988. An Overview of Cognitive Grammar. I: Rudzka-Ostyn, Brygida (ed.), *Topics in Cognitive Linguistics*. Amsterdam & Philadelphia: John Benjamins, s. 3–48.
- Langlotz, Andreas 2006. *Idiomatic Creativity*. Amsterdam & Philadelphia: John Benjamins.
- Lenneberg, Eric 1967. *Biological Foundations of Language*. New York: Wiley & Sons.
- Lindberg, Inger 2004. Språkanvändning – kreativt skapande eller förutsägbara mönster? I: Olofsson, Mikael (red.), *Symposium 2003, Arena Andraspråk*. Stockholm: HLS Förlag, s. 61–81.
- Lindberg, Inger 2009. I det nya mångspråkiga Sverige. *Utbildning och Demokrati* 18, s. 9–37.
- Lindberg, Inger, Boyd, Sally, Ekberg, Lena, Fraurud, Kari & Källström, Roger 2001. Språk och språkbruk bland ungdomar i flerspråkiga storstadsmiljöer. Projektansökan till Riksbankens Jubileumsfond, juni 2001.

- Lorentzen, Henrik, 2008: Korpusbaseret lemma-eksklusion eller ”hvor gik ordene hen da de gik ud?”. I: Jóhannesson, Kristinn, Landquist, Hans, Lundqvist, Aina, Rogström, Lena, Sköldberg, Emma & Wallgren Hemlin, Barbro (red.), *Nog ordat? Festskrift till Sven-Göran Malmgren*. (Meijerbergs Arkiv för svensk ordforskning 34.) Göteborg: Göteborgs universitet, s. 259–267.
- Makkai, Adam 1972. *Idiom structure in English*. The Hague & Paris: Mouton.
- Mackey, William F. 1970. The description of bilingualism. I: Fishman, Joshua A. (ed.), *Readings in the Sociology of Language*. The Hague: Mouton, s. 554–84.
- Malmgren, Sven-Göran 2003. Begå eller ta självmord? Om svenska kollokationer och deras förändringsbenägenhet. *Språk och stil* 13, s. 123–168.
- Meara, Paul 1983. Word associations in a foreign language. A report of the Birkbeck Vocabulary Project. *Nottingham Linguistic Circular* 11, s. 29–38.
- Milton, James 2009. *Measuring Second Language Vocabulary Acquisition*. Bristol, Buffalo & Toronto: Multilingual Matters.
- Moon, Rosamund 1998. *Fixed Expressions and Idioms in English. A Corpus-Based Approach*. Oxford: Oxford University Press.
- Moon, Rosamund 2008. Conventionalized as-similes in English. A problem case*. *International Journal of Corpus Linguistics* 13, s. 3–37.
- Moyer, Alene 1999. Ultimate attainment in L2 phonology. The critical factors of age, motivation, and instruction. *Studies in Second Language Acquisition* 21, s. 81–108.
- Nagy, William E. & Herman Patricia A. 1987. Breadth and depth of vocabulary knowledge: Implications for acquisition and instruction. I: McKeown, Margret G. & Curtis, Mary E. (eds.), *The nature of vocabulary acquisition*. Hillsdale NJ: Lawrence Erlbaum, s. 19–35.
- Namei, Shidrokh 2002. *The bilingual lexicon from a developmental perspective*. Stockholm: Center for Research on Bilingualism, Stockholm University.
- Nation, Paul 2001. *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.
- Nattinger, James R. & DeCarrico, Jeanette S. 1992. *Lexical Phrases and Language Teaching*. Oxford: Oxford University Press.
- O’Keefe, Anne, McCarthy, Michael & Carter, Ronald 2007. *From Corpus to Classroom: Language Use and Language Teaching*. Cambridge: Cambridge University Press.
- Otterup, Tore 2005. ”Jag känner mig begåvad bara”. Om flerspråkighet och identitetskonstruktion bland ungdomar i ett multietniskt förortsområde (Göteborgsstudier i nordisk språkvetenskap 2.) Göteborg: Institutionen för svenska språket, Göteborgs universitet.

- Palmer, Harold E. 1933. *Second Interim Report on English Collocations*. Tokyo: Kaitakusha.
- Partee, Barbara H. 1984. Compositionality. I: Landman, Fred & Veltman, Frank (eds.), *Varieties of Formal Semantics*. Dordrecht: Foris, s. 281–311.
- Pavlenko, Aneta 2009. Conceptual Representation in the Bilingual Lexicon and Second Language Vocabulary Learning. I: Pavlenko, Aneta (ed.), *The Bilingual Mental Lexicon. Interdisciplinary Approaches*. Bristol, Buffalo & Toronto: Multilingual Matters, s. 125–160.
- Pawley, Andrew & Syder, Frances H. 1983. Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. I: Richards, Jack & Schmidt, Richard (eds.), *Language and Communication*. London: Longman s. 191–221.
- Penfield, Wilder & Roberts, Lamar 1959. *Speech and Brain Mechanisms*. New York: Atheneum.
- Prentice, Julia 2010. Käppen i hjulen. Behärskning av svenska konventionaliserade uttryck bland gymnasieelever med varierande språklig bakgrund. Rapporter i svenska som andraspråk (ROSA12.) Göteborg: Institutet för svenska som andraspråk, Göteborgs universitet.
- Prentice, Julia & Emma Sköldberg und. utg. *Klättra på väggarna eller bara vara ett med soffan?* Om figurativa ordförbindelser bland ungdomar i flerspråkiga skolmiljöer. *Språk och stil* 20, s. 5–35.
- Sinclair, John 1966. Beginning the Study of Lexis. I: Bazell, Charles E., Catford, John C., Halliday, M. A. K. & Robins, Robert H. (eds.), *In Memory of J.R. Firth*. London: Longmans, s. 410–430.
- Sinclair, John 1987. Collocation: A progress report. I: Steele, Ross & Treadgold, Terry (eds.), *Language Topics. Essays in honour of Michael Halliday*. Amsterdam & Philadelphia: John Benjamins, s. 319–331.
- Skehan, Peter 1998. *A Cognitive Approach to Language Learning*. Oxford: Oxford University Press.
- Skutnabb-Kangas, Tuve 1981. *Tvåspråkighet*. Lund: Liber Läromedel.
- Sköldberg, Emma 2004. *Korten på bordet. Innehålls- och uttrycksmässig variation hos svenska idiom*. (Meijerbergs Arkiv för svensk ordforskning 31.) Göteborg: Göteborgs universitet.
- Stubbs, Michael 2001. *Words and Phrases. Corpus Studies of Lexical Semantics*. Oxford & Malden, MA: Blackwell.
- Svanlund, Jan 2001. *Metaforen som konvention. Graden av bildlighet i svenskans vikt- och tyngdmetaforer*. (Stockholm Studies in Scandinavian Philology. New series. 23.) Stockholm: Almqvist & Wiksell.

- Svanlund, Jan 2009. *Lexikal etablering. En korpusundersökning av hur nya sammansättningar konventionaliseras och får sin betydelse*. (Stockholm Studies in Scandinavian Philology. New series. 52.) Stockholm, Stockholms universitet.
- Svenskt språkbruk. Ordbok över konstruktioner och fraser* 2003. Utgiven av Svenska språknämnden Stockholm: Norstedts ordbok.
- Svensson, Gudrun 2009. *Diskurspartiklar hos ungdomar i mångspråkiga miljöer i Malmö*. (Lundastudier i nordisk språkvetenskap A:67.) Lund: Språk- och litteraturcentrum, Lunds Universitet.
- Thomas, Wayne & Collier, Virginia 2002. *A National Study of School Effectiveness for Language Minority Students' Long-Term Academic Achievement*. Santa Cruz CA: Center for Research on Education, Diversity & Excellence (CREDE), University of California-Santa Cruz.
- Tingsell, Sofia 2007. *Reflexivt och personligt pronomen. Anaforisk syftning hos ungdomar i flerspråkiga storstadsmiljöer*. (Göteborgsstudier i nordisk språkvetenskap 8.) Göteborg: Institutionen för svenska språket, Göteborgs universitet.
- Trost, Jan 2007. *Enkätboken*. Lund: Studentlitteratur.
- Turner, Mark & Fauconnier, Gilles 2000. Metaphor, metonymy, and binding. I: Barcelona, Antonio (ed.), *Metaphor and Metonymy at the Crossroads*. Berlin & New York: Mouton de Gruyter, s. 133–145.
- Viberg, Åke 1993. Andraspråksinläring i olika åldrar. I: Cerú, Eva (red.), *Lärarbok 2. Svenska som andraspråk. Mera om språket och inläringen*. Stockholm: Natur och Kultur & Utbildningsradion, s. 13–83.
- Viberg, Åke 1994. Vocabularies. I: Ahlgren, Inger & Hyltenstam, Kenneth (eds.), *Bilingualism in Deaf Education*. Hamburg: Signim-Verlag, s. 169–199.
- Viberg, Åke 2004. Lexikal utveckling i ett andraspråk. I: Hyltenstam, Kenneth & Lindberg, Inger (red.), *Svenska som andraspråk – i forskning, undervisning och samhälle*. Lund: Studentlitteratur, s. 197–220.
- Wiktorsson, Maria 2003. *Learning Idiomaticity. A corpus-based study of idiomatic expressions in learners' written production*. Lund: Lund University.
- Wray, Alison 2002. *Formulaic Language and the Lexicon*. Cambridge: Cambridge University press.
- Wray, Alison 2008. *Formulaic Language: Pushing the Boundaries*. Oxford: Oxford University Press.
- Wray, Alison & Perkins, Michael R. 2000. The functions of formulaic language: An integrated model. *Language and Communication* 20, s. 1–28.
- Wulff, Stefanie 2008. *Rethinking Idiomaticity. A Usage-based Approach*. London & New York: Continuum.

Yorio, Carlos A. 1989. Idiomaticity as an indicator of second language proficiency. I: Hyltenstam, Kenneth & Obler, Loraine K. (eds.), *Bilingualism across the lifespan*. Cambridge: Cambridge University Press, s. 55–72.

Övriga källor

Google, <www.google.se>

Skolverket, <www.skolverket.se>

Skolverket, SALSA, <<http://salsa.artisan.se/definitioner.htm>>

Språkbanken, <<http://spraakbanken.gu.se/>>

Språkbanken, PAROLE, <<http://spraakbanken.gu.se/parole/>>

I

FRASEOLOGISKA ENHETER
PÅ FLERSPRÅKIG GRUND



Julia Große

Göteborgs universitet

Fraseologiska enheter på flerspråkig grund

En jämförelse mellan verklighet och fiktion

Sammanfattning

Användningen av *fraseologiska enheter* är ett av de språkliga dragen nämnda i tidigare forskning som har stor betydelse för ett idiomatiskt språkbruk. Bruket är därför intressant att studera i flerspråkiga sammanhang, t.ex. bland ungdomar i flerspråkiga storstadsområden. Debatten kring multietniska ungdomsvarieteter har stundtals varit livlig, även i samband med romanen *Ett öga rött* (Khemiri 2003), där författaren använder förvrängningar av fraseologiska enheter för att skapa huvudpersonens idiolekt. Romanen kan därför tjäna som intressant referenspunkt i beskrivningen av fraseologiska enheter i ungdomsspråk i flerspråkiga miljöer. I detta arbete jämförs data från projektet *Språk och språkbruk i flerspråkiga storstadsmiljöer* (SUF), dels med ett standardspråkligt referensmaterial, dels med ett utdrag ur Khemiris roman, med avseende på användningen av fraseologiska enheter. För att beskriva variationen i förhållande till ett standardspråkligt jämförelsematerial har begreppet *idiom schemas* (Moon 1998) tillämpats. I den analyserade delen av romanen visar sig avvikelserna i första hand ligga på lexiko-grammatisk nivå. Avvikelserna i SUF-materialet är däremot delvis mer komplexa då de i vissa fall ligger på både lexiko-grammatisk och semantisk-pragmatisk nivå.

Nyckelord: fraseologiska enheter, konventionaliserat språk, idiomscheman, flerspråkighet och språkvariation, ungdomsspråk i flerspråkiga miljöer

Inledning

Idiomatiska uttryck, kollokationer och andra fraseologiska enheter anses ha stor betydelse för vad som identifieras som så kallat infött (native-like) språkbruk (se t.ex. Pawley & Syder 1983, Ekberg 2004). I föreliggande artikel studeras och jämförs fraseologiska enheter i två olika material som båda potentiellt kan uppvisa drag av inflytande från andra språk. Det ena är insamlat bland ungdomar i flerspråkiga

miljöer och det andra är hämtat från romanen *Ett öga rött* av Jonas Hassen Khemiri (2003), som uppmärksammas inte minst för det speciella språk som författaren använder. Det har bl.a. påståtts att boken är skriven på »tvättäkta rinkebysvenska,« en karakterisering som författaren dock själv vänder sig emot. Det autentiska materialet har samlats in inom ramen för projektet *Språk och språkbruk bland ungdomar i flerspråkiga storstadsmiljöer* (hädanefter SUF), ett projekt som baseras på språkliga data från sammanlagt 222 gymnasieelever från flerspråkiga skolmiljöer i Stockholm, Göteborg och Malmö.¹

Utgångspunkten för jämförelsen mellan SUF-materialet och romanen är ett från standardspråket delvis avvikande bruk av fraseologiska enheter som observerats i språkbruket hos ungdomarna i de flerspråkiga miljöerna. Det är också uppenbart att fraseologiska enheter är ett av de språkliga drag som Khemiri har tagit fasta på och som han använder på ett avvikande sätt för att konstruera romanfiguren Halims idiolekt (se vidare Källström 2006). Syftet med dessa medvetna avvikelser från svenskt standardspråk är att betona Halims identitet som flerspråkig och uppväxt i förorten. Tillsammans med en konsekvent användning av rak ordföljd i huvudsatser bidrar dessa avvikelser kanske allra mest till intrycket av en för litteraturen helt ny och annorlunda användning av svenska, som också kan förknippas med vissa stereotypa föreställningar om vad som kännetecknar svenska på mångspråkig grund.

Syfte och frågeställningar

Det är rimligt att anta att det finns vissa skillnader i användningen av fraseologiska enheter mellan SUF-materialet och den konstruerade idiolekten i romanen, även om båda uppvisar avvikelser i förhållande till svenskt enspråkigt standardspråk. En av frågeställningarna för denna studie är alltså hur de två materialen skiljer sig åt i detta avseende. Det huvudsakliga syftet med studien är dock att med hjälp av en jämförande analys synliggöra och beskriva avvikelser som kan förekomma i bruket av fraseologiska enheter bland några av ungdomarna i SUF-projektet. Här fungerar Khemiris litterära lek med just sådana fraser som en intressant referenspunkt.

Eftersom den här aktuella studien bygger på ett begränsat material får den ses som explorativ. Det betyder att den knappast kan ge underlag för några generella slutsatser om språkbruk bland ungdomar i flerspråkiga miljöer. Resultaten kan däremot visa på intressanta tendenser ifråga om avvikelser från konventionaliserade mönster vid användningen av fraseologiska enheter i flerspråkiga miljöer, bl.a. med tanke på hur sådana avvikelser kan beskrivas i fortsatta studier.

Frågeställningarna för denna undersökning är följande:

- Hur kan man beskriva olika avvikelser i förhållande till standard-språket i bruket av fraseologiska enheter hos ungdomarna i SUF-materialet?
- Hur kan man beskriva olika avvikelser i förhållande till standard-språket i bruket av fraseologiska enheter hos protagonisten Halim i Khemiris roman *Ett öga rött* (2003)?
- Vilka likheter och skillnader kan observeras mellan det avvikande bruket av fraseologiska enheter i ungdomarnas och romanfigurens språkbruk?

Avvikelserna i användningen av fraseologiska enheter analyseras bl.a. med hjälp av Moons modell för beskrivningen av lexiko-grammatisk variation hos så kallade fasta uttryck (1998). Ett sekundärt syfte med undersökningen är att anpassa denna modell till den aktuella typen av material, t.ex. för att även kunna beskriva avvikelser på semantisk-pragmatisk nivå. En annan teoretisk utgångspunkt utgörs av Hyltenstams studier av *nästan-ifödda drag* i svenskan (1992)

I det följande kommer jag först att redogöra för relevant forskning på områdena ungdomsspråk i flerspråkiga miljöer samt forskning kring idiomacitet och konventionaliserat språk. Därefter presenteras undersökningens teoretiska utgångspunkter som i huvudsak går tillbaka på Hyltenstam (1992) och Moon (1998). I anslutning till detta redovisas de olika delmaterialen samt exempelurvalet och analysmetoden. I resultatavsnittet följer en redogörelse för en kvantitativ och en kvalitativ analys av de båda delmaterialen (SUF- samt romanmaterialet) utifrån ett antal exempel. Slutligen jämförs och sammanfattas resultaten för de båda delmaterialen. I den avslutande diskussionen behandlas några problematiska aspekter i samband med denna undersökning samt utgångspunkter för fortsatt forskning.

Tidigare forskning och teoretiska utgångspunkter

Språkanvändning bland ungdomar i flerspråkiga storstadsmiljöer

Språkbruket bland ungdomar i flerspråkiga storstadsmiljöer är ett forskningsområde som under de senaste decennierna fått allt större uppmärksamhet, inte minst i Nordeuropa. Förutom det redan nämnda svenska projektet (*Språk och språkbruk bland ungdomar i flerspråkiga storstadsmiljöer*) kan flera andra forskningsprojekt med detta fokus i Norden nämnas. Hit hör bl.a. UPUS-projektet (Utviklingspro-

sessar i urbane språkmiljø) i Norge (Universitetet i Oslo 2008), men även forskning initierad av enskilda forskare som bl.a. Kotsinas (1988), Fraurud & Bijvoet (2004) och Quist (2004).

En fråga som man måste förhålla sig till i samband med studier av ungdomsspråk i multietniska områden är i vilken utsträckning flerspråkiga ungdomar ska betraktas som förstaspråks- respektive andraspråkstalare av svenska. Denna problematik diskuteras bl.a. av Boyd & Fraurud (under utg.) i samband med en analys av ungdomarnas bakgrundsdata, där författarna kommer fram till ett 80-tal olika profiler för de 222 informanter i SUF-projektet. Quist (2004) anser att den traditionella distinktionen mellan infödda enspråkiga och flerspråkiga, där den enspråkiga gruppen ska representera en språklig norm, i många fall är en förenkling. Utmärkande för informanterna i hennes studie av en flerspråkig skolmiljö i Köpenhamn är, enligt författaren, i stället att de i vissa situationer överskrider gränserna mellan lingvistisk, etnisk och social tillhörighet (2004:102f).

Informanterna i denna studie och i SUF-projektet som helhet representerar, oavsett språklig bakgrund, i första hand en grupp ungdomar i en flerspråkig miljö. Eftersom dessa ungdomar under sin uppväxt i varierande grad kommit i kontakt med svenska kan det emellertid också vara relevant att fråga sig i vilken utsträckning den variation de eventuellt uppvisar i bruket av fraseologiska enheter bör relateras till sociolingvistiska eller psykolingvistiska/kognitiva förklaringsgrunder. Medan en sådan variation utifrån ett sociolingvistiskt perspektiv kan relateras till konstruktionen av en flerspråkig identitet och tolkas som uttryck för tillhörighet i en specifik multietnisk subkultur (se t.ex. Fraurud & Bijvoet 2004 och Quist 2004), kan den utifrån ett psykolingvistiskt/kognitivt perspektiv associeras med den typ av icke helt etablerade och automatiserade språkliga system som förknippas med avancerad andraspråksanvändning (se t.ex. Hyltenstam 1992, Ekberg 1997, 1998 och Wray 2002). Även om den relativt fylliga bakgrundsinformation med avseende på födelseland och startålder etc., som finns tillgänglig för informanterna i SUF-projektet, i viss mån kan ge viktiga ledtrådar för valet av tolkningsram i de enskilda fallen, kan isolerade analyser av enskilda språkdrag knappast förväntas ge några entydiga svar i detta avseende. För detta krävs systematiska analyser av strategiskt utvalda språkliga drag och ett betydligt större material än det som står i fokus för denna explorativa studie. Därför kommer denna fråga endast i ringa utsträckning att behandlas. I det följande redogörs dock kort för några aspekter på fraseologiska enheter ur ett mer psykolingvistiskt perspektiv.

Fraseologiska enheter, idiomatiskt språkbruk och språklig kreativitet

Den terminologi som används inom fraseologins område är tämligen svårhanterlig. Det finns t.ex. en lång rad olika benämningar för det fenomen som i detta arbete kallas för *fraseologisk enhet* liksom för olika begrepp som hänger samman med det. Det är därför svårt att hitta termer som inte redan bär med sig ett antal teoretiska antaganden (Wray 2002:8f). Jag har valt att använda termen *fraseologisk enhet*, då den, såvitt jag vet, inte för med sig några direkta associationer till andra forskares teoretiska antaganden. Med fraseologisk enhet menas här en flerordsenhet som kännetecknas av vissa inskränkningar i de ingående ordens kombinerbarhet. De fraseologiska enheter som beskrivs i detta arbete är semantiskt-pragmatiskt och/eller grammatiskt inte helt kompositionella, det vill säga inte fullständigt härledbara enbart utifrån sina semantiska eller grammatiska komponenter. En utgångspunkt här är att fraseologiska enheter är flerordsuttryck som rutinemässigt används i en konventionaliserad, *kanonisk* form. Man bör dock komma ihåg att den kanoniska formen, som oftast är den form som presenteras i ordböcker, i många fall bör ses som en förenkling med tanke på hur uttrycken faktiskt används. I själva verket använder infödda, enspråkiga talare fraseologiska enheter långt ifrån alltid i sin kanoniska form (se bl.a. Moon 1998). För svenskans del har Sköldberg (2004) visat att även de uttryck som allmänt anses som de mest fixerade bland fasta flerordskombinationer, nämligen idiomerna, varierar avsevärt bland infödda talare.

Synen på språkanvändning som ett ständigt kreativt nyskapande där man som språkbrukare fritt utnyttjar möjligheterna inom ramen för den inre grammatiken har sedan den generativa grammatikens genombrott haft ett starkt stöd. Under senare år har man dock inom andraspråksforskning riktat alltmer uppmärksamhet på det faktum att modersmålstalare i relativt stor utsträckning använder sig av återkommande, delvis stereotypa mönster och förutsägbara konstruktioner och därmed långt ifrån alltid utnyttjar det grammatiska systemets kreativa möjligheter (Lindberg 2004, se Ekberg 1998 nedan, jf. även bl.a. Hakuta 1974, Schmidt 1983 och Schmitt 2005). Pawley & Syder (1983:193) menar att en mycket begränsad andel av de potentiellt grammatiska uttryckssätten betraktas som naturliga för infödda talares språkbruk: »native speakers do *not* exercise the creative potential of syntactic rule to anything like their full extent [...] indeed, if they did do so they would not be accepted as exhibiting nativelike control of the language« (se även Wray 2002:11). Wray (2002:12) resonerar på följande sätt kring språklig stereotypi respektive kreativitet:

...capacity for handling novelty, both ideational and grammatical, is sufficient to rule out the possibility that language knowledge consist *only* of a set of prefabricated phrases and sentences memorized from previous encounters with them (Bloom 1973:17). Whatever determines our preferences for certain phrases – their storage in prefabricated form, or something else – the most we could argue is that this process *co-exists* with our ability to create and understand entirely novel strings. Although we customarily say, *Hi, how are you doing?* [...] there is nothing at all to stop us saying, *what a pleasant event it is to see you. Tell me how your life is progressing at the moment.*

Användningen av konventionaliserade strukturer är, som antytts tidigare, högst relevant ur ett flerspråkighetsperspektiv. Ekberg har studerat språkbruket bland högstadiel elever från Rosengård – en flerspråkig stadsdel i Malmö – ur detta perspektiv (1997, 1998, 2004:261). Hennes resultat visar att både enspråkiga och flerspråkiga elever har aktiv kunskap om de undersökta konventionaliserade lexikala och grammatiska konstruktionerna. De enspråkiga eleverna använder däremot fler konventionaliserade strukturer på alla språkliga nivåer (Ekberg 1998:261, se även 1997:106ff). En annan skillnad är att variationen i användningen av dessa frasliknande strukturer är större bland de flerspråkiga eleverna, vilket innebär att deras språkbruk i den givna talsituationen är mindre stereotypt än de enspråkiga elevernas. Det kan, enligt Ekberg, vara ett tecken på att de tvåspråkiga barnen har en mindre uppsättning färdiga konstruktioner att välja bland, vilket i sin tur kan bero på en relativt låg grad av kontakt med målspråket (1998:261). Det är alltså »kanske just stereotypin som i vissa fall utmärker ett idiomatiskt språkbruk!« (Ekberg 2004:274).

Approximationer och kontaminationer

Fraseologiska enheter har också studerats av Hyltenstam (1992) som har undersökt icke-infödda drag (*non-native features*) hos mycket avancerade eller så kallade *inföddlika* andraspråksbrukare (*near-native speakers*) av svenska (1992:351). I hans studie jämförs två grupper av unga andraspråksinlärare, som har finska respektive spanska som förstaspråk, med en grupp infödda talare av svenska. Informanterna testades både i muntlig och skriftlig produktion. Materialet för undersökningen är begränsat och Hyltenstam talar därför om hypoteser snarare än färdiga resultat. Studien visar dock några intressanta tendenser med avseende på fraseologiska enheter. En av de mest frekventa avvikelserna från ett enspråkigt svenskt standard-

språkbruk bland de tvåspråkiga informanterna i Hyltenstams material är vad han kallar för »*close approximations to the target form* [of a lexical phrase]« (1992:360), det vill säga fraser som används på ett sätt som bara skiljer sig marginellt från den kanoniska formen. Ett exempel från Hyltenstams material är *den fungerar nog inte på längden*. Den idiomatiska form som ligger närmast till hands i detta fall är *något fungerar inte i längden* med den konventionaliserade betydelsen »i det långa loppet, under en längre tid«. *På längden* har i stället den mer konkreta betydelsen »i längdriktningen« (Svenskt språkbruk 2003:734). Denna typ av avvikelse är betydligt mer frekvent bland de inföddlika informanterna än bland informanterna i den enspråkiga kontrollgruppen. Ett annat intressant fenomen i materialet är *kontaminationer*, det vill säga sammanblandningar av två eller flera olika fraser, som i »*man beräknar med att...*« (1992:360f) där element från fraserna *räkna med någ ot* »förvänta sig, göra sig hopp om något« och *beräkna något*, med den konventionaliserade betydelsen »företse, bedöma något« blandas ihop (Svenskt språkbruk 2003:964, 105). Intressant nog visar sig kontaminationer vara lika frekventa bland de enspråkiga informanterna som bland de inföddlika andraspråksbrukarna (Hyltenstam 1992:360f). Begreppen *approximation* och *kontamination* tillämpas i denna studie för att beskriva olika typer av avvikelse vid användningen av fraseologiska enheter.

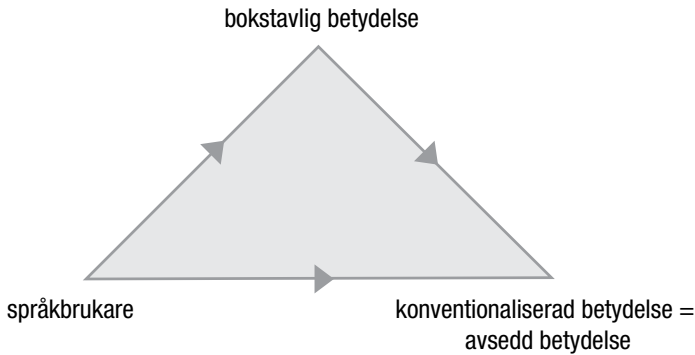
Idiomscheman

Begreppet *fraseologisk enhet*, som det används i detta arbete, motsvarar begreppet *fixed expression and idioms* (FEI) hos Moon (1998) och omfattar flera olika typer av flerledade uttryck som t.ex. kollokationer och idiom.² Moon påpekar dock att termen *fixed expressions* inte är helt tillfredställande eftersom många fasta uttryck av den typ som hon diskuterar i själva verket inte är fasta (ibid.). Hon beskriver variationen hos FEI:s enligt begreppet *idiomschema*. Ett idiomschema grundar sig på en metafor eller ett tankemönster som är gemensamt för olika realiseringar av en fraseologisk enhet. Såväl strukturen som de lexikala konstituenterna kan alltså i vissa fall variera rätt avsevärt inom ett sådant schema, så länge tankemönstret förblir intakt, det vill säga så länge den fraseologiska enhetens olika realiseringar kan föras tillbaka till samma metaforiska grundtanke. En sådan modell för variationen hos fraseologiska enheter förklarar, enligt Moon, också varför det oftast inte är något problem att avkoda även sådana realiseringar som ligger relativt långt ifrån ett uttrycks kanoniska form (1998:163).

Med utgångspunkt i Moons resonemang har jag valt att använda begreppet *tankemönster* i min analys. Det kan noteras att tankemönster, som det används här, uppvisar likheter med begreppet *källdomän* i konceptuell metafor-teori (Lakoff 1993), som är en del av den kognitiva semantiken. I korta ordalag tänker man sig inom denna teori att metaforiska uttryck är baserade på olika erfarenhetsområden, föreställningar och koncept – domäner – inom en kulturkrets. En metafor har en konceptuell bas i en källdomän. Metaforen bildas genom att en viss aspekt av denna källdomän överförs kognitivt till en annan erfarenhetsdomän (måldomänen). Överföringar av olika aspekter av samma källdomän resulterar språkligt i olika metaforiska uttryck. Ett exempel är en sådan överföring mellan källdomänen *resande* och måldomänen *kärlek*, som språkligt finner uttryck i konventionaliserade fraseologiska enheter som *stå vid ett vägskäl* eller *köra fast* i sin relation (Svanlund 2001:14ff, se vidare Lakoff 1993).

Moon beskriver variationen som ett kontinuum, från relativt fasta konstruktioner med endast ett variabelt led till sådana uttryck som hon kallar för *fria realiseringar* och *idiomscheman*. Som ett exempel nämner Moon ett schema som varierar ifrån *shake in one's shoes* (»att vara extremt nervös eller rädd för något«) till bl.a. *quake in one's boots and quake in one's Doc Marten's*³ (1998:161). Ett annat exempel är vattnets enkelriktande rörelse som metafor för något som ligger i det förflutna och som inte kan göras ogjort i frasen *something is water under the bridge*. Lexiko-grammatiska realiseringar inom samma idiomschema kan i detta fall t.ex. vara *water over the dam*, *water under the dyke*, *water under the thingy* eller betydligt friare realiseringar som »... *so much has happened under the bridge*«, där vattnet över huvudet inte ingår som lexikal komponent (Moon 1998:164).

Moon diskuterar även kognitiva processer i samband med användningen av metaforiska uttryck (jf. Searle 1979:122, Svanlund 2001:67). Figur 1 återspeglar den kognitiva processen vid användning av idiomscheman. Språkbrukaren har här således direkt tillgång till den konventionaliserade betydelsen (som är den av språkbrukaren avsedda betydelsen), samtidigt som det bakomliggande tankemönstret och därmed även medvetenheten om de lexikala komponenternas bokstavliga betydelse bevaras. »In this way metaphorical FEIs can be seen as 'at the same time holistic and analyzable'« (Coulmas 1979:149, citerat i Moon 1998:164).



Figur 1 Modell för processer verksamma vid användningen av idiomscheman (jf. Moon 1998:164)⁴

Denna modell för användningen av idiomscheman kan alltså användas för att hantera variation hos fraseologiska enheter på lexiko-grammatisk nivå. Däremot redogör modellen inte för en annan relevant nivå i samband med min undersökning, nämligen den semantisk-pragmatiska. Moon (1998:163) utgår ifrån ett tankemönster, »associated with particular lexicalizations« och ser därmed semantiken som i viss mån given. Jag kommer emellertid i det följande att argumentera för att fraseologiska enheter även kan varieras på semantisk-pragmatisk nivå inom samma idiomschema. Med andra ord kan själva tankemönstret aktiveras, trots relativt grova avvikelser eller extrem variation i förhållande till den konventionaliserade användningen eller betydelsen (se vidare under *Analysmetod*).

Material och metod

Det empiriska underlaget för denna undersökning består alltså av två olika material. Det ena utgörs av både talspråkliga och skriftliga data som har samlats in inom ramen för SUF-projektet och det andra utgörs av ett utdrag ur romanen *Ett öga rött* (Khemiri 2003).

SUF-materialet består av sammanlagt drygt 300 timmar inspelat tal och mer än 200 skrivna texter från det nationella provet i gymnasiekurs B i svenska. Föreliggande studie baseras av förklarliga skäl på en avgränsad del av detta material, nämligen på det som har samlats in i Göteborg. Den delkorpus varifrån exemplen excerperats består av åtta transkriberade inspelningar av åtta olika informanter. I tillägg har dessa informanternas uppsatser från det nationella provet⁵ använts.

Sju av åtta informanter har flerspråkig bakgrund, det vill säga de har utlandsfödda föräldrar. En av dem har en svensk och en utlandsfödd förälder, är själv född i Sverige och anger svenska som sitt första språk⁶. Övriga sex har föräldrar som båda är födda utanför Sverige. Tre av dessa informanter är födda i Sverige och tre utomlands. Alla sex anger ett annat språk än svenska som sitt förstaspråk (en anger både tigrinja och svenska) och de har alla börjat lära sig svenska i förskoleåldern. Den återstående informanten har enspråkig svensk bakgrund, vilket innebär att både hon och hennes föräldrar är födda i Sverige och hon har angett svenska som sitt första språk. Som jag tidigare antytt bygger denna studie inte på något systematisk urval av informanter med hänsyn till språklig bakgrund. Urvalet är istället datadrivet, det vill säga jag har utgått ifrån ett antal avvikelser som observerats i materialet och sedan tittat närmare på muntligt och skriftligt material som producerats av samma informanter.

Romanen *Ett öga rött* (Khemiri 2003) är skriven i dagboksform och kan beskrivas som ett skriftspråkligt material med många tal-språkliga drag. Även här har materialet begränsats så att endast exempel från romanens första 5 kapitel excerperats, vilket motsvarar ett textmaterial på ca. 6 700 ord.

Ett urvalskriterium för de fraseologiska enheter som beskrivs i detta arbete är att de inte är helt kompositionella på semantisk-pragmatisk eller grammatisk nivå. Som jag nämnde inledningsvis, fokuseras undersökningen dessutom på sådana exempel som uppvisar någon typ av avvikelse från en standardspråklig användning på någon av de relevanta nivåerna. De fraseologiska enheter som ingår i materialet har identifierats som avvikande från ett standardenligt bruk med min språkliga intuition och ett standardspråkligt referensmaterial (se nästföljande avsnitt) som utgångspunkt. Jag har alltså excerperat fraseologiska enheter som jag spontant bedömt som avvikande och sedan jämfört dessa med användningen av samma fraseologiska enheter i det standardspråkliga referensmaterialet. Utifrån de tidigare nämnda urvalskriterierna excerperades sammanlagt 19 exempel ifrån SUF-materialet. 28 exempel i romanmaterialet matchade urvalskriterierna och excerperades därmed också för undersökningen.

Tabell 1 Det undersökta materialets omfång.

Delmaterial	Antal ord	Antal exempel
SUF-material	ca. 16 800	19
a) Talspråk	ca. 9 300	12
b) Nationella prov	ca. 7 500	7
Romanmaterial	ca. 6 500	28
Totalt	ca. 23 300	47

Standardspråkligt referensmaterial

För att avvikelser från standardspråket vid bruket fraseologiska enheter ska kunna studeras måste materialet sättas i relation till ett material som kan sägas representera svenskt standardspråk. I detta syfte har jag använt frasordboken *Svenskt språkbruk. Ordbok över konstruktioner och fraser* (hädanefter *Svenskt språkbruk*, 2003) samt korpuserna PAROLE och SUC2 i Språkbanken⁷ (Språkbanken 2008) som referensmaterial. Jag har valt att använda *Svenskt språkbruk* eftersom det är den senaste och mest heltäckande frasordboken över svenska språket. Den innehåller ca. 85 000 konstruktioner och fraser och drygt 70 000 exempel. PAROLE-korpusen består av 19.4 miljoner löpord och SUC2 av ca. 1.2 miljoner löpord. Dessa korpuser har valts ut eftersom de kan sägas representera ett standardliknande bruk av det svenska (skrift)språket och är sammansatta av olika typer av textmaterial (dagstidningar samt bl.a. romaner, tidskrifter och webbtexter) (Språkbanken 2006b). Med tanke på att den övervägande delen av materialet i studien består av talat språk är referensmaterialet naturligtvis inte helt idealiskt. Tillgången till relevanta talspråkskorpuser av det slag som skulle behövas här är emellertid än så länge mycket begränsad. (Existerande korpuser, som t.ex. *Göteborg Spoken Language Corpus*⁸ (GSLC), är för begränsade i storlek för att fånga upp den typen av långfrekventa fraseologiska enheter som diskuteras här.) Ett sätt att kompensera denna brist hade varit att göra sökningar på Google, där man med stor sannolikhet kan hitta belägg för vissa exempel från mitt material. Det är dock inte heller oproblematiskt att använda webben som jämförelsekorpus. Det är inte helt självklart i vilken mån man kan betrakta texterna som man får fram via t.ex. Google som representativa för någon slags standardspråklig norm. Dessutom är frekvensangivelserna på Google inte alltid helt pålitliga (se bl.a. Lorentzen 2008:260).

Exempel

Av utrymmesskäl är det här inte möjligt att redovisa analysen för varje enskilt exempel. Därför kommer jag i det följande att utgå från åtta typ-exempel som diskuteras i detalj. Dessa exempel representerar olika typer av avvikelser som förekommer i materialet och lämpar sig dessutom för att förtydliga analysmodellens tillämpning. En sammanfattande beskrivning av de åtta exemplen ges i tabell 2.

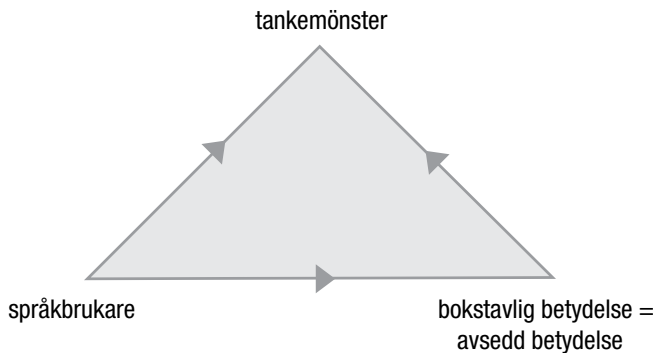
Tabell 2 Exempel på olika typer av avvikelser i materialen.

Exempel	Kanonisk form och betydelse enligt Svenskt språkbruk	Kontext	Informant och delmaterial
a) <i>lägga allting på bordet</i>	lägga korten på bordet om att säga sanningen el. erkänna något	Informanten påminner en klasskamrat om att hon har blivit inspelad medan hon drar snuskiga skämt. Han säger att han lät henne berätta klart utan att säga något eftersom han ville lägga allting på bordet	S23 manlig född utomlands, två utlandsfödda föräldrar, L1: kurdiska. Startålder svenska: 5 år SUF, tal
b) <i>ligga efter på något</i>	ligga efter : 1 (vara försenad, inte vara i kapp) ligga efter med ngt hon låg rejält efter med skolarbetet efter den långa sjukfrånvaron	...så måste man plugga när man kommer hem och på väg till sin träning när man sitter på bussen. Och man känner press på sig själv när man ligger efter på något	P10 manlig född utomlands, två utlandsfödda föräldrar, L1: syrianska. Startålder svenska: 5 år SUF, skrift
c) <i>visa hur snacket går</i>	–	Diskussion kring språket i Leiva Wengers novell Elixir (2001) som är skriven på en multietnisk ungdomsvariant. Författaren skriver på detta sätt eftersom han »vill visa hur snacket går »	S37 kvinnlig född i Sverige, två svenskfödda föräldrar, L1: svenska. SUF, tal
d) <i>det är värsta barnleken</i>	ngt är en barnlek om mycket enkel uppgift	Grupparbete, analys av kärleksdikt. Se exempel (2) nedan	P38 kvinnlig född i Sverige, två utlandsfödda föräldrar, L1: gujarati. Startålder svenska: 2,5 år SUF, tal
e) <i>skriva rakt på pucken</i>	* hårda/raka pucker när ngt är rakt på sak, utan krusiduller	Därför jag rev sidan och istället provade skriva rakt på pucken om vad som hänt idag	Halim (Khemiri 2003:13), roman

f) <i>oskyldig som värsta lamm</i>	<i>oskyldig som ett lamm</i> om ngn som är helt oskyldig	...på vägen dit jag fick lite ågren och kände fet skuld fast jag var <i>oskyldig som värsta lamm</i>	Halim (Khemiri 2003:19), roman
g) <i>samla på damm</i>	<i>samla damm</i> [...] äv. överfört om saker som inte använts på länge och därför har fallit i glömska	...maskinen bara stod i Nourdins rum och <i>samlade på damm</i>	Halim (Khemiri 2003:10), roman
h) <i>det går som kallaste kår i ryggen</i>	<i>*få kalla kårar [...]</i> <i>få kalla kårar längs ryggen/ryggraden</i> <i>el. det går kalla kårar längs ryggen/ryggraden</i> när ngn ryser till av rädsla eller obehag	Också hon har sagt att vi araber är inte som andra blattar utan mera civiliserade och när hon säger nästan varje gång det <i>går som kallaste kår i ryggen</i>	Halim (Khemiri 2003:11), roman

Analysmetod

För att kunna beskriva avvikelserna på semantisk-pragmatisk nivå har jag vidareutvecklat Moons kognitiva modell (se figur 1) till att omfatta de olika typer av avvikelser på semantisk-pragmatisk nivå som finns representerade i materialet. I det mest extrema fallet används en fraseologisk enhet som associeras med en konventionaliserad betydelse i en kontext där enheten endast kan tolkas bokstavligt. Man skulle alltså kunna säga att idiomschemat konkretiseras. Detta skulle innebära att avsedd och bokstavlig betydelse (eller *utterance meaning* och *sentence meaning* i Moons terminologi) sammanfaller, samtidigt som användningen går tillbaka på samma tanke-mönster som den idiomatiska användningen av frasen:



Figur 2 Konkretisering i idiomschemat.

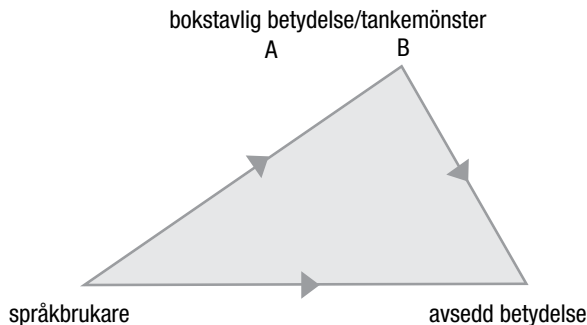
Detta kan förtydligas med följande exempel som är hämtat från SUF-materialet⁹. Här är det fråga om en bildbeskrivning. På intervjuarens fråga om informanten (P29)¹⁰ tror att de två personerna på bilden känner varandra svarar informanten:

(1)

P29: »Nä, jag tror inte det [...] för han tar avstånd i början [...] ja han tar avstånd tar avstånd i början han sitter ganska långt bort.«

Här används alltså den fraseologiska enheten *ta avstånd* med den konventionaliserade betydelsen »att visa att man ogillar ngn eller ngt« (Svenskt språkbruk 2003:74) i bokstavlig betydelse (»hålla sig på fysiskt avstånd från någon«).

En annan aspekt att ta hänsyn till i detta sammanhang är att ett tankemönster i de flesta fall har olika aspekter, det vill säga flera betydelsekomponenter – även om det bara är en av dessa som fokuseras i samband med den fraseologiska enhetens konventionaliserade betydelse. En typ av semantisk-pragmatisk avvikelser kan alltså innebära att språkbrukaren fokuserar en annan aspekt av tankemönstret ifråga än den som fokuseras i den konventionaliserade användningen av samma fras. Med Moons kognitiva modell som utgångspunkt skulle detta kunna beskrivas enligt figur 3, där A står för den betydelsekomponent i tankemönstret som står i fokus för den konventionaliserade betydelsen. B står här för den betydelsekomponent som fokuseras vid avvikelser från den konventionaliserade användningen.



Figur 3 Förskjutet fokus i idiomschemat.

Denna typ av avvikelser kan exemplifieras med informanten P38:s användning av den fraseologiska enheten »något är en barnlek« som, enligt Svenskt språkbruk (2003:84), har den konventionaliserade betydelsen »[något är en] mycket enkel uppgift«. Exemplet är taget

från ett gruppsamtal i skolan där eleverna har i uppgift att diskutera ett antal olika kärleksdikter (se även exempel 8 i resultatavsnittet):

(2)

P29: asså den hä(r) no(g) ganska så hä(r) kanske lite erotiskt på sin tid
så hä(r)

P38: DEN

P29: a: va(r) de(t) inte de(t) i starten

P31: xx mycket natur å: djur å

P38: a i dagens läge så ä(r)de(t) värsta BARNLEKEN eller nånting
herregud

Informanten (P38) använder sig här av samma övergripande tankemönster som ligger till grund för den konventionaliserade användningen av frasen, nämligen »barn som leker«. Men istället för att fokusera lekens enkelhet, tar hon fasta på en annan aspekt av detta tankemönster, nämligen barnen som »barnsliga« eller »oskyldiga« (för att avvisa påståendet om erotik i dikten).

Analysen innehåller en kvantitativ del, men är i huvudsak kvalitativ. Vad gäller den kvalitativa analysen kan Moons modell för idiomscheman ses som en länk mellan den lexiko-grammatiska och semantisk-pragmatiska beskrivningsnivån, då den kan förklara avvikelser från konventionaliserade former med hjälp av en gemensam nämnare i deras betydelse. Samtidigt kan man med hjälp av en utvecklad version av analysmodellen förklara även avvikelser på den semantisk-pragmatiska nivån, så länge användningen av den fraseologiska enheten ifråga fortfarande kan knytas till någon aspekt av det tankemönster som ligger till grund för den konventionaliserade användningen (se *konkretisering* i figur 2 och *förskjutet fokus* i figur 3).

Den kvantitativa analysen kan beskrivas stegvis som följer:

- Exempelen från materialet har jämförts med de aktuella fraseologiska enheternas form och betydelse som de listas i Svenskt språkbruk (2003).
- Beläggen på den standardenliga formen i Språkbanken har räknats.
- Antalet belägg på det avvikande exemplet i Språkbanken har räknats.
- Resultaten från de två föregående stegen har jämförts.

Den kvalitativa delen av analysen har gått ut på att undersöka bruket av aktuella idiomscheman i referensmaterialet och jämföra med mitt material för att se på vilket sätt mina exempel avviker från ett stan-

dardspråkligt bruk av de fraseologiska enheterna ifråga. Den kvalitativa beskrivningen av de fraseologiska enheterna i båda delmaterialen sker alltså utifrån två olika frågeställningar:

- Avviker exemplet ifråga från referensmaterialet på lexiko-grammatisk nivå?
- Avviker exemplet ifråga från referensmaterialet på semantisk-pragmatisk nivå?

I ett sista steg jämförs exemplen från SUF-materialet med exemplen från romanen i en kvalitativ, deskriptiv analys.

Resultat

Som jag tidigare påpekat är det den kvalitativa beskrivningen av materialet som är mest relevant för detta arbete. En hypotes angående avvikande användning av fraserna i materialet, som ställdes upp efter en första materialgenomgång, är att avvikelserna i vissa fall tycks handla om relationen mellan konventionaliserad och bokstavig betydelse hos fraseologiska enheter. Först presenteras dock några kvantitativa resultat för hela materialet. I samband med den kvalitativa analysen diskuteras sedan några enstaka exempel som illustrerar skillnaderna mellan de olika delmaterialen.

Kvantitativ analys

Som redan framgår av materialbeskrivningen är ett resultat av föreliggande undersökning att avvikelser vid bruket av fraseologiska enheter är mindre frekventa hos informanterna i SUF-projektet än i romanmaterialet. Resultatet sammanfattas i tabell 3 nedan.

Tabell 3 Frekvens för avvikande användningar av fraseologiska enheter i de olika delmaterialen.

Delmaterial	Antal ord	Antal exempel	Antal exempel per 1000 ord
SUF-material	ca. 16 800	19 (totalt)	1,1
a) Talspråk	ca. 9 300	12	1,3
b) Nationella prov	ca. 7 500	7	0,93
Romanmaterial	ca. 6 500	28	4,31
Totalt	ca. 23 300	47	2,02

Att avvikelserna är mer frekventa i romanmaterialet är inget oväntat resultat med tanke på att författaren medvetet leker med fraseologiska enheter (Khemiri 2006, samtal på seminarium, se även Dagens Nyheter, 2 augusti 2004). Det framgår vidare av tabell 3 att det inte finns någon betydande skillnad i frekvens när man jämför tal- och skriftspråk i SUF-materialet. På grund av urvalsprocessen är avvikelser i förhållande till det standardspråkliga referensmaterialet en given faktor. Avvikelser på såväl lexiko-grammatisk som semantisk-pragmatisk nivå förekommer i båda delmaterialen, dock i olika utsträckning, vilket sammanfattas i tabell 4.

Tabell 4 **Avvikelser på lexiko-grammatisk respektive semantisk-pragmatisk nivå i de olika materialen.**

Typ av avvikelse	SUF-material	Romanmaterial
Lexiko-grammatisk	16/19 exempel	28/28 exempel
Semantisk-pragmatisk	6/19 exempel	1/28 exempel

Som framgår av tabell 4 förekommer avvikelser på lexiko-grammatisk nivå i något lägre utsträckning i SUF-materialet, medan avvikelser på semantisk-pragmatisk nivå är vanligare i SUF-materialet än i romanmaterialet. Man kan säga att avvikelserna i förhållande till referensmaterialet är påfallande homogena i romanmaterialet. Alla exempel avviker på lexiko-grammatisk nivå varav ett även på semantisk-pragmatisk nivå. Även om det ibland förekommer något mer komplexa avvikelser i Halims användning av fraseologiska enheter, är det dock framförallt på lexikal och grammatisk nivå som Khemiri låter sin huvudperson experimentera med språket. Detta är naturligtvis inget överraskande resultat. För en skönlitterär författare ligger det nära till hands att variera den språkliga formen för att uppnå vissa effekter som läsaren lätt kan ta till sig.

Kvalitativ analys

I det följande beskrivs relevanta kvalitativa aspekter av analysen av SUF-materialet utifrån ett mindre antal exempel, som redovisas i tabell 5 nedan.

Tabell 5 Exempel på analys av SUF-materialet.

Exempel	Lexiko-grammatisk avvikelse	Icke-infött drag	Semantisk-pragmatisk avvikelse
a) ni vet om att dom spelar in? Ja(g) visste det hela tiden men jag ville få ut allting, du vet, <i>lägga allting på bordet</i>	+	–	–
b) Och man känner press på sig själv när man <i>ligger efter på något</i>	+	approximation	–
c) visa hur <i>snacket går</i>	–	–	+ (konkretisering)
d) i dagens läge är det <i>värsta barnleken</i>	+	–	+ (förskjutet fokus)

Exempel (a) – *lägga allting på bordet* – kan ses som avvikande från den kanoniska formen *lägga korten på bordet* som finns listad i Svenskt språkbruk med betydelsen »att säga sanningen eller erkänna något« (2003:644). I korpusmaterialet varierar denna fraseologiska enhet en hel del; man lägger bl.a. »alla sina kort«, »alla tolkningskort« eller »partimotionerna på bordet« (Språkbanken 2006a). Även verbet varierar i vissa fall som i »Därför kräver jag pappren på bordet...« (ibid.). Det finns däremot inga belägg på den form som SUF-informanten använder sig av, vilket innebär att det föreligger lexiko-grammatisk variation i förhållande till referensmaterialet. På semantisk-pragmatisk nivå stämmer betydelsen hos *lägga allting på bordet* i den aktuella kontexten däremot väl överens med betydelsen »att säga sanningen eller erkänna något« (Svenskt språkbruk 2003:644) och beläggen i Språkbanken (se tabell 5).

Exempel (b) – *ligga efter på något* – avviker från standardenligt bruk av frasen *ligga efter med något* på lexiko-grammatisk nivå genom en prepositionsavvikelse. I Svenskt språkbruk anges den fraseologiska enheten *ligga efter (med något)* med betydelsen »vara försenad, inte vara i kapp« (2003:699). Korpusmaterialet uppvisar inga avvikelser från den konventionaliserade formen *ligga efter med + NP* (frasen har dock med sammanlagt 10 belägg relativt låg frekvens). I termer av icke-infödda drag (Hyltenstam 1992) kan exemplet beskrivas som en approximation till den konventionaliserade formen *ligga efter med något*. Kontexten visar att exemplet stämmer överens med referensmaterialet på semantisk-pragmatisk nivå.

Till skillnad från flertalet exempel i SUF-materialet avviker exemplet (c) – *snacket går* – inte från referensmaterialet på lexiko-gramma-

tisk nivå. Denna fraseologiska enhet finns inte representerad i Svenskt språkbruk (2003). Jag har ändå valt att inkludera exemplet i materialet, då det lyfter fram komplexiteten på semantisk-pragmatisk nivå som kan finnas inom vissa idiomscheman. I referensmaterialet i Språkbanken finns denna flerordsenhet belagd 20 gånger, vilket tyder på att den är någorlunda etablerad som en fraseologisk enhet. På lexiko-grammatisk nivå finns det ingen nämnvärd inbördes variation bland dessa belägg. På semantisk-pragmatisk nivå förekommer däremot vissa betydelseskiftningar, som dock kan sägas rymmas inom samma idiomschema. Det går att utskilja tre olika betydelsenysanser av enheten *snacket går* inom detta schema, vilket exemplifieras i (3) nedan, nämligen (3a) »det skvallras, rykten sprids«, (3b) »det (som) är den allmänna opinionen, diskursen, konsensus eller det folk pratar om i ett visst sammanhang« och (3c) »något diskuteras livligt«.

(3)

- a) ...men så såg jag att han hade en kvinna med sej. Hon kanske kom till honom precis när jag stuckit. Då var det sant som *snacket gått*. Ja, nu skulle det väl gå ännu mer.
- b) Det är otäckt för man vet hur det går till allting när man slavat där [på sjukhus]. Man vet liksom hur *snacket går* och man tänker att vad som helst får hända bara man slipper komma hit som patient.
- c) ...om man har tur, kan även de som inte har parabolantenn få in MTV. Då *går snacket* dagen därpå.
De som inte såg sändningen grämer sig... (Språkbanken 2006a).

Kontexten i SUF-materialet tyder däremot på att frasen här används i närmast bokstavlig betydelse. Bakgrunden till den aktuella språksituationen är ett grupparbete om Leiva Wengers novell *Elixir* som är skriven på multietniskt ungdomsspråk. Frågan som eleverna diskuterar är varför författaren väljer att skriva på det viset.

(4)

- S30: ja(g) tror han har gjort så bara för att vis+ a å och för att visa hur e+ a hur *snacket+* a hu+ hur hur hur de(t) E ute
- S35: för att vara lite annorlunda a: visa också den här SNACKEN
- S37: hur *snacket GÅR* så att säga

Kontexten och själva uppgiften tyder således på att exemplet snarare syftar på den språkliga formen och inte – som i frasens konventionaliserade betydelse – på innehållet. Den avsedda betydelsen (»det sätt som ungdomarna i multietniska miljöer pratar på«) sammanfaller med en mer konkret tolkning av denna fraseologiska enhet. Använd-

ningen av *snacket går* i (4) kan alltså beskrivas som konkretisering i idiomschemat (se figur 2).

Kontexten till exempel (d) *det är värsta barnleken* har redan beskrivits ovan. I Svenskt språkbruk (2003) anges formen *något är en barnlek* med betydelsen »om mycket enkel uppgift«. Dessutom finns formen *något är rena barnleken* belagd (Språkbanken 2006a). Formen *värsta barnleken* som används i exempel (d) finns däremot inte belagd i referensmaterialet. Denna variation på lexiko-grammatisk nivå är mycket typisk för ungdomsspråk och i dagens läge knappast att betrakta som avvikande. Bristen på belägg kan i detta fall snarare hänga ihop med referensmaterialets ålder och natur. Det är inte alls omöjligt att man skulle hitta belägg på denna form i ett mer aktuellt talspråkligt referensmaterial. Som framgår av tidigare resonemang avviker exemplet från referensmaterialet på semantisk-pragmatisk nivå, då informanten avser en annan aspekt av tankemönstret »lekande barn« än vid den konventionaliserade användningen av denna fraseologiska enhet, nämligen det barnsliga och oskyldiga i leken.

Nedan redogörs för den kvalitativa analysen av romanmaterialet, även här med hjälp av fyra exempel som ska förtydliga de olika nivåerna i analysen.

Tabell 6 Exempel på analys av romanmaterialet.

Exempel	Lexiko-grammatisk avvikelse	Icke-infött drag	Semantisk-pragmatisk avvikelse
e) <i>skriva rakt på pucken</i> (om vad som hade hänt)	+	kontamination: (<i>vara</i>) <i>rakt på sak</i> och <i>raka puckar</i>	–
f) <i>oskyldig som värsta lamm</i>	+	–	–
g) maskinen stod [...] och <i>samlade på damm</i>	+	kontamination: <i>samla damm</i> och <i>samla på ngt</i>	–
h) <i>det går som kallaste kår i ryggen</i>	+	–	+(förskjutet fokus)

Exempel (e) i tabell 6 – *skriva rakt på pucken* – kan beskrivas som en kontamination av den fraseologiska enheten *raka puckar*¹¹ och frasen *rakt på sak*. I Svenskt språkbruk finns *hårda/raka puckar* listad med betydelsen »något är rakt på sak, utan krusiduller« (2003:900). Den semantiskt härledbara (kompositionella) frasen *rakt på sak*

anges alltså som betydelse för den icke kompositionella enheten *raka puckar*, vilket gör att exemplet lämpar sig väl för att illustrera Moons modell för den kognitiva processen i samband med idiomscheman. Kontaminationen av *rakar puckar* och den kompositionella frasen *rakt på sak* i den aktuella kontexten gör att både den konventionaliserade och den bokstavliga betydelsen¹² är relevanta, vilket kan stödja hypotesen att språkbrukaren har tillgång till båda dessa betydelser samtidigt (se Moons ursprungliga modell i figur 1 ovan). I korpusmaterialet är det t.ex. »*raka puckar* som gäller« eller man kan vara »*rakt på sak*« (Språkbanken 2006a). Exempel (e) avviker alltså tydligt ifrån referensmaterialet på lexiko-grammatisk nivå. Vad gäller den semantisk-pragmatiska nivån kan man däremot säga att Halims användning av den kontaminerade formen följer samma tankemönster som de båda standardspråkliga formerna, vilket innebär att »något görs, sägs eller skrivs rakt ut«.

Även exempel (f) i tabell 6 – *oskyldig som värsta lamm* – avviker på lexiko-grammatisk nivå från användningen i referensmaterialet. I frasordboken finns två fraseologiska enheter representerade som kan anses bygga på samma idiomschema, nämligen *oskyldig som ett lamm* (»om någon som är helt oskyldig«) och *ngn är from/snäll som ett lamm* (»om någon som är mycket snäll el. beskedlig«) (Svenskt språkbruk 2003:686). I förhållande till dessa varierar exempel (f) alltså genom den för ungdomsvarieteter typiska förstärkningen *värsta*. I referensmaterialet visar det sig däremot att denna fraseologiska enhet varieras starkt och att *lamm* verkar vara den enda lexikala komponenten som krävs för att aktivera det aktuella idiomschemat. Variationen inom idiomschemat *oskyldig eller from/snäll som ett lamm* (ibid.) kan illustreras med följande realiseringar från referensmaterialet:

(5)

- a) De som bombade oss har nu smugit sig in här *som lamm* som vill ha fred...
- b) Mary var *inget lamm*. [förslag på titeln på ett filmmanus]
- c) Han var *som ett sårat lamm*.
- d) När jag kom tillbaka så var pappa *som ett lamm*...
- e) så fort Maria kom in på ungarlshärbärgen blev fyllona *som små lamm*... (Språkbanken 2006a)

Exempel (f) ligger alltså relativt nära den fras som finns representerad i Svenskt språkbruk, men finns inte representerat i referenskorpusarna, där tankemönstret (lamm som ett oskyldigt, mjukt, litet och snällt väsen) realiseras på olika sätt som alla ligger relativt långt från formerna i Svenskt språkbruk. Själva tankemönstret är dock detsamma i roma-

nexemplet och i referensmaterialet, varför exemplet endast bedöms som avvikande på lexiko-grammatisk nivå.

Exempel (g) i tabell 6 – *samlar på damm* – avviker från referensmaterialet på lexiko-grammatisk nivå och kan beskrivas som en kontamination av den fraseologiska enheten *samlar damm* och den kompositionella frasen *samlar på något*. Det finns inga belägg på den kontaminerade formen i Språkbanken. Enligt Svenskt språkbruk (2003) används *samlar damm*, både i bokstavlig betydelse och i konventionaliserad betydelse, »om saker som inte använts på länge och därför har fallit i glömska«. Kontexten tyder på att Khemiri avser den senare när han skriver:

(6)

Nourdine lovade han kunde laga [Pokemonmaskinen] och sen hela sommaren han har sagt: »Snart är den färdig, i övermorgon, nästa vecka, snart...« Fast egentligen maskinen bara stod och *samlade på damm* (2003:10, min kursivering).

På semantisk-pragmatisk nivå föreligger alltså ingen avvikelse från den betydelse som finns listad i Svenskt språkbruk för den (icke kompositionella) fraseologiska enheten *samlar damm*. Detta gäller även i jämförelse med de belägg som finns i referenskorpusarna i Språkbanken, även om det bör påpekas att frasen inte är särskilt frekvent här (sammanlagt 6 belägg). Liksom exempel (e) – *skriva rakt på pucken* – illustrerar detta exempel Moons tankemodell enligt figur 1 väl.

Alla de tre tidigare nämnda exemplen kan beskrivas enligt Moons ursprungliga modell för idiomscheman, där en fraseologisk enhet används i konventionaliserad betydelse, samtidigt som språkbrukaren även har tillgång till den bokstavliga betydelsen (se figur 1). I samband med exempel (h) – *det går som kallaste kår i ryggen* – måste däremot även figur 3 (*förskjutet fokus i idiomschemat*) inkluderas i beskrivningen, eftersom exemplet avviker från referensmaterialet på såväl lexiko-grammatisk som semantisk-pragmatisk nivå. De former som finns listade i Svenskt språkbruk är *få kalla kårar*¹³, *få kalla kårar längs ryggen/ryggraden* och *det går kalla kårar längs ryggen/ryggraden* (2003:672). I korpusarna är variationen inom detta idiomschema relativt stor, särskilt vad gäller verbanvändningen; *kalla kårar* kan t.ex. *framkallas*, *ges*, *fås* eller *jagas längs ryggen* (Språkbanken 2006a). Den form som Khemiri använder finns inte belagd i referensmaterialet. På lexiko-grammatisk nivå kan man notera att Halim använder *kår* (jf. *brandkår*, *studentkår* och liknande) som singularisform av *kårar*, istället för *kåre* som är formen listad i *Svenska Akademiens ordlista* (2006:498). När det

gäller den semantisk-pragmatiska nivån är den konventionaliserade betydelsen listad i Svenskt språkbruk »när någon ryser till av rädsla eller obehag« (2003:672f). Även i referensmaterialet implicerar 17 av totalt 18 belägg en känsla av obehag medan kontexten hos Khemiri snarare tyder på att Halim få rysningar i positiv mening:

(7)

Dalanda kan allt om arabernas historia och det är hon som har berättat det är vi som har bäst filosofer och smartast matematiker och grymmast krigare. Också hon har sagt vi araber är inte som andra blattar utan mera civiliserade och när hon säger nästan varje gång det går som kallaste kår i ryggen (Khemiri 2003:11)

I (7) förskjuts alltså fokus i tankemönstret (se figur 3) »rysningar i kroppen« från »att rysa av obehag« till »rysningar som utlöses av positiva känslor«.

Jämförelse och sammanfattning

Nedan jämförs resultaten från den kvalitativa analysen för SUF- och romanmaterialet. Dessutom ges en sammanfattning av undersökningens resultat som helhet. Ett resultat som följer ur den kvantitativa jämförelsen av de båda delmaterialen (som presenteras i tabell 2) är att beskrivningsmodellen i figur 3 är mer relevant för SUF-materialet, då det finns fler avvikelser på semantisk-pragmatisk nivå i denna del av materialet. De flesta exemplen från romanmaterialet kan beskrivas enligt Moons ursprungliga modell (se figur 1). Detta innebär att avvikelserna från referensmaterialet hos ungdomarna på det hela taget är mer komplexa än i Khemiris roman. Tittar man på hela materialet är semantisk-pragmatiska avvikelser enligt figur 3 (5/47 exempel) något vanligare än de enligt figur 2 (2/47 exempel). Granskar man skillnaderna mellan tal- och skriftspråk inom SUF-materialet framgår det att alla avvikelser på semantisk-pragmatisk nivå förekommer i talspråksmaterialet. Dessa resultat sammanfattas i tabell 7.

Tabell 7 Fördelning av exemplen på de olika beskrivningsmodellerna.

Tillämpad beskrivningsmodell	SUF: tal	SUF: skrift	Roman
Ursprunglig modell (Moon 1998)	6	7	28
Konkretisering	2	–	–
Förskjutet fokus	4	–	1
Totalt antal exempel	12	7	28

När det gäller beskrivningen i termer av icke-infödda drag (Hylténstam 1992) är det i stort sett lika vanligt med kontamination och approximation i materialen. I SUF-materialet förekommer sju kontaminationer, sex approximationer och ett exempel där båda begrepp kan appliceras. I romanmaterialet finns fem kontaminationer, sju approximationer och tre exempel där båda begreppen kan tillämpas. Här bör dock påpekas att endast hälften av exemplen kan beskrivas i termer av dessa icke-infödda drag. Detta sammanfattas i tabell 8.

Tabell 8 Fördelning av icke-infödda drag.

Icke-infött drag	SUF: tal	SUF: skrift	Roman
Kontamination	5	2	5
Approximation	1	5	7
Båda	1	–	3
Ingen	5	–	13
Totalt antal exempel	12	7	28

Sammanfattningsvis kan man konstatera följande: Liksom huvudpersonen Halim i Khemiris roman avviker de åtta informanterna i SUF-projektet i sin användning av fraseologiska enheter från det standard-språkliga referensmaterialet men på ett delvis annorlunda och mer komplext sätt. Avvikelser i användningen av fraseologiska enheter är mer frekventa i den analyserade delen av *Ett öga rött* än hos de åtta informanterna från SUF-projektet. I båda materialen är lexiko-grammatiska avvikelser betydligt vanligare än avvikelser på semantisk-pragmatisk nivå, även om de senare är fler i SUF-materialet än hos Khemiri.

I romanmaterialet avviker alla exempel förutom ett endast på lexiko-grammatisk nivå från referensmaterialet, vilket innebär att den kognitiva processen vid användningen av dessa fraser kan beskrivas

enbart enligt Moons ursprungliga modell (1998) som redovisas i figur 1 ovan. I SUF-materialet finns det däremot sex exempel som avviker från referensmaterialet på båda beskrivningsnivåerna. Dessa exempel analyseras enligt den utvidgade variant av Moons modell som redovisas i figurerna 2 och 3.

När det gäller Halims språkbruk kan man, i likhet med Källström (2006), konstatera att leken med fraseologiska enheter är ett klart utmärkande drag. De avvikelser som observerats i romanmaterialet kan alltså antas bero på en medveten förvrängning av språket i litterärt syfte (jf. Källström 2005:153 om ordföljd), vilket också överensstämmer med Khemiris egen beskrivning av språket i *Ett öga rött* (se bl.a. *Dagens Nyheter* 2004).

Avslutande diskussion

Begreppet idiomschema och Moons kognitiva modell för variation vid användning av fraseologiska enheter (1998:164) har visat sig vara en lämplig utgångspunkt för analysen av föreliggande material. Som tidigare nämnts kan flertalet exempel beskrivas med hjälp av Moons ursprungliga modell (se figur 1). För att man ska kunna beskriva mer komplexa avvikelser i materialet, som konkretisering och förskjutning av fokus i idiomschemat, krävs dock den komplettering av modellen som illustreras i figur 2 och 3.

Ett metodiskt problem i samband med denna studie är utan tvekan det skriftspråkliga referensmaterialet. När det gäller talspråksexemplen från SUF-materialet är det mycket möjligt att flera av dem i jämförelse med ett mer adekvat referensmaterial skulle framstå som mindre avvikande. Problemet är dock, som tidigare påpekats, avsaknaden av relevanta talspråkscorpor.

Denna undersökning baseras på ett litet material och de resultat som redovisas bör därför närmast betraktas som tendenser och utgångspunkter för fortsatt forskning i större skala. Detta gäller naturligtvis även resultatens relevans i förhållande till tidigare forskning; utifrån ett material med detta omfång kan resultaten av tidigare forskning alltså knappast vare sig bekräftas eller falsifieras. De tendenser som visar sig i mitt material pekar dock i flera avseenden i samma riktning som en del tidigare forskning på relevanta områden.

Detta arbete tillför dock ny kunskap om hur avvikelser från konventionaliserade mönster vid bruket av fraseologiska enheter kan se ut och beskrivas, i synnerhet ur ett mångspråkigt perspektiv, något som kan fungera som intressanta utgångspunkter för fortsatt forsk-

ning. En relevant forskningsfråga skulle t.ex. vara huruvida den mer komplexa variationen i användningen av fraseologiska enheter – med avvikelser på både lexiko-grammatisk och semantisk-pragmatisk nivå – är mer frekvent i flerspråkiga ungdomars språkbruk än hos ungdomar med enspråkig svensk bakgrund. Det skulle också vara intressant att – i anknytning till Ekbergs studier (1997, 1998, se även Ekberg 2004) – undersöka om ungdomarna med enspråkig bakgrund generellt är mer stereotypa i sin användning av fraseologiska enheter än ungdomar med en mer heterogen språklig bakgrund.

Med en större informantgrupp och ett mer omfattande material skulle man naturligtvis också kunna undersöka om, och i så fall i vilken mån, avvikelser från konventionaliserade mönster i svenskans fraseologi av den typen som beskrivits i detta arbete är typiska för multietniska ungdomsvarianteter mer generellt. En intressant fråga i detta sammanhang är om sådan eventuell sociolingvistiskt betingat variation i bruket av fraseologiska enheter uppvisar principiella skillnader från variation som snarast skulle förknippas med andraspråksutveckling och avancerad andraspråksanvändning. I det avseende är förekomsten av kontaminationer och approximationer (Hyltenstam 1992), två fenomen som i viss utsträckning iakttagits även i föreliggande material, en intressant aspekt att studera.

Relevant för fortsatt forskning är även frågan om hur pass medvetet informanterna producerar fraseologiska enheter som avviker från svenskans konventionaliserade mönster. Ganuza (2008) har i en studie av ordföljdsvariation inom ramen för SUF-projektet visat att vissa ungdomar använder avvikelser från ordföljdmönster i standardsvenskan, det vill säga rak istället för omvänd ordföljd, medvetet som en aktiv språklig resurs i olika kommunikativa funktioner. Ganuza nämner i detta sammanhang bl.a. informanternas identifikation med den flerspråkiga miljön, solidaritet med kamrater och familj samt motstånd mot standardspråkliga normer och den officiella diskursen i skolan som tänkbara påverkande faktorer (2008:152). Det vore alltså, i anknytning till dessa resultat intressant att se om ungdomarna i SUF-projektet, i likhet med Khemiris romanfigur Halim, även använder avvikelser från standardsvenskans fraseologi i den typen av socio-pragmatiska funktioner som Ganuza beskriver.

En annan intressant fråga är också i vilken utsträckning, och i så fall på vilket sätt, en eventuell högre grad av kreativitet i flerspråkiga ungdomars språkbruk kan komma att påverka svenskans fraseologi även ur ett standardspråkligt perspektiv. Tillgången till flera språk och kulturer är en resurs som på olika sätt kan bidra till ett rikt och varie-

rat svenskt språkbruk, även om man som andraspråksanvändare kan ha tillgång till en relativt sett mer begränsad uppsättning färdiga konstruktioner och fasta fraser. Detta kan naturligtvis resultera i ett mindre idiomatiskt och mindre stereotyp språkbruk (jf. Ekberg 2004). Det ger dock samtidigt utrymme för kreativitet i form av nybildade uttryck med ursprung i andra språk och kulturella tankemönster (jf. Moon 1998) som med tiden kan konventionaliseras och komma att användas som fraseologiska enheter i svenskan. Det är knappast rimligt att anta att fraseologin i svenskt ungdomsspråk och i svenskan i övrigt i längden inte skulle påverkas av en växande flerspråkighet och kulturell mångfald bland ungdomar och i samhället i övrigt.¹⁴

En gemensam nämnare för ungdomarna i SUF-projektet och den fiktiva huvudpersonen i Khemiris roman är uppväxten i, och den dagliga kontakten med, en flerspråkig, multietnisk miljö. Undersökningen visar vissa likheter men framför allt några tydliga skillnader mellan ett autentiskt ungdomsspråk och det konstruerade språket i romanen. Här handlar det dels om kvantitativa skillnader, som tar sig uttryck i förhållandevis fler belägg på avvikelser på fraseologisk nivå i romanmaterialet, dels om kvalitativa skillnader i form av mer komplexa avvikelser på såväl lexiko-grammatisk som semantisk-pragmatisk nivå i SUF-materialet. Sådana skillnader är naturligtvis förväntade med tanke på de båda materialens olika natur. Jämförelsen med de betydligt mer homogena och överdrivet konsekventa avvikelserna från standardspråket i Khemiris konstruerade idiolekt synliggör dock bl.a. komplexiteten i det avvikande bruket av fraseologiska enheter hos de ungdomar som studerats i detta arbete.

Om artikelförfattaren

Julia Große är doktorand i svenska som andraspråk vid Institutionen för svenska språket, Göteborgs universitet. Kontaktadress: julia.grosse@svenska.gu.se

Notar

- ¹ Projektet finansierades av Riksbankens jubileumsfond 2002 till 2006 och drevs i samarbete mellan Göteborgs, Stockholms och Lunds universitet. Den studie som redovisas i denna artikel är en del av ett avhandlingsprojekt där fraseologin i språkanvändningen bland ungdomarna i projektet studeras mera i detalj.
- ² För en mer detaljerad diskussion kring olika typer av fraseologiska enheter, se bl.a. Moon (1998:2ff, 19ff).
- ³ Märkesskor som liknar militärstövlar.

- ⁴ Moon använder begreppen *speaker*, *utterance meaning* och *sentence meaning* i originalfiguren. Begreppen som används här är alltså inga direktöversättningar utan har anpassats för att förtydliga hur modellen tillämpas i föreliggande arbete.
- ⁵ Textmaterial saknas från en av informanterna.
- ⁶ För denna informant saknas uppgifter om startåldern för inläringen av svenska.
- ⁷ Språkbanken är en databas bestående av språkliga referensmaterial i form av olika textkorporer samt lexikaliska data vid Institutionen för svenska språket, Göteborgs universitet.
- ⁸ Institutionen för lingvistik, Göteborgs universitet, <<http://www.ling.gu.se/projekt/tal/index.cgi?PAGE=3>>
- ⁹ Denna del av materialet består av data som Sofia Tingsell (2007) har samlat in i samband med sin forskning.
- ¹⁰ Kodnamn inom SUF-projektet som innehåller information om informantens kön samt vilken skola han eller hon går på. Manliga informanter betecknas med siffror upp till 25, kvinnliga med siffror över 25.
- ¹¹ Tankemönstret som ligger bakom denna fraseologiska enhet är ishockey puckar som skjuts rakt mot mål.
- ¹² Man kan naturligtvis diskutera om *rakt på sak* vanligen används i bokstavlig betydelse. Det är kanske snarare frågan om en graduell skillnad i semantisk kompositionalitet mellan de kontaminerade enheterna i detta exempel.
- ¹³ *Kåre* (ibland *kår*) betyder »vindfläkt« eller »rysning« (Svenska Akademiens ordlista 2006:498)
- ¹⁴ Große och Sköldberg (u. a.) undersöker bl.a. graden av konventionalisering i ungdomars figurativa skriftspråk och huruvida nybildningar på frasnivå kan kopplas till ungdomarnas språkliga och kulturella bakgrund.

Referenser

- Boyd, Sally och Kari Fraurud (u.u.): Challenging the Homogeneity Assumption in Language Variation Analysis. Findings from a Study of Multilingual Urban Spaces. I Auer, Peter och Jürgen Erich Schmidt (red.), *Language and Space – an International Handbook of Language Variation*. Berlin/New York: Mouton de Gruyter.
- Ekberg, Lena 1997: Diskursiva och syntaktiska mönster i återberättelser hos invandrarbarn i Rosengård. I: Gisela Håkansson (utg.). *Svenskans beskrivning 22*. Lund: Lund University Press. s. 97–110.
- Ekberg, Lena 1998: Regeltillämpning kontra lexikonkunskap i svenskan hos invandrarbarn i Rosengård. I: Janus Møller, Pia Quist, Anne Holmen & Jens Norrman Jørgensen (utg.). *Nordiska sprog som andetsprog*. Københavnstudier i tosprogethed 30. Köpenhamn: Københavns Universitet. s. 247–263.
- Ekberg, Lena 2004: Grammatik och lexikon i svenska som andraspråk på nästan infödd nivå. I: Kenneth Hyltenstam och Inger Lindberg (red.). *Svenska som andraspråk – i forskning, undervisning och samhälle*. Lund: Studentlitteratur. s. 259–276.
- Fraurud, Kari och Ellen Bijvoet 2004: Multietniskt ungdomsspråk och andra varieteter av svenska i flerspråkiga miljöer. I: Kenneth Hyltenstam och Inger Lindberg (red.). *Svenska som andraspråk – i forskning, undervisning och samhälle*. Lund: Studentlitteratur. s. 389–418.

- Ganuza, Natalia 2008: *Syntactic Variation in the Swedish of Adolescents in Multilingual Urban Settings. Subject-verb Order in Declaratives, Questions and Subordinate Clauses*. Stockholm: Stockholms universitet, Centrum för tvåspråkighetsforskning.
- Große, Julia och Emma Sköldberg (u. a.): *Figurative language among adolescents in multilingual urban environments*. Föredrag på AILA 2008, s. 24–29 augusti i Essen.
- Hakuta, Kenji 1974: Prefabricated patterns and the emergence of structure in second language acquisition. I: *Language Learning* 24 (2). s. 289–297.
- Hyltenstam, Kenneth 1992: Non-native features of Near-native speakers: On the Ultimate Attainment of Childhood L2 Learners. I: Richard Harris (ed.). *Cognitive processing in Bilinguals*. Amsterdam: Elsevier. s. 351–368.
- Khemiri, Jonas Hassen 2003: *Ett öga rätt*. Stockholm: Norstedts.
- Kotsinas, Ulla-Britt 1988: Rinkebyvenska – en dialekt? I: Per Linell, Viveka Adelsvärd, Torbjörn Nilsson & Per A. Pettersson (red.). *Svenskans beskrivning* 16. Linköping: Universitetet i Linköping. s. 264–278.
- Källström, Roger 2005: Litterärt språk på tvärs. Lite om språket hos Leiva Wenger och Hassen Khemiri. I: Boel De Geer och Anna Malmbjer (utg.). *Språk på tvärs. Rapport från ASLA:s höstsymposium Södertörn*, 11–12 november 2004. (ASLA:s skriftserie 18.) Uppsala. s. 147–158.
- Källström, Roger 2006: »Flygande blattesvenska« – recensenter om språket i *Ett öga rätt*. I: Per Ledin, Lena Lind Palicki, Christina Melin, Gunvor Nilsson, Karolina Wirdenäs och Håkan Åbrink (utg.). *Svenskans beskrivning* 28. *Förhandlingar vid Tjugoåttonde sammankomsten för svenskans beskrivning*, Örebro den 14–15 oktober. Örebro: Örebro universitet. Humanistiska institutionen. s. 147–158.
- Lakoff, George 1993. The contemporary theory of metaphor. I: Andrew Ortony (red.). *Metaphor and Thought*. Cambridge: Cambridge University Press. s. 202–251.
- Lindberg, Inger 2004: Språkanvändning – kreativt skapande eller förutsägbara mönster? I: Mikael Olofsson (red.). *Symposium 2003, Arena Andraspråk*. Stockholm.
- Lorentzen, Henrik 2008: Korpusbaseret lemma-eksklusion eller »Hvor gik ordene hen da de gik ud?«. I: Kristinn Jóhannesson, Hans Landqvist, Aina Lundqvist, Lena Rogström, Emma Sköldberg och Barbro Wallgren Hemlin (utg.). *Nog ordat? Festskrift till Sven-Göran Malmgren*. Göteborg: Meijerbergs Arkiv för svensk ordforskning. s. 259–267.
- Moon, Rosamund 1998: *Fixed expressions and Idioms in English. A Corpus-Based Approach*. Oxford: Oxford University Press.
- Pawley, Andrew och Frances Hodgetts Syder 1983: Two puzzles for linguistic theory: nativelike selection and nativelike fluency. I: Jack Richards och Richard Schmidt (eds.). *Language and Communication*. London: Longman. s. 191–221.
- Quist, Pia 2004. Crossing i storbyens heterogene skole. I: Christine B. Dabelsteen & Juni S. Arnfast (red.). *Taler de dansk? Aktuel forskning i dansk som andetsprog*. Københavnerstudier i tosprogethed 37. København: Københavns Universitet. s. 99–118.
- Searle, John R. 1979: *Expression and Meaning. Studies in the Theory of Speech Acts*. Cambridge: Cambridge University Press.
- Sköldberg, Emma 2004: *Korten på bordet. Innehålls- och uttrycksmässig variation hos svenska idiom*. Göteborg: Meijerbergs arkiv för svensk ordforskning.

- Schmidt, Richard W. 1983: Interaction, acculturation, and the acquisition of communicative competence: a case study of an adult. I Nessa Wolfson & Elliot Judd (eds.). *Sociolinguistics and Language Acquisition*. Rowley MA: Newbury House. s. 137–174.
- Schmitt, Norbert 2005: Formulaic Sequences in action. An Introduction. I: Norbert Schmitt, (ed.). *Formulaic Sequences*. Amsterdam: John Benjamins. s. 1–22.
- Svanlund, Jan 2001: *Metaforen som konvention. Graden av bildlighet i svenskans vikt- och tyngdmetaforer*. Stockholm: Almqvist & Wiksell International.
- Tingsell, Sofia 2007: *Reflexivt och personligt pronomen. Anaforisk syftning hos ungdomar i flerspråkiga stadsmiljöer*. Göteborg: Göteborgs universitet, Institutionen för svenska språket.
- Wray, Alison 2002: *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.

Övriga referenser

- Dagens Nyheter* 2004. Måndag 2 augusti 2004.
- Göteborg Spoken Language Corpus (GSLC) 2008: Institutionen för lingvistik, Göteborgs universitet <<http://www.ling.gu.se/projekt/tal/index.cgi?PAGE=3>>.
- Språkbanken 2006a: <http://spraakbanken.gu.se/cgi-bin/lb/parole/parole.cgi>.
- Språkbanken 2006b: <http://spraakbanken.gu.se/parole/parole_material_kommentar.phtml>.
- Språkbanken 2008. <<http://spraakbanken.gu.se>>
- Språk och språkbruk bland ungdomar i flerspråkiga stadsmiljöer* (SUF) 2008: Institutionen för svenska språket/ISA, Göteborgs universitet, <http://hum.gu.se/institutioner/svenska-spraket/isa/verk/projekt/pag/pg_forsk2>.
- Svenska Akademiens ordlista* 2006: 13:e upplagan. Stockholm: Norstedts Akademiska Förlag.
- Svenskt språkbruk. Ordbok över konstruktioner och fraser* 2003: Stockholm: Norstedts ordbok.
- Universitetet i Oslo 2008: <<http://www.hf.uio.no/iln/forskning/forskningsprosjekter/upus/>>

Appendix

Transkriptionsnyckel

- + avbrutet ord
- xx ohörbart tal
- (t), (g) etc. reducerad (tyst) konsonant i slutet av ord
- VERSALbetonat ord eller stavelse

II

*KLÄTTRA PÅ VÄGGARNA ELLER
BARA VARA ETT MED SOFFAN*

Klättra på väggarna eller bara vara ett med soffan?

Om figurativa ordförbindelser bland ungdomar i flerspråkiga skolmiljöer

AV JULIA PRENTICE och EMMA SKÖLDBERG

Abstract

Prentice, Julia, julia.prentice@svenska.gu.se, Ph.D student, Dept. of Swedish Language, Institute for Swedish as a Second Language, University of Gothenburg, Sweden; *Sköldbberg, Emma*, emma.skoldberg@svenska.gu.se, Ph.D, Researcher, Dept. of Swedish Language, University of Gothenburg, Sweden: “*Climb the walls or become one with the sofa? On the use of figurative word combinations by adolescents in multilingual school environments*”. *Språk och stil* 20, 2010, pp. 5–xx.

This article reports on a survey on the use of figurative word combinations in texts produced by high school students in multilingual urban school environments in Sweden. The data consist of essays written by 175 students and have been collected for the research project *Language and language use among young people in multilingual urban settings*. The texts have been divided into five different categories according to the students' linguistic background. It was found that the figurative word combinations in the data can be classified into three main categories according to their degree of conventionalization: (1) *conventionalized word combinations*, (2) *partially modified conventionalized word combinations* and (3) *novel word combinations*. Conventionalized word combinations appear to be the most frequently used type in the data, followed by partially modified and novel figurative word combinations, respectively. The results also support the initial hypothesis that the salience and types of figurative expressions in the participants' texts can, at least to some degree, be linked to the participants' linguistic background; L1 students use more conventionalized word combinations than L2 students, whereas modifications of conventionalized figurative expressions are more frequent in the L2 students' texts. The study also shows that the students create novel figurative word combinations based on a great variety of source domains.

Keywords: phraseology, youth language, figurative language, conventionalization, Swedish.

1. Inledning

Ungdomars språkbruk i flerspråkiga storstadsmiljöer har diskuterats flitigt i svenska medier. Åsikterna som framförts, bl.a. i en debatt på Dagens Nyheters kultursidor (se t.ex. DN 19.4.2006), har dock ofta saknat stöd i språkvetenskap-

lig forskning. Debatten har dessutom präglats av en grundläggande förvirring angående skillnaden mellan sociolingvistiska fenomen som multietniska ungdomsvarieteter och inlärafenomen som bristfällig svenska eller brytning (Bjvoet & Fraurud 2006).

I denna artikel redogör vi för en undersökning som genomförts inom ramen för forskningsprojektet Språk och språkbruk hos ungdomar i flerspråkiga stadsmiljöer (hädanefter SUF).¹ Syftet med projektet har varit att på olika språkliga plan undersöka och beskriva hur språkbruket bland ungdomar i svenska storstadsområden, som präglas av stor språklig mångfald bland invånarna, faktiskt ser ut.

Föremålet för just denna studie är figurativa ordförbindelser i texter skrivna i samband med nationella prov i svenska och svenska som andraspråk. Nedan återges två textavsnitt ur elevtexterna som exempel (1) och (2). Dessa, och övriga textavsnitt i artikeln, återges oredigerade, dvs. exakt såsom det står i elevernas texter. Exempelen är intressanta eftersom de innehåller figurativa ordförbindelser av det slag som undersöks här. För att förtydliga har de aktuella fraserna kursiverats.

(1) Det finns en dokusåpa som jag personligen inte skulle *sortera i samma fack* som andra dokusåpor här i Sverige. Programmet är Fame Factory. [...] Det anser jag vara mer underhållande än att se isolerade människor i ett hus som *klättrar på väggarna* för att dem är uttråkade.

(2) Mycket av det hon presenterar i sin insändare är just det att hon vill känna att hon har under veckan varit så stressad så när fredagen kommer ska hon bara få pusta ut, vila och *bara vara ett med soffan*.

I materialet finns det för det första välkända figurativa ordförbindelser som *klättra på väggarna*. För det andra finns det fraser som påminner om väletablerade uttryck men i viss mån avviker från dessa till form eller innehåll, t.ex. *sortera i samma fack*, som kan jämföras med betydligt vanligare varianter som *hamna* eller *placera i samma fack*. Slutligen finns det formuleringar som är mer nybildade och kreativa, t.ex. *bara vara ett med soffan*.

För vår studie har vi utgått från följande frågeställningar:

- I vilken omfattning använder eleverna figurativa ordförbindelser i de analyserade texterna?
- Vilka typer av figurativa ordförbindelser finns det i materialet och hur konventionaliserade är de?

¹ SUF-projektet finansierades av Riksbankens Jubileumsfond (2002–2006) och bedrevs i samarbete mellan Göteborgs, Stockholms och Lunds universitet. Emma Sköldberg har genomfört sin del av undersökningen inom ramen för sin forskartjänst som finansieras av Svenska Akademien.

- Finns det skillnader mellan olika informanternas användning av figurativa ordförbindelser, och kan de eventuella skillnaderna kopplas till informanternas språkliga och kulturella bakgrund?

En undersökning av figurativa ordförbindelser, som producerats i en skolmiljö som präglas av språklig mångfald, är intressant då såväl språkliga bilder som sätten att uttrycka dem lingvistiskt ofta är (mer eller mindre) konventionaliserade på olika vis inom olika språkgemenskaper (jfr Moon 1998 s. 163). De resultat som presenteras här baseras på både kvalitativa och kvantitativa metoder. Kvalitativa analyser har använts för att kategorisera och beskriva ordförbindelserna i materialet. De kvantitativa analyserna tydliggör frekvensen av olika typer av ordförbindelser i texterna och skillnader i bruk mellan informantgrupper med olika språklig bakgrund.

Artikeln är disponerad på följande sätt: I avsnitt 2 redogör vi kortfattat för forskning om figurativa ordförbindelser och deras betydelse vid andrapråksanvändning. Vi redogör också för svensk forskning om figurativt språk i såväl enspråkiga som flerspråkiga elevers språkbruk. I avsnitt 3 presenteras vårt material, dvs. våra informanter, de texter och – framför allt – de ordförbindelser som vi undersökt närmare. I avsnitt 4 presenteras de tre huvudtyper av ordförbindelser som framträder i materialet och i avsnitt 5 redogörs för hur dessa typer används. Här diskuteras också eventuella skillnader i bruket av de aktuella ordförbindelserna bland olika elevgrupper. Artikeln avslutas med en diskussion av resultat och utgångspunkter för fortsatt forskning på området.

2. Teoretiska utgångspunkter

2.1 Figurativa fraseologiska enheter

I ett inledande skede led fraseologi som forskningsområde av begreppsförvirring. Något som bidrog till förvirringen, och som i viss mån fortfarande är aktuellt, är att forskare inte lade samma betydelse i begrepp som användes vid beskrivningen av olika uttrycks egenskaper, som t.ex. *idiomaticitet*, *kompositionaltitet*, *analyserbarhet* och *konventionalisering* (jfr Wray 2002 s. 8 f.). Nuförtiden råder dock i stort sett konsensus vad gäller den övergripande terminologin inom ämnesområdet, och huvudtyperna av flerordsuttryck definieras på liknande sätt av de flesta fraseologiforskare (Burger 1998 s. 33).

Gläser (1988 s. 271) konstaterar att typologisering av uttryck länge har varit grundläggande inom det fraseologiska arbetet. Detta har resulterat i en rad olika typologier över fraseologiska enheter (för presentationer av ett antal sådana, se t.ex. Fleischer 1997 s. 110 ff., Sköldbberg 2004 s. 29 ff.). Typologierna baseras på olika kriterier, bl.a. uttryckens grammatiska struktur (se t.ex. Anward & Linell 1976, Burger 1998) och den bokstavliga betydelsen hos konstituenterna som ingår i fraserna (se t.ex. Krohn 1994). Vidare finns det typologier som baseras på uttryckens överförda betydelse (se t.ex. Schottman & Petersson 1989). Som Burger lyfter fram är det syftet med klassificeringen och studien ifråga som avgör vilken typologi som är lämpligast och hur detaljerad den bör vara. Denna pragmatiska syn på typologier över flerordsuttryck innebär att klassificeringen av vissa fraser kan ske mer eller mindre ad hoc och att detta i viss mån måste accepteras som en del i klassificeringsprocessen (Burger 1998 s. 33). I denna studie har vi funnit det lämpligt att använda en tredelad typologi som baseras på hur konventionaliserade uttrycken är. Begreppet *konventionaliserad* definieras här som ett inom en språkgemenskap etablerat sätt att uttrycka ett språkligt innehåll (jfr Svanlund 2001 s. 56 f., 2002 s. 8; jfr också termen *institutionaliserad* som används med liknande innebörd; se vidare Sköldbberg 2004 s. 23 f.). Den typologi som vi arbetat utifrån beskrivs närmare i avsnitt 4.

I denna studie ligger fokus på figurativa ordförbindelser, dvs. flerordskombinationer som innehåller någon form av metaforik, metonymi eller liknande (jfr Nunberg m.fl. 1994 s. 492). Ordförbindelserna kännetecknas av att de semantiskt inte är helt kompositionella, dvs. deras betydelse kan inte härledas från en ord-för-ord-tolkning av de ingående konstituenterna (Moon 1998 s. 8 f., Langlotz 2006 m.fl., jfr även Svanlund 2002).

Grundläggande för figurativt språk är som sagt *metaforen*. Litteraturen kring begreppet är omfattande och det finns ingen vedertagen definition av fenomenet. Det beror bl.a. på att olika forskare koncentrerat sig på olika typer av metaforer (Svanlund 2001 s. 10). Men enligt Ekberg (1993 s. 107) innebär en metaforisk överföring att man använder sig av en semantisk struktur som hör till en viss erfarenhetsdomän för att uttrycka en betydelse som hör hemma i en annan erfarenhetsdomän (se även Lakoff & Johnson 1980 och Svanlund 2001 s. 14 ff.). Ekberg skriver vidare:

Inom den kognitiva semantiken benämner man de båda domäner som är involverade i en metaforisk överföring för käll- respektive måldomänen. Den »bokstavliga» betydelsen hos ett metaforiskt uttryck är alltså den betydelse som uttrycket har i källdomänen, medan den »överförda» betydelsen är den betydelse som uttrycket har i måldomänen. (Ekberg 1993 s. 107)

En uppsättning projiceringar mellan två domäner kallas ofta kollektivt för en *domänmappning* (Svanlund 2001 s. 14). Det kan alltså vara flera metaforiska uttryck som projiceras, dvs. hämtas, från en och samma källdomän. Ett exempel på en sådan överföring är metaforiska uttryck för måldomänen *kärlek* som hämtas från källdomänen *resande*. Detta resulterar i svenska uttryckssätt som *stå vid ett vägskäl* eller *köra fast* i sin relation (Svanlund 2001 s. 14 ff., se även Lakoff & Johnson 1980 s. 44). Idén om domänmappning kan sägas ha vidareutvecklats inom den s.k. *blendteorin*, som lanserades av Fauconnier & Turner på sent 1980-tal. Inom denna teori betraktas domänerna i större utsträckning som interagerande – till skillnad från en traditionellare syn på domänmappning som innebär en mer ensidig överföring från källdomän till måldomän (se Fauconnier & Turner 2002; för kortare sammanfattningar av teorin, se t.ex. Svanlund 2001 s. 16–19 och Knowles & Moon 2006). Några exempel på senare metaforstudier, som också tar upp blendteorin, är Müller 2008 och Semino 2008.

En central typ av ordförbindelse bland flerordsuttrycken är kategorin *idiom*. Också den har definierats på olika sätt (se Sköldberg 2004 s. 21–29 med anförd litteratur). I det här arbetet använder vi termen i enlighet med Langlotz, som utgår ifrån följande definition:

An idiom is an institutionalised construction that is composed of two or more lexical items and has the composite structure of a phrase [...], which may feature constructional idiosyncrasy. An idiom primarily has an ideational discourse-function and features figuration, i.e. its semantic structure is derivationally non-compositional. Moreover, it is considerably fixed and collocationally restricted. (Langlotz 2006 s. 5)

Denna definition är vid och inkluderar så pass olika typer av figurativa konstruktioner som *red herring* 'villospår' och *take the bull by the horns* 'ta tjuren vid hornen' (ibid.). Langlotz är dock noggrann med att påpeka att gränstragningen mellan idiomerna och intilliggande fraseologiska enheter på intet vis är skarp.

Ett annat relevant begrepp för denna studie är *liknelse* som enligt Burger (1998 s. 44) och Moon (2008) är ett flerordsuttryck som innehåller en jämförelse. På svenska uttrycks denna grammatiskt med hjälp av en komparativ subjunktionsfras med *som*. Jämförelsen fungerar ofta som förstärkning till ett verb eller adjektiv (t.ex. *äta som en häst* och *ful som stryk*). Liknelser kan vara allt från helt nyskapade till konventionaliserade (jfr Wellander 1939 s. 624–631, Svensén 2004 s. 244).

2.2 Variation av fraseologiska enheter

Ett återkommande drag i beskrivningar av fraseologiska enheter är att de kännetecknas av något slags *fixering*. Fixeringen innebär här att dessa enheter, till skillnad från fria syntaktiska konstruktioner, typiskt förekommer i ett begränsat antal syntaktiska former eller konstruktioner (jfr Nunberg m.fl. 1994 s. 492). Såsom Burger (1998 s. 25) påpekar är dock denna fixering relativ. Det är i själva verket väldigt få flerordsenheter som inte tillåter någon lexikal eller annan strukturell variation.

Burger skiljer också mellan *variation* och *modifiering* av fraseologiska enheter. Med variation avses användningen av olika konventionaliserade varianter av ett och samma fraseologiska uttryck (Burger 1998 s. 25, Moon 1998 s. 161). Två varianter av ett och samma idiom är t.ex. *strö salt i såren/såret* (Svenskt språkbruk 2003 s. 984). Modifiering är däremot en medveten avvikelser från den konventionaliserade formen i avsikt att uppnå något slag av stilistisk effekt (Burger 1998 s. 27 f.). Exempel på modifieringar är formuleringar som *strö ett par kilo salt i såren* och *strö salt i öppna sår* (ur materialen Press 96 och GP 01 i Språkbanken, <<http://spraakbanken.gu.se>>), där inskjutna bestämningar, troligen medvetet, använts av tidningsskribenter. Icke avsiktliga avvikelser från uttryckens standardform klassificerar Burger som *fel*. Författaren konstaterar dock att det kan vara svårt att skilja på varianter, modifieringar och sådana användningar som ska betraktas som felaktiga (1998 s. 28). I föreliggande arbete försöker vi heller inte skilja modifieringar och fel från varandra, eftersom det är omöjligt för oss att veta om en avvikelse från standardbruket är medveten eller inte. Vidare är det inte alltid så enkelt att klassificera en användning som felaktig. Formuleringen kan vara grammatiskt korrekt, men icke typisk och därmed uppfattas som mindre idiomatisk (jfr Hultman & Westman 1977 s. 246).

I jämförelse med lexiko-grammatiska avvikelser har variation och avvikelser från konventionaliserat bruk vad gäller uttryckens semantik fått mindre uppmärksamhet. Sköldberg (2004 s. 76) skriver att det inte råder konsensus bland forskare vad gäller t.ex. idioms semantiska varierbarhet. Vissa forskare, som t.ex. Lindfors Viklund (1991), konstaterar att idiom i princip är semantiskt fixerade enheter, medan andra, som t.ex. Koller (1977) och Clausén (1996), menar att uttryckens idiomatiska betydelse ofta är vag och att den styrs av den kontext som idiomerna används i. Sköldberg (2004) visar dock att det finns stora skillnader mellan idiom vad gäller semantisk variation; det rör sig om ett kontinuum från sådana som varierar mycket lite till sådana som varierar relativt

mycket. Ett svenskt uttryck som uppvisar starkt begränsad semantisk variation är enligt Sköldberg *lägga sista handen vid något* med betydelsen 'till slut bli färdig med något'. I hennes material är den som lägger sista handen vid något i stort sett alltid en person, som till slut bli färdig med någon form av kreativt, ofta skriftligt, arbete. Ett idiom som varierar mer på det semantiska planet är *sätta käppar i hjulet för någon/något* med betydelsen '(avsiktligt) vålla (oöverstigit) hinder för någon eller något'. Den semantiska variationen består i huvudsak i att olika betydelsefaktorer realiseras i olika kontexter. För det första kan hindret vållas avsiktligt eller inte och för det andra kan hindret vara allt från oöverstigit till endast lätt besvärande (se Sköldberg 2004 s. 140, 156–158 och 2005 för exempel).

Ett antagande för denna studie är sålunda att ordförbindelsernas konventionaliserade betydelse i viss mån kan variera och att betydelsen delvis beror på den kontext som uttrycken används i. Detta gäller förstås inte bara figurativa ordförbindelser av det slag som undersöks här utan även många andra typer av lexikala enheter (se vidare Sköldberg 2004 s. 76).

2.3 Användningen av figurativt språk i svenska elevers skrivande

Elevtexter har studerats ur en rad olika perspektiv. För en översikt över tidigare svenska studier, se t.ex. Nyström (2000 s. 21–22) och Norberg Brorsson (2007 s. 34 ff.) med anförda referenser. Forskare har valt att fokusera på bl.a. elevernas ordval, syntax och referensbindning, men också på elevtexternas disposition och på lärarnas bedömning av arbetena (se Hultman & Westman 1977, Josephson m.fl. 1990, Ledin 1998, Nyström 2000, Östlund-Stjärnegårdh 2002, Palmér & Östlund-Stjärnegårdh 2005, Wiksten Folkeryd 2006, Norberg Brorsson 2007 m.fl.). Ett aktuellt exempel är också Magnussons undersökning av *grammatiska metaforer* i texter skrivna av informanterna som ingår i SUF-projektet (se t.ex. Magnusson 2009). Med grammatiska metaforer menas, i enlighet med den systemisk-funktionella grammatikens modell, betydelser som uttrycks på ett lexiko-grammatiskt inkongruent sätt. Detta sker t.ex. vid nominalisering, där en process, som typiskt uttrycks genom ett verb, i stället uttrycks genom ett substantiv. Magnusson är alltså i viss mån intresserad av överförd betydelse i elevernas texter. Det är dock *grammatiska* metaforer och inte figurativa uttryck i sig som är i fokus för hennes studie.

Överlag har dock bruket av figurativt språk i elevtexter fått mindre uppmärk-

samhet. Josephson m.fl. (1990 s. 54), som undersöker skoluppsatser från år 1–9, konstaterar att bilder är sällsynta i elevtexter och att sådant bruk därför bör uppmuntras av lärarna. Liknande resonemang återfinns hos Hultman & Westman som undersöker gymnasietexter skrivna på tidigt 1970-tal. Hultman & Westman (1977 s. 253) skriver att det är oroande att gymnasisterna som grupp visar få prov på stilistiska experiment, vilket leder till att »deras språk så ofta är en grå, trögflytande och halvljummen välling som knappast retar aptiten». Dessa iakttagelser kan jämföras med t.ex. Wennerholm (2003) som funnit relativt många exempel på figurativt språk i texter skrivna av gymnasielever år 2002. Han konstaterar dock att det stora antalet döda och lexikaliserade metaforer i hans material antyder att bildspråk ofta används utan medvetet stilistiskt syfte (2003 s. 31). Denna slutsats kan dock diskuteras. För det första är det inte helt okomplicerat att avgöra om en metafor är död eller levande (jfr exempelvis Svanlund 2001), och för det andra torde det vara omöjligt att avgöra om metaforerna används medvetet eller inte utan att intervjua skribenterna.

Man måste dock vara medveten om att det är svårt att dra några generella slutsatser om bruket av figurativt språk i elevtexter med tanke på att det är så många individuella och situationsbundna faktorer som spelar in, t.ex. elevernas språkliga bakgrund, vilken texttyp eller genre det gäller och ämnet för texten.

2.4 Fraseologi i flerspråkiga elevers språkbruk

Ett annat forskningsfält som är relevant för denna studie rör eventuella skillnader i bruket av fraseologiska enheter i enspråkiga respektive flerspråkiga elevers språkanvändning, ett område som bl.a. Ekberg har utforskat (1997, 1998, 2004). Hennes resultat visar att såväl enspråkiga som flerspråkiga elever i årskurserna 5, 6, 7 och 9 har aktiv kunskap om konventionaliserade ordförbindelser. Det finns emellertid en skillnad i fråga om frekvens i användning: de enspråkiga eleverna använder fler konventionaliserade uttryck än de flerspråkiga eleverna. Konsekvensen av detta blir att de flerspråkiga elevernas språkbruk är mindre stereotyp – vilket också gör att det många gånger upplevs som mindre idiomatiskt, dvs. mindre likt ett infött språkbruk.

Också Hyltenstam (1992) har undersökt konventionaliserade ordförbindelser i tvåspråkiga ungdomars språkbruk. Han beskriver s.k. *icke-infödda drag* i både talat och skrivet språk hos gymnasieelever som talar svenska på nästan infödd nivå. Hyltenstam jämför två grupper av tvåspråkiga informanter, med finska respektive spanska som modersmål, med en grupp som har svenska som

modersmål. En relativt vanlig typ av icke infödda drag som han beskriver är s.k. *approximationer* till fraseologiska enheters standardform. Ett exempel är »den fungerar nog inte på längden», där »på längden» beskrivs som en approximation till uttrycket *i längden* (a.a. 1992 s. 360). Hyltenstam har funnit att denna typ av avvikelse från standardanvändningen av fraseologiska enheter är vanligare bland de tvåspråkiga eleverna än bland de enspråkiga i kontrollgruppen. Ett annat fenomen i materialet är så kallade *kontaminationer*, dvs. uttryck där element från två eller flera olika konventionaliserade uttryck kombineras (s. 361). Som exempel anförs »man beräknar med att» där element från ordförbindelserna *räkna med något* och *beräkna något* har kombinerats. Intressant nog är kontaminationerna lika vanliga i den enspråkiga kontrollgruppens som i den tvåspråkiga gruppens språkbruk (s. 361).

Även Große (2008) beskriver olika typer av avvikelser från standardspråket vid användningen av fraseologiska enheter. Materialet för hennes undersökning består dels av talat och skrivet språk, hämtat från SUF-projektet, dels av utdrag ur romanen *Ett öga rött* (Khemiri 2003). De avvikelser som hon beskriver kan uppträda i form av lexiko-grammatiska modifieringar eller fel, som i *det går som kallaste kår i ryggen* (jfr *Svensk språkbruk* 2003 s. 672), eller när ett fraseologiskt uttryck används i en kontext där det inte kan tolkas på konventionaliserat sätt. Den senare typen exemplifierar Große (2008 s. 116 f.) med ett utdrag ur ett gruppsamtal från SUF-materialet, där eleverna diskuterar kärleksdikter.

P29: asså den här nog ganska så här kanske lite erotiskt på sin tid så här
[...]

P38: DEN!... i dagens läge så är det värsta BARNLEKEN eller nånting herregud

Enligt *Svenskt språkbruk* (2003 s. 84) har uttrycket *något är en barnlek* den konventionaliserade betydelsen '[något är en] mycket enkel uppgift'. I kontexten ovan framhävs däremot det barnliga och oskyldiga i barnens lek, i motsats till påståendet om diktens erotiska innebörd.

3. Material och metod

I det här avsnittet beskrivs hur informantgruppen har delats in med hänsyn till elevernas språkliga bakgrund. Vidare presenterar vi elevtexterna och själva studieobjekten, nämligen de figurativa uttryck som förekommer i materialet.

3.1 Informanter och informantgrupper

I denna studie har vi utgått ifrån 175 av de 222 informanter som finns med i SUF-projektet då vi endast inkluderat de elever som har gjort det nationella provet och som vi har tillräcklig bakgrundsinformation om. Samtliga elever går det andra året på gymnasiet och deras skolor ligger i flerspråkiga områden i Stockholm, Göteborg och Malmö.

För att kunna undersöka likheter och skillnader mellan olika elevgruppers språkbruk har vi delat in informanterna i fem grupper, baserade på den bakgrundsinformation som informanterna själva har lämnat. De faktorer som ligger till grund för grupperingen är vilka språk som informanterna anger vara deras första- respektive andraspråk, om de har vuxit upp i en-, två- eller flerspråkiga hem, vilket land deras föräldrar är födda i samt deras startålder för inläring av svenska. De fem informantgrupperna är:

- (1) SVE: informanter med enspråkig svensk bakgrund (54 elever)
- (2) SVE+: informanter med svenska som förstaspråk och med en eller två utlandsfödda föräldrar (28 elever)
- (3) TVÅ: informanter som angivit att de har svenska och något eller några andra språk som förstaspråk, dvs. informanter som vuxit upp i ett två- eller flerspråkigt hem (10 elever)
- (4) START < 6: informanter med svenska som andraspråk med en startålder för svenska före skolstarten, dvs. före 6–7 års ålder (60 elever)
- (5) START > 6: informanter med svenska som andraspråk och med en startålder efter att de börjat grundskolan, dvs. efter 6–7 års ålder (23 elever).

Vi är medvetna om att denna indelning inte helt och hållet speglar den mångfald som finns när det gäller informanternas språkliga bakgrund (jfr Boyd & Fraurud under utg.), men har funnit att ovan nämnda faktorer och denna indelning är de mest relevanta för vårt syfte.

De flesta informanterna (60 elever) ingår i grupp 4, dvs. de har alltså svenska som andraspråk och har lärt sig svenska före skolstart, framför allt i samband med att de börjat förskolan. I undersökningen ingår också ett förhållandevis stort antal elever (54 personer) med enspråkig svensk bakgrund, dvs. grupp 1. Grupp 2, med elever som har svenska som förstaspråk men en eller två utlandsfödda föräldrar, och grupp 5, med informanter som lärt sig svenska efter att de börjat skolan, innehåller 28 respektive 23 elever. Däremot är det bara 10 elever som har angivit att de har svenska och något annat/några andra språk som förstaspråk, dvs. som har vuxit upp i ett två- eller flerspråkigt hem (grupp 3).

Det faktum att grupperna är olika stora är givetvis ett problem vid jämförelsen mellan informanternas språkbruk. För att testa om de kvantitativa skillnaderna i materialet är statistiskt signifikanta, har vi använt ett icke-parametriskt test (²) och Pearson Correlation-test (se vidare Körner & Wahlgren 2006).²

3.2 Textmaterialet

I denna studie har vi granskat texter som producerats inom ramen för de nationella proven i svenska och svenska som andraspråk på ungdomsgymnasiet och komvux under höstterminen 2003 och vårterminen 2004 (svenska B/svenska som andraspråk B). Uppsatsdelen i provet består av två separata skrivuppgifter och utgångspunkten för dessa uppgifter är ett häfte på ca 30 A4-sidor. Häftet innehåller olika typer av text, bl.a. artiklar, dikter, skönlitterära texter, men också bildmaterial såsom serieteckningar och reklambilder. De texter och bilder som ingår i häftet knyter an till ett visst tema, i vårt fall »Tid». De två texthäften som våra informanter har utgått från bär nämligen rubrikerna »Har du tid?» (ht 03) och »Människan och tiden» (vt 04). (För en mer ingående beskrivning av nationella proven, se Garne 2002 och <www.skolverket.se>. Se t.ex. Nyström 2000, Nyström & Ohlsson 2002 och Östlund-Stjärnegård 2002 för andra språkvetenskapliga undersökningar av texter skrivna inom ramen för nationella proven i svenska.)

Elevuppsatserna varierar avsevärt i längd. Den kortaste är på 168 löpande ord och den längsta på 2103 löpande ord. I tabell 1 sammanfattas hur materialet fördelar sig på de fem informantgrupperna.

Tabell 1. Korpusens storlek och fördelningen på informantgrupper.

Informantgrupp	Antal informanter	Antal ord/informantgrupp	Antal ord/informant
1 (SVE)	54	45 812	848
2 (SVE+)	28	25 218	901
3 (TVÅ)	10	8 090	809
4 (START < 6)	60	48 981	816
5 (START > 6)	23	18 856	820
Totalt	175	146 957	840

Som framgår av tabell 1 rör det sig om ett textmaterial på sammanlagt närmare 147 000 löpande ord. Granger (1998 s. 11) menar att om man undersöker hög-

² Ett varmt tack till FD Sofie Johansson Kokkinakis för hjälp med signifikansprövningen av vårt material.

frekventa ord eller strukturer i språket kan relativt små korpusar på ca 20 000 ord räcka. Vid mer lågfrekventa språkliga enheter – som i vårt fall – krävs det däremot en större korpus, vilket även visar sig vara ett problem i samband med våra kvantitativa resultat (se vidare avsnitt 5). Storleken på vårt material kan dock jämföras med t.ex. det gymnasistmaterial på nästan 90 000 löpande ord som Hultman & Westman (1977 s. 52) undersökte. Olevard (2002) undersöker drygt 155 000 ord i texter producerade av elever i år 9.

Vidare framgår det av tabell 1 att det är eleverna i grupp SVE+, dvs. informanterna med svenska som förstaspråk med en eller två utlandsfödda föräldrar, som i genomsnitt har skrivit de längsta texterna med i medeltal 901 löpande ord. Men för övrigt har eleverna i de olika grupperna i genomsnitt producerat ungefär lika mycket text; det rör sig om mellan 816 och 848 löpande ord/elev.

3.3 Figurativa ordförbindelser i materialet

De ordförbindelser som excerperats ur materialet har figurativt innehåll. För att klarare avgränsa materialet har vi valt att exkludera enkla metaforiska lexikala enheter (t.ex. *källa*) och partikelverb (t.ex. *sticka ut*). Till skillnad från några tidigare studier, se Magnusson 2009, i vilka forskarna studerat elevtexter ur ett mer funktionellt perspektiv, är vår analys främst strukturellt baserad (jfr Josephson m.fl. 1990 s. 8–10). Vi strävar efter att beskriva de figurativa ordförbindelsernas form och betydelse i de aktuella kontexterna. Däremot fokuserar vi inte på texterna i sin helhet.

De excerperade figurativa ordförbindelserna har delats in i tre huvudtyper utifrån deras grad av konventionalisering. De tre typerna är:

- a) konventionaliserade figurativa ordförbindelser
- b) modifierade konventionaliserade ordförbindelser
- c) nybildade figurativa ordförbindelser.

För att avgöra i vilken mån en ordförbindelse är konventionaliserad i fråga om form och betydelse har vi i första hand använt konstruktionsordboken Svenskt språkbruk. Ordbok över konstruktioner och fraser (2003, hädanefter SSB), som är den senaste och mest heltäckande ordboken över svenska fraseologiska enheter. Man bör dock komma ihåg att ordböcker, och då i synnerhet tryckta ordböcker som SSB, har begränsningar vad gäller omfånget. De innehåller ett urval av verkligt språkbruk (se Svensén 2004 s. 79–87 om några faktorer att ta hänsyn till vid urvalet av data i ordböcker). I denna studie har vi därför också

jämfört de excerperade ordförbindelserna med språkanvändningen i de autentiska texter som ingår i Språkbanken vid Institutionen för svenska språket, Göteborgs universitet (<<http://spraakbanken.gu.se/konk/>>). I vissa fall, när ordförbindelserna enligt vår egen språkliga intuition är konventionaliserade, men varken SSB eller Språkbanken ger något entydigt svar, har vi dessutom använt oss av sökmotorn Google (<www.google.se>). Som exempel kan en formulering som finns i flera av elevtexterna anföras, nämligen *hålla näsan över vattnet* i betydelsen 'undgå att gå under'. Konstruktionen är inte upptagen i SSB. I Språkbankens texter (som innehåller drygt 100 miljoner löpande ord) används uttrycksättet *hålla näsan över vattnet* endast en handfull gånger. Men det finns liknande uttrycksätt som används med samma betydelse, t.ex. *ha näsan över vattnet* och *hålla näsan över vattenytan*. Vidare får man mer än 1 000 träffar på strängen »näsan över vattnet» på Google (12.6.2009). Även om det går att ifrågasätta tillförlitligheten hos Googles frekvensangivelser och i vad mån de Internettexter som man får fram via sökmotorn alltid är representativa för något slags standardspråklig norm (se bl.a. Lorentzen 2008 s. 260, Moon 2008), styrker den relativt frekventa användningen vår intuitiva uppfattning om att ordförbindelsen *hålla näsan över vattnet* är konventionaliserad på svenska. Man kan åtminstone konstatera att den är en av flera likartade ordförbindelser som alla går tillbaka på en viss källdomän, nämligen djupt vatten, och som alla har en gemensam måldomän, nämligen att nått och jämnt klara av sin tillvaro. Exemplet illustrerar dock att det rör sig om ett kontinuum från icke konventionaliserade ordförbindelser till konventionaliserade sådana och att det ibland är svårt att avgöra vilken av de tre kategorierna ovan en viss ordförbindelse tillhör. Något som också inverkar och försvårar en striktare klassificering är att konventionalisering inte är ett statistiskt fenomen utan en process och att en ordförbindelses konventionalisering kan förändras över tid (jfr Svanlund 2001 s. 10).

4. Tre typer av figurativa ordförbindelser

Som tidigare påpekats har vår analys resulterat i en tredelad typologi, dvs. tre typer av figurativa ordförbindelser som uppvisar olika grad av konventionalisering. De tre typerna av figurativa ordförbindelser beskrivs och exemplifieras nedan.

4.1 Konventionaliserade figurativa ordförbindelser

Den första typen av ordförbindelse består främst av idiom och lexikaliserade liknelser (se avsnitt 2.1 ovan). Dessa exemplifieras i (3) och (4) nedan:

(3) Alla dessa dokusåpor *står mig upp i halsen!* (C34, TVÅ)

(4) Fast dessa [dokusåpadeltagarna] som gör detta, de har också någon talang eller är skicklig inom vissa saker men det kommer inte fram till folket för det enda som visas är intriger efter intriger och lite fest »*som grädden på moset*». (C06, SVE+)³

I exempel (3) och (4) används idiomat *något står någon upp i halsen* respektive liknelsen *som grädden på moset* (jfr SSB 2003 s. 455, 433). Andra exempel på konventionaliserade figurativa ordförbindelser i materialet är det redan nämnda *klättra på väggarna* (D46, SVE) men också *av hela mitt hjärta* (S37, SVE) och *dra något över en kam* (B13, START < 6) (jfr SSB 2003 s. 1366, 489, 576).

4.2 Modifierade konventionaliserade ordförbindelser

Den andra typen av figurativa ordförbindelser består av formuleringar som avviker från konventionaliserat bruk. De är därmed en form av modifieringar. Denna kategori är något mer komplex, då den innehåller tre olika typer av modifieringar. De kan vara: a) lexiko-grammatiska, b) semantisk-pragmatiska eller c) både lexiko-grammatiska och semantisk-pragmatiska.

Den första typen, dvs. en lexiko-grammatisk modifiering, återfinns i exempel (5) nedan:

(5) Att fundera över tid, dåtid, nutid och framtid kan som sagt göra en snurrig. Man *vet varken in eller ut*. (P28, SVE)

I exemplet används ordförbindelsen *veta varken in eller ut*. Den konventionaliserade formen är, enligt SSB (2003 s. 1335), *veta varken ut eller in*. Som synes har alltså de två konstituenterna *ut* och *in* en annan ordning i elevtexten än i ordboken.

³ Som synes är dessa exempel, liksom de som följer hädanefter, försedda med koder som C34, C06 etc. Bokstäverna står för vilka skolor eleverna går på och numren innehåller information om informantens kön. Manliga informanter betecknas med siffror upp till 25, kvinnliga med siffror över 25. Exempel (3) har alltså skrivits av en kvinnlig elev och exempel (4) av en manlig. Vi anger också vid varje exempel vilken grupp eleven som producerat frasen tillhör, t.ex. TVÅ i exempel (3) och SVE+ i exempel (4).

Ordförbindelsen i (6) är ett exempel på typ b, dvs. en semantisk-pragmatisk modifiering:

(6) Nu vet jag att ni läsare har *blivit varma i kläderna* av igenkännande, och av att ni i mig har funnit en likasinnad. Men som inte nog med det ämnar jag att på resterande rader göra er saliga. (L02, SVE)

Den konventionaliserade betydelsen av det figurativa uttrycket *bli varm i kläderna* är 'hinna vänja sig vid något' eller 'hinna få nödvändig erfarenhet av något' (jfr SSB 2003 s. 607). I exempel (6) används ordförbindelsen i kontexten *blivit varma i kläderna av igenkännande*. Det som eleven vill uttrycka är, av kontexten att döma, att läsarna bör ha blivit varma inombords av att känna igen något (och av att ha hittat någon som har gjort liknande erfarenheter som de själva). Ordförbindelsen används alltså här i en icke typisk kontext, vilket i sin tur orsakar en semantisk modifiering. Man skulle också kunna hävda att modifieringen i exempel (6) orsakas av en kontamination av betydelseerna hos uttrycken *bli varm i kläderna*, *bli varm om hjärtat* och *bli varm inombords (av något)* (jfr SSB 2003 s. 489, 543).

Exempel (7) nedan innehåller en kombination av de två typerna av modifiering, nämligen både en lexiko-grammatisk och en semantisk-pragmatisk avvikelser:

(7) Jag fruktar mina nära och käras död. Sorgen är en oändlig påminnelse om att något tragiskt hänt. Tiden då allting står stilla. Tiden då ens tankar vandrar iväg. *Tiden då inget läker ens sår.* (K35, START<6)

Det är ett rimligt antagande att *tiden då inget läker ens sår* är en modifiering av det konventionaliserade uttrycket *tiden läker alla sår* (jfr SSB 2003 s. 1197). Man kan konstatera att formen som används av informanten delvis innehåller samma konstituenten som den konventionaliserade ordförbindelsen ifråga. Rent semantiskt är den modifierade formen baserad på samma metafor som den konventionaliserade ordförbindelsen, nämligen känslomässiga sår som läker (respektive inte läker). Innebörden är alltså att man lär sig att hantera sina känslor och negativa erfarenheter på något sätt. Den semantiska avvikelser från konventionaliserat bruk uppstår i första hand genom betoningen av en annan betydelsen hos *tid*. Den konventionaliserade ordförbindelsen syftar på den abstrakta, mätbara storheten *tid* i sig. I exempel (7) avser informanten däremot en viss tidsperiod, tiden då något händer eller, i detta fall, *inte* händer (jfr de två huvudbetydelseerna hos uppslagsordet *tid* i Svensk ordbok utgiven av Svenska Akademien 2009).

4.3 Nybildade figurativa ordförbindelser

Den tredje typen av ordförbindelser utgörs av figurativt språk som inte är konventionaliserat. Exempel (8) och (9) nedan illustrerar vad vi i denna studie klassificerar som *nybildade figurativa ordförbindelser*:

(8) Så varför upplever vi tid så olika? [...] Jag letar efter en gemensam nämnare för den upplevda tiden som vid det aktuella tillfället tycktes gå snabbt. [...] Nämnaren blir de många aktiviteterna och aktiviteternas mångfald; jag hade mycket för mig och jag hade kul [...]. Den andra tiden då, den som gick sakta? Återigen *aktiverar jag sökarljusen* och finner en gemensam nämnare; bristen på aktiviteter och aktiviteternas brist på intressegrad. (L02, SVE)

(9) Texten 'Eufori' blir man glatt av. Man kommer in i en slags skön och fridfull stämning. Helt ärligt så fattar jag inte ett enda skit av den texten, innebörden eller vad den vill säga. Det enda jag kan säga om den är att man blir *mjuk som en muffins* när man läser den. (B14, SVE)

I exempel (8) resonerar eleven kring erfarenheten att man ibland känner att tiden går väldigt fort, medan man vid andra tillfällen upplever att den går extremt långsamt. Informanten använder frasen *Återigen aktiverar jag sökarljusen* som, enligt kontexten, betyder att han letar efter svaret på frågan som han har ställt i föregående mening, nämligen vad som kännetecknar de stunder då man tycker att tiden kryper fram. I exempel (9) används en nybildad liknelse, *mjuk som en muffins*, som känns igen på den typiska konstruktionen ADJ *som en* NP (jfr avsnitt 2.1 ovan). Det är inte helt självklart vad eleven menar med denna formulering, men det är uppenbart att han försöker beskriva en positiv känsla.

Liksom vid de andra två huvudtyperna av ordförbindelser, uppstår det även här gränsdragningsproblem genom att det råder ett kontinuum vad gäller graden av konventionalisering. I vissa fall är det avgörande vilka element som fokuseras i analysen. För att förtydliga kan följande exempel nämnas:

(10) Mycket av det hon presenterar i sin insändare är just det att hon vill känna att hon har under veckan varit så stressad så när fredagen kommer ska hon bara få pusta ut, vila och *bara vara ett med soffan*. (L03, SVE).

Man skulle kunna kategorisera *vara ett med soffan* som en nybildad figurativ ordförbindelse. I så fall fokuseras bilden av en persons fysiska förening med en soffa, vilket är en rätt uppenbar, dock inte i denna form konventionaliserad, metafor för *att vila*. Vi väljer dock, efter jämförelse med referensmaterialen, att i stället utgå från den fraseologiska enheten *vara ett med något*. Det är en figurativ ordförbindelse med en variabel konstituent. Med andra ord kan man *vara*

ett med en rad olika saker, även om substantivet *soffa* inte är den mest typiska konstituenten i detta sammanhang. *Vara ett med soffan* klassas alltså som en konventionaliserad ordförbindelse.

5. Användningen av figurativa ordförbindelser i materialet

Hittills har vi diskuterat övergripande om de olika typer av figurativa ordförbindelser som finns i materialet. I det följande kommer vi att resonera kring hur dessa typer används – och av vilka informanter. Några av de kvantitativa resultaten sammanfattas i tabell 2. Vid signifikansprövning, med hjälp av ett icke-parametriskt test (χ^2), visar det sig att de flesta skillnader mellan grupperna inte är statistiskt signifikanta (se vidare nedan). Den typ av ordförbindelser vi undersöker är helt enkelt inte tillräckligt frekvent i alla informantgrupperns skrivande. Icke desto mindre visar analysen av de kvantitativa data några intressanta tendenser. I tabellen anges absoluta tal med en decimal. Procentvärdena är däremot avrundade till heltal.

Tabell 2. Antal ordförbindelser och deras fördelning över informantgrupperna.

Grupp	Antal informanter	Antal ordförbindelser	Antal ordförbindelser/informant	Konventionaliserade ordförbindelser	Modifierade ordförbindelser	Nybildade ordförbindelser
1 (SVE)	54	293	5,4	192 (65 %)	64 (22 %)	37 (13 %)
2 (SVE+)	28	109	3,9	62 (57 %)	33 (30 %)	14 (13 %)
3 (TVÅ)	10	40	4,0	18 (45 %)	12 (30 %)	10 (25 %)
4 (START<6)	60	249	4,2	123 (49 %)	94 (38 %)	32 (13 %)
5 (START>6)	23	148	6,4	62 (42 %)	60 (40 %)	26 (17 %)
Totalt:	175	839	4,8	457 (54 %)	263 (31 %)	119 (14 %)

Som framgår av tabell 2 använder eleverna figurativt språk i sitt skrivande, i genomsnitt ca 5 ordförbindelser per informant (jfr t.ex. Josephson m.fl. 1990). Det finns dock stor individuell variation i informantgruppen när det gäller förekomsten av figurativa ordförbindelser. Medan de är relativt vanliga i vissa elevers skrivande finns det också en del texter som i princip helt saknar sådana ordförbindelser. Det är i synnerhet en av de kvinnliga informanterna i grupp 5 (START > 6) som ligger långt över genomsnittet, då hon använder totalt 29 figurativa ordförbindelser i sina texter. Sammanlagt 19 informanter använder

mellan 10 och 18 ordförbindelser. Däremot är det totalt 17 informanter som inte använder figurativa ordförbindelser över huvud taget i sina texter.

Vidare framgår det några intressanta skillnader mellan de olika informantgrupperna av procentvärdena i tabell 2. Det är eleverna i grupp 1 (SVE), dvs. de med enspråkig svensk bakgrund, som använder klart flest konventionaliserade ordförbindelser. Även hos de övriga informanterna med svenska som enda förstaspråk, dvs. grupp 2 (SVE+), är andelen konventionaliserade ordförbindelser högre än i de övriga grupperna. De enda skillnader mellan informantgrupperna som visat sig vara statistiskt signifikanta finns dock mellan grupp 1 (SVE) och 4 (START < 6): De enspråkiga informanterna i grupp 1 använder signifikant fler konventionaliserade figurativa ordförbindelser än eleverna med svenska som andraspråk i grupp 4. De sistnämnda använder i sin tur signifikant fler modifieringar av konventionaliserade ordförbindelser (jfr Ekberg 1997 och 1998, se vidare nedan). Både de signifikanta skillnader och de tendenser som beskrivs ovan bekräftar alltså tidigare forskningsresultat av Ekberg från 1997 och 1998.

Modifierade figurativa ordförbindelser är mest frekventa i grupp 4 (START < 6) och 5 (START > 6). Men avvikelser från konventionaliserat bruk av figurativa ordförbindelser förekommer även hos informanterna med enspråkigt svensk bakgrund. Även detta mönster ligger alltså i linje med resultat från tidigare studier av Ekberg (1997, 1998) och Hyltenstam (1992). Här bör dock påpekas att denna kategori kan innehålla både medvetna och omedvetna modifieringar av konventionaliserade ordförbindelser. Det är för oss omöjligt att säkert avgöra vilket det rör sig om (även om vi ibland har våra aningar).

När det gäller de nybildade figurativa fraserna, är andelen ungefär densamma i samliga grupperna 1 (SVE), 2 (SVE+), 4 (START < 6) och 5 (START > 6).

Ett intressant resultat är att grupp 5, där informanterna har en startålder för inläringen av svenska som är 6 år eller senare, i genomsnitt har det största antalet figurativa ordförbindelser per informant. Med andra ord finns det inget i vårt material som tyder på att flerspråkiga elever använder mindre figurativt språk i sitt skrivande än elever med enspråkig svensk bakgrund.

Vi har som tidigare nämnts inte kunnat påvisa statistiskt signifikanta skillnader mellan de flesta av informantgrupperna. Däremot visar ett Pearson Correlation-test att det finns signifikanta korrelationer på individnivå vad gäller användningen av de olika typerna av figurativa ordförbindelser i materialet. De informanter som använder fler konventionaliserade ordförbindelser tenderar att också använda fler modifierade och nybildade ordförbindelser. Det verkar

alltså som att figurativt språk är ett återkommande element i dessa elevers skrivande.

5.1 Konventionaliserade ordförbindelser

Bland de konventionaliserade figurativa ordförbindelser som används i vårt material finns många välkända idiom. De har delvis olika struktur. Bland de använda uttrycken återfinns t.ex. sådana som utgörs av en hel sats (*när det kommer till kritan*, L01, SVE). Vidare finns det sådana som utgörs av en nominalfras, t.ex. *en dans på rosor* (L38, SVE) eller en prepositionsfras: *i all sin prakt* (L37, SVE). De flesta uttrycken består dock av ett verb följt av något slags komplement, t.ex. *dra sitt strå till stacken* (D36, SVE), *rinna ut i sanden* och *ha ett finger med i spelet* (L39 och C06, SVE+).

Vissa av de konventionaliserade ordförbindelserna återkommer i flera texter i materialet. Detta gäller bl.a. *i första hand* (D28, SVE; B25, P26 och L40 i START < 6) och *kännas som en evighet* (L06 och B32 i SVE). Det är dessutom uppenbart att elevernas bruk av figurativa ordförbindelser i viss mån påverkats av språket i texterna som skrivuppgifterna utgått från. Flera av eleverna använder som sagt frasen *hålla näsan över vattnet* (L27, L44 och L45 i SVE; B06 i START < 6) och den finns redan i en av utgångstexterna i texthäftet från höstterminen 2003. Ytterligare ett exempel är *tid är pengar*, vilket också används i en av texterna i samma utgångshäfte. Det senare uttrycket används av bl.a. D10, P25 och B28 (i START < 6) samt L10 (i SVE+). Det faktum att elever påverkas av utgångstexter i liknande skrivsituationer har redan iakttagits av t.ex. Hultman & Westman (1977 s. 23, 251) och Josephson m.fl. (1990 s. 47). De menar att detta kan vara ett sätt för eleverna att utvidga sitt ord- och frasförråd. Men den här typen av avskrivning kan också leda till stilbrott i elevernas texter. För övrigt kan man notera att det finns flera modifieringar av uttrycket *tid är pengar* i materialet, t.ex. i texter skrivna av B33 och P33 (i SVE+) och L36 (i START > 6), där konstituenterna har kastats om. Modifieringen *pengar är tid* används dock i texthäftet, så även här kan användningen av en viss typ av figurativt uttryck antas ha inspirerats av utgångstexterna.

Exemplen *kännas som en evighet* och *tid är pengar* illustrerar också att flera av de konventionaliserade ordförbindelserna i elevtexterna är relaterade till temat *tid*. I vissa fall uttrycker dessa fraser att tiden går snabbt, som i *tiden rinner iväg* (B29, START > 6; S49, SVE+). Andra uttryck lyfter i stället fram att tiden

kan gå väldigt långsamt eller knappt avancera alls: *tiden sniglar sig fram* (B32, SVE) och *det känns som om tiden står stilla* (B02, TVÅ).

Flertalet av exemplen ovan används av informanter i grupperna 1 och 2, dvs. elever med svenska som förstaspråk. Som framgår av tabell 2 använder dock även eleverna i grupperna 4 (START < 6) och 5 (START > 6), där eleverna har angett andra språk än svenska som sina modersmål, konventionaliserade figurativa ordförbindelser, om än i mindre utsträckning. Nedan följer några exempel från dessa två grupper:

(11) Kan du lova mig att du strävar efter nå alla dina mål som du har här i livet? Att inte kasta bort din tid, för »*tid är pengar*». (D10, START < 6)

(12) Det sista jag vill tillägga är att det är vi ungdomar som får *ta det första steget*. [...] Få dem att förstå att vi är nästa generation och som ska styra samhället i framtiden. Få en ändring att ske! (S08, START < 6)

(13) Som tonåring är allting upp och ner och man vet inte riktigt vad man vill. Det ställs alltför höga krav och snart måste man även lära sig att *stå på egna ben*. (D47, START > 6)

(14) När vuxenlivet knacker på dörren en morgon och man öppnar dörren i klädd sina rosa Nasse pyjamas och Nalle Puh tofflor inser man snart att det är tid för att *lägga sitt ångestfyllda och dramatiska ungdomsliv på hyllan*. (D40, START > 6)

5.2 Modifierade konventionaliserade ordförbindelse

Det som Hyltenstam kallar för icke infödda drag är relevanta i samband med kategorin modifierade konventionaliserade ordförbindelser (se avsnitt 2.4). För det första finns det ett antal exempel på approximationer i materialet. Några återkommande typer av detta fenomen visar sig uppstå genom a) omkastning av konstituent, b) utbyte av konstituent eller c) tillägg eller bortfall av konstituent i ordförbindelserna.

Exempel på approximation av typ a är de redan nämnda formuleringarna *veta varken in eller ut* och *pengar är tid*. Modifieringarna sker här alltså genom att konstituenterna *in* och *ut* respektive *tid* och *pengar* har bytt plats med varandra.

Den andra typen av approximation, b, uppstår, som tidigare nämnts, genom att en eller flera lexikala konstituent som ingår i ordförbindelsen har ersatts av någon annan eller några andra lexikala enheter. Ofta gäller det funktionsord, t.ex. prepositioner, vilka endast i begränsad omfattning bidrar till uttryckens semantik men som har större betydelse för om bruket av ordförbindelsen ska upp-

fattas som idiomatiskt. Ett exempel från materialet är *haka upp sig i något* (C41, START < 6), som avviker från det konventionaliserade uttrycket *haka upp sig på ngt* (jfr SSB 2003 s. 454) genom användningen av prepositionen *i* i stället för *på*. Men även innehållsord i vissa ordförbindelser byts ut. Som exempel kan *de har blivit utelåsta från dagens samhälle*, som finns i en text av C04 (SVE), anföras. I detta fall har *utelåsta* använts i stället för standardformen *utestängda*. Genom de lexikala utbytena får ordförbindelserna ofta en mer konkret innebörd. Eventuellt kan man också säga att utbytena ofta leder till ökad grad av bildlighet hos annars inte helt levande metaforer (jfr Svanlund 2001 s. 15).

Som exempel på en approximation av typen c, dvs. de som uppstått genom tillägg av ett ord, väljer vi frasen *rulla på tummarna* (C39, START < 6). Denna approximation har alltså uppstått genom ett tillägg av prepositionen *på* till den konventionaliserade formen *rulla tummarna* (jfr SSB 2003 s. 1243). Ett exempel, som uppstått genom bortfall av en konstituent, är *det tror jag inte på fem öre* (C06, TVÅ), där den konventionaliserade formen är *inte för fem öre* (SSB 2003 s. 1421). Denna förekommer bl.a. i kontexten *det tror jag inte på för fem öre*. Det går emellertid inte att utesluta att exemplet helt enkelt är resultatet av en felskrivning.

I materialet finns också det som Hyltenstam (1992) benämner kontamination. Några av dessa fall återges i exempel (15)–(17) nedan:

(15) I G.Forsells teckning, cell i fängelset på långholmen, står en dystert ensam man i en trång lite cell. De höga väggarna i rummet *får håret på ryggraden att resa sig*. (S30, TVÅ)

(16) Hamnar man i fel umgänge *hänger framtiden på en lös tråd*. (D26, START > 6)

(17) *Tiden rinner ut för oss alla*, så talade William Shakespeare, visst kan man tycka att det stämmer (L31, START < 6)

Den målande ordförbindelsen i (15) kan tänkas vara en kontamination av de konventionaliserade uttrycken *håren reser sig* och *få rysningar/kalla kårar längs ryggraden* (SSB 2003 s. 506, 672). I (16) återfinns varianten *något hänger på en lös tråd*, vilken kan ses som en kontamination av det konventionaliserade uttryckssättet *något hänger på en skör tråd* med uttryck som *lösa trådar* och *hänga löst* (SSB 2003 s. 1235 och Språkbankens konkordanser). Eftersom det endast är bruket av konstituenten *lös* som gör att elevens uttryckssätt avviker från den konventionaliserade formen *något hänger på en skör tråd*, skulle man dock kunna hävda att det i stället är fråga om en approximation. Men med

tanke på att adjektivet *lös* är en konstituent i uttryckssätten *lösa trådar* och *hänga löst*, känns det dock mer naturligt att betrakta exemplet som kontamination. Slutligen kan ordförbindelsen som används i (17) ses som en kontamination av de konventionaliserade uttrycken *tiden rinner iväg* och *tiden är ute för någon* (jfr SSB 2003 s. 1198 och Språkbankens konkordanser). Detta fall är dock något mer komplext och återspeglar ännu en gång de gränsdragningsproblem som kan uppstå vid kategorisering av ordförbindelserna som fokuseras här. Formen *tiden rinner ut* (för ngn) skulle också kunna betraktas som ett översättningslån av engelskans *time is running out*. Man skulle dock kunna göra bilden ännu mer komplex genom att också beakta möjligheten av en kontamination mellan den svenska formen *tiden rinner iväg* och det engelska uttrycket, som ligger nära varandra i både form och betydelse. Det är heller inte otänkbart att formuleringen *tiden rinner ut för oss alla* är baserad på en mer konkret bild av hur sanden rinner ur den övre delen av ett timglas.

I tabell 3 sammanfattas förekomsten av approximationer och kontaminationer i de olika informantgruppernas texter.

Tabell 3. Frekvens av approximationer och kontaminationer per informantgrupp.

Grupp	Antal elever	Antal modifierade ordförbindelser totalt	Antal approximationer	Antal kontaminationer
1. SVE	54	64	4 (6 %)	13 (20%)
2. SVE+	28	33	2 (6 %)	4 (12%)
3. TVÅ	10	12	–	3 (25%)
4. START < 6	60	94	21 (22 %)	16 (17%)
5. START > 6	23	60	1 (2 %)	11 (18%)
Total	175	263	28 (11 %)	47 (18%)

Tabell 3 visar att approximationer och kontaminationer är relativt ovanliga i materialet, vilket gör det svårt att dra några slutsatser av denna del av analysen. Ett viktigt resultat är dock att det som Hyltenstam kallar för icke-infödda drag också förekommer i grupp 1 (SVE), dvs. i texter som är skrivna av elever med enspråkigt svensk bakgrund. Detta stämmer överens med Hyltenstams analys, som beskrivits i avsnitt 2.4. Man kan även jämföra med de exempel på kontaminationer som redovisas i Hultman & Westman 1977 s. 245 f. Kontaminationer är sammanlagt något mer frekventa än approximationer i materialet.

Hittills har vi i första hand diskuterat lexiko-grammatiska avvikelser från konventionaliserat bruk. Informanterna avviker även semantiskt och pragma-

tiskt från ordförbindelsernas konventionaliserade former. Dessa modifieringar är dock mindre frekventa än de modifieringar som direkt kan kopplas till ordförbindelsernas form. I avsnitt 4.2 nämndes den icke typiska användningen av uttrycket *bli varm i kläderna*. En annan konventionaliserad ordförbindelse som används i en avvikande kontext i vårt material är *fastna i minnet*.

(18) Vad varje människa har gjort och har gemensamt sen gamla tider och generationer är att man tittar tillbaka i sitt liv. *Funderingar fastnar i minnet* och man vill, vare sig man har lidit och lyckats i livet övertyga andra men speciellt sig själv att man har gjort sitt bästa. (C10, START < 6)

Vid typisk användning är det en bild eller händelse som *fastnar i någons minne*, dvs. de gör ett starkt intryck på personen ifråga och han eller hon kan lätt återkalla dem igen för en tid framöver. Formuleringen *funderingar fastnar i minnet* torde här betyda att man grubblar, eller rent av ältar något. I det här fallet försöker man tänka ut vad man har gjort med sitt liv och om man verkligen tagit vara på sin tid på jorden.

Vidare finns det flera exempel på både lexiko-grammatiska och semantisk-pragmatiska modifieringar i materialet. Flera av informanterna använder sig av metaforer som kan beskrivas som överföringar mellan källdomänen *huvud* till måldomänen *tankeverksamhet*. Den underliggande idén i dessa fall är att olika typer av tankeverksamhet utspelar sig i avgränsade områden i huvudet. Det som man inte tänker aktivt på, men som ändå finns i medvetandet, antas vara lokaliserat i bakre delen av huvudet. Detta uttrycks språkligt i figurativa uttryck som *ha något* eller *något finns* eller *något ligger i bakhuvudet* (jfr SSB 2003 s. 80). I konventionaliserat bruk används dessa ordförbindelser vanligen om framtidsplaner eller något man oroar sig över. I vårt material hittar vi däremot exempel som de följande:

(19) Man längtar efter att sitta i en mysig stuga framför kakelugnen och dricka en kopp te och bara glömma *alla vardagliga problem som sitter i bakhuvudet*. (S18, SVE)

(20) Men är inte hennes text överdriven till den grad att du skulle läst den med ironi i tankarna? Kanske är det just det som också är meningen, att vi ska ifrågasätta varför vi *läser den med ironi i bakhuvudet*. (D33, SVE)

(21) [...] jag är en av de många som föredrar kompisarna framför plugget och även den som alltid börjar sista kvällen med alla inlämningar [...]. Och för att det inte ska fortsätta med att jag sitter hela natten lång och *skriver med stressen och pressen i bakhuvudet* så har jag kommit på ett avgörande sätt. (L28, START < 6)

I exempel (19) används uttrycket i en relativt typisk kontext men däremot är valet av verbet *sitta* inte konventionaliserat. I exemplen (20) och (21) använder informanterna formerna *läsa* och *skriva med något i bakhuvudet*, som lexikogrammatiskt avviker från de konventionaliserade ordförbindelser som diskuteras ovan. Dessutom är bruket pragmatiskt avvikande, dvs. att man har *ironi*, *stress* eller *press* i bakhuvudet är inte typiska kontexter som ordförbindelsen brukar användas i.

5.3 Nybildade figurativa ordförbindelser

En intressant fråga i samband med nybildade figurativa ordförbindelser i materialet är vilka källdomäner eleverna utgår ifrån när de skapar nya metaforer och liknelser. Wellander (1939 s. 634, 628) konstaterar att valet av bild eller liknelse i en text i hög grad bestäms av författarens livserfarenhet och intressen. Ett mindre överraskande resultat, med tanke på texthäftenas tema, är att *tid* är en frekvent källdomän i materialet (se avsnitt 3.2 ovan). Detta exemplifieras i (22) och (23) nedan.

(22) Vi satt på buss, det hände nästan ingenting, vägen var monoton, horisonten alldeles för långt borta – vi färdades *på ett helt hav av tid*, men vi satt fast i en alldeles för trång båt. (L02, SVE)

(23) För föraren som sitter i favoritracerbilen som just under det varvet tänker slå ett nytt varvrekord *flyger sekunderna iväg som mittsträcken på banan* i trehundra-kilometers hastighet. (L08, TVÅ)

Men i övrigt kan man se att de källdomäner som eleverna utgått från när de skapat sina metaforer är av mycket olika karaktär:

(24) Jag tänkte nu i ett antal hundra ord lämna mitt bidrag som antagligen (förhoppningsvis) reder ut en del eller bara tar oss djupare ner i *träsket av oklarheter*. (C05, SVE)

(25) Tiden efter AIK;s 1-0 mål har man blicken fäst på matchuret större delen av tiden samtidigt som man tuggar nagelband *likt Musse Pigg och Kalle Anka käkar majs* på den berömda husvagnsemester. (S12, SVE)

(26) Att man som tonåring känner en plågsam känsla av att tiden och livet är en ända stort problem är det ingen tvivel om »Hur länge ska det ta innan bröstet har vuxit ut». »Hur länge ska det ta innan jag får min mens, de andra har ju fått sin. Oooh, jag vill bara döööö». Jag vet hur otäck tiden kan vara som en tonåring och man känner att problemen bara *staplar upp sig som en pyramid i Egypten*. (D40, START > 6)

I exempel (24) används källdomänen *träsk* för att skapa den figurativa frasen *ett träsk av oklarheter*. Metaforen bygger här på vissa aspekter av konceptet *träsk* – en svårdefinierad massa som man lätt kan fastna i. I (25) baseras metaforen på den tecknade filmens värld, där vissa handlingar, t.ex. att äta, görs på ett speciellt sätt, dvs. ofta med stor intensitet och i hög fart. Tecknade Disney-figurer är en källdomän som uppenbarligen har sina rötter i västvärlden, medan liknelser som i exempel (26) bygger på ett koncept från en annan kultur, nämligen det antika Egyptens pyramider.

Något som man bör komma ihåg i samband med de »nybildade» figurativa fraserna i denna studie, är det potentiella inflytandet från andra språk än svenska. Detta är speciellt viktigt vid beskrivningen av de figurativa ordförbindelser som används av eleverna i grupperna 3 (TVÅ), 4 (START < 6) och 5 (START > 6), även om också elever med enspråkigt svensk bakgrund naturligtvis kan påverkas av andra språk i sitt skrivande, inte minst engelskan. De fraser som kategoriseras som nybildade i vårt material kan alltså i vissa fall vara översättningar, varianter eller modifieringar av konventionaliserade figurativa ordförbindelser på ett annat språk. »Nybildad» ska därför här tolkas som 'inte konventionaliserad i svenskan'. Oftast är det omöjligt för oss att avgöra om de figurativa fraser som eleverna använder är lån från andra språk som de möter i sin omgivning. I enstaka fall, som i exempel (27), finns det dock tydliga indikationer på att så är fallet:

(27) Som ett arabiskt ordspråk säger »Tiden är som ett svärd, ifall du inte kan använda den skär den dig.» (C41, START < 6)

Informanten C41 anger arabiska som sitt förstaspråk och man kan därför utgå ifrån att det rör sig om en direktöversättning av uttrycket ifråga. I andra fall, som i exempel (28) nedan, är vi däremot tämligen säkra på att ordförbindelsen är ett översättningslån från engelskan:

(28) Dockusåpamedlemmarna dricker sig redlöst fulla, *kallar varandra namn* utan gränser eller helt enkelt bara baktalar varandra. (C36, SVE+)

Slutligen kan vi konstatera att informanterna ofta skapar nya figurativa ordförbindelser i form av liknelser. Här är alltså liknelsen den mest frekventa typen av figurativa ordförbindelser, medan den är relativt ovanlig bland de konventionaliserade och modifierade figurativa ordförbindelserna i materialet.

6. Sammanfattning och diskussion

I denna artikel har vi presenterat en undersökning av hur figurativa ordförbindelser används i texter skrivna av 175 elever i gymnasiet i samband med de nationella proven i svenska och svenska som andraspråk. Materialet är insamlat i skolor som ligger i flerspråkiga storstadsområden. De excerperade figurativa ordförbindelserna har delats in i tre huvudtyper utifrån deras grad av konventionalisering. De tre typerna är a) konventionaliserade figurativa ordförbindelser, b) modifierade konventionaliserade ordförbindelser och c) nybildade figurativa ordförbindelser. Problemen som diskuterats i tidigare avsnitt, t.ex. svårigheterna vid excerpering och kategorisering av olika typer av ordförbindelser samt skillnaderna i gruppernas storlek, bidrar till denna studies begränsningar. Trots detta visar sig intressanta tendenser i materialet.

Ett resultat av undersökningen är att de konventionaliserade figurativa ordförbindelserna används mest. Därefter kommer modifierade ordförbindelser och nybildade sådana. Det finns visserligen relativt stora individuella frekvensskillnader i informanternas användning av figurativa ordförbindelser. Av exemplen som diskuterats i detta arbete framgår dock att det språkbruk som undersökts inte kan betraktas som en »grå, trögflytande och halvljummen välling som knappast retar aptiten» (jfr Hultman & Westman 1977 s. 253). Signifikanta korrelationer i materialet tyder dessutom på att figurativt språk är ett återkommande element i vissa elevers skrivande.

Vad gäller användningen av olika typer av figurativa ordförbindelser visar denna undersökning att konventionaliserade och modifierade liknelser är relativt lågfrekventa i elevernas skrivande. Intressant nog är liknelserna däremot den mest använda formen när det gäller nybildade figurativa ordförbindelser. Ytterligare en intressant fråga angående nybildningar hänger ihop med de aspekter inom metafor teorin som refereras i avsnitt 2.1 (se Lakoff & Johnson 1980, Ekberg 1993 och Svanlund 2001). Med utgångspunkt i Wellanders antagande (1939 s. 634, 628) borde det finnas en viss koppling mellan elevers kulturella och språkliga bakgrund och de erfarenhetsdomäner som eleverna använder när de skapar figurativa ordförbindelser. För att bekräfta denna hypotes krävs ett mer omfattande material. Tittar man på exemplen som återges i avsnitt 5.3 kan man dock ana vissa influenser från informanternas modersmål eller kulturella erfarenheter (se t.ex. exempel 25, 27 och 28).

Vidare bekräftar vår undersökning tidigare forskningsresultat av Ekberg (1997, 1998). Till exempel framgår det att både de enspråkiga och de flerspråkiga eleverna har kunskap om konventionaliserade figurativa ordförbindelser

på svenska. Liksom i Ekbergs undersökningar använder dock de enspråkiga informanterna fler konventionaliserade ordförbindelser än de flerspråkiga eleverna; analysen visar på en signifikant skillnad mellan grupperna 1 (SVE) och 4 (START < 6) och de övriga resultaten visar på en liknande tendens.

I anknytning till Hyltenstam 1992 har vi även undersökt förekomsten av approximationer och kontaminationer i materialet. Vår korpus har visat sig vara för begränsad för att tillåta några generella slutsatser vad gäller korrelationer mellan t.ex. startålder för språkinläringen och användningen av just dessa modifieringar. I likhet med Hyltenstam kan vi dock konstatera att approximationer och kontaminationer förekommer i både de enspråkiga och de flerspråkiga informanternas skrivande. Vi har dessutom funnit att det går att urskilja tre återkommande typer av approximationer i vårt material, nämligen approximation till en konventionaliserad ordförbindelse i form av a) omkastning av konstituenten, b) utbyte av konstituenten och c) tillägg eller bortfall av konstituenten. Förekomsten av kontaminationer verkar också vara i linje med Hyltenstams resultat, i den meningen att detta drag är ungefär lika frekvent i informanternas texter, oavsett språklig bakgrund.

I denna studie är inte skillnaderna i språkbruk mellan samtliga informantgrupper statistiskt signifikanta. Tendenserna som visats här kan dock vara en utgångspunkt för en större undersökning av huruvida dessa skillnader kan kopplas till faktorer som startålder för inläring eller påverkan från andra språk än svenska. För detta krävs ett större och mer balanserat material. Man skulle också kunna tänka sig att faktorer som kön, sociokulturell bakgrund och skolmiljö har betydelse för elevernas bruk av figurativt språk. Dessa faktorer har vi inte tagit hänsyn till i denna undersökning. Det vore därför intressant att undersöka om en annan indelning av informanterna skulle leda till fler signifikanta skillnader i materialet.

Det är också tänkbart att titta närmare på vissa av informantgrupperna i en mer omfattande studie. Särskilt elever som växt upp med två eller flera modersmål parallellt (med svenska som ett av dem, dvs. grupp 3) kan denna studie knappast göra rättvisa. En intressant fråga är om det finns skillnader i användning och behärskning av figurativa konventionaliserade flerordsenheter, mellan enspråkigt och flerspråkigt uppväxta informanter. Cummins & Gulutsan (1974) har t.ex. visat att balanserat tvåspråkiga informanter uppvisar större språklig originalitet än enspråkiga (se även Baker & Hornberger 2001, Otterup 2005). Med detta som utgångspunkt vore det intressant att undersöka i vilken mån en balanserad tvåspråkighet gynnar kreativiteten i samband med användningen av figurativa ordförbindelser.

En annan möjlig uppföljning vore att undersöka hur elevtexterna har bedömts och huruvida bruk av figurativt språk uppmuntras eller premieras av lärarna. Man skulle kunna tänka sig att eleverna i de olika grupperna inte har samma uppfattning om vilken norm som gäller och vilka stilideal de ska eftersträva. Vissa vill skriva mer expressivt och använda mycket bildspråk medan andra i stället strävar efter en mer saklig stil.

Litteratur

- Anward, Jan & Linell, Per, 1976: Om lexikaliserade fraser i svenskan. I: *Nysvenska studier* 55/56. S. 77–119.
- Baker, Colin & Hornberger, Nancy H., 2001: *An Introductory Reader to the Writings of Jim Cummins*. Clevedon.
- Bijvoet, Ellen & Fraurud, Kari, 2006: »Svenska med något utländskt». I: *Språkvård* 3. S. 4–10.
- Boyd, Sally & Fraurud, Kari, under utg.: *Challenging the Homogeneity Assumption in Language Variation Analysis. Findings from a Study of Multilingual Urban Spaces*. I: *Language and Space – an International Handbook of Language Variation*, ed. by Peter Auer & Jürgen Erich Schmidt. Berlin/New York.
- Burger, Harald, 1998: *Phraseologie. Eine Einführung am Beispiel des Deutschen*. Berlin.
- Clausén, Ulla, 1996: *Idiom i bruk*. I: *Språket lever! Festskrift till Margareta Westman den 27 mars 1996*. (Skrifter utgivna av svenska språknämnden 80.) Stockholm. S. 36–42.
- Cummins, Jim & Gulutsan, Metro, 1974: *Bilingual education and cognition*. I: *The Alberta Journal of Educational Research* 20. S. 259–269.
- Ekberg, Lena, 1993: *Verbet ta i metaforisk och grammatikaliserad användning*. I: *Språk och stil* NF 3. S. 105–139.
- Ekberg, Lena, 1997: *Diskursiva och syntaktiska mönster i återberättelser hos invandrabarn i Rosengård*. I: *Svenskans beskrivning 22*, red. av Gisela Håkansson m.fl. Lund. S. 97–110.
- Ekberg, Lena, 1998: *Regeltillämpning kontra lexikonkunskap i svenskan hos invandrabarn i Rosengård*. I: *Nordiska sprog som andetsprog*, utg. av Janus Møller m.fl. (Københavnstudier i tosprogethed 30.) Köpenhamn. S. 247–263.
- Ekberg, Lena, 2004: *Grammatik och lexikon i svenska som andraspråk på nästan infödd nivå*. I: *Svenska som andraspråk – i forskning, undervisning och samhälle*, red. av Kenneth Hylltenstam & Inger Lindberg. Lund. S. 259–276.
- Fauconnier, Gilles & Turner, Mark, 2002: *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities*. New York.
- Fleischer, Wolfgang, 1997: *Phraseologie der deutschen Gegenwartssprache (2. durchgesehene und ergänzte Auflage)*. Tübingen.
- Garne, Birgitta, 2002: *Skolan, språket och de nationella proven*. I: *Språk på väg. Om elevers språk och skolans möjligheter*, red. av Birgitta Garne. (Ord och stil. Språkvårdssamfundets skrifter 33.) Uppsala. S. 9–21.

- Gläser, Rosemarie, 1988: The grading of idiomaticity as a presupposition for a taxonomy of idioms. I: Understanding the lexicon. Meaning, sense and world knowledge in lexical semantics, ed. by Werner Hülsen & Rainer Schulze. Tübingen. S. 264–279.
- Granger, Sylviane, 1998: The computerized learner corpus: a versatile new source of data for SLA research. I: Learner English on computer, ed. by Sylviane Granger. London/New York. S. 3–18.
- Große, Julia, 2008: Fraseologiska enheter på flerspråkig grund. En jämförelse mellan verklighet och fiktion. I: Nordand. Nordisk tidskrift för andrespråksforskning 2. S. 103–132.
- Hyltenstam, Kenneth, 1992: Non-native features of Near-native speakers: On the Ultimate Attainment of Childhood L2 Learners. I: Cognitive processing in Bilinguals, ed. by Richard Harris. Amsterdam. S. 351–368.
- Hultman, Tor G. & Westman, Margareta, 1977: Gymnasistsvenska. Lund.
- Josephson, Olle, Melin, Lars & Oliv, Tomas, 1990: Elevtext. Lund.
- Knowles, Murray & Moon, Rosamund, 2006: Introducing Metaphor. London/New York.
- Koller, Werner, 1977: Redensarten. Linguistische Aspekte, Vorkommensanalysen, Sprachspiel. Tübingen.
- Krohn, Karin, 1994: Hand und Fuß. Eine kontrastive Analyse von Phraseologismen im Deutschen und Schwedischen. (Göteborgs germanistiska Forschungen 36.) Göteborg.
- Körner, Svante & Wahlgren, Lars, 2006: Statistisk dataanalys. Lund.
- Lakoff, George & Johnson, Mark, 1980: Metaphors we live by. Chicago/London.
- Langlotz, Andreas, 2006: Idiomatic Creativity. A cognitive-linguistic model of idiom-representation and idiom-variation in English. Amsterdam/Philadelphia.
- Ledin, Per, 1998: Att sätta punkt: hur elever på låg- och mellanstadiet använder mening i sina uppsatser. I: Språk och stil NF 8. S. 5–47.
- Lindfors Viklund, Maja, 1991: Svenska idiom: ur ett internationellt perspektiv. (Nordrapport. Rapporter från institutionen för nordiska språk/svenska språket vid Göteborgs universitet 7.) Göteborg.
- Lorentzen, Henrik, 2008: Korpusbaseret lemma-eksklusion eller »hur gick ordene hen da de gik ud?». I: Nog ordat? Festskrift till Sven-Göran Malmgren, utg. av Kristinn Jóhannesson m.fl. (Meijerbergs Arkiv för svensk ordforskning 34.) Göteborg. S. 259–267.
- Magnusson, Ulrika, 2009: Systemisk-funktionell lingvistik och andraspråksforskning – skriftspråksutveckling i olika forskningstraditioner. I: Nordand. Nordisk tidskrift för andrespråksforskning 1. S. 67–94.
- Moon, Rosamund, 1998: Fixed Expressions and Idioms in English: a corpus-based approach. Oxford.
- Moon, Rosamund, 2008: Conventionalized *as*-similes in English. A problem case. I: International Journal of Corpus Linguistics 13:1. S. 3–37.
- Müller, Cornelia, 2008: Metaphors Dead and Alive, Sleeping and Waking. A Dynamic View. Chicago/London.
- Norberg Brorsson, Birgitta, 2007: Man liksom bara skriver. Skrivande och skrivkontexter i grundskolans år 7 och 8. (Studier från Örebro i svenska språket 2.) Örebro.

- Nunberg, Geoffrey, Sag, Ivan, A. & Wasow, Thomas, 1994: Idioms. I: *Language. Journal of the Linguistic Society of America* 3. S. 492–538.
- Nyström, Catharina, 2000: *Gymnasisters skrivande. En studie av genre, textstruktur och sammanhang.* (Skrifter utgivna av Institutionen för nordiska språk vid Uppsala universitet 51.) Uppsala.
- Nyström, Catharina & Ohlsson, Maria, 2002: *Analys – skrivuppgift utan ämnesgränser. I: Språk på väg. Om elevers språk och skolans möjligheter, red. av Birgitta Garne. (Ord och stil. Språkvårdssamfundets skrifter 33.) Uppsala. S. 22–51.*
- Olevar, Helena, 2002: *Skrivuppgift i repris – presentation av ett forskningsprojekt. I: Språk på väg. Om elevers språk och skolans möjligheter, red. av Birgitta Garne. (Ord och stil. Språkvårdssamfundets skrifter 33.) Uppsala. S. 86–100.*
- Otterup, Tore, 2005: »Jag känner mig begåvad bara». Om flerspråkighet och identitetskonstruktion bland ungdomar i ett multietniskt förortsområde. Göteborg.
- Palmér, Anne & Östlund-Stjärnegårdh, Eva, 2005: *Bedömning av elevtext. En modell för analys.* Stockholm.
- Pawley, Andrew & Syder, Frances H., 1983: Two puzzles for linguistic theory: native-like selection and nativelike fluency. I: *Language and Communication*, ed. by Jack Richards & Richard Schmidt. London. S. 191–221.
- Schottman, Hans & Petersson, Rikke, 1989: *Wörterbuch der schwedischen Phraseologie in Sachgruppen.* Münster.
- Semino, Elena, 2008: *Metaphor in Discourse.* Cambridge.
- Sköldbberg, Emma, 2004: *Korten på bordet. Innehålls- och uttrycksmässig variation hos svenska idiom.* Göteborg.
- Sköldbberg, Emma, 2005: *När faktorer lyser med sin frånvaro: om några typer av betydelsevariation hos svenska idiom. I: Hermes. Journal of Linguistics* 35. S. 119–134.
- SSB = Svenskt språkbruk 2003. *Svenskt språkbruk. Ordbok över konstruktioner och fraser.* Utarbetad av Svenska språknämnden. Stockholm.
- Svanlund, Jan, 2001: *Metaforen som konvention. Graden av bildlighet i svenskans vikt- och tyngdmetaforer.* Stockholm.
- Svanlund, Jan, 2002: *Lexikalisering. I: Språk och stil NF* 12. S. 7–45.
- Svensén, Bo, 2004: *Handbok i lexikografi: ordböcker och ordboksarbete i teori och praktik. 2 uppl.* Stockholm.
- Svensk ordbok utgiven av Svenska Akademien 2009. Stockholm.
- Wennerholm, Peter, 2003: *Det är som att köra bil med endast en ratt. En undersökning av bildspråket i gymnasisters skrivande. (MISS. Meddelanden från Institutionen för svenska språket. 45)* Göteborg.
- Wellander, Erik, 1939: *Riktig svenska. En handledning i svenska språkets vård.* Stockholm.
- Wiksten Folkeryd, Jenny, 2006: *Writing with an Attitude. Appraisal and student texts in the school subject of Swedish. (Acta Universitatis Upsaliensis. Studia Linguistica Upsaliensia 5.) Uppsala.*
- Wray, Alison, 2002: *Formulaic Language and the Lexicon.* Cambridge.
- Östlund-Stjärnegårdh, Eva, 2002: *Godkänd i svenska? Bedömning och analys av gymnasieelevers svenska. (Skrifter utgivna av Institutionen för nordiska språk vid Uppsala universitet 57.) Uppsala.*

Webb-källor

<<http://www.dn.se/kultur-noje/vem-ager-svenskan-1.643868>>, publicerad 19.4.2006.

<<http://www.google.com>>

<<http://www.skolverket.se>>

<<http://spraakbanken.gu.se/konk/>>

III

KÄPPEN I HJULEN

Innehåll

1. Inledning.....	1
1.1. Syfte och frågeställningar	2
2. Idiomaticitet, språklig konvention och konventionaliserade uttryck..	3
2.1. Idiomaticitet	3
2.2. Konventionalisering, kompositionalitet och genomskinlighet..	5
2.3. Idiom.....	8
2.4. Kollokationer	9
2.4.1. Kontextualistisk kollokationsforskning.....	10
2.4.2. Systemorienterad kollokationsforskning.....	12
3. Lexikal och fraseologisk kunskap	
– vad kan man när man ”kan” ett uttryck?.....	15
3.1. Vad kan man när man kan ett ord?	16
3.2. Att mäta lexikonkunskap	21
4. Konventionaliserade uttryck i språkutveckling	25
4.1. Förstaspråsutveckling.....	25
4.2. Andraspråkutveckling.....	27
4.3. Mental representation av konventionaliserade uttryck	29
4.4. Konventionaliserade uttryck i andra- och främmandespråksinläring	32
4.4.1. Studier inom andraspråksforskning	33
4.4.2. Korpuslingvistiska studier inom främmandespråksinläring.....	38
4.5. Konventionaliserade uttryck i utbildningsrelaterat språk.....	39

5. Material och metod	41
5.1. Etnisk och social bakgrund som påverkande faktorer i samband med elevers skolresultat	41
5.2. Skolor och elevsammansättning i deltagarklasserna.....	43
5.3. Informantgrupperna	46
5.4. Enkäten	48
5.5. Analysmetod.....	52
6. Resultat	55
6.1. Resultaten för de deltagande gymnasieklasserna.....	55
6.2. Resultaten i förhållande till språklig bakgrund	58
6.2.1. Resultatspridning.....	60
6.3. De olika deltesten	64
6.3.1. Skillnader mellan delresultaten inom varje grupp	67
6.4. De enskilda uttrycken	68
6.4.1. Förståelse	68
6.4.2. Lucktest i kontext	71
6.4.3. Lucktest utanför kontext	75
6.4.4. Kollokationer.....	79
6.5. Validitet	83
6.6. Kvalitativ analys	84
6.6.1. De svåraste uttrycken	86
6.6.1.1. hålla någon i schack	87
6.6.1.2. Sätta käppar i hjulen	89
6.6.1.3. få upp ögonen för något	91
6.6.1.4. på rak arm	93
6.6.1.5. livet leker för någon	95
6.6.1.6. med något i högsta hugg	97
6.6.1.7. klättra på väggarna	98
6.6.1.8. få någon/något på halsen	100
6.6.1.9. lyckligt ovetande	101
6.6.1.10. inte för fem öre	103
6.6.2. De vanligaste avvikelserna	105
6.6.3. Enkäterna med de högsta resultaten	107

7. Sammanfattning och slutsatser.....	111
7.1. Kvantitativ analys	111
7.1.1. Skillnader mellan deltesten.....	112
7.2. De enskilda uttrycken	113
7.3. Den kvalitativa analysen.....	115
7.3.1. Lexikala komponenter, kontext och mental representation	116
8. Avslutande diskussion.....	121
Referenser.....	125
Övriga källor	134
Appendix.....	135

Tabeller och figurer

TABELL 1:	Olika aspekter av lexikonkunskap	18
TABELL 2:	Bakgrundsdata för deltagarklasserna.....	44
TABELL 3:	Studieresultat för relevanta utbildningar i årskurs 3 för hela landet i procent	45
TABELL 4:	Antal elever från varje årskurs i grupperna	47
TABELL 5:	Testresultat per gymnasieklass	57
TABELL 6:	Resultat för hela enkäten) per startåldersgrupp	59
TABELL 6.1:	Detaljerad redovisning av resultat för informerterna med svenska som L2	60
TABELL 7:	Resultat för enkätens olika deltester (% av maximalt resultat)	65
TABELL 7.1:	Delresultaten i fallande ordning (% av maximalt resultat) per grupp.	67
TABELL 8:	Resultat på förståelsetest i del A per uttryck (i fallande ordning)	72
TABELL 8.1:	Resultat på lucktestet i del B per uttryck (i fallande ordning)	74
TABELL 8.2:	Resultat på lucktestet i del C per uttryck (i fallande ordning)	77

TABELL 8.3:	Resultat på matchningstestet i del D per uttryck (i fallande ordning)	80
TABELL 9:	Antal olika typer av avvikelser i samband med uttrycket <i>hålla någon i schack</i> per informantgrupp	87
TABELL 9.1:	Antal olika typer av avvikelser i samband med uttrycket <i>sätta käppar i hjulen</i> per informantgrupp	89
TABELL 9.2:	Antal olika typer av avvikelser i samband med uttrycket <i>få upp ögonen för något</i> per informantgrupp	91
TABELL 9.3:	Antal olika typer av avvikelser i samband med uttrycket <i>på rak arm</i> per informantgrupp	93
TABELL 9.4:	Antal olika typer av avvikelser i samband med uttrycket <i>livet leker för någon</i> per informantgrupp	95
TABELL 9.5:	Antal olika typer av avvikelser i samband med uttrycket <i>med något i högsta hugg</i> per informantgrupp	97
TABELL 9.6:	Antal olika typer av avvikelser i samband med uttrycket <i>klättra på väggarna</i> per informantgrupp.....	98
TABELL 9.7:	Antal olika typer av avvikelser i samband med uttrycket <i>få någon/något på halsen</i> per informantgrupp	100
TABELL 9.8:	Antal olika typer av avvikelser i samband med uttrycket <i>lyckligt ovetande</i> per informantgrupp	101
TABELL 9.9:	Antal olika typer av avvikelser i samband med uttrycket <i>inte för fem öre</i> per informantgrupp	103
TABELL 10:	Antal olika typer av avvikelser för samtliga tio uttryck per informantgrupp	105
TABELL 10.I:	Antal olika typer av avvikelser för samtliga tio uttryck per informantgrupp – utökad material	106
TABELL II:	Typer av avvikelser i de 20 enkäterna med högst resultat.....	109

FIGUR 1:	Exempel på relationer mellan ett idiom och dess lexikala komponenter.....	30
FIGUR 2:	Jämförelse av resultatspridningen mellan informantgrupperna	61
FIGUR 3.1:	Antal informanter som nått olika resultat i grupp L1	63
FIGUR 3.2:	Antal informanter som nått olika resultat i grupp L2/tidig	63
FIGUR 3.3:	Antal informanter som nått olika resultat i grupp L1	64
FIGUR 4:	Nätverk mellan de kontaminerade uttrycken och deras lexikala komponenter	117
FIGUR 4.1:	Nätverk mellan olika konventionaliserade uttryck och deras lexikala och semantiska komponenter	118

1. Inledning

Användningen av konventionaliserade uttryck utgör ofta ett av de sista stora hindren på vägen mot en inföddlik språkbehärskning för andraspråksinlärare och behärskning av olika typer av fasta fraser och konventionaliserade uttryck har alltmer lyfts fram som central i samband med både första- och andraspråksinlärning (Wray 2002, Abrahamsson & Hyltenstam 2009, Prentice & Sköldberg und.utg.).

Rimligtvis är det viktigt för inlärare att behärska de mest frekventa uttrycken i vardagligt språk såväl som i ett mer domänspecifikt språkbruk – t.ex. utbildningsrelaterat språk. Vilka konventionaliserade uttryck som är intressanta och relevanta ur språkbeskrivnings- och språkinlärningsperspektiv är emellertid inte bara en fråga om frekvens. Även mer lågfrekventa uttryck, som t.ex. idiom, fyller viktiga funktioner för språkanvändaren och förtjänar större uppmärksamhet än att endast betraktas och läras ut som språkspecifika kuriositeter (McCarthy 1998, O’Keefe, McCarthy & Carter 2007:80ff). O’Keefe, McCarthy & Carter påpekar att idiomatiska uttryck inte i första hand är ”colourful alternatives to their literal counterparts” (2007:82) utan att de innehåller viktig kulturell information samtidigt som de fyller diskursiva funktioner, som t.ex. att uttrycka författarens värderingar (McCarthy 1998: 131f).

Detta arbete fokuserar olika typer av det som här kallas för *konventionaliserade uttryck* (se avsnitt 2.2). En relevant typ av uttryck är den som bl.a. Ellis, Simpson-Vlach & Maynard (2008) kallar *formulaic expression* och som kan vara relativt högfrekvent i allmänspråket såväl som

i det språkbruk som elever möter i skolan. Som exempel kan uttryck som *på egen hand* eller *dra en slutsats* nämnas. I undersökningen tas dock också generellt mer lågfrekventa språkliga fenomen upp – t.ex. idiom som *prata i nattmössan* eller *sätta käppar i hjulen*.

1.1. Syfte och frågeställningar

Den undersökning som redovisas i denna rapport är en enkätstudie som genomförts bland gymnasieelever i flerspråkiga skolmiljöer. Studiens övergripande syfte är att undersöka i vilken mån elever med varierande språklig bakgrund behärskar, dvs. känner igen, förstår och kan producera, några olika typer av svenska konventionaliserade uttryck som de kan förväntas möta och använda. De frågeställningar som ligger till grund för undersökningen är:

- Finns det skillnader i behärskning av konventionaliserade uttryck mellan olika elevgrupper
 - som kan relateras till elevernas språkliga bakgrund, dvs. modersmål och startålder för inlärning e.d.?
 - vad gäller olika aspekter av lexikal kompetens, t.ex. förståelse och produktion?
- Finns det skillnader i behärskning av enskilda konventionaliserade uttryck mellan olika elevgrupper?
- På vilket sätt avviker eleverna från konventionaliserade mönster i sin produktion av konventionaliserade uttryck i samband med denna undersökning?

Frågorna under första punkten kommer här att undersökas med hjälp av en kvantitativ analys. Även den andra punkten utforskas med hjälp av en kvantitativ analys, som dock samtidigt ligger till grund för en kvalitativ analys av ett antal enskilda uttryck. Denna kvalitativa analys svarar mot den tredje frågeställningen.

2. Idiomaticitet, språklig konvention och konventionaliserade uttryck

I detta avsnitt introduceras några fraseologiska begrepp som även har haft viss relevans för forskning om andraspråksinlärning och är grundläggande för de resonemang som förs i detta arbete. Inledningsvis tas begreppet *idiomaticitet* upp, som är en viktig utgångspunkt i analysen av konventionaliserade språkliga strukturer. *Konventionalisering*, *kompositionalitet* och *genomskinlighet* kan sägas utgöra grundläggande begrepp för beskrivningen av konventionaliserade uttryck. Dessutom introduceras även två huvudtyper av konventionaliserade uttryck, nämligen *idiom* och *kollokationer*.

2.1. Idiomaticitet

Den kanske mest grundläggande aspekten i samband med behärskning och användning av konventionaliserade uttryck är det som särskilt i engelskspråkig litteratur ofta kallas *idiomaticitet* (eng. *idiomaticity*). Vissa forskare använder även termer som *inföddlik* (eng. *nativelike språkanvändning* med liknande innebörd (se Pawley & Syder 1983). Idiomaticitet är ett tämligen svårdefinierat begrepp, vilket framgår av följande citat:

Idiomaticity is a non-phonological “accent”, not always attributable to surface language errors, but to a certain undefined quality which many frustrated L2 composition teachers define as “I don’t know what’s wrong with this, but we just don’t say that in English” (Yorio 1989:64).

Detta definitionsförsök gör tydligt att det är svårt att komma åt idiomacitetens kärna och vad exakt det är som avgör om ett språkligt yttrande betraktas som idiomatiskt. Yorio framhäver här dessutom att grammatisk korrekthet inte är en lämplig grund för bedömning av idiomacitet i språkbruk. Även Wiktorsson (2003:3) betonar vikten av att skilja idiomacitet från grammatisk korrekthet i samband med sin korpuslingvistiska studie av idiomatiska uttryck i texter skrivna av svenska inlärare av engelska. En diskussion kring denna distinktion förs även av Pawley & Syder (1983:196), som anför en rad grammatiskt korrekta men icke idiomatiska sätt att uttrycka samma innehåll som i den konventionaliserade frasen *I want to marry you*, t.ex. *I wish to be wedded to you*, *I want marriage with you* och *my becoming your spouse is what I want*. Pawley & Syder skriver vidare att endast en liten andel av de grammatiskt korrekta yttrandena i ett språk också är idiomatiska och därmed skulle accepteras som naturliga av infödda talare. De påpekar också att infödda talare inte på något sätt till fullo utnyttjar sin förmåga att skapa språkliga yttranden genom kreativ användning av syntaktiska regler. Det verkar alltså just vara återanvändningen av vissa konventionaliserade strukturer som utmärker en inföddlik – eller idiomatisk – språkanvändning (jfr Ekberg 1999, 2004).

Wray skriver om språkbrukarens kapacitet att hantera nya, kreativt skapade yttranden, både produktivt och receptivt. Hon betonar att man som mest kan hävda att språkbrukarens benägenhet att använda sig av en relativt stor andel färdigbildade, konventionaliserade fraser existerar parallellt med hans eller hennes förmåga att skapa och förstå nya ordförbindelser:

Although we customarily say, *Hi, how are you doing*, or some other idiomatic greeting on meeting a friend, there is nothing at all to stop us from saying, *What a pleasant event it is to see you. Tell me, how is your life progressing at the moment* (2002:12).

Ett återkommande problem i samband med behandlingen av begreppet idiomacitet är dock att det används med något olika innebörd av olika

forskare. Ofta används det i den betydelse som beskrivs ovan, dvs. avseende den kvalitet i våra språkliga val som gör att de accepteras som naturliga av infödda talare (Yorio 1989:64). I andra fall skriver forskare om idiomacitet som en egenskap hos ett uttryck som anger dess fixering till sin form och brist på genomskinlighet i betydelse. Med andra ord skulle man kunna hävda att ju mer idiomatiskt ett uttryck är, desto mer fixerat är det till sin form och desto mindre genomskinligt är det i sin betydelse (jfr Coulmas 1979, Gläser 1986, 1988, Fleischer 1997). Fleischer beskriver olika grader av idiomacitet hos fraseologiska enheter som han kallar för *fraseologismer* (ty. *Phraseologismen*). Han ser graden av idiomacitet hos ett givet uttryck som ett mått på konstituenternas semantiska förenlighet (kongruens) inom ramen för detta uttryck. Enligt detta synsätt uppvisar uttrycket *dra alla över en kam* 'bedöma alla enligt samma mall' (Svenskt språkbruk 2003:576) en hög grad av idiomacitet, då konstituenterna inte är semantiskt kongruenta med avseende på deras bokstavliga¹ betydelser. Dessutom bestäms graden av idiomacitet hos ett uttryck av huruvida en eller flera av konstituenterna används i sin bokstavliga betydelse (se avsnitt 2.2 nedan). Som exempel kan här nämnas *prata i nattmössan* 'inte veta vad man talar om' (Svenskt språkbruk 2003:804f), där verbet *prata* används i sin bokstavliga betydelse, vilket, ur denna synvinkel, innebär att uttrycket uppvisar en lägre grad av idiomacitet än *dra alla över en kam* (jfr Fleischer 1997:30f). I detta arbete används emellertid termen idiomacitet i den betydelsen som antyds hos bl.a. Yorio (1989:64); idiomatiskt språkbruk är ett sätt att använda språket, som är naturligt ur den infödda talarens perspektiv.

2.2. Konventionalisering, kompositionalitet och genomskinlighet

I likhet med tidigare arbeten (Große 2008, Prentice & Sköldberg und. utg.) använder jag termen *konventionalisering* för att beskriva språkbru-

¹ Fleischer använder termerna *wendungsintern* och *wendungsextern*, som man skulle kunna översätta med 'uttrycksintern' och 'uttrycksextern', dvs. de betydelser som konstituenterna har inom respektive utom det idiomatiska uttrycket ifråga.

karnas benägenhet att uttrycka ett visst språkligt innehåll med en viss (mer eller mindre fixerad) språklig form. Tillämpningen av begreppet är, som Svanlund påpekar, inte helt oproblematiskt med tanke på att konventionalisering måste ses som en ständigt pågående, dynamisk process, vilket innebär att uppsättningen av etablerade mönster kontinuerligt förändras (Svanlund 2001:33). Dessutom skriver Svanlund att innehållet i begreppet *språklig konvention* eller *konventionalisering* inte är lika självklart som många språkvetare verkar anta. Vad man menar beror delvis på vilka motsatsrelationer man utgår ifrån. Svanlund nämner bl.a. språklig konvention som ett kollektivt, socialt styrt fenomen i motsats till de psykologiskt betingade, individuellt särpräglade aspekterna av språkanvändningen. Den motsatsrelation som är relevant i samband med detta arbete är språklig konvention som etablerade mönster i språkbruket gentemot nyskapande eller tillfälliga, dvs. ”mer perifert eller hittills inte alls nyttjade uttrycksmöjligheter” (Svanlund 2001:11). Om man anammar det sistnämnda perspektivet, har användningen av begreppet konventionalisering, enligt min åsikt, flera fördelar. Dels riskerar man ingen begreppsförvirring i samband med idiomaticitetsbegreppet, dels framhäver begreppet även den diakrona aspekten som har betydelse för användningen av fraseologiska enheter; konventionaliserade uttryck uppstår genom att språkbrukarna inom en språkgemenskap upprepade gånger gör samma språkliga val som därmed så småningom etableras som ett naturligt sätt att uttrycka ett visst innehåll. I anslutning till detta resonemang väljer jag även att använda termen *konventionaliserat uttryck* för att beteckna föremålet för denna undersökning. Begreppet används i en vid bemärkelse och innefattar olika typer av uttryck som idiom, kollokationer och andra typer av i viss mån fixerade flerordsenheter.

Det är i praktiken knappast meningsfullt att betrakta olika typer av fraseologiska enheter som helt olika, från varandra isolerade fenomen, då skillnaderna emellan dem i många fall måste betraktas som graduel-la. Olika typer av konventionaliserade uttryck befinner sig snarare längs ett kontinuum; i ena ändan finns det som av vissa forskare har kallats för *rena* eller *prototypiska idiom* (t.ex. Cowie 1988, Sköldberg 2004), som är starkt fixerade i sin form och semantiskt icke *kompositionella* (se vidare i detta avsnitt nedan) – t.ex. *ta tjuren vid hornen* ’direkt och oförskräckt ta itu med något som är svårt eller obehagligt’ (Svenskt språk-

bruk 2003:1211). I andra ändan ligger sådana uttryck som Moon kallar för *frames* (1998:145ff). Dessa har en viss återkommande struktur men kan lexikalt realiseras relativt fritt. Här kan *få kalla kårar*, *få kalla kårar längs ryggen/rygggraden* eller *det går kalla kårar längs ryggen/rygggraden* 'när någon ryser till av rädsla eller obehag' nämnas som exempel (Svenskt Språkbruk 2003:672, jfr Moon 1998:120ff, Sköldberg 2004:174ff). Däremellan finns det bl.a. relativt fasta konstruktioner med ett variabelt led, t.ex. *hålla modet/humöret uppe* 'stå ut med svårigheter, vara vid gott mod' (Luthman 2002:56).

Ett annat relevant begrepp i samband med denna undersökning är *kompositionalitet*. Partee beskriver kompositionalitetsprincipen som följer:

The meaning of an expression is a function of the meaning of its parts and of the way they are syntactically combined (Partee 1984:281).

Svanlund (2009:34) nämner andra, betydligt striktare formuleringar av denna princip, som innebär att betydelsen hos ett flerordsuttryck enbart utgör summan av de ingående elementens betydelse samt de strukturella relationerna emellan dem. Ett exempel på en sådan radikal formulering återfinns hos Costello & Keane:

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).

Kompositionalitetsbegreppet ligger nära det som i svensk litteratur ofta beskrivs som grad av semantisk genomskinlighet, dvs. i vilken mån det är möjligt att härleda frasens betydelse utifrån de enskilda konstituenternas betydelse. Svanlund (2009:36f) diskuterar begreppet *genomskinlighet* i förhållande till en sådan strikt syn på kompositionalitet. Genomskinlighet anges ofta som en av de centrala egenskaper hos ett uttryck (i Svanlunds fall sammansättning) som är nära kopplat till adjektiv av typen *förutsägbar* eller *härledbar*. Svanlund själv använder begreppet genomskinlighet i sin undersökning av hur nya sammansättningar etableras i språket på ett, som han anser, mer preciserat sätt:

Om vi uppfattar en sammansättning som genomskinlig så innebär det att vi lätt kan generera en föredragen defaultbetydelse. Detta innebär inte att andra betydelser är otänkbara, men en viss betydelse uppfattas som mer naturlig än andra kandidater och aktiveras därmed först – om ingenting i sammanhanget indikerar annorlunda (Svanlund 2009:36f).

Om ett uttryck är genomskinligt eller inte är alltså, enligt detta synsätt, en fråga om språkbrukarens upplevelse, vilket bl.a. innebär att bedömningen av ett uttrycks genomskinlighet – som är en del av den kvalitativa analysen i detta arbete – inte är helt oproblematiskt. Det är, som Svanlund påpekar, inte möjligt att komma åt den ”genuina upplevelsen” (2009:36) av genomskinlighet som en läsare eller lyssnare kan ha vid första mötet med ett visst uttryck, om man redan vet vad uttrycket betyder. Man måste alltså använda sin språkliga intuition för att avgöra hur pass genomskinligt uttrycket är för någon som läser eller hör det för första gången. Det som ska avgöras är ofta, inte minst i samband med denna undersökning, om läsaren, lyssnaren eller informanten kan förväntas tolka uttrycket på det konventionaliserade sättet. I detta sammanhang bör det också påpekas att genomskinlighet – liksom kompositionalitet – i enlighet med Svanlund bör ses som ett graduellt fenomen. Med andra ord kan den konventionaliserade betydelsen hos ett uttryck vara mer eller mindre uppenbar (Svanlund 2009:37).

2.3. Idiom

En typ av konventionaliserade uttryck som har beskrivits och definierats på olika sätt är *idiom*. O’Keefe, McCarthy & Carter(2007:62) påpekar att en stor del av forskningen kring idiom har fokuserat de mest icke-kompositionella idiomatiska uttrycken. Dessa kallas ibland för *prototypiska* idiom (jfr Fernando & Flavell 1981, Sköldberg 2004). Som svenska exempel på sådana idiom nämner Sköldberg *väcka ont blod* och *i ett nötskal* (2004:26). I engelskspråkig litteratur anförs gärna bl.a. *kick the bucket* och *pull someone’s leg* som typexempel för idiom (Moon 1998:4, Wray 2002:4). Prototypiska idiom är sådana som har fler av de egenskaper som man betraktar som utmärkande för just idiom än andra, mindre prototypiska idiom (Sköldberg 2004:26). Även här blir det alltså

tydligt att idiom är ett svåravgränsat fenomen med mer eller mindre prototypiska exempel snarare än en homogen kategori. De för idiomerna utmärkande egenskaper som Sköldberg framhäver i sin definition är bl.a. en relativt fixerad form, figurativ betydelse, icke-kompositionalitet och konventionalisering som hon kallar för *institutionalisering* (jfr Moon 1998:7). Vissa forskare har dock använt idiombegreppet i så pass vid bemärkelse att det innefattar olika typer av konstruktioner som inte är fritt formade, som kollokationer (se avsnitt 2.4 nedan), men även icke-kompositionella sammansättningar som *blackbird* (Makkai 1972, jfr Moon 1998:4). O’Keefe, McCarthy & Carter(2007:62) refererar till ”the ordinary idioms of every day” och avser bl.a. öppningsfraser i samtal (eng. *gambits*) av typen *guess what*, och diskurspartiklar (jfr Coulmas 1979).

Jag utgår däremot i samband med denna undersökning, liksom i tidigare arbeten (Prentice & Sköldberg und. utg.), från en definition som återfinns hos Langlotz (2006):

An idiom is an institutionalised construction that is composed of two or more lexical items and has the composite structure of a phrase or semi-clause, which may feature constructional idiosyncrasy. An idiom primarily has an ideational discourse-function and features figuration, i.e. its semantic structure is derivationally non-compositional [or not fully compositional]. Moreover, it is considerably fixed and collocationally restricted (Langlotz 2006:5)

Langlotz kallar denna vida definition för preliminär (ibid.) och även han är noggrann med att påpeka att gränsdragningen mellan idiom och intilliggande fraseologiska enheter på intet vis är skarp (jfr O’Keefe, McCarthy & Carter2007:81f). Langlotz (2006:6) menar vidare att en sådan definition kan inkludera så pass olika konstruktioner som *red herring* (’villospår’), *take the bull by the horns* (’ta tjuren vid hornen’) och *fall flat* (’misslyckas’) (Prentice & Sköldberg und. utg.).

2.4. Kollokationer

En annan typ av fraseologiska enheter som fokuseras i denna undersökning är den som brukar kallas för *kollokation*. Även detta begrepp be-

traktas på olika sätt av forskare inom olika traditioner, vilket refereras i de två följande avsnitten.

2.4.1. Kontextualistisk kollokationsforskning

I litteraturen framställs ofta Firth (bl.a. 1957) som pionjär inom kollokationsforskningen, även om Palmer (1933), som Nation påpekar (2001: 317), använde begreppet långt tidigare. Själva kollokationsbegreppet förblir dock något vagt i Firths egna arbeten (jfr 1957:195). Han kan sägas i hand ha inspirerat i första hand den typen av kollokationsforskning som definierar kollokation som frekvent samförekomst (eng. *co-occurrence*) av två eller flera ord i en textkorpus (Malmgren 2003:124, jfr Stubbs 2001:29). Både Halliday, McIntosh & Strevens (1964, se även Halliday 1966) och Sinclair (bl.a. 1966, 1987) har sedan byggt vidare på Firths kollokationsbegrepp i sin forskning. Halliday et al. argumenterar utifrån sannolikheten av ords samförekomst och konstaterar: "Given the item 'chair', we are more likely to find in the same utterance the items 'sit' or 'comfortable' [...] than, say, 'haddock'..." (1964:33). Det lexikala element som studeras med avseende på dess "kollokationsbeteende" kallar Halliday (1966) för *nod* (eng. *node*). De lexikala element som kollokerar med en viss nod kallar han *kollokat* (eng. *collocate*). Sinclair skiljer i sin tur mellan två olika typer av kollokationer: *upward collocations* och *downward collocations*. Distinktionen är relaterad till nodens respektive kollokatets frekvens; om noden är det mer frekventa ordet i en kollokation med ett mindre frekvent kollokat har vi att göra med en *downward collocation*. Om noden däremot kollokerar med ett mer frekvent kollokat talar Sinclair om *upward collocation*. Han illustrerar denna distinktion med en rad kollokationer där han utgår ifrån ordet *back* som nod. Som typiska exempel på downward collocates nämner Sinclair bl.a. verb som *arrive* (to arrive back on thursday), *cut* (to cut back on something) eller *fall* (to fall back on something). Han nämner även substantiv som *hotel* och adjektiv som *normal* (back to the hotel, back to normal), adverb som *straight* och *further* (straight/further back) och prepositioner som *behind* och *towards* (back behind/towards the church) som möjliga *downward collocates* (Sinclair 1987:328f). Som exempel på *back* (både i betydelsen 'tillbaka' och 'rygg') i kollokation med mer frekventa kollokat (*upward collocates*) nämner han bl.a. *being back at school*, *follow*

someone back into the wood och a *slap on the back*. Sinclair ser en systematisk skillnad mellan dessa två kollokationstyper:

Upward collocation of course is the weaker pattern in statistical terms, and the words tend to be elements of grammatical frames, or superordinates. Downward collocation by contrast gives us a semantic analysis of a word (Sinclair 1987:326).

För svenskans del har bl.a. Svanlund (2001) använt sig av frekvensbaserade kollokationsanalyser i sin korpusbaserade studie av konventionella metaforer. Han påpekar dock att lågfrekventa ordkombinationer inte nödvändigtvis är ointressanta:

... ibland kan de spegla en viss semantisk kombinationstyp utan att för den skull alla instantieringar av typen behöver vara särskilt vanliga (Stubbs 1996). Risken att missa väsentliga typer på grund av frekvensspärren är dock liten; sådana typer representeras som regel ändå bland de frekventa kollokaten (Svanlund 2001:107f).

Ett av problemen i samband med kollokationsbegreppet är att frekvens är en relevant aspekt, men att den i sig inte är tillräcklig för att definiera kollokationer som ”meningsfull kategori för en analys av lexikala mönster i språkssystemet” (Lindberg 2004:68, jfr Malmgren 2003:125). Malmgren framhäver de framsteg inom kollokationsforskningen som ledde till en åtskillnad mellan ”grammatiskt välformade och rent tillfälliga kollokationer” (2003:125). Dessa framsteg nåddes redan på 1970- och 80-talet, inte minst med hjälp av stora empiriska studier i svenska (Allén 1975, jfr även Kjellmer 1994 för engelska). Inom den datoriserade korpuslingvistik har man utvecklat statistiska metoder för att kunna skilja mellan visserligen frekventa, men tillfälliga samförekomster av ord, t.ex. *och det är...* (jfr Biber & Conrad 1999 om *lexical bundles*), och ordförbindelser där de ingående orden är starkt associerade med varandra (jfr Ellis, Simpson-Vlach & Maynard 2008), t.ex. *fatta beslut*. Med hjälp av datorer kan man inte bara ta reda på en ordförbindelses frekvens i en korpus, utan också mäta det som kallas för *Mutual Information* (MI). MI är ett statistiskt mått på i vilken mån konstituenterna i en ordförbindelse förekommer oftare tillsammans än man utifrån slumpen kan förutsäga (jfr Church & Hanks 1990); ju högre frasens MI är desto starkare associeras de ingående orden med varandra. På det sättet

kan man alltså skilja kollokationer och andra fraseologiska enheter som har specifika betydelser och funktioner från sådana ordkombinationer som är frekventa endast till följd av att de ingående orden är frekventa i den aktuella korpusen (Ellis, Simpson-Vlach & Maynard 2008:380, jfr Svanlund 2001:108f).

2.4.2. Systemorienterad kollokationsforskning

Något som helt saknas i den frekvensbaserade, kontextualistiska traditionen inom kollokationsforskningen är direktionalitetsperspektivet. Detta sätt att tänka kring kollokationer utvecklades av bl.a. Hausmann (1985) inom ramen för den systemorienterade kollokationsforskningen i Tyskland (Malmgren 2003:127). Hausmann, som är en hård kritiker av den kontextualistiska skolan, anser att konstruktioner som *köpa en bok* inte är kollokationer, då de är semantisk triviala. Enligt ett systemorienterat synsätt kännetecknas en kollokation av att en av konstituenterna i viss mån är oförutsägbar och begränsad i sin kombinerbarhet (Malmgren 2003:127). Denna skillnad i konstituenternas status är en viktig förutsättning för tillämpningen av direktionalitetsbegreppet. Hausmann använder begreppen *bas* och *kollokator*, som alltså skiljer sig från kontextualisternas nod och kollokat just genom betoningen på direktionalitet, dvs. ett slags hierarkiskt förhållande emellan dem; i kollokationen *dra en slutsats* är *slutsats* basen och *dra* kollokator. Direktionaliteten kan, enligt Hausmann, motiveras med att man vid produktion av kollokationer (i text) går från bas till kollokator (1985:119). Med andra ord kan man som språkbrukare tänkas leta efter det rätta verbet att använda tillsammans med substantivet *slutsats*, om man vill uttrycka innehållet i *dra en slutsats*. Däremot skulle språkbrukaren knappast utifrån verbet *dra* leta efter rätt substantiv att använda i syfte att uttrycka samma innehåll på ett idiomatiskt sätt (jfr Hausmann 1985:119ff, Malmgren 2003:127).

Kollokationsbegreppet i systemorienterad mening introducerades i svensk forskning av Svensén (1987). Huvudsyftet med hans arbete inom detta område har varit att utveckla begreppet för tillämpning inom lexikografisk forskning (Malmgren 2003:127). Inom detta fält har Sköldberg & Toporowska Gronostaj (2008) introducerat en avgränsning av kollokationsbegreppet, som de tillämpar i samband med beskriv-

ningen av kollokationer i ett elektronisk medicinsk lexikon (MedLex, Kokkinakis 2006). Författarna ansluter sig i det väsentliga till den systemorienterade forskningstraditionen. De skriver om prototypiska kollokationer och syftar till att ge en i första hand strukturell och funktionell beskrivning av kollokationer i medicinskt sammanhang (Sköldberg & Toporowska Gronostaj 2008:434f, jfr Bartsch 2004). De konstaterar att det finns många prototypiska kollokationer med liknande strukturer i materialet, bl.a. sådana med samma semantiska förhållande mellan bas och kollokator. Som exempel nämns kollokationer som uttrycker gradförhållanden, t.ex. *hög feber* och *djup depression* (Sköldberg & Toporowska Gronostaj 2008:435). De vanligast förekommande relationerna mellan komponenterna i en kollokation beskriver författarna i termer av lexikala funktioner (se Mel'čuk 1998). De tar även upp direktionaliteten i sin definition och nämner som exempel kollokationen *ta tempen*, som består av basen *tempen* och kollokatorn *ta*. De betonar dock också att det finns kollokationer där det är betydligt svårare att bestämma direktionaliteten. Exempel på sådana kollokationer är, enligt Sköldberg & Toporowska Gronostaj, *ordinera vila* och *ordinera läkemedel* (2008:434).

En tidigare svensk studie med strukturell och funktionell inriktning om en viss typ av just kollokationer är Ekbergs undersökning av det hon kallar för abstrakta övergångsfraser (1989). Med det avses sådana kollokationer som med Ekbergs terminologi utgör ”syntaktisk komplexa predikat som betecknar övergången från ett tillstånd till ett annat tillstånd och är semantiskt ekvivalenta med enkla verb eller adjektivkonstruktioner”. Några av Ekbergs exempel är *falla i sömn*, *gå till attack* och *komma till besinning*.

Om man ska försöka ringa in kollokationsbegreppet på ett sätt som motsvarar denna undersöknings syfte, kan man säga att en kollokation består av en bas som starkt associeras med en, eller ett fåtal kollokatorer. Dessutom kännetecknas kollokationer av ett visst mått av icke-kompositionalitet. Med det menas att det utifrån basens betydelse inte är helt förutsägbart vilka som är de, enligt konventionen, prefererade kollokatorerna. Enligt en sådan definition är *fatta ett beslut* en relevant kollokation, med substantivet *beslut* som bas och verbet *fatta* kollokator. *Läsa en bok* matchar däremot inte kriterierna för en kollokation; för det första finns inga starka restriktioner vad gäller konstituenternas kombinerbar-

het med andra ord – man kan göra olika saker med en *bok* och man kan *läsa* olika saker. För det andra är frasen kompositionell, dvs. utifrån betydelsen av *bok* kan förutsägas vilka ord det kan kombineras med, vilket också gäller för verbet *läsa* (jfr Malmgren 2003:127, Hausman 1985).

Vidare kan man konstatera att frekvens är en relevant, men inte avgörande aspekt i samband med kollokationsbegreppet, som O’Keefe, McCarthy & Carter(2007) påpekar i citatet nedan.

We are all familiar with the situation in class where we fall back on the statement ‘that’s just the way we say it’, when faced with an awkward question from a student about why something is expressed the way it is, and often, what we are really explaining is a strong statistical preference which can be powerfully demonstrated by the use of corpus data. The answer, in the final analysis, is still that collocation shows us ‘the way we say it’, but we can gain considerable confidence as teachers if we can present something as a widespread frequent collocation rather than a one-off occurrence in the particular text we are working with (O’Keefe, McCarthy & Carter2007:59).

Författarna menar att det ofta är de kollokationer vars konstituenten består av banala, vardagliga ord, som är svåra att ta fram och belysa endast utifrån den egna språkliga intuitionen. Dessa fraseologiska enheter kan med fördel tas fram och undersökas med hjälp av kvantitativa, korpuslingvistiska metoder, vilket jag till viss del även har gjort i samband med denna undersökning (se vidare avsnitt 5.3). Frekvensen i en korpus kan dessutom bekräfta intuitionen om ett uttrycks konventionalisering, vilket kan komma till nytta, bl.a. i andra- eller främmandespråksundervisning.

3. Lexikal och fraseologisk kunskap – vad kan man när man ”kan” ett uttryck?

Inom kognitiva teorier utgår man ifrån att vår kunskap om världen består av konceptuella strukturer som är lagrade i minnet, och som vid behov kan överföras till fenomen som upplevs i den reella världen för att förstå dem (Langlotz 2006:61). Langlotz nämner som exempel att en människa kan föreställa sig en katt utan att varken se, höra eller röra vid en katt i verkligheten. Detta tyder på att människan har ett begrepp lagrat i minnet, som består i information om bl.a. hur en katt ser ut, låter och känns (jfr även Lamb 2000:183). Dessa begrepp fungerar, enligt Langlotz, som kategorier som hjälper oss att hantera och organisera de enorma mängder information om vår omvärld som vi utsätts för:

Working as a piece of meaningful knowledge, we can project this conceptual CAT-schema as a categorization-standard to make sense of any concretely perceived target-creature that features a certain degree of ‘catness’. Thus, meaningful experience is not only related to our immediate perception of the physical world, but draws upon conceptual routines [...]. Concepts allow us to make sense of our environments [...], conceptual units work as [...] categorisation-standards that can be mapped onto the flow of experience to group and organize it (Langlotz 2006:61f).

Bybee (1998, 2006) är en i en rad funktionellt och kognitivt inriktade lingvister som bedriver språkforskning inom en teoretisk ram som går under benämningen *usage-based approach* (se Greenberg 1966, Hopper & Thomson 1980, Langacker 2000). Hennes utgångspunkt är att lexikonets struktur är baserad på det hon kallar språkliga erfarenheter (jfr Langlotz 2006).

While all linguists are likely to agree that grammar is the cognitive organization of language, a usage-based theorist would make the more specific proposal that grammar is the cognitive organization of one's experience with language. [...] [C]ertain facets of linguistic experience, such as frequency of use of particular instances of constructions, have an impact on representation that we can see in evidence in various ways, for example, in speakers recognition of what is conventionalized and what is not. [...] The proposal presented here is that the general cognitive capabilities of the human brain, which allow it to categorize and sort for identity, similarity, and difference, go to work on the language events a person encounters, categorizing and entering in memory these experiences (Bybee 2006:711).

Detta synsätt på grammatik har bl.a. en del konsekvenser för hur man ser på representationen av konventionaliserade språkliga strukturer i det mentala lexikonet, vilket diskuteras vidare i avsnitt 4.1.1 nedan.

3.1. Vad kan man när man kan ett ord?

Av den språkliga kompetensens många olika aspekter är det ordförrådet som utvecklas mest under en människas liv (Viberg 1993:35). Denna utveckling framskrider särskilt snabbt under skolåren. En allmän uppskattning är att ett barn vid skolstarten har ett ordförråd på mellan 8 000 och 10 000 ord på modersmålet och att det lär in ca 3 000 nya ord per år. Här ska dock nämnas att dessa siffror i första hand bygger på läsförståelsetest och att det naturligtvis också finns stor individuell variation (Viberg 1993:35f). Även om forskarna inte är helt överens om de exakta siffrorna innebär skolstarten alltså en drastisk expansion av ordförrådet för förstaspråkselever och samtidigt en stor utmaning för elever som har undervisningsspråket som andraspråk. Andraspråkseleverna behöver komma i kapp de jämnåriga förstaspråkseleverna när det gäller basordförrådet som dessa redan har vid skolstarten, samtidigt som de ska

utöka ordförrådet på andraspråket med flera tusen ord om året (Viberg 1993:36).

Att ha ett ordförråd som räcker till för att hantera de språkliga krav som ungdomar ställs inför, bl.a. i skolan, är dock inte enbart en fråga om mängden ord man hunnit lära in. Lexikonkunskap är en av de mest komplexa och mångfacetterade aspekter av språkbehärskningen, eller med Nations ord:

Words are not isolated units of language, but fit into many interlocking systems and levels. Because of this, there are many things to know about any particular word and there are many degrees of knowing (Nation 2001:23).

Nation utgår ifrån en modell för lexikonkunskap som innebär en generell distinktion mellan receptiv och produktiv lexikonkunskap. Denna modell innefattar en rad olika aspekter av kunskap och språkanvändning (Nation 2001:28) och sammanfattas i tabell 1.

TABELL I: Olika aspekter av lexikonkunskap (efter Nation 2001:27)

		Receptiv/ produktiv kunskap	
Form	tal	R	Hur låter ordet?
		P	Hur uttalas ordet?
	skrift	R	Hur ser ordet ut ?
		P	Hur skrivs och stavas ordet?
morfologi	R	Vilka delar av ordet kan urskiljas?	
	P	Vilka delar behövs för att uttrycka den avsedda betydelsen?	
Betydelse	form och betydelse	R	Vilken betydelse förmedlar formen ifråga?
		P	Vilken form kan användas för att uttrycka den avsedda betydelsen?
	begrepp och referenter	R	Vad ingår i det aktuella begreppet?
		P	Vilka föremål i den reala världen kan begreppet referera till?
	associationer	R	Vilka andra ord kommer vi att tänka på i samband med det aktuella ordet?
		P	Vilka andra ord (synonymer) skulle vi kunna använda istället?
Språkanvändning	grammatisk funktion	R	I vilka konstruktioner och mönster förekommer ordet?
		P	I vilka konstruktioner ska ordet användas?
	kollokationer	R	Vilka ord eller typer av ord förekommer i kombination med det aktuella ordet?
		P	Vilka ord eller typer av ord ska användas i kombination med det aktuella ordet?
	begränsningar i användning register, frekvens	R	Var, när och hur ofta kan vi förvänta oss att möta ordet?
		P	Var, när och hur ofta kan vi använda ordet?

De flesta beskrivningar av språkinläringen utgår ifrån att receptiv inläring och språkanvändning är enklare än den produktiva behärskningen av språket. Däremot råder det oklarhet över orsakerna till denna skillnad. Nation refererar ett antal möjliga förklaringar som ges i littera-

turen, vilka, enligt honom, inte nödvändigtvis utesluter, utan snarare kompletterar varandra. En förklaring är den som Nation kallar för "the amount of knowledge explanation" (2001:28ff), som bygger på antagandet att produktiv behärskning av lexikonet kräver kunskap utöver den som krävs för receptiv språkbehärskning. Språklig produktion, muntlig eller skriftlig, kräver, enligt detta synsätt, mer exakta kunskaper om ordets eller uttryckets form och betydelse samt dess plats i olika delar av språkssystemet, än receptiv hantering av det språkliga inflödet. I samband med det senare kan det räcka långt att känna till ett mindre antal utmärkande drag hos den lexikala enheten ifråga. En annan möjlig förklaring är enligt Nation "the 'practice' explanation" (2001:29). Enligt denna teori får inlärare vanligtvis fler tillfällen att öva sina receptiva än sina produktiva färdigheter. Nation hänvisar i det sammanhanget till forskning som tyder på att receptiva och produktiva språkfärdigheter måste övas var för sig för att leda till full språkbehärskning (DeKeyser & Sokalski 1996). Denna förklaring strider mot den nyss nämnda synen på produktiv språkbehärskning som summan av alla receptiva, och ytterligare, specifikt produktiva, kunskaper (Nation 2001:29). Ytterligare en förklaring är "the 'access' explanation" som bygger på tanken att ett nytt inlärt ord i andra- eller främmandespråket i ett tidigt inlärningsstadium är direkt länkat till ekvivalenten i inlärarens förstaspråk (Ellis & Beaton 1993, jfr Bialystok 2001, se vidare avsnitt 4.3). Detta innebär att ordet kan hanteras receptivt och förstås med hjälp av förstaspråkskompetensen. På förstaspråket har inläraren även tillgång till en rad olika associationer i samband med det aktuella ordet (bl.a. synonymer, kollokatorer och antonymer). För att kunna använda ordet produktivt på ett korrekt och idiomatiskt sätt måste inläraren välja rätt bland alla dessa associationer och sedan överföra dem till andraspråket, vilket gör den produktiva användningen mer komplex (Nation 2001:28f).

Även Henriksen (1999), och andra forskare före henne (t.ex. N. Ellis 1995 och R. Ellis 1995) har betonat att lexikonkunskap är sammansatt av flera olika typer av kunskap som tillsammans innebär full behärskning av ett ord (eller ett konventionaliserat uttryck). Hon föreslår en tredimensionell modell för lexikal utveckling, enligt vilken lexikal kompetens består av "(a) *partial to precise knowledge*, (b) *depth of knowledge*, and (c) *receptive to productive use ability*" (1999:303). Modellen är dels tänkt som underlag för fortsatt forskning i form av longitudinella

studier av utveckling av lexikal kompetens, dels för utveckling av olika metoder för att testa inlärares lexikala kompetens utifrån alla tre dimensionerna vid en viss tidpunkt (Henriksen 1999:315). De tre ovannämnda dimensionerna berörs i viss mån även i samband med min undersökning. Dimensionerna (a) och (c) kan på samma sätt appliceras på flerordsenheter som på enskilda ord i lexikonet. Depth-of-knowledge-perspektivet (b) är relevant både utifrån en holistisk syn på flerordsenheter och i relation till uttryckens enskilda konstituent; djupare kunskap om ett ords betydelse handlar inte endast om relationen mellan ordet – det begrepp som det betecknar (se avsnitt 3) – och dess referent i den reella världen, utan även om ordets relationer till andra ord i lexikonet. Härmed avser Henriksen bl.a. paradigmatiske relationer som synonymi eller hyponymi och syntagmatiska relationer som bl.a. innebär restriktioner ifråga om vilka andra ord ett givet ord kan kombineras med (Henriksen 1999:305f, jfr Nation 2001). Det senare är i hög grad relevant för denna undersökning.

I detta avsnitt har ordkunskapsforskningen hittills diskuterats utifrån det enskilda ordet. En anledning till detta är att mycket av ordinlärningsforskningen i stor utsträckning fokuserat kunskap på ordnivå (Schmitt 2008:340). Beskrivningarna av olika aspekter av lexikonkunskap som tagits upp här kan dock i princip tillämpas på både ettords- och flerordsenheter, även om det naturligtvis finns ytterligare aspekter som är speciellt relevanta för kunskap om och behärskning av flerordsenheter av olika slag. Utvecklingen av lexikal kunskap på frasnivå och de konventionaliserade uttryckens roll i det mentala lexikonet diskuteras mer utförligt, utifrån tidigare fraseologisk forskning, i avsnitt 4.

En relevant aspekt i samband med kunskap om kollokationers betydelser tas upp av Bartsch (2004), som poängterar att prototypiska kollokationer bär på implicit kulturell eller domänspecifik information:

Beyond influencing and altering the meanings of the constituent words, collocations often carry an additional element of meaning that is not overtly introduced through the individual meanings of their constituents. E.g. the collocation *consenting adult* is defined in the CED [Collins English Dictionary, 1991] as 'a male person over the age of twenty-one, who may legally engage in homosexual behaviour in private'. Taken at face value, all the word combination denotes is an adult who can consent to some unspecified action or event. The additional meaning components are culturally

determined stereotypes implicitly conveyed by this particular collocation in a British English context (Bartsch 2004:75).

Sköldberg & Toporowska Gronostaj (2008:435) nämner *fri abort* och *medicinsk* respektive *kirurgisk abort* som exempel som kan vålla vissa problem, inte minst för inlärare av svenska, som inte enbart kan förlita sig på sina språkliga kunskaper när det gäller att använda eller tolka dessa kollokationer på ett idiomatiskt sätt. I fallet av *fri abort* krävs det kunskap om samhällskontexten, och för att förstå skillnaden mellan *medicinsk* och *kirurgisk abort* måste språkbrukaren använda sig av domänspecifik (medicinsk) metakunskap. Bartsch betonar att "[s]uch implicit or inherent meanings regarding cultural or domain specific stereotypes or encyclopaedic knowledge about the world must be taken as critical for collocations" (2004:76).

3.2. Att mäta lexikonkunskap

Lexikonkunskapens komplexitet medför bl.a. att det kan vara svårt att välja rätt testmetod för att mäta inlärares behärskning av lexikonet i målspråket. Vilket testformat som väljs beror givetvis till stor del på vilken aspekt av lexikonkunskapen som ska testas. Nation diskuterar olika typer av ordförrådstest med utgångspunkt i tidigare forskning (2001:344ff). Flervalstester har, enligt Nation (2001:349ff med anförda referenser), en del fördelar framför andra testformat. Inläraren kan t.ex. ha nytta av partiell behärskning av ett ord (2001:349). Han hänvisar också till forskning som tyder på att vissa problem – som t.ex. att inlärares kan gissa sig fram till det korrekta svaret – inte är av särskilt stor betydelse när det gäller flervalssuppgifter i praktiken:

Paul, Stallman and O'Rourke (1990) interviewed first language learners about the strategies they used to answer particular multiple-choice items [...]. Paul et al. found that for both high-ability and low-ability readers, over 50% of the answers were chosen because of association. [...] The high-ability group guessed only 8% of the answers (with about a 50% success rate) while the low-ability group guessed 21% (with a 35% success rate). This indicates that guessing is not a major problem with multiple choice items and that learners responses are generally not random but largely driven by some knowledge of the words (Nation 2001:349f).

Med andra ord verkar det alltså inte vara möjligt att i någon nämnvärd grad kompensera brist på ordkunskap med gissning i ett flervalstest. Man bör dock i detta sammanhang komma ihåg att den undersökning som gjorts av Paul et al. baseras på självrapporterade data. Forskning som har jämfört flervalstester med andra testformat har visat att flervalstuppgifter verkar vara betydligt lättare att klara av för inlärarna än uppgifter där de t.ex. ombeds att förklara ordet i en mening eller att komplettera en mening med ordet ifråga (Nist & Olejnik 1995). Det är dock möjligt att konstruera flervalstest av varierande svårighetsgrad, som i första hand verkar vara relaterad till hur pass nära distraktorerna ligger den korrekta formen i betydelse (Nagy, Herman & Anderson 1985, se även Nation 2001:349). Denna fråga har även varit relevant att ta ställning till i samband med den enkät som jag konstruerat för min undersökning (se vidare avsnitt 5.4).

Andra testformat som är relevanta i detta sammanhang är det som Nation kallar för *sentence-completion*, vilket i samband med denna undersökning innebär att informanterna får komplettera ett ofullständigt konventionaliserat uttryck i kontext, och *supply item types* (2001:349), där uppgiften består i att komplettera konventionaliserade uttryck utan kontext (se vidare avsnitt 5). Båda dessa testformat har beskrivits som relativt svåra i jämförelse med t.ex. matchningsuppgifter, där ordet (eller uttrycket) ifråga är givet och uppgiften består i att matcha ihop det med dess korrekta betydelse. En skillnad i svårighetsgrad består i att inläraren, utöver att känna till ordens betydelse, också måste ha kunskap om deras kollokationspotential (Nation 2001:352f), vilket i hög grad är relevant för min undersökning.

Golden (2005:88ff) har i sin studie av andraspråkselevs förståelse av metaforiska uttryck i läroböcker använt en form av matchningstest, bestående av flervalstuppgifter som dels testar elevernas förståelse av metaforiska uttryck, dels deras behärskning av konstituenters bokstavliga betydelse. Hennes resultat visar bl.a. att det inte nödvändigtvis är en förutsättning att känna till konstituenternas bokstavliga betydelse för att förstå det metaforiska uttrycket ifråga. Golden förklarar detta resultat med att en del av uttrycken är föremål för automatiserad receptiv användning, dvs. eleverna förstår dem ”direkte uten å gå veien om den konkrete betydningen” (2005:285, jfr Moon 1998:164).

Studier av ordförrådet i svenska läromedel har, med hjälp av en rad olika metoder, genomförts inom ramen för projektet *Ordförrådet i läromedel* (OrdIL). Förutom korpuslingvistiska analyser av hur ordförrådet i läromedel ser ut har man inom ramen för detta projekt också testat elevers behärskning av delar av detta ordförråd med hjälp av olika testformat. Holmegaard (2007:135ff) har bl.a. använt sig av matchning och flervalstest för att testa elevers förståelse av sammansatta ord som *förhandlingsvilja*, *majoritetsbeslut* eller *justitiedepartement*. De infödda eleverna i undersökningen klarade dessa typer av uppgifter betydligt bättre än andraspråkseleverna, vilket måste ses som ett förväntat resultat. Eleverna fick också i uppgift att dela orden i dess huvudled (t.ex. *förhandlings/vilja*) och att bedöma sin egen kunskap om de aktuella orden i ett självskattningstest (ibid.). De ord som var svårast att förstå för samtliga elever var, enligt Holmegaard (2007:158) ”betydelsestunga substantiv som kräver sociokulturell förståelse” (t.ex. *justitiedepartement*). Tittar man på andraspråkseleverna som separat grupp är det däremot abstrakta ”läroboksord” (Holmegaard 2007:158) som *inflytande* och *företräda* samt polysema ord som *rörelse* och *utskott* som verkar vålla störst problem.

4. Konventionaliserade uttryck i språkutveckling

4.1. Förstaspråksutveckling

Wray (2002:105) påpekar att man inom forskningen kring förstaspråksutveckling sedan länge är överens om att barn ofta verkar hantera flerordskombinationer som enheter i början av sin språkutveckling (Bolinger 1975, Crystal 1997). Hon skriver vidare att forskarna i stort sett är eniga om att barn använder relativt komplexa ordförbindelser innan de behärskar dessa enheters grammatiska struktur. Däremot råder ingen konsensus kring dessa fraseologiska enheters betydelse för förstaspråksutvecklingen (2002:105). Medan vissa forskare inte tillmäter inläringen och användningen av oanalyserade flerordsenheter hos barn någon större betydelse (t.ex. Brown 1973), ser andra helfrasinläringen som en av de grundläggande aspekterna av tidig förstaspråksinläring (t.ex. Cruttenden 1981, se Wray 2002:105f). Cruttenden antar att förstaspråksinläring sker i två stadier, nämligen *item-learning* och *system-learning*. Detta gäller, enligt Cruttenden, alla aspekter av språket, dvs. fonologi, morfologi, syntax och semantik (1981:79ff, se även Wray 2002:106). Cruttenden menar vidare att man kan utgå ifrån att ett yttrande har lärts in som en helhet om ingen av de ingående enheterna används i andra ordkombinationer (1981:84). Locke menar att småbarn i tidiga stadier av sin språkutveckling knappt verkar urskilja enskilda ord som enheter över huvud taget (1997:272, se även Wray 2002:106).

Inom barnspråksforskningen utgår man ifrån att barn i tidiga inlärningsstadier använder två olika typer av flerordskombinationer; (1) ord-

förbindelser som barnet har tagit in via inflödet och som kallas för *underanalyzed strings* och (2) flerordsenheter som barnet själv antas ha skapat och sedan lagrat i minnet som en enhet. Dessa kallas för *fused strings*. Den senare typen kan vara grammatiskt välformad eller inte. Här handlar det om idiosynkratiska ordförbindelser som smälts samman till enheter och på så sätt görs lättillgängliga vid on-line kommunikation (Wray 2002:106ff, jfr Wray & Perkins 2000:2, jfr Peters 1983:82). Eftersom föreliggande arbete fokuserar uttryck som är konventionaliserade i (vuxna) infödda talares språkbruk, är det bara den första typen som är relevant här.

Wray & Perkins presenterar en psykolingvistisk modell för första-språkutveckling, i syfte att förklara den frekventa användningen av (idiomatiska eller idiosynkratiskt bildade) fraseologiska enheter i början av barns språkutveckling. Användningen har funnits minska under en period, för att sedan öka igen, genom en mer automatiserad användning av idiomatiska, konventionaliserade uttryck (2000:19ff). Modellen bygger på Lockes teori (1997) om två olika neurologiska mekanismer som ger upphov till en sådan utveckling. Enligt denna teori är den tidiga språkinläringen stimulerad av barnets sociala omgivning, som sätter ramar för det språkliga inflödet, i vilket komplexa språkliga konstruktioner identifieras som meningsfulla och lagras i minnet. Eftersom barnets kognitiva förmåga ännu inte är tillräckligt utvecklad för en språklig analys av dessa konstruktioner, tas de in som oanalyzerade enheter. Den mekanism som ligger till grund för denna process kallas *specialization in social cognition* (SSC) (Wray & Perkins 2000:19). För att språkutvecklingen ska framskrida måste barnet börja analysera de språkliga enheter som samlats i det mentala lexikonet. För att detta ska kunna ske, aktiveras i nästa utvecklingsfas det som Locke kallar för *grammatical analysis module* (GAM). Förklaringen till att användningen av fraseologiska enheter återigen ökar i nästa utvecklingsfas och nu även i högre grad inkluderar konventionaliserade uttryck är, enligt Wray & Perkins (2000:20), att GAM i ökande grad trängs undan av SSC till förmån för en mer effektiv, ekonomisk språkanvändning. Förmågan att tolka och använda metaforiska uttryck som idiom på ett idiomatiskt sätt antas utvecklas i ett ännu senare stadium, då de processer som styrs av GAM och SSC integreras i allt högre grad (Wray & Perkins 2000:19ff, jfr Skehan 1998 om andraspråkutveckling nedan och Bybee 1998). Ef-

fektivitetsskålet, som framstår som mest grundläggande för en ökad användning av flerordsenheter i Wray & Perkins modell, kan dock knappast vara hela sanningen när det gäller användningen av konventionaliserade uttryck. Detta påpekas bl.a. av Wiktorsson (2003), som framhäver uttryckens semantiska funktion (jfr Bartsch 2004:75):

When certain expressions are used repeatedly in the same context, they attract meaning components from that context. These meaning components are part of the meaning of the expressions, and the expressions therefore function as efficient means of communicating quite complex situations (Wiktorsson 2003:149).

4.2. Andraspråksutveckling

Kognitivt inriktade andraspråksforskare som Skehan (1998) antar att andraspråksinläring fungerar med hjälp av två parallella, interagerande system – ett *exempelsystem* och ett *regelbaserat system* – och att det är samspelet mellan dessa system som möjliggör en effektiv och kontinuerlig utveckling av inlärares språkbehärskning (Skehan 1998:88ff, jfr Cruttenden 1981 och Wray & Perkins 2000 om förstaspråksinläring). Skehan beskriver utvecklingen av de två systemen utifrån termerna *lexikalisering*, *syntaktikalisering* och *relexikalisering*. Med hjälp av lexikaliseringen byggs inlärares exempelsystem, som tjänar till att tillhandahålla färdigbildade fraser i kommunikationssituationer där tid och kapacitet för processande är begränsade. För att inlärares kontinuerligt ska kunna utveckla sitt språk måste så småningom de inlärdade fraseologiska enheterna analyseras genom syntaktikalisering, som möjliggör kreativ användning utifrån ett regelbaserat system. Det krävs dock ytterligare ett steg för att så småningom kunna klara av effektiv on-linekommunikation på en mer avancerad nivå, nämligen att de ordförbindelser som genomgått syntaktikalisering relexikaliserar för att återigen bli tillgängliga som oanalyserade enheter för inlärares vid automatiserad användning (Skehan 1998:90, se även Lindberg 2004:65, jfr Wray & Perkins 2000). Den samverkan mellan de olika systemen som en effektiv språk-användning kräver sätts, enligt Skehan, vid förstaspråksinläring automatiskt igång med hjälp av individens *language acquisition devise* (LAD). Andraspråksinlärare med sen startålder antas däremot inte längre ha till-

gång till LAD och riskerar därför att stagnera på lexikaliseringsstadiet (Skehan 1998:90). En liknande utveckling av användningen av konventionaliserade uttryck i ett andraspråk – från holistisk inläring över en period av analytiskt processande, tillbaka till en mer automatiserad språkanvändning – beskrivs av Wray & Perkins (2000, jfr Bybee 1998, 2006 och Wray 2008 nedan). Som nämndes tidigare, är en anledning till barns frekventa användning av oanalyserade flerordsenheter, enligt Wray & Perkins, att deras kognitiva förmåga inte är tillräckligt utvecklad för att klara av en mer analytisk hantering av det språkliga inflödet. Då detta inte kan förklara liknande mönster i kognitivt mer utvecklade, äldre inlärares andraspråksutveckling, antar författarna, liksom Skehan (1998), att det är kraven på en någorlunda smidig on-linekommunikation som motiverar inlärare till att använda komplexa flerordsenheter innan de har förmågan att analysera deras inre struktur (Wray & Perkins 2000:19ff, se även Wiktorsson 2003:49). Wray (2002:ix) konstaterar att samtidigt som såväl första- som andraspråksinlärare i hög grad förlitar sig på flerordsenheter i början av sin språkutveckling, verkar en idiomatisk användning av konventionaliserade flerordsuttryck vara en av de största utmaningarna på vägen mot en inföddlik språkanvändning (jfr Wray & Perkins 2000:2). I det sammanhanget ska också nämnas att utvecklingen av förstaspråksinlärares förmåga att använda konventionaliserade strukturer på ett idiomatiskt sätt, enligt Wray & Perkins modell, inte avslutas förrän i artonårsåldern (2000:19ff). För föreliggande undersökning innebär detta alltså att även förstaspråkssevernans förmåga att använda konventionaliserade språkliga strukturer på ett idiomatiskt sätt fortfarande befinner sig under utveckling.

Wray (2002, 2008) argumenterar för att det mentala lexikonets innehåll bestäms av det hon kallar för *Needs Only Analysis* (NOA):

The process of analysis which the native speaker child engages in is not that of breaking down as much linguistic material as possible into its smallest components. Rather, nothing is broken down unless there is a specific reason (Wray 2008:17).

Hon betonar vidare att både andraspråksinlärares behov och kognitiva möjligheter skiljer sig från förstaspråksinlärares ifråga om i vilken mån språklig inflöde, och därmed även konventionaliserade flerordsuttryck, analyseras. Hennes hypotes är att sena andraspråksinlärare överlag tende-

rar att i högre grad analysera inlärd uttryck och att lagra mindre lexikala enheter i minnet, vilket innebär ett större behov av regelanvändning (Wray 2008:20). Wray menar att detta kan vara en förklaring till bristen på idiomicitet i många vuxna inlärares språkbruk (jfr Pawley & Syder 1983, Ekberg 1998, 1999).

This proposal offers an explanation for the difficulties that are routinely reported for adult learners in relation to restricting their output to only the idiomatic patterns that they have encountered [jfr Bybee 1998, 2006] in input – that is, there is a tendency to express messages in alternative ways, which are meaningful but not nativelike (Wray 2008:20).

Frågan om flerordsenheters representation i det mentala lexikonet diskuteras vidare i avsnitt 4.3 nedan.

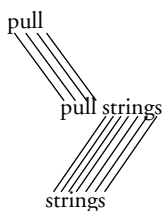
4.3. Mental representation av konventionaliserade uttryck

Som tidigare nämnts, diskuterar också Bybee (bl.a. 1998, 2006) frågan på vilket sätt konventionaliserade flerordsenheter är representerade i språkbrukarnas lexikon och hur denna representation utvecklas och påverkas. Som antyds i citatet i avsnitt 3 ovan, innebär Bybees antaganden om grammatik, som inkluderar lexikon, att konventionaliserade flerordsuttryck förs in i minnet och lagras där som lexikala enheter. Denna lexikala representation av idiom och andra konventionaliserade flerordsuttryck är nödvändig på grund av de oförutsägbara aspekterna av deras betydelser (Bybee 2006:713), dvs. på grund av deras brist på kompositionalitet. Hon betonar dock samtidigt att detta inte innebär att dessa typer av uttryck inte är analyserbara för språkbrukaren:

...they are not completely isolated from related words and constructions since many aspects of their meaning and form derive from more general constructions and the meaning of the component words in other contexts (Bybee 2006:713).

Bybee antar alltså att de konventionaliserade uttryck som inlärares ”erfar” lagras som enheter i minnet, samtidigt som de är länkade till de en-

skilda orden som utgör deras lexikala komponenter och andra relaterade lexikala enheter som innehåller delvis samma lexikala, semantiska eller grammatikaliska komponenter (1998:425, 2006:713). Detta innebär alltså att det bildas ett nätverk som integrerar större enheter och deras enskilda komponenter på ett sätt som ger språkbrukaren tillgång till en uppsättning färdigbildade enheter som, inom samma system, kan analyseras i sina beståndsdelar. Bybee illustrerar relationerna mellan idiomatiska uttryck och deras lexikala komponenter som i figur 1 nedan.



FIGUR 1: Exempel på relationer mellan ett idiom och dess lexikala komponenter (Bybee 1998:425, 2006:713)

Att uttala sig om hur lexikala enheter lagras och organiseras i det mentala lexikonet är i samband med föreliggande undersökning ett i hög grad spekulativt företagande. Det finns dessutom en hel del överlappningar mellan de olika modeller som presenterats ovan, samtidigt som de skiljer sig ifråga om vissa grundläggande teoretiska antaganden. Ett antagande som ligger till grund för denna undersökning är att kompositionalitet och därmed gränsen mellan olika typer av konventionaliserade uttryck, är ett graduellt fenomen. Detta förutsätter en kognitiv representation av både enskilda ord (mindre enheter) och större enheter som är sammansatta av dessa enskilda ord i ett slags integrerat system, där de större enheterna är mer eller mindre kompositionella, genomskinliga och fixerade. Med andra ord är Bybees nätverksmodell lättare att förena med övriga grundläggande antaganden som görs i samband med denna undersökning än en modell som ser *exemplarrepresentation* (eng. *exemplar representation*, se Skehan 1998:60ff) och regel användning som helt åtskilda fenomen. Dessutom säger varken Wray &

Perkins (2000) eller Skehan (1998) mycket om hur interaktionen mellan de olika mekanismerna eller de olika systemen skulle kunna se ut.

En viktig faktor för att ett uttryck ska vara representerat i lexikonet som en enhet är, enligt Bybee, frekvensen. Denna påverkar representationen på ett avgörande sätt. Med andra ord har inte alla enheter som är representerade i nätverket samma status: "those that are used frequently have very strong representations, or possibly even multiple representations, but those that are used infrequently, or are potential but have not been used, may have no representation" (1998:423). Hur nätverket ser ut och i vilken mån ett uttryck som lagras som en helhet kan länkas till olika enheter och därmed bli analyserbart beror alltså på inlärares tidigare språkliga erfarenheter, dvs. på i vilken mån dessa enheter i sin tur finns lagrade i minnet.

Liknande antaganden görs av Wray (2008) som argumenterar för ett heteromorft lexikon, där både holistiska fraser och olika typer av mindre komponenter lagras. Även hennes modell kan sägas vara *usage-based*:

The determination of which components are separately stored and, thus, how much flexibility there is for varying the form of the complete string, depends on input evidence and user need (Wray 2008:15).

Wray beskriver flerordsenheter som *morphem equivalent units* (MEUs), men föreslår samtidigt att många av dessa enheter inte är lagrade i helt fixerad form utan som "partly-fixed frames", vilket, enligt henne, är en förutsättning för en effektiv integration av det hon kallar för *formulaicity* och kreativ regel användning (2008:16).

Vad gäller den lexikala representationen av två eller flera lexikon hos flerspråkiga individer är forskarna inte helt eniga i frågan om lexikala enheter representeras i ett integrerat eller två oberoende system (Bialystok 2001:102f). Bialystok påpekar att forskningen har gett en del indikationer på att ordförråden är oberoende av varandra men refererar samtidigt forskning som har tagit fram en, som hon anser, trovärdig modell för ett integrerat flerspråkigt ordförråd (Dijkstra & van Heuven 1998). Bialystok påpekar att forskarna däremot är eniga om att båda ordförråden är aktiva när en tvåspråkig individ använder ett av sina språk (2001:102). Ett annat problem i sammanhanget är relationen mellan lexikala enheter på de olika språken och deras betydelser och referenter i den reella

världen. Bialystock betonar att enkla dikotomier som den mellan oberoende och integrerade ordförråd eller den mellan direktlänkar från begrepp till ord i andraspråket å ena sidan och vägen via förstaspråket (jfr Nation 2001) å andra sidan inte styrks av empirin. Det flerspråkiga mentala lexikonet kan i stället antas vara mer komplext.

Adequate descriptions of the organization of mental representations include the effect of such factors as level of proficiency and the circumstances of second-language learning. Moreover it appears likely that multiple arrangements can even coexist in the mind of an individual speaker... (Bialystock 2001:103).

Bialystock utgår vidare ifrån att flerspråkig lexikal utveckling hos barn involverar en lika komplex mental representation som hos vuxna andraspråksinlärare. Språk och begrepp antas vara sammanlänkande på ett dynamiskt sätt och strukturen ses som öppen för förändring i samband med ändrade språkliga behov och ökad språkfärdighet (jfr Wray 2008). Frågan är naturligtvis intressant med tanke på andraspråkinformanterna i undersökningen, vilket är anledningen till att den kommenteras här. Då denna undersökning dock inte behandlar frågor om transfer, lämnas denna diskussion i fortsättningen därhän.

4.4. Konventionaliserade uttryck i andra- och främmandespråksinlärning

Lindberg (2004:69) skriver att det verkar råda stor överensstämmelse mellan infödda talare och mycket avancerade andraspråkstalare av ett språk, när det gäller att bedöma om en ordförbindelse är idiomatisk eller inte. Däremot är andraspråkstalare på mindre avancerad nivå i betydligt mindre grad överens (med infödda talare såväl som sinsemellan) i sina bedömningar av vilka ordförbindelser som känns naturliga i andraspråket. Van Lier (1995) demonstrerar detta när han ger läsaren i uppgift att para ihop de för engelskan mest idiomatiska ordförbindelserna genom att välja från en lista med (förstärkande) adverb och en lista med adjektiv (jfr Granger 1998). Enligt Van Lier blir resultatet kollokationer av typen *solidly built* och *carefully crafted*, vilket han alltså menar att de

allra flesta infödda talare av engelska skulle hålla med om, medan även relativt avancerade andraspråkinlärare uppvisar betydligt större variation i sina svar (jfr Ekberg 1999). Lindberg ger svenska exempel av typen *osedvanligt korkad* och *komplett galen* (2004:81). Denna övning är väl lämpad att illustrera idiomicitetsbegreppet; det finns inget i det grammatiska regelsystemet eller i ordens semantik som hindrar en infödd talare av engelska att t.ex. välja ordkombinationen *solidly crafted* istället. En förstaspråkstalarare av svenska skulle på samma sätt kunna tänkas bilda sekvenserna *osedvanligt galen* och *komplett korkad* utifrån Lindbergs lista. Det verkar dock finnas en samstämmighet, som i första hand verkar råda bland infödda talare inom en språkgemenskap, om vilka kollokationer som är de mest etablerade, de mest konventionaliserade och därmed de mest idiomatiska i det aktuella språket (jfr Wiktorsson 2003:56). Varken van Lier eller Lindberg baserar sina resonemang på empiriska studier, vilket gör det desto mer intressant att testa hypotesen om en tydlig skillnad mellan första- och andraspråksanvändares resultat när uppgiften är att sätta ihop de mest idiomatiska kollokationerna utifrån en uppsättning givna ord. Denna hypotes kan ses som utgångspunkt för kollokationstestet i föreliggande undersökning (se vidare avsnitt 5.4).

4.4.1. Studier inom andraspråksforskning

Abrahamsson & Hyltenstam (2009) har undersökt behärsknigen av idiom och ordspråk hos andraspråksinlärare på inföddlik nivå. Undersökningen ingår i en större studie om startålder och *uppfattad inföddhet*. Abrahamsson & Hyltenstam använder sammanlagt tio olika testinstrument för att undersöka i vilken mån andraspråksinlärare, som till vardags passerar som infödda talare av svenska, i själva verket behärskar språket på samma nivå som infödda talare. Förutom behärskning av idiom och ordspråk testas informanterna på bl.a. grammatikalitetsbedömning och perception av tal i brus. När det gäller idiom och ordspråk består metoden av två separata datoriserade test (innehållande 50 idiom respektive 50 ordspråk), där informanterna ställs inför uppgiften att inom 10 sekunder muntligt komplettera ett idiom respektive ordspråk som presenteras ofullständigt på datorskärmen. Exempel från testen är *Hon löpte verkligen [linan] ut* och *Ju fler kockar [desto sämre soppa]* (or-

den inom hakparentes är de som utelämnats på skärmen). I analysen jämförs de avancerade inlärnarnas svar med resultaten från en infödd kontrollgrupp. Resultaten för idiomtestet visar tydliga skillnader mellan informanter med olika startålder för andraspråkinläringen. 58 % av informanterna med startålder 1–11 år når samma eller högre resultat än de lägst presterande infödda informanterna, men bara 20 % av informanterna med startålder 13–19 år når upp till dessa resultat. I ordspråks-testet var motsvarande resultat 16 % och 10 % (Abrahamsson & Hyltenstam 2009:279ff).

Ekberg har studerat förhållandet mellan regeltillämpning och lexikonkunskap hos enspråkiga och flerspråkiga elever i årskurserna 5, 6, 7 och 9 i Rosengård (1997, 1998, 1999, se även 2004). En del av hennes arbete består i att kartlägga frekventa syntaktiska och lexikala strukturer som enspråkiga svenska elever använder sig av i en viss språksituation (återgivning av bildberättelser), varefter flerspråkiga elevers bruk av konventionaliserade strukturer i samma språksituation har undersökts (1998:260f). Utöver detta har Ekberg i motsvarande grupper bl.a. undersökt användningen av *komplexa predikat* som pseudosamordningar – t.ex. *hansitter och läser* – samt partikelverb, reflexiva verb och flerordskombinationer som är lexikaliserade på frasnivå – t.ex. nominalfraser som *Vita huset* eller verbfraser som *spela kort* (Ekberg 1998:249f, se även 1997:97ff).

Studierna visar att både enspråkiga och flerspråkiga elever har aktiv kunskap om de undersökta konventionaliserade lexikala och grammatiska konstruktionerna. Skillnaderna mellan de båda grupperna gäller i första hand frekvens och variation i användningen av konventionaliserade strukturer. De enspråkiga eleverna använder fler konventionaliserade strukturer på alla språkliga nivåer (Ekberg 1998:261, se även 1997:106ff). Dessutom är variationen i användningen av dessa frasliknande strukturer större bland de flerspråkiga eleverna, vilket innebär att deras språkbruk i den givna talsituationen är mindre stereotyp, och i många fall även mindre idiomatiskt, än de enspråkiga elevernas. Resultaten kan, enligt Ekberg, vara ett tecken på att de flerspråkiga barnen har en mindre uppsättning färdiga konstruktioner att välja bland, vilket i sin tur kan vara en följd av en relativt låg grad av kontakt med målspråket (1999:94, se även Prentice & Sköldberg und. utg.).

Vad gäller svensk kollokationsforskning på andraspråksområdet har Enström studerat andraspråksinlärares verbanvändning. Hennes material består av uppsatser i svenska som andraspråk, skrivna av gymnasieelever med olika modersmål och varierande språkbehärskning (1996: 108f). Enström har bl.a. beskrivit ”de felanvändningar av verb som i första hand [...] rör rent syntagmatiska relationer” (1996:130) och kategoriserat dessa som *kombinatoriska fel* inom ramen för det hon kallar för *lexikalisk-semantisk felkategorisering*. Denna kategorisering innebär en primär uppdelning i *konstruktionsfel* och *primärt semantiska fel*, som sedan indelas vidare i ett antal subkategorier (Enström 1996:108ff). De kombinatoriska felen (som alltså tillhör kategorin konstruktionsfel) innefattar bl.a. det som Enström kallar för *lexikaliska verbkollokationer*. Hon påpekar att de ordkombinationer i hennes material som ”kategoriserats som lexikaliska verbkollokationer utmärks av att de ingående ordens betydelser är förenliga, varför felen inte kan ses som brott mot gällande selektionsrestriktioner utan som att ordkombinationen inte är idiomatisk” (1996:149). Som exempel på icke idiomatiska användningar av verbkollokationer ges bl.a. följande meningar (Enström 1996: 150ff):

Lärarna måste **ge** större krav på **läxorna**.
För att motverka detta måste vi **göra** större **åtgärder**.
Barnen hade redan gjort snö-gubbar på vår gård och **lekte** **snöboll**

Enström beskriver också kombinatoriska fel i samband med det hon kallar för *lexikaliserade verbalfraser*. Hon ger exempel på en rad konventionaliserade uttryck som vållar problem för informanterna i hennes undersökning. De flesta exempel från hennes material består av *kontaminationer*, dvs. en sammanblandning av två eller flera uttryck, som i

Om man ligger riktigt stilla och **inte gör ett väsen ifrån sig**, då kan man höra hur fåglarna kvittrar,

som, enligt Enström, kan antas vara en kontamination av *inte ge ett ljud ifrån sig* och *inte göra något väsen av sig* (1996:158f).

Även Hyltenstam (1992) har beskrivit avvikelser vid användning av konventionaliserade uttryck hos flerspråkiga och enspråkiga elever. Han framhäver *approximationer* (1992:360) till den konventionaliserade for-

men (som i ”*den fungerar nog inte på längden*”) och kontaminationer av element från två eller flera olika uttryck (som i ”*man beräknar med att...*”) som vanliga feltyper i elevernas muntliga och skriftliga produktion. Analysen visar att approximationer är betydligt vanligare i de flerspråkiga informanternas språkbruk, medan kontaminationer visar sig vara lika frekventa bland enspråkiga elever som i den flerspråkiga gruppen (Hyltenstam 1992:360ff, Große¹ 2008, Prentice & Sköldberg und. utg.). Enströms (1996) och Hyltenstams (1992) beskrivning av avvikelser från konventionaliserade mönster vid andraspråksinlärares användning av konventionaliserade uttryck är relevant för den kvalitativa analysen i samband med min undersökning.

I en studie av receptiv behärskning av konventionaliserade uttryck i *English for Academic Purposes* (EAP) har Ellis, Simpson-Vlach & Maynard (2008) i ett antal experimentella undersökningar visat att det är olika faktorer för första- och avancerade andraspråkstalare av engelska som påverkar hur snabbt de klarar av att processa en fraseologisk enhet. Testresultaten visar att processningstiden för förstaspråksanvändare i första hand påverkas av den aktuella fraseologiska enhetens *MI-värde* (Mutual Information) (se avsnitt 2.4.1) och inte av dess frekvens i inflödet (ibid.). Andraspråksanvändarnas reaktionstider påverkas däremot i första hand av den fraseologiska enhetens frekvens (Ellis, Simpson-Vlach & Maynard 2008:382ff). Den slutsats som författarna drar av sina resultat kan betraktas som högst relevant för forskningen om lexikal kompetens hos avancerade andraspråksinlärare såväl som för andraspråksundervisning:

...native speakers and advanced ESL learners have become sensitive from their usage histories to these [formulaic] expressions so that they process them preferentially. But native speakers and learners are sensitive to different determinants of fluency – learners to [...] frequency, fluent natives to MI (Ellis, Simpson-Vlach & Maynard 2008:389).

En av de avgörande faktorerna för *slutlig nivå* (eng. *ultimate attainment*) i andraspråksinläring, har visat sig vara startåldern för inläring (Birdsong 1999, Hyltenstam & Abrahamsson 2000, Abrahamsson & Hyl-

¹ Författaren heter numera Julia Prentice

tenstam 2004, 2009). Det finns många kontroverser inom forskningen om startålder och olika aspekter av inlärares språkbehärskning. Dessa handlar bl.a. om att forskare sätter olika åldersgränser och utgår ifrån olika modeller för biologiska mognadsbegränsningar i samband med studier av inlärares möjligheter att nå en inföddlik slutnivå i andraspråket (Hyltenstam & Abrahamsson 2000, Munro & Mann 2005). Denna diskussion kommer inte att fördjupas här. Faktum är dock att man i många studier, liksom även i denna, delar in informanter efter startålder för sin andraspråksinläring (eng. *age of onset*, AO eller *age of immersion*, AOI). Vilka åldersgränser man utgår från vid denna indelning varierar däremot en hel. I den ursprungliga versionen av det som kallas för *the critical period hypothesis* (CPH), dvs. en kritisk period för språkinläring (Lenneberg 1967), sätts gränsen för påbörjande av språkinläring som kan leda upp till infödd slutnivå vid puberteten. Senare studier tyder däremot på att möjligheterna att uppnå en infödd slutnivå på andraspråket avtar successivt långt före puberteten (Hyltenstam & Abrahamsson 2000, jfr Pallier., Dehaene, Poline, LeBihan, Argentié, Dupoux & Mehler 2003:160). Abrahamsson & Hyltenstam (2009) väljer en indelning i fem grupper, baserad på AO: *tidig barndom* (AO ≤ 5 år), *sen barndom* (AO 6–11 år), *ungdom* (AO 12–17 år), *tidigt vuxenliv* (AO 18–23) och *senare vuxenliv* (AO ≥ 24 år). Indelningen kan, enligt författarna, delvis motiveras med utgångspunkt från olika språkutvecklingsfaser, samtidigt som den återspeglar mönster som träder fram i deras analys av sambandet mellan startålder för andraspråksinläring och i vilken mån en inlärare uppfattas som infödd talare (Abrahamsson & Hyltenstam 2009:268). Butler (2000) delar i sin studie av andraspråksbehärskning hos kinesiska andraspråkstalare av engelska in informanterna i tre startåldersgrupper samt en grupp med infödda talare. Grupperna består av informanterna med startålder, 3–5 år, 5–10 år samt 10–15 år (Abrahamsson & Hyltenstam 2004). Den indelning som har valts i samband med denna studie kan i stort sett sägas motsvara indelningarna i de två sistnämnda studierna (se vidare i avsnitt 5.3)

4.4.2. Korpuslingvistiska studier inom främmandespråksinlärning

Betydelsen av konventionaliserade uttryck för andra- och främmandespråksinlärning har också framhävts ur ett kontrastivt perspektiv i en rad korpuslingvistiska undersökningar. Ett relativt stort antal korpusbaserade studier har under de senaste två decennierna visat att även inlärare på avancerad nivå underanvänder idiom och andra konventionaliserade uttryck i jämförelse med infödda talare (jfr O’Keefe, McCarthy & Carter 2007). I det sammanhanget har bl.a. inlärarkorpusen ICLE, the International Corpus of Learner English, spelat en framträdande roll. Korpusprojektet lanserades 1990 av Granger, med målet att bygga upp en databas med texter skrivna i engelska som främmandespråk av inlärare med olika modersmål. Materialet tjänar som underlag för kontrastiva studier i större skala (se Granger 1993, 1996, 1998, 2003, Granger, Hung & Petch Tyson 2002). Granger själv har bl.a. studerat användningen av kollokationer med förstärkande adjektiv som *bas* (t.ex. *perfectly natural* eller *deeply in love*) hos fransktalande inlärare av engelska (1998:146f). Hennes resultat tyder på en tydlig underanvändning av det hon kallar för *native-like collocations* bland inlärarna. Granger förklarar sina resultat dels med interferens från inlärarnas L1, då de idiomatiska kollokationer som inlärarna använder ofta har en direkt ekvivalent i deras modersmål (1998:151), dels med inlärarnas outvecklade känsla för vilka ordkombinationer som är mer framträdande (eller mer konventionaliserade) än andra i målspråket (Granger 1998:152). Som tidigare nämnts, har Wiktorsson (2003) studerat användningen av konventionaliserade flerordsuttryck (eng. *prefabs*) i svenska inlärares skriftliga produktion i engelska. Även hennes arbete baseras på ICLE, närmare bestämt den svenska subkorpusen SWICLE. Hon jämför data från två olika grupper av inlärare – gymnasieelever och universitetsstudenter i engelska – med infödda skribenters texter. Resultaten visar att skillnaden i användningen av konventionaliserade uttryck mellan den mer avancerade inlärargruppen och de infödda skribenterna snarare är kvalitativ än kvantitativ, dvs. de avancerade inlärarna (svenska universitetsstudenter) använder inte färre, utan andra typer av konventionaliserade uttryck. Den huvudsakliga skillnaden tycks ligga i att inlärarna använder fraser som är utmärkande för ett mer informellt språkbruk, vilket gör att

de i det skriftspråkliga sammanhanget framstår som mindre idiomatiska (2003:108f). Mellan de mindre avancerade inlärarna (svenska gymnasielever) och de infödda skribenterna har Wiktorsson däremot funnit en frekvensskillnad i användningen av konventionaliserade uttryck. Inlärarnas texter innehåller färre konventionaliserade uttryck än de inföddas uppsatser (jfr Ekberg 1997, 1998, 1999). Den kvalitativa skillnaden visar sig vara större mellan de mindre avancerade inlärarna och de infödda skribenterna än mellan de senare och den mer avancerade inlärarygruppen. Med andra ord använder de mindre avancerade inlärarna fler uttryck i sitt skriftspråk som bland infödda talare är vanligare i talspråk än de avancerade inlärarna. Dessutom uppvisar gymnasieelevernas texter en högre grad av interferens från förstaspråket än de mer avancerade inlärarnas texter (Wiktorsson 2003:145ff)

4.5. Konventionaliserade uttryck i utbildningsrelaterat språk

O’Keefe, McCarthy & Carter(2007:92) poängterar att konventionaliserade uttryck, liksom flerordsenheter generellt, kan kopplas till olika register och hänvisar bl.a. till forskning om talat språk i akademiska sammanhang som har visat att idiom utgör ”a not-insignificant feature of the lexical landscape of academic speech” (Simpson & Mendis 2003: 427). Som nämndes inledningsvis är vissa typer av konventionaliserade uttryck inte bara vanligt förekommande, utan även högst relevanta ur ett funktionellt perspektiv i det som brukar kallas för akademiskt språk.

Cognitive science demonstrates that knowledge of these formulas is crucial for fluent processing. And current trends in SLA [Second Language Acquisition] and EAP [English for Academic Purposes] demand ecologically valid instruction that identifies and prioritizes the most important formulas in different genres (Ellis, Simpson-Vlach & Maynard 2008:379).

Som antytts tidigare, belyser Ellis, Simpson-Vlach & Maynard frågan om receptiv behärskning av konventionaliserade uttryck i EAP från såväl ett korpuslingvistiskt som ett psykolingvistiskt och didaktiskt perspektiv (2008:375ff). Författarna har med hjälp av korpuslingvistiska

metoder tagit fram fraseologiska enheter som är signifikant mer frekventa i akademisk engelska än i icke-akademiska diskurser (2008:379) och som därmed också antas vara relevanta för undervisning i EAP. För att bekräfta det senare har Ellis, Simpson-Vlach & Maynard, utifrån detta material, låtit ett antal språktexperter och lärare i EAP göra en av följande bedömningar: (1) om de ansåg ordkombinationerna ifråga utgöra fasta fraser, konventionaliserade uttryck eller flerordsenheter (eng. *chunks*), (2) om de ansåg fraserna ha en betydelse och funktion som en enhet (som inte motsvarar de enskilda konstituenternas betydelse och funktion), eller (3) om de ansåg fraserna vara värda att läras ut som idiomatiska uttryck (Ellis, Simpson-Vlach & Maynard 2008:381). Resultatet visar att de enheter som bedöms som mest konventionaliserade och värda att lära ut var de mest frekventa, och de med de högsta värdena för Mutual Information (MI) i korpusen. Lärarnas bedömningar kan alltså på så sätt anses ha bekräftat den korpuslingvistiska analysen.

5. Material och metod

Materialet för denna undersökning har samlats in med hjälp av en enkät som delats ut till sammanlagt 10 gymnasieklasser i olika årskurser, på två olika gymnasieskolor. Materialet samlades in under vårterminen 2009. Sammanlagt har 219 elever deltagit i enkätundersökningen. 212 elever har inkluderats i analysen (sju enkäter har alltså av olika anledningar exkluderats). 186 av dessa elever kan sägas studera i en språkligt heterogen klassmiljö. De övriga 26 eleverna tillhör en gymnasieklass (årskurs 1) med enbart elever med helsvensk bakgrund som fungerar som jämförelseklass. Resultaten från den enspråkiga klassmiljön används alltså i första hand som referenspunkt i analysen (se vidare i 5.2).

5.1. Etnisk och social bakgrund som påverkande faktorer i samband med elevers skolresultat

Det är knappast förvånande att vistelsetiden i målspråkslandet är en faktor som påverkar andraspråkelevernas prestationer i skolan. Thomas & Collier (2002) har visat att det tar mellan fem och tio år tid för invandrarelever att nå en nivå i sin skolrelaterade andraspråksbehärskning som motsvarar infödda elevers behärskning av skolrelaterat språk (Lindberg 2007:82, jfr Viberg 1993). Skolverket visar i en rapport att varje år senare som en elev anlände i grundskolan i genomsnitt motsvarar ca fem poäng lägre meritvärde i årskurs 9 (Skolverket 2004:47).

Med tanke på att den forskning som har bedrivits sedan mitten på 1960-talet på detta område (t.ex. Coleman 1966) har visat tydliga samband mellan elevers socioekonomiska status och deras skolresultat, kan man anta att även resultaten i min undersökning i viss mån kan påverkas av socioekonomiska faktorer (se vidare i avsnitt 5.2 nedan). Här finns dock ett antal metodologiska problem som gör att en del tidigare forskningsresultat är svåra att tolka. Det har bl.a. att göra med att socioekonomisk status är ett flerdimensionellt begrepp som innefattar bl.a. ekonomiska, sociala och kulturella faktorer. Dessutom är forskningsresultaten också avhängiga de typer av skolresultat som fokuserats, t.ex. vilka ämnen som undersökts och huruvida man tittat på individuella resultat eller resultat på gruppnivå (Skolverket 2009:98f). I senare studier har man dock tagit hänsyn till dessa definitionsproblem. Bl.a. har man börjat skilja mellan ekonomiskt kapital (föräldrarnas inkomst) och kulturellt kapital (som, enligt skolverket, innebär föräldrarnas utbildningsbakgrund) och kommit fram till att det i första hand är det senare som uppvisar tydliga samband med elevers resultat i skolan (t.ex. Gustafsson 1997, Yang & Gustafsson 2003, se Skolverket 2009:100). Även boendesegregationen har pekats ut som en faktor som påverkar skolresultaten negativt för elever med utländsk bakgrund i segregerade områden. Här handlar det ofta också om andra samverkande faktorer som t.ex. ekonomiska problem, etnisk minoritetsstatus, bristfällig behärskning av majoritetsspråket och föräldrarnas utbildningsnivå liksom även skolrelaterade faktorer som skolresurser och lärarkompetens (Skolverket 2004: 47, 2009:96). Elever med utländsk bakgrund är också överrepresenterade bland elever med låg socioekonomisk status och i skolor med låg andel elever med utländsk bakgrund är familjernas socioekonomiska status således ofta högre i termer av ekonomiska resurser, föräldrarnas utbildningsbakgrund och arbetsmarknadsanknytning (Skolverket 2009: 50). Från Skolverkets sida har man tolkat detta som att det i första hand är socioekonomiska faktorer, snarare än språkliga och andra faktorer i samband med elevernas härkomst, som ligger bakom skillnaderna i skolresultat mellan infödda elever och de med utländsk bakgrund (2005: 10). Denna tolkning kan ses som delvis motsägelsefull i relation till den tidigare nämnda rapporten om ankomstålder och skolresultat (2004:47). Lindberg påpekar också problem med att basera prognoser om skolframgång på uppgifter om föräldrarnas utbildningsbakgrund när det

gäller elever med utländsk bakgrund och understryker att skillnaderna i prestation mellan elever som anländer efter skolstarten och infödda elever kvarstår även efter att de socioekonomiska faktorerna har tagits med i beräkningen (Lindberg 2007:78f).

Som framgår ovan är sambanden mellan elevernas sociala bakgrund och deras prestationer i skolan (eller i undersökningar som denna) tämligen komplexa och därför svårbeforskade. Jag har därför i denna undersökning avstått från att relatera resultaten till uppgifter om de individuella elevernas bakgrund i detta avseende och valt att endast kontrollera för deltagarnas språkliga bakgrund. Däremot ges en kort beskrivning av skillnader i skolornas elevsammansättning, bl.a. även med avseende på etnisk och social bakgrund, i nästa avsnitt.

5.2. Skolor och elevsammansättning i deltagarklasserna

De två gymnasieskolorna som har deltagit i undersökningen skiljer sig på flera olika sätt från varandra, vilket man bör beakta vid tolkning av resultaten. Det ena gymnasiet, hädanefter kallat Prismagymnasiet (P), ligger i en förortskommun till Göteborg. Den andra skolan, hädanefter kallad Kampusgymnasiet (K), är en innerstadsskola i Göteborg. Även om det inte längre finns fasta upptagningsområden för skolorna, är det inte orimligt att anta att skolans geografiska läge spelar viss roll för elevsammansättningen. Man kan t.ex. tänka sig att en skola som ligger centralt drar till sig elever från flera olika områden i staden, vilket resulterar i en mer heterogen elevgrupp än i en skola som Prismagymnasiet, som ligger i en förortskommun. Som framgår av tabell 2 har Prismagymnasiet också betydligt lägre andel elever med utländsk bakgrund än Kampusgymnasiet på de aktuella gymnasieprogrammen. Detta gäller även om man tittar på skolorna som helhet. Andelen elever med utländsk bakgrund på Kampusgymnasiet är 40 %, medan endast 17 % av eleverna på Prismagymnasiet har utländsk bakgrund (skolverket 2010). Med anknytning till diskussionen kring socioekonomisk bakgrund i föregående avsnitt kan man alltså anta att eleverna på Kampusgymnasiet i högre uträkning kommer från familjer med relativt låg socioekonomisk

TABELL 2: Bakgrundsdata för deltagarklasserna

	KLASS 1	KLASS 2	KLASS 3	KLASS 4	KLASS 5	KLASS 6	KLASS 7	KLASS 8	KLASS 9	Jämförelse- klass
Skola	Prisma	Kampus	Kampus	Kampus	Prisma	Kampus	Kampus	Kampus	Kampus	Prisma
Program och årskurs	SP 1	SP 1	OP 1	SP 1	SP 2	SP 2	OP 2	OP/ NP 2	OP 3	SP/ NP 1
Andel elever med utländsk bakgrund på gymnasieprogram	27 %	51 %	42 %	51 %	27 %	51 %	42 %	62 %/ 42 %	42 %	27 %/ 28 %
Antal elever	28	19	14	24	25	22	16	22	23	26
Antal (andel) flickor	14 (50 %)	11 (58 %)	12 (86 %)	16 (67 %)	10 (40 %)	15 (68 %)	13 (81 %)	16 (73 %)	17 (74 %)	17 (65 %)
Antal (andel) elever med svenska som L2	11 (39 %)	5 (26 %)	14 (100 %)	24 (100 %)	7 (28 %)	14 (64 %)	2 (13 %)	22 (100 %)	14 (61 %)	0
Typ av svensk- undervisning	sv	sv	sva	sva	sv	sv	sv	sva	19 sv/ 4 sva	sv

status än eleverna på Prismagymnasiet. Liknande förhållanden kan tänkas råda när det gäller andelen elever som bor i så kallade segregerade områden, som alltså torde vara högre på Kampusgymnasiet (se föregående avsnitt, jfr Skolverket 2004:47, 2009:96).

Tabell 2 sammanfattar bakgrundsdata för alla de tio gymnasieklasser som deltagit i undersökningen. Bl.a. framgår det att flickor är i majoritet i alla klasser förutom i klasserna 1 och 5. Dessutom visar tabellen hur många andraspråkselever som det finns i varje klass, om klassen läser svenska eller svenska som andraspråk och vilket gymnasieprogram eleverna går på. Skillnaderna mellan olika gymnasieprogram är också en faktor som i viss mån bör beaktas vid tolkning av undersökningens resultat. Informantgruppen består av elever som läser på omvårdnadsprogrammet (OP), samhällsvetenskapliga programmet (SP) och naturvetenskapliga programmet (NP), vilket alltså innebär att informantgruppen är sammansatt av både elever från i huvudsak yrkesförberedande program, liksom från i huvudsak studieförberedande utbildningar (jfr Skolverket 2008:18). Enligt Skolverket (2008:35) når elever som går på studieförberedande utbildningarna i högre grad behörighetskraven. Man betonar dock att skillnaderna mellan olika program är relativt stora. Skillnaderna i studieresultat mellan de här aktuella programmen i årskurs 3 för läsåret 2006/07 (för hela landet) återges i tabell 3. Dessutom anges hur många elever från varje program som ingår i informantgruppen som helhet.

TABELL 3: Studieresultat för relevanta utbildningar i årskurs 3 för hela landet i procent (se Skolverket 2008:25)

Program	Antal elever från respektive program i informantgruppen	Endast samlat betygsdokument (ej slutbetyg)	Slutbetyg men ej behörig	Slutbetyg med grundläggande behörighet
OP	59	ca 21 %	ca 10%	ca 69 %
SP	112	ca 13 %	ca 8 %	ca 79 %
NP	15	ca 8 %	ca 4 %	ca 88 %

Som nämndes i avsnitt 5.1 är det naturligtvis en rad olika faktorer som ligger till grund för dessa skillnader mellan olika gymnasieutbildningars resultat. Sambandet mellan elevernas studieresultat och resultaten av

denna undersökning är heller inte i fokus för detta arbete. Statistiken som redovisas i tabell 3 kan dock antas ha viss betydelse för vilka slutsatser som kan dras av analysen.

5.3. Informantgrupperna

Informanterna har delats in i tre huvudgrupper, efter startålder för sin inläring av svenska. Indelningen visas i tabell 4. Grupperna är grupp *L1* (elever som har angett svenska som sitt förstaspråk och därmed antas att ha börjat inläringen vid födseln), *L2/tidig* (elever med startålder ≤ 6) samt *L2/sen* (informanter med startålder > 6). Den sistnämnda gruppen har i sin tur delats in i två undergrupper, nämligen informanter med en startålder till och med 10 år och informanter med en startålder efter 10 år¹. Indelningen motiveras dels av språkutvecklingsmönster som visats i tidigare forskning (Abrahamsson & Hyltenstam 2009, Butler 2000), dels av ålder för skolstarten i Sverige; informanterna i grupp *L2/tidig* har påbörjat sin inläring av svenska före skolstarten. De har alltså påbörjat sin andraspråkutveckling under den tidiga barndomen och därutöver fått hela sin skolgång på andraspråket. Grupp *L2/sen* består av informanter som påbörjat sin andraspråkutveckling efter sex års ålder, dvs. samtidigt med eller efter skolstarten. Detta innebär en relativt sen startålder, med tanke på biologiska mognadsbegränsningar för språkinläring (Hyltenstam & Abrahamsson 2000), och därmed heller inte börjat utveckla andraspråket innan de ställts inför skolans språkliga krav. Den sena inlärargruppen har delats in i två undergrupper för att göra det möjligt att säga något om eventuella skillnader mellan elever som har fått hela, eller nästan hela, sin skolgång på andraspråket, och sådana som först i något högre ålder har kommit i kontakt med både vardaglig och skolrelaterad svenska. Startåldersgränsen för den senare gruppen (> 10) har alltså valts för att den ligger några år efter skolstarten, men samtidigt före inträdet i puberteten, som i tidigare undersökningar

¹ Informanterna med en startålder efter 10 år kan i regel antas att under några år ha gått i skolan på sitt modersmål. Denna faktor har dock inte kontrollerats för i samband med detta arbete.

har visat sig vara en känslig ålder ur språkutvecklingssynpunkt (Abrahamsson & Hyltenstam 2004).

Utöver gruppindelningen visar tabell 4 hur många elever från varje årskurs som ingår i respektive grupp. Som framgår av tabellen är fördelningen av elever från årskurserna 1 och 2 relativt jämn i respektive grupp. Från årskurs 3 finns det däremot bara en klass bland deltagarklasserna och därmed relativt få elever från denna årskurs i varje grupp. Vilken årskurs eleverna går i kan antas ha viss relevans för deras behärskning av konventionaliserat språk då elever som har gått igenom årskurserna 1 och 2 borde ha kommit längre i sin språkutveckling generellt och också ha ett mer utvecklat skolrelaterat språkbruk (jfr avsnitt 4.4.1). Frågan är dock om denna faktor väger tyngre än t.ex. elevernas språkliga bakgrund.

TABELL 4: Antal elever från varje årskurs i grupperna

	GRUPP L1	GRUPP L2/tidig (startålder, ≤ 6)	GRUPP L2/sen			JK
			Informanter med start- ålder ≤10	Informanter med start- ålder >10	Totalt	
ÅK1	28	35	10	8	18	26
ÅK2	38	30	6	9	15	0
ÅK3	8	10	2	2	4	0
Totalt	74	75	18	19	37	26

Utöver de elever som delats in i startåldersgrupper deltog, som tidigare nämnts, även en jämförelseklass (JK) i enkätundersökningen. Som framgår av tabell 4 består denna klass av enbart infödda svenska elever i årskurs 1 på Prismagymnasiet. I denna klass följer elever från såväl naturvetenskapligt som samhällsvetenskapligt program en utbildning med musikinriktning. Klassen har tagits med i undersökningen i syfte att möjliggöra jämförelser mellan resultaten för informantgrupperna där samtliga elever kan sägas studera i en flerspråkig miljö, med resultat för elever som läser i en mer enspråkig skolmiljö.

5.4. Enkäten

Under hösten 2006 genomfördes en pilotundersökning på Kampusgymnasiet för att, utifrån en första version av enkäten, undersöka vilka problem som kan uppstå, dels vid genomförandet av en enkätundersökning som den föreliggande, dels vid analysen av elevernas svar. Pilotundersökningen fokuserade endast förståelsen, dvs. elevernas receptiva behärskning av de konventionaliserade uttrycken ifråga. Analysen av provenkätens genomförande och resultat visade bl.a. att flervalssformatet är att föredra framför öppna frågor av typen ”*inte sticka under stol med betyder: _____*” när det gäller att testa elevernas kunskap om uttrycks betydelse. Ett återkommande problem visade sig nämligen vara att vissa svar på olika sätt kunde ligga nära den konventionaliserade, efterlysta betydelsen, utan att vara helt korrekta. Om informanterna måste formulera betydelsen självständigt blir det alltså svårt att dra nytta av partiell kunskap om ett uttryck (jfr Nation 2001), då det är omöjligt att analysera ett sådant test i termer av ’mer eller mindre korrekt’. Dessutom blev det tydligt under arbetet med pilotundersökningen, att en undersökning av endast receptiv kunskap i samband med ungdomars behärskning av konventionaliserade uttryck skulle bli för endimensionell. Det finns, som nämnts tidigare, många andra aspekter av lexikonkunskap som är viktiga för en idiomatisk användning av dessa typer av uttryck (jfr bl.a. Nation 2001, Henriksen 1999).

Den slutgiltiga versionen av den enkät som informanterna ombads att fylla i, består av fyra olika delar (se appendix 1). Två av dem, A och D, består av flervalstest, medan delarna B och C har formen av lucktest. På enkätens första sida ombeds informanterna att besvara ett antal frågor om sig själva och sin språkanvändning. Informationen används för analyser av enkätundersökningens resultat i relation till olika bakgrundsvariabler som kön, språklig bakgrund och startålder för inläring av svenska.

I A-delen testas elevernas receptiva kunskap om sammanlagt tio olika konventionaliserade uttrycks betydelse. Till varje uttryck finns fyra svarsalternativ, ett korrekt svar, två distraktorer samt möjligheten att avvisa de övriga tre alternativen och själv föreslå uttryckets korrekta betydelse. Det sistnämnda fungerar dock endast som ytterligare distraktor, då den konventionaliserade betydelsen i samtliga fall finns bland de giv-

na alternativen (se appendix 1). Uppgifterna ser exempelvis ut på följande sätt:

Exempel 1:

ha händerna fulla

- vara rik
- ha mycket att göra
- inte kunna röra sig
- Inget av alternativen stämmer, uttrycket betyder följande:

Som nämndes i avsnitt 3.2 beror svårighetsgraden i ett flervalstest till stor del på hur nära distraktorerna ligger det korrekta svaret i betydelse (Nation 2001). Tanken bakom A-delens utformning är att en av distraktorerna ska kunna kopplas till uttryckets, eller någon av konstituenternas, bokstavliga betydelse medan den andra är tänkt att ligga så långt ifrån det korrekta svaret som möjligt, utan att för den delen kunna utslutas direkt. Möjligheten att ge ett eget förslag tjänar alltså dels som ytterligare distraktor, dels till att göra det möjligt för eleverna att visa eventuell partiell kunskap om uttryckets betydelse som kan vara relevant för den kvalitativa analysen av materialet.

B-delen består i ett lucktest eller, som Nation (2001) kallar det, ett *sentence-completion test*. Informanterna ställs inför uppgiften att komplettera ett konventionaliserat uttryck i kontexten av en mening genom att fylla i det ord som saknas. Med Nations modell om lexikonkunskap som utgångspunkt (se avsnitt 3.1 ovan), testas på så sätt dels receptiv kunskap om ords och uttrycks betydelse i kontexten, dels produktiv kunskap om kollokationspotential och uttryckets konventionaliserade form. En viktig aspekt när det gäller B-delens utformning har varit frågan om det finns en skillnad i uppgiftens svårighetsgrad beroende på vilket led i frasen som utelämnas. För att kunna utläsa eventuella mönster i detta sammanhang har i vissa fall substantivet och i andra verbet utelämnats, vilket illustreras i exemplet nedan:

Exempel 2:

I Expedition Robinson måste man jobba för att överleva. Det funkar inte att åka snålskjuts för då blir man utröstad direkt.

Samma sak gäller dokusåpan Farmen. Vill man vara kvar där så måste man dra sitt strå till stacken.

Del C i enkäten kan, med Nations terminologi, betecknas som ett *supply-item test*. Här består uppgiften i att komplettera sammanlagt femton uttryck utanför en kontext på följande sätt:

Exempel 3:

gå ner i varv

Även här testas alltså produktiv kunskap om bl.a. ords kombinationsmöjligheter, men framför allt om konventionalisering; det finns naturligtvis flera ord som kan kombineras med de givna orden i testet, utan att det uppstår rent grammatiska fel eller semantisk oförenlighet. Informanten måste alltså (receptivt) känna igen ordkombinationen som ett fragment av ett konventionaliserat uttryck, samt (produktivt) välja det ord som fattas för att det konventionaliserade uttrycket ifråga ska bli komplett.

Även i enkätens sista deltest (D) ställs informanterna inför uppgiften att fylla i det utelämnade ordet i ett uttryck. Denna gång ges det däremot ett antal alternativa ord att välja bland. Här handlar det, med andra ord, om ett matchningstest. Sammanlagt ska 20 kollokationer bildas. Kollokationerna är indelade i fyra grupper med fem liknande konstruktioner (t.ex. *verb + substantiv* och *adverb + adjektiv*). Till varje grupp tillhandahålls åtta alternativa ord som kan paras ihop med de fem givna orden, av vilka tre alltså enbart tjänar som distraktorer. Testformatet exemplifieras nedan.

Exempel 4:

Fyll i det ord som saknas i följande uttryck. Välj bland följande alternativ:
bilda, dra, föra, hålla, leda, lägga, bryta, ta

föra ett samtal

hålla ett tal

bryta ett beroende

bilda en union

dra en slutsats

Informanterna görs uppmärksamma på att alla ord inte behöver användas. Däremot nämns inte i uppgiften att det kan finnas två likvärdiga alternativ att välja bland orden, vilket dock bara är fallet i samband med en kollokation – nämligen *vara/ komma till nytta* (se appendix 1). Informanterna får heller inte veta om samma ord kan användas i flera av de fem kollokationerna (vilket inte är fallet). Återigen är alltså syftet att testa informanternas produktiva kunskap om ords kollokationspotential och konventionaliserade ordförbindelser. Som antytts tidigare har utgångspunkten för denna del av undersökning varit att testa hypotesen om skillnaden mellan första- och andraspråksanvändare (även på avancerad nivå) när det gäller uppgiften att bilda kollokationerna genom att para ihop ett antal givna ord (se avsnitt 4.4).

Urvalet av konventionaliserade uttryck i delarna A, B och C är baserat på en tidigare undersökning av Prentice & Sköldbberg (und. utg.), där materialet består av olika typer av figurativa ordförbindelser hämtade från uppsatser som samlats in inom ramen för projektet *Språk och språkbruk bland ungdomar i flerspråkiga storstadsmiljöer*.² Tanken bakom urvalet är att testa konventionaliserade uttryck som förekommer i texter skrivna av gymnasieelever i motsvarande ålder och skolmiljö. Det handlar alltså om uttryck som använts i skolrelaterat språkbruk, även om de inte är av den typen som är mest frekvent i t.ex. läromedel. Kollokationstestet i D är däremot sammansatt med utgångspunkt från två olika korpusar, nämligen *OrdiL-korpusen* (Lindberg & Johansson Kokkinakis 2007) och korpusen *LäSBarT* (Mühlenbock 2009). *OrdiL*-materialet består av läromedelstexter i NO, SO och matematik för grundskolans senare år. Korpusen innehåller ca 980.000 löpord (Lindberg 2007). *LäSBarT* är sammansatt av lättlästa nyhetstexter, informationstexter från kommunen, skönlitteratur och barnlitteratur. Materialet består av sammanlagt 1,5 miljoner löpord (Mühlenbock & Johansson Kokkinakis 2009). Fyra olika typer av kollokationer har hämtats från de frekvenslistor som bygger på *OrdiL* och *LäSBarT*³, nämligen *verb + substantiv*, *verb + preposition + substantiv*, *adjektiv + substantiv* samt *adverb +*

² Projektet genomfördes i samarbete mellan Göteborgs, Stockholm och Lunds universitet och finansierades av Riksbankens jubileumsfond år 2002–2006 .

³ Ett varmt tack till FD Sofie Johansson Kokkinakis för hjälp med framtagning av speciella frekvenslistor ur *OrdiL* och *LäSBarT*.

adjektiv, i syfte att undersöka om det finns vissa typer av kollokationer som vållar mer problem för informanterna än andra. Då de nämnda korpusarna inte är speciellt stora har även kollokationer som förekommer endast en gång valts ut för undersökningen. Urvalet har alltså snarare varit kvalitativt än kvantitativt. Uttrycken har dock sökts i konkordansmaterialet i Språkbanken (2010), och i sökmotorn Google (2010a) för en bedömning av hur pass vanliga de är i allmänspråket. De flesta kollokationerna som valts ut kan sägas vara relativt frekventa, enligt referensmaterialen. Jag har dock även valt att ta med ett mindre antal ovanliga kollokationer, för att undersöka i vilken mån frekvens i allmänspråket kan tänkas påverka elevernas resultat. Dessa kollokationer är *bilda en union*, *dra en suck*, *hamna i missbruk* och *bryta ett beroende*. Kollokationerna ur korpusmaterialet har sedan kompletterats med ytterligare ett antal uttryck med liknande lexikal och syntaktisk struktur, för att varje typ ska vara representerad med samma antal exempel. Syftet har i detta sammanhang varit att få ihop fem liknande konstruktioner samt åtta alternativa ord att välja bland, där inget skulle kunna uteslutas för enkelt. Tanken är naturligtvis att eleverna ska använda sin lexikala kunskap, snarare än uteslutningsmetoden, för att lösa uppgiften.

Uppgifterna i enkäten har formulerats för att framstå så tydliga och lättbegripliga som möjligt, utan att erbjuda några ledtrådar för hur uppgiften ska lösas. Begreppet *konventionalisering* nämns inte i uppgifterna. I stället betecknas uttrycken som "vanliga" (del B). Informanterna ombeds fylla i de ord som de tycker saknas (del B och C) alternativt para ihop de ord som de anser "passar bäst ihop" (del D). Informanterna har inte fått någon tidsbegränsning för att fylla i enkäten. De har däremot uppmanats att inte fundera för länge på de enskilda uppgifterna, då man inte kan resonera sig fram till lösningen, utan helt enkelt måste känna till uttrycket ifråga. Det har tagit mellan 10 och 20 minuter för de flesta informanter att fylla i enkäten. Författaren var själv närvarande vid undersökningens genomförande i samtliga grupper.

5.5. Analysmetod

Den jämförande analysen av enkätsvaren är i första hand kvantitativ. Varje svar i enkäterna har antingen bedömts som korrekt eller som avvi-

kande från det konventionaliserade uttrycket ifråga. Varje korrekt svar har givit ett poäng. Informanternas resultat har sedan jämförts på olika sätt och utifrån olika grupperingar (se vidare avsnitt 4). Dessutom har resultaten för de olika deltesten jämförts med varandra, liksom resultaten för de olika uttrycken. I den senare delen av analysen spelar alltså även kvalitativa aspekter en viss roll. Resultatskillnaderna mellan de olika informantgrupperna och enkätens olika komponenter har sedan testats för statistisk signifikans⁴. Skillnaderna har testats med hjälp av *en-vägs variansanalys* (ANOVA) samt *Newman-Keuls-testet* för multipla jämförelser (när endast två grupper jämförs sammanfaller ANOVA med ett två-sample t-test och inget multipeltest behövs). I samband med denna undersökning anses en skillnad vara signifikant om p-värdet är mindre eller lika med 0.05. Normalfördelningsbaserade parametriska metoder har använts, eftersom antalet observationer är så stort att den centrala gränsvärdesatsen kan anses ge approximativt normalfördelade medelvärden och frekvenser. De ovannämnda metoderna samt *Pearson korrelationskoefficient* har dessutom använts för att testa undersökningsmetodens validitet. Frågan är i det sammanhanget i vilken utsträckning de enskilda uppgifterna i enkäten skiljer informanter med bra resultat på hela enkäten ifrån dem med sämre resultat. Detta kan ses som en indikation på hur bra de enskilda uppgifterna (dvs. varje uttryck som skulle kompletteras) har fungerat när det gäller att påvisa skillnader mellan olika elevgrupper (jfr Körner & Wahlgren 2005).⁵

En del av materialet har även analyserats kvalitativt. Utgångspunkten här har dels varit svaren från de informanter som har lägst resultat i varje grupp, dels de tio uttryck i enkäten, som genererat lägst antal korrekta svar. Till att börja med har enkätsvaren från de 45 av de lägst presterande informanterna analyserats för de uttryck som visat sig vara svårast för informanterna. För att bekräfta eventuella mönster har sedan ytterligare tio av enkäterna med lägst antal poäng analyserats. Dessutom har enkäterna med högsta resultat, dvs. minst 45 av 55 möjliga poäng

⁴ Ett varmt tack till FD Tommy Johnsson från Institutionen för filosofi, lingvistik och vetenskapsteori, Göteborgs universitet, för hjälp med de statistiska analyserna.

⁵ Pga en felaktig dubbelkodning har de statistiska analyserna gjorts på 187 informanter (75 i grupp L1). Detta har dock ingen större betydelse för resultaten. Siffrorna har rättats till i tabellerna.

analyserats kvalitativt. Syftet har i detta fall varit att undersöka i vilken mån avvikelserna från de konventionaliserade formerna bland de högpresterande informanterna skiljer sig från avvikelserna bland informanterna med lägre resultat. Med utgångspunkt i den tidigare nämnda undersökningen av Ellis et al. (2008) har dessutom eventuella samband mellan uttryckens frekvens och andraspråkselevernas avvikelser från dessa konventionaliserade flerordsenheter undersökts, med avseende på kvantitativa och kvalitativa resultat. PAROLE-korpusen i Språkbanken, som består av 19.4 miljoner löpord (Språkbanken 2010) och innehåller olika typer av textmaterial, har använts som referenspunkter när det gäller uttryckens frekvens i det allmänna språkbruket. I viss mån även Google (2010a) använts i samma syfte.

6. Resultat

I detta avsnitt redovisas resultaten för en rad olika analyser av materialet. Som ett första steg i den kvantitativa analysen har de deltagande gymnasieklassernas resultat jämförts med varandra, dels som hela klasser (se tabell 5), dels utifrån en indelning i första- och andraspråkselever (se tabell 6). Denna grova indelning förfinas sedan stegvis. Vissa indelningar har dock inte följts upp genom hela analysen då det visat sig inte finnas några relevanta skillnader mellan testresultaten för vissa undergrupper (se vidare avsnitt 6.2). Bl.a. har testresultaten analyserats i förhållande till kön (även här utifrån en indelning i första- och andraspråkstalare). Då det inte finns några större skillnader mellan testresultaten för kvinnliga och manliga elever, ges dessa resultat dock inte något större utrymme här. Vidare har de olika informantgruppernas resultat för varje deltest jämförts, liksom resultaten för de enskilda uttrycken i enkäten. Avslutningsvis redovisas den kvalitativa delen av analysen, där eventuella mönster i avvikelser från de konventionaliserade uttrycken ifråga har undersökts.

6.1. Resultaten för de deltagande gymnasieklasserna

Tittar man på resultatet för de sammanlagt 9 gymnasieklasser som har deltagit i undersökningen, samt jämförelseklassen, är det de tre svenska

som andraspråksklasserna på Kampusgymnasiet, klasserna 3, 4 och 8, vars resultat (24 %, 40 % respektive 43 % av maximal poäng) som skiljer sig signifikant från de övriga (p-värde <0.0001). Klass 3 når dessutom ett signifikant lägre resultat än klasserna 4 och 8 (p-värden <0.0001). De övriga klasserna når resultat mellan 61 % (klass 2 på Kampusgymnasiet) och 71% (klass 5 på Prismagymnasiet) av det maximala resultatet (se tabell 5).

Jämförelseklassens resultat ligger med 82 % av det maximala antalet poäng signifikant högre än samtliga övriga deltagarklassers resultat (p-värde <0.0001). Överlag kan de resultat som sammanfattas i tabell 5 alltså anses vara förväntade. Om man bortser från de rena andraspråksklasserna och jämförelseklassen, verkar det däremot inte finnas något direkt samband mellan klassernas resultat och andelen andraspråkslever i klasserna; resultatskillnaderna mellan klasserna 1, 2, 5, 6, 7 och 9 är inte särskilt stora och inte heller statistiskt signifikanta (se tabell 5).

Det verkar inte heller finnas något uppenbart samband mellan årskurs och resultat. Klasserna i årskurs 1 når visserligen något lägre resultat och som framgår av tabell 5, är det klass 3, en andraspråksklass i årskurs 1, som når det signifikant lägsta resultatet. Det finns dock ingen signifikant skillnad i resultat mellan de övriga andraspråksklasserna eller mellan klass 4 (årskurs 1) och 8 (årskurs 2). Klass 9, som är den enda klassen i årskurs 3, skiljer sig heller inte signifikant från de lägre årskurserna. Dessutom går jämförelseklassen, som tidigare nämnts, i årskurs 1 samtidigt som den når det signifikant högsta resultatet.

Det bör i detta sammanhang poängteras att de signifikanta skillnaderna mellan de deltagarklasserna som läser svenska respektive svenska som andraspråk inte bör tolkas som en följd av skillnader i fråga om undervisningsämne. Att elever i de klasser som läser svenska som andraspråk når signifikant lägre resultat än elever som läser svenska, hänger i första hand samman med andelen elever med utländsk bakgrund i klassen, som är 100 % i andraspråksklasserna. Det bör dock noteras att det inte finns några signifikanta skillnader som kan kopplas till andelen andraspråkslever i respektive klass mellan de klasser som läser svenska som modersmål. Däremot är majoriteten av andraspråksleverna i dessa klasser tidiga inlärare medan det stora flertalet av de elever som har påbörjat sin andraspråksinläring efter sexårsåldern går i andraspråksklasserna 3, 4 och 8. Skillnaden mellan klasserna som läser svenska och de

TABELL 5: Testresultat per gymnasieklass

	KLASS 3	KLASS 4	KLASS 8	KLASS 2	KLASS 9	KLASS 7	KLASS 6	KLASS 1	KLASS 5	JK
Skola:årskurs	Kampus:1	Kampus:1	Kampus:2	Kampus:1	Kampus:3	Kampus:2	Kampus:2	Prisma:1	Prisma:2	Prisma:1
Typ av svensk- undervisning	sva	sva	sva	sv	19 sv/ 4 sva	sv	sv	sv	sv	sv
Antal elever	14	23	22	17	22	16	23	27	22	26
Antal elever med Svenska som L2	14 (100 %)	23 (100 %)	22 (100 %)	5 (26 %)	14 (61 %)	2 (13 %)	14 (64 %)	11 (41 %)	7 (28 %)	0 (0 %)
Resultat	24 %	40 %	43 %	61 %	63 %	68 %	69 %	62 %	71 %	82 %

som läser svenska som andraspråk lär således i första hand bero på startålder för andraspråksinlärning, snarare än på faktorerna utländsk bakgrund eller typ av svenskundervisning i sig.

6.2. Resultaten i förhållande till språklig bakgrund

I följande avsnitt redovisas och jämförs testresultaten för hela enkäten per informantgrupp. Dessutom har andraspråkselevernas resultat i de olika startåldersgrupperna jämförts. Som framgår av tabell 6, når eleverna med svenska som förstaspråk (grupp L1), helt enligt förväntningarna, det högsta resultatet bland de sammanlagt fyra informantgrupperna (72 % av maximal poäng). Jämförelseklassens resultat är med 82 % korrekta svar signifikant högre än L1-gruppens (p-värde 0.0031). Av grupperna med svenska som andraspråk når grupp L2/tidig det näst högsta resultatet (50 %), medan resultatet för grupp L2/sen (39 %) ligger tydligt under den förstnämnda. Skillnaderna mellan de tre huvudgrupperna är statistiskt signifikanta (p-värde <0.0001). Tittar man närmare på de två undergrupperna i gruppen L2/sen, ser man ett mönster som överensstämmer med tidigare forskning (jfr bl.a. Thomas & Collier 2002, Hyltenstam & Abrahamsson 2000) som visat att senare startålder och en påbörjad inlärning efter skolstart har relativt stor betydelse för andraspråksbehärsksningen såväl generellt som för utvecklingen av ett mer avancerat skolrelaterat språk. Det finns ingen statistisk signifikant skillnad mellan grupp L2/tidig och informanterna med en startålder till och med 10 år i grupp L2/sen. Däremot finns en signifikant skillnad (p-värde <0.0001) mellan informanterna med senare startålder (>10 år) i grupp L2/sen, som når 33 % av maximalt resultat, och de övriga andraspråksgrupperna, dvs. grupp L2/tidigt samt informanter med startålder ≤ 10, som båda når högre resultat (50 respektive 46 %).

Mest relevanta, vad gäller de resultat som sammanfattas i tabell 6 nedan, är alltså skillnaderna

- mellan förstaspråkselever och andraspråkselever med en startålder för inlärningen ≤10 år,

- mellan förstaspråkselever och andraspråkselever med en startålder >10 år,
- mellan andraspråkselever med en startålder ≤10 år och andraspråkselever med en startålder >10 år.

Det handlar, med andra ord, om en generell skillnad dels mellan första- och andraspråkselever, dels mellan andraspråkselever med olika startålder för inläringen, där 10 år verkar vara en viktig åldergräns.

TABELL 6: Resultat för hela enkäten) per startåldersgrupp

	GRUPP L1	GRUPP L2/tidig (startålder ≤ 6)	GRUPP L2/sen			JK
			start- ålder ≤10	start- ålder >10	totalt	
Antal elever	74	75	19	18	37	26
Resultat	72%	50%	46%	33%	39%	82%

I tabell 6.1 redovisas enbart testresultaten för de informanter som har angett svenska som sitt andraspråk. I denna del av analysen har både grupp L2/tidig och grupp L2/sen delats upp i två undergrupper, beroende på födelseland och startålder. Analysen visar inga nämnvärda skillnader mellan informanter födda i Sverige och informanter födda utomlands i grupp L2/tidig, vilket är anledningen till att denna uppdelning inte följs upp vidare i detta arbete.

TABELL 6.1: Detaljerad redovisning av resultat för informanterna med svenska som L2

	GRUPPL2/tidig (startålder ≤6)			GRUPP L2/sen (start i skolådern)			JK
	födda i Sverige	födda utoml.	totalt	startålder ≤ 6	startålder >10	totalt	
Antal elever	45	30	75	18	19	37	26
Antal elever som läser sv	29	14 *	43 *	8	1 *	9 *	26
Antal elever som läser sva	16	17 *	33 *	10	17**	27 *	0
Antal elever som läst modersmål	33	25	58	16	17	33 *	0
Resultat	50%	49%	50%	46%	33%	39%	82%

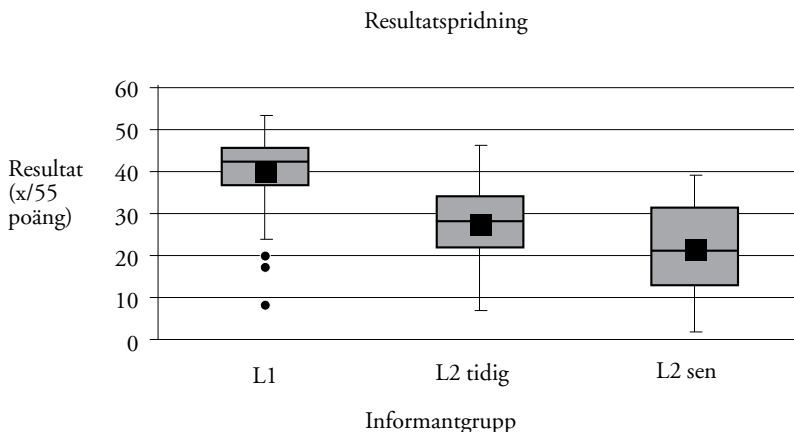
* 1 av informanterna i gruppen läser både sv och sva

** 1 av informanterna i gruppen läser både sv och sva, 1 har kryssat för "vet ej" och 1 har inte besvarat frågan.

6.2.1. Resultatspridning

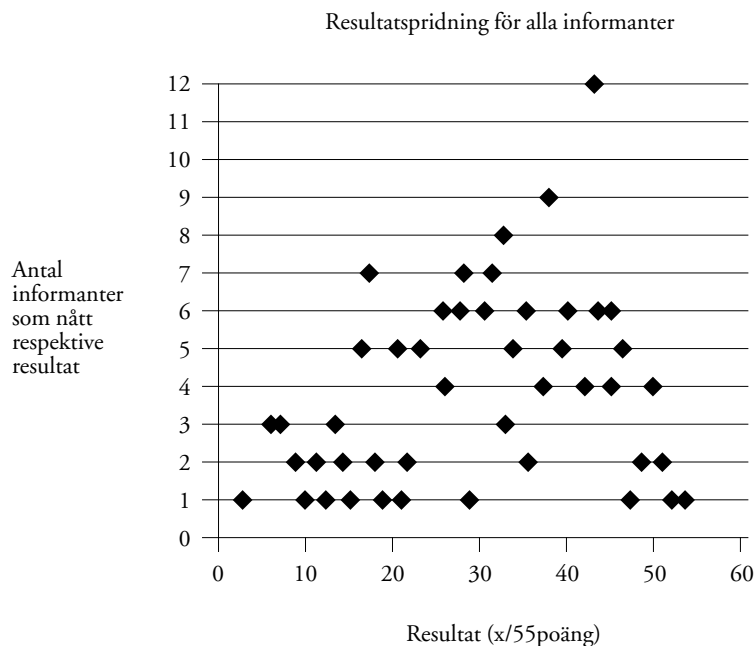
Huvudfokus i den kvantitativa analysen ligger av naturliga skäl på resultatet på gruppnivå. Det är likväl intressant att visa hur resultatet är fördelade på individnivå, dvs. hur resultaten sprider sig mellan 0 och 55 i informantgrupperna. Figur 2 visar skillnaderna i resultatspridning mellan informantgrupperna. Lådorna som visas i diagrammet (eller *lådagrammet*, se Wahlgren 2005) visar det spann inom vilken 50 % av alla resultat ligger. Den lodräta linjen visar hur resten av värdena sprider sig ovan och nedanför medianen, som representeras av det vågräta strecket i lådorna. Fyrkanterna i respektive låda representerar medelvärden. Extremt avvikande värden (eng. *outliers*) i förhållande till medianen markeras med egna symboler (se grupp L1 i figur 2). Det framgår tydligt av lådagrammet att den minsta spridningen finns i grupp L1 och den största i grupp L2/sen. Resultatspridningen blir alltså större, ju senare informanterna påbörjat sin inläring av svenska. Lådagrammet visar också

att medelvärdet sjunker med stigande startålder; detsamma gäller de maximala och minimala resultatvärdena. Det finns dock samtidigt en del andraspråkselever som når betydligt högre resultat än en del av eleverna i grupp L1. Avvikande värden (som markeras med prickar i figur 2) förekommer enbart i grupp L1, där det endast förekommer 3 resultat på 20 poäng eller lägre (se även figur 3.1 nedan).



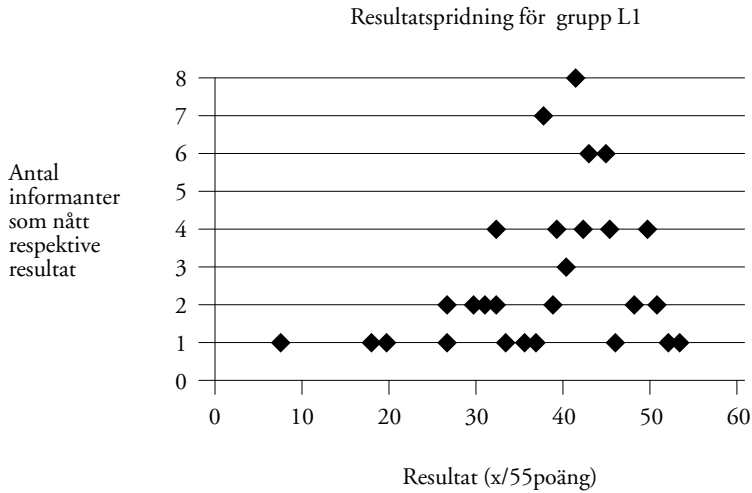
FIGUR 2: Jämförelse av resultatspridningen mellan informantgrupperna

Figurerna 3-3.3 nedan visar något mer detaljerat vilka resultat som förekommer i informantgrupperna och hur många av informanterna som nått ett visst resultat. I figur 3 visas denna resultatspridning för samtliga 186 informanter. Här framgår bl.a. att det vanligaste resultatet är 42 av 55 poäng. De resultat som förekommer endast en enstaka gång ligger antingen under 30 eller runt 50 poäng. Det högsta resultatet är 53 av 55 poäng. Detta resultat förekommer endast en gång bland de 186 informanterna. Det lägsta resultatet är 2 av 55 poäng och förekommer också endast en gång i hela informantgruppen.



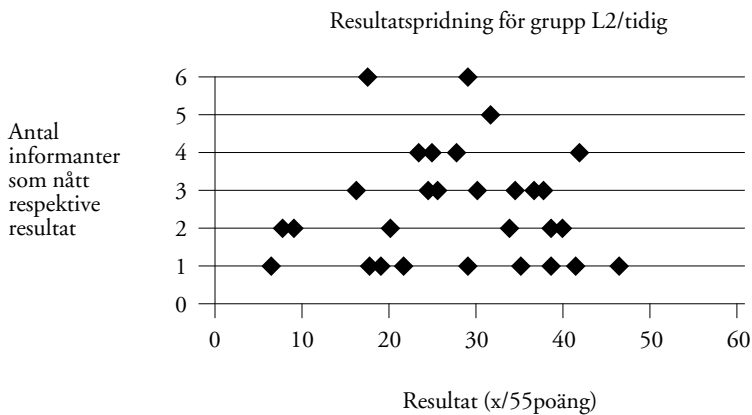
FIGUR 3: Antal informanter som nått olika resultat i hela informantgruppen

Figur 3.1 visar resultatspridningen i grupp L1. Som framgår här är det vanligaste resultatet (som förekommer 8 gånger) i denna grupp 42 av 55 poäng (precis som för hela informantgruppen, jfr figur 2). Majoriteten (42 av 74) av resultaten i grupp L1 ligger mellan 40 och 50 poäng och endast 6 resultat ligger under 30 poäng i denna grupp. Det högsta resultatet (som förekommer en gång) är 53 av 55 poäng och det lägsta resultatet bland dessa elever är 8 av 55 poäng. Även detta resultat förekommer endast en gång. Det finns alltså enstaka resultat under 20 och ett under 10 poäng även i grupp L1. Överlag kan man dock konstatera att spridningen är relativt liten, då det stora flertalet av resultaten ligger mellan 30 och 53 (se figur 3.1).



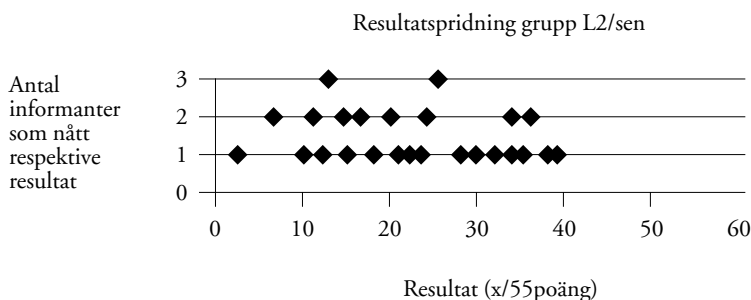
FIGUR 3.1: Antal informanter som nått olika resultat i grupp L1

Av figur 3.2 framgår att de resultat som är vanligast (med 6 förekomster var) i grupp L2/tidig är 17 respektive 28 poäng. Flertalet av resultaten ligger mellan 20 och 46. Här finns alltså en något större spridning än i grupp L1. I grupp L2/tidig är 46 – ett resultat som förekommer endast en gång – det högsta resultatet. Det lägsta resultat i denna grupp är 7 poäng och förekommer också endast en gång.



FIGUR 3.2: Antal informanter som nått olika resultat i grupp L2/tidig

Spridningen i grupp L2/sen skiljer sig tydligt från den i de övriga grupperna (se även figur 2); som framgår av figur 3.3 förekommer samma resultat maximalt tre gånger i denna informantgrupp, och detta gäller endast 2 olika resultat. De allra flesta resultaten förekommer endast en eller två gånger. Som redan framgått av figur 2 ovan, förefaller den individuella variationen alltså vara större hos de sena andraspråksinlärarna än hos övriga informanterna. Större delen av resultaten (21 av 37) ligger mellan 20 och 39. Det högsta resultatet i grupp L2/sen är 39 av 55 poäng och det lägsta 2 av 55 poäng.



FIGUR 3.3: Antal informanter som nått olika resultat i grupp L1

6.3. De olika deltesten

Syftet med denna del av analysen är dels att undersöka eventuella mönster i fråga om skillnader mellan delresultaten inom samma grupp, dels att undersöka skillnader i fråga om resultat för de olika deltesten mellan informantgrupperna. Analyserna syftar alltså till att besvara frågan om vilken typ av uppgifter informanterna har klarat bättre respektive sämre. Resultaten sammanfattas i tabell 7.

TABELL 7: Resultat för enkätens olika deltester (% av maximalt resultat)

	GRUPP L1	GRUPP L2/tidig (startålder, ≤ 6)	GRUPP L2/sen			JK
			start- ålder ≤10	start- ålder >10	totalt	
Antal elever	74	75	18	19	37	26
A Förståelse	84%	72%	72%	57%	64%	91%
B Produktion i kontext	72%	41%	33%	16%	24%	83%
C Produktion utanför kontext	64%	34%	24%	14%	19%	74%
D Kollokationer	74%	55%	54%	43%	49%	83%

Som framgår av tabell 7, når alla informantgrupper (inklusive jämförelseklassen) sina högsta resultat i del A. Med tanke på tidigare refererad forskning är detta inget oväntat resultat. Uppgifter som i första hand kräver receptiv behärskning av ordförrådet har i många studier visat sig vara lättare för både första- och andraspråkstalare, än sådana där man måste använda sig av produktiva språkfärdigheter (Nation 2001). Dessutom har Golden (2005), som tidigare nämnts, visat att det inte nödvändigtvis krävs full behärskning av ett uttrycks alla lexikala komponenter för att kunna förstå ett konventionaliserat uttryck. Av tabell 7 framgår vidare att skillnaden mellan grupp L1 och andraspråksgrupperna inte är lika stor i del A som i de övriga deltesten. Icke desto mindre är skillnaderna mellan grupp L1 och samtliga andraspråksgrupperna statistiskt signifikanta. Detta gäller också skillnaden mellan gruppen med den senaste startåldern (>10) och de övriga informantgrupperna (p-värde <0.0001). Resultatskillnaden mellan jämförelseklassen och grupp L1 är däremot inte statistiskt signifikant för denna del av enkäten.

Resultaten i B-delen är betydligt mer spridda än i A-delen. Resultaten sträcker sig från 72 % i grupp L1 till endast 16% i gruppen med den senaste startåldern (se tabell 7). Mönstret liknar dock det i förståelsetestet i det avseendet att grupp L1 når det högsta resultatet (72 %), följt av grupp L2/tidig (41%) och de tidigare inlärarna i grupp L2/sen (33%). De senare inlärarna i grupp L2/sen når med 16 % det klart lägsta resultatet även här. Jämförelseklassens 83 % är signifikant högre än resultatet för grupp L1 (p -värde = 0.0080). Som i del A är det grupp L1 och de senare inlärarna i grupp L2/sen som skiljer sig signifikant från samtliga övriga grupper i sina resultat (p -värde <0.0001). Skillnaden mellan dessa två informantgrupper är störst i denna del av enkäten. Som nämndes i avsnitt 5.4 har en aspekt vid utformningen av del B varit att undersöka om det finns resultatskillnader beroende på vilket led i uttrycket som utelämnas. Analysen visar dock inga nämnvärda skillnader i resultat mellan de uppgifter där verbet utelämnas och de där informanterna ska komplettera uttrycket med ett substantiv.

Samtliga informantgrupper når sina lägsta resultat i enkätens tredje del (C). Alla resultatskillnader mellan informantgrupperna är statistiskt signifikanta (p -värde > 0.0001) i denna del av enkäten. Mönstret för skillnaderna är detsamma som i de övriga delarna, dvs. grupp L1 når det högsta resultatet (64 %) och de senaste inlärarna i grupp L2/sen når det lägsta resultatet (14 %). Det verkar alltså som om det är svårast för informanterna att komplettera ett ofullständigt konventionaliserat uttryck utan kontext. Även dessa resultat ligger i linje med tidigare forskning (Nation 2001), som har visat att uppgifter av det slag som ställts i del C, kräver en rad olika typer av kunskap som språkbrukaren till fullo måste utnyttja när han eller hon inte har hjälp av kontexten.

Den sista delen (D) i enkäten är ett kollokationstest. Utgångspunkten för detta test har varit en hypotes om att det råder stor enighet bland förstaspråkstalare om hur ett antal givna ord ska paras ihop för att bilda de mest idiomatiska kollokationerna, medan andraspråkstalare uppvisar betydligt mer variation i sina lösningar av en sådan uppgift (van Lier 1995, Lindberg 2004). Resultaten av analysen av den sista delen i enkäten kan sägas ligga i linje med denna hypotes och visar att det finns en signifikant skillnad (p -värde <0.0001) mellan grupp L1 och båda L2-grupperna. Skillnaden mellan grupp L2/tidig och grupp L2/sen är däremot inte statistiskt signifikant. Det gäller även skillnaden mellan de se-

naste inlärarna (>10) i grupp L2/sen och inlärarna i grupp L2/tidig. Även om siffrorna i tabell 7 antyder vissa tendenser till att andraspråksinlärare med tidigare startålder har lättare för kollokationsuppgiften än informanter med mycket senare startålder, är alltså den enda signifikanta resultatskillnaden den mellan första- och andraspråkstalare. Jämförelseklassen når även i det sista deltestet ett signifikant högre resultat än för samtliga övriga grupper (p-värde = 0.0315).

6.3.1. Skillnader mellan delresultaten inom varje grupp

En annan intressant fråga i samband med de olika deltesten i enkäten är vilka resultatskillnader som finns mellan deltesten inom respektive informantgrupp. Som tidigare nämnts visar resultaten ett tydligt mönster, nämligen att samtliga grupper når sitt högsta resultat i A-delen och sitt lägsta i C-delen. Därutöver visar sig ett liknande mönster när det gäller de övriga delresultaten i samtliga grupper. Med andra ord är rangordningen mellan delresultaten densamma för varje grupp, vilket framgår i tabell 7.1.

TABELL 7.1: Delresultaten i fallande ordning (% av maximalt resultat) per grupp.

	GRUPP L1	GRUPP L2/tidig (startålder, ≤ 6)	GRUPP L2/sen			JK
			start- ålder ≤10	start- ålder >10	totalt	
Antal elever	74	75	18	19	37	26
A	84%	72%	72%	57%	64%	91%
D	74%	55%	54%	43%	49%	83%
B	72%	41%	33%	16%	24%	83%
C	64%	34%	24%	14%	19%	74%

När det gäller resultatskillnadernas signifikans finns det däremot viss variation grupperna emellan. Jämför man delresultaten för grupp L1, finns det signifikanta skillnader mellan alla deltest, utom mellan del B och D; resultaten för båda dessa deltest är signifikant lägre än delresultatet för grupp L2/sen.

tatet för A-delen och signifikant högre än delresultatet för C-delen (p-värde <0.0001). I grupp L2/tidig är däremot alla skillnader mellan delresultaten statistiskt signifikanta (p-värde <0.0001). I båda startåldersgrupperna inom grupp L2/sen är alla skillnader, förutom den mellan delresultaten för del B och C, statistiskt signifikanta (p-värde <0.0001). I jämförelsegruppen skiljer sig endast det lägsta delresultatet (C) signifikant från de övriga (p-värde = 0.0001).

Trots att inte alla skillnader mellan delresultaten är statistiskt signifikanta kan man alltså konstatera att det finns en tendens till ett enhetligt mönster bland informantgrupperna, när det gäller vilken typ av uppgifter i enkäten man klarar bättre respektive sämre.

6.4. De enskilda uttrycken

För att undersöka vilka typer av uttryck som vållar mer respektive mindre problem för informanterna – i synnerhet med tanke på andraspråkseleverna – har resultaten för de enskilda uttrycken analyserats och jämförts med varandra i informantgrupperna. Dessa resultat redovisas här för de olika delarna i enkäten.

6.4.1. Förståelse

Som redan nämnts, när samliga informantgrupper de högsta resultaten i enkätens första del, förståelsetestet A, där uttryckets betydelse ska anges. I tabell 8 redovisas resultaten för de enskilda uttrycken i förståelsetestet i fallande ordning, dvs. de uttryck som genererat högst poäng (om man räknar ihop resultaten från alla informantgrupper) står högst upp i tabellen. En del av resultaten kan betraktas som förväntade. Om man utgår från faktorer som t.ex. genomskinlighet, är det inte särskilt överraskande att fler elever förstår uttrycket *sätta färg på något* än uttrycket *prata i nattmössan*. I andra fall kan man, särskilt med tanke på andraspråksgруппernas resultat, hitta möjliga förklaringar i uttryckets frekvens i språkbruket (jfr Ellis et al. 2008). Uttrycket *innan man vet ordet av*, som kan beskrivas som icke-kompositionellt då de enda semantiskt genomskinliga komponenterna är formorden *innan* och *man*, genererar

exempelvis relativt högt resultat i alla grupper. Här finns heller inga signifikanta skillnader mellan startåldersgrupperna. I PAROLE förekommer detta uttryck 19 gånger (på söksträngen "v. * ordet av"), medan det endast finns 8 belägg (på söksträngen "i nattmössan"¹) på *prata i nattmössan* (Språkbanken 2010a). Google visar ca 134 000 träffar på söksträngen *innan man "vet ordet av"* samt ca 66 000 på *innan man "visste ordet av"* – gentemot ca 3 800 belägg av *prata "i nattmössan"* (Google 2010a)². Ser man däremot till den totala bilden som framgår av tabell 8, finns det dock en del resultat som inte enbart kan förklaras med faktorer som hänger ihop med frekvens eller graden av kompositionalitet och genomskinlighet.

Det finns relativt få signifikanta skillnader mellan informantgrupperna när det gäller de uttryck som informantgruppen som helhet klarar bäst. 7 av 10 uttryck i A-delen får sammanlagt mellan 80 % och 90 % av det maximala resultatet. I endast 2 av dessa 7 fall når de senaste inlärnarna i grupp L2/sen ett signifikant lägre resultat än de övriga startåldersgrupperna. Detta gäller uttrycken *fastna för något* (p-värde <0.0001) och *ha mage att göra ngt* (p-värde <0.0001). Jämförelsegruppens resultat ligger mellan 96% och 100 % för de sju uttryck som ligger överst i tabell 8. Inget av dessa sju resultat skiljer sig signifikant från resultaten för grupp L1.

Resultaten för uttrycket *något står en upp i halsen* är något iögonfallande, med avseende på resultatskillnader mellan startåldersgrupperna; i detta fall är det startåldersgruppen ≤ 10 som med 94 % når det signifikant högsta resultatet (p-värde = 0.0227). Därutöver finns inga signifikanta resultatskillnader mellan grupp L1, som når 61 % av det maximala antalet poäng, och de övriga andraspråksgrupperna eller jämförelsegruppen. Skillnaden mellan grupp L2/tidig (71 %) och informanterna med startålder ≥ 10 i grupp L2/sen (53 %) är heller inte signifikant för uttrycket *något står en upp i halsen*. Ett problem i samband med detta resultat är naturligtvis att undergrupperna i L2/sen är relativt små i förhållande till de övriga informantgrupperna. Intressant är dock att det

¹ För att minimera risken att missa relevanta träffar har jag i vissa fall valt den minsta möjliga söksträngen och sedan räknat de relevanta träffarna manuellt.

² Samtliga frekvenssökningar i PAROLE och Google genomfördes den 11 januari 2010.

fortfarande finns en tendens till ett högre resultat i andraspråksgrupperna än i förstaspråksgruppen, om man jämför hela grupp L2/sen (73 %) och grupp L2/tidig (71 %) med förstaspråkseleverna i grupp L1 (61%). Dessutom är andraspråkgruppernas resultat högre än jämförelsegruppens, som också når 61 % av maximalt antal poäng. Detta förekommer endast i samband med just uttrycket *något står en upp i halsen*. Dessa skillnader är dock inte statistiskt signifikanta.

Uttrycket *något är/blir en ond cirkel* är det enda i A-delen som genererar ett signifikant högre resultat (86 %) i grupp L1, jämfört med samtliga L2-grupper (p-värde <0.0001). Jämförelseklassens resultat (88 %) skiljer sig inte signifikant från resultatet för grupp L1. Icke-kompositionalitet skulle kunna vara en påverkande faktor, som i detta fall dock verkar ha större betydelse för andraspråkselevernas resultat än för förstaspråkselevernas. Frekvensuppgifter ger inga entydiga svar med avseende på en möjlig förklaring till detta resultat; med 34 förekomster i PAROLE (söksträng: "*en ond cirkel*" Språkbanken 2010a) kan uttrycket varken beskrivas som särskilt hög- eller lågfrekvent. Googles frekvensuppgifter är i detta fall inte användbara, då sökstängen "*ond cirkel*" ger ca 72 000 träffar, "*en ond cirkel*" däremot drygt 500 000 (Google 2010a). Som påpekats i tidigare arbeten (Große 2008, jfr Prentice & Skölberg und. utg.), är Google inte helt oproblematiskt som referensmaterial, bl.a. eftersom man inte har någon kontroll över vilka slags texter som finns presenterade där. Dessutom är frekvensuppgifterna inte alltid konstanta och därmed svåra att tolka. Webben kan däremot, med tanke på dess storlek och texternas mångfald, ge en indikation om ett visst uttryck använts vid en viss tidpunkt och, åtminstone relativt sett, i vilken utsträckning (Moon 2008:28, jfr även Lorentzen 2008).

När det gäller uttrycket *prata i nattnössan*, som får det lägsta resultatet i samtliga informantgrupper, finns inga signifikanta resultatskillnader mellan startåldersgrupperna. Jämförelsegruppens resultat (69 %) ligger däremot signifikant högre än alla övriga resultat för detta uttryck (p-värde = 0.0266).

Det kan alltså konstateras att de signifikanta skillnaderna mellan informantgrupperna är få när det gäller resultaten för de enskilda konventionaliserade uttrycken i förståelsetestet. I de sammanlagt 4 fall där det finns en signifikant skillnad är det dessutom endast en av startåldersgrupperna som skiljer sig signifikant från de övriga. Jämförelsegruppens

resultat, som i 7 av 10 fall ligger på en jämn, hög nivå mellan 96 % och 100 %, skiljer sig i de flesta fall inte heller signifikant från L1-gruppens resultat.

Det finns 2 fall där eleverna med den högsta startåldern i grupp L2/sen som enda grupp skiljer sig signifikant från de övriga informanterna; för uttrycken *fastna för något* och *ha mage att göra något* när de sena inlärarna med 53 % respektive 42 % de signifikant lägsta resultaten. Orsakerna till detta kan här bara spekuleras i. *Fastna för något*, med 109 träffar (söksträngar "fastna.* för" och "fastna.* [1 eller 2 ord emellan] för" Språkbanken 2010a) respektive ca 550 000 träffar (söksträngar "fastna [samtliga verbformer] för" och "fastna [samtliga verbformer] du/hon/han/vi/de för", Google 2010a), kan ses som relativt högfrekvent. *Ha mage att göra något* är däremot mindre frekvent i det använda referensmaterialet (22 förekomster i PAROLE på "mage att" och ca 51 000 träffar på Google på *ha "mage att"*). Skillnaden i frekvens skulle kunna anföras som förklaring till att hela grupp L2/sen når ett lägre resultat för det mer lågfrekventa uttrycket, medan förstaspråkselevernas resultat ligger på ungefär samma höga nivå för båda uttrycken (jfr Ellis et al. 2008). Skillnaden mellan grupp L1 och gruppen med den högsta startåldern lär däremot bero på andra faktorer. Både *fastna för något* och *ha mage att göra något* är semantiskt relativt ogenomskinliga. Icke-kompositionalitet kan alltså tänkas vara åtminstone en del av förklaringen till att just dessa uttryck verkar vara svårare för andraspråkselever med sen startålder än för förstaspråkselever och andraspråkselever med en tidigare startålder. Ett sätt att försöka komma något närmare en förklaring är genom en kvalitativ analys, som i avsnitt 6.5 nedan. Motsvarande gäller för det tidigare nämnda uttrycket *något är en ond cirkel*. Vissa av dessa resultat kommer dock att förbli svåra att förklara, inte minst det faktum att andraspråkseleverna når högre resultat än förstaspråkseleverna (inklusive jämförelsegruppen) för just uttrycket *något står en upp i halsen*.

6.4.2. Lucktest i kontext

I jämförelse med resultaten i förståelsetestet, finns det större variation i resultaten mellan och inom informantgrupperna i B-delen i fråga om de enskilda konventionaliserade uttrycken. Ser man på summan av resultaten för samtliga startåldersgrupper, varierar resultaten mellan 90 % av

maximalt resultat, för uttrycket *på egen hand* och 19 % för *sätta käppar i hjulen*. Ett annat iögonfallande resultat som framgår av tabell 8.1 är att det finns flera nollresultat i gruppen med startålder >10, samt ett i hela grupp L2/sen, något som inte förekommer i förståelsetestet. Även jämförelsegruppen uppvisar större variation i resultaten för de enskilda uttrycken i detta *sentence-completion-test* (mellan 46 % och 100 %) än i den första delen av enkäten. Det finns få (2 av tio) signifikanta skillnader mellan jämförelsegruppen och grupp L1 även i den detaljerade analysen av B-delen.

Samtliga startåldersgrupper (förutom grupp L1, som når ett marginellt högre resultat för uttrycket *ta första steget*) når alltså sitt högsta resultat i B-delen för uttrycket *(göra något) på egen hand*. När det gäller skillnaderna grupperna emellan, uppvisar de senare inlärarna i grupp L2/sen (53 %), ett signifikant lägre resultat (p-värde <0.0001). Övriga resultat ligger mellan 89 % och 96 % och skillnaderna här är inte statistiskt signifikanta. Ett liknande mönster gäller för uttrycket *ta första steget*, som ligger på andra plats i tabell 8.1. Som framgår av tabellen är variationen mellan gruppernas resultat större för de övriga uttrycken i detta deltest. I den nedre delen av tabell 8.1 framträder en allt tydligare skillnad mellan grupp L1 och andraspråksgrupperna. Ser man på resultatskillnaderna mellan startåldersgrupperna för de sex uttrycken som står nederst i tabellen, är resultatet för grupp L1 det signifikant högsta i alla sex fallen (p-värden <0.0001). De sex konventionaliserade uttrycken är (i fallande ordning) *komma till kritan*, *falla för frestelsen*, *hålla modet/humöret uppe*, *få någon/något på halsen*, *klättra på väggarna* och *sätta käppar i hjulen*. För de senare 4 av dessa 6 uttryck uppnår de senare inlärarna i grupp L2/sen 0 % av det maximala resultatet. Det är med andra ord ingen informant i denna grupp som kompletterar dessa konventionaliserade uttryck på det förväntade sättet. För uttrycket *sätta käppar i hjulen* gäller detta för hela grupp L2/sen. Det är också detta uttryck som har fått lägst poäng i B-delen överlag. Intressant i samband med detta uttryck är också att jämförelsegruppens resultat (81 %) skiljer sig tydligt och signifikant från L1-gruppens, som endast når 43% av maximalt resultat. Med tanke på faktorn kompositionalitet, är det inte helt oväntat att ett semantiskt ogenomskinligt uttryck som *sätta käppar i hjulen* är bland dem som relativt få elever behärskar i en undersökning som denna. Mot bakgrund av de resultat som presenteras i Ellis et al.

(2008) kan även uttryckets jämförelsevis låga frekvens tänkas påverka framför allt andraspråkslevernas behärskning av *sätta käppar i hjulen* negativt; uttrycket förekommer sammanlagt 30 gånger i PAROLE (på ”*käppar i hjulen*” och *käppar i hjulet*” Språkbanken 2010a), medan det finns 470 belägg på högstplacerade *på egen hand* (söksträng ”*på egen hand*”) i korpusen. Använder man Google som referenspunkt får man ca 290 000 träffar på uttrycket *sätta käppar i hjulen*, (sammanlagt på söksträngarna *sätta ”käppar i hjulen*” och *sätta ”käppar i hjulet*”) och ca 758 000 på söksträngen ”*på egen hand*” (Google 2010a). Ser man i första hand på frekvenserna i PAROLE, finns det vissa tendenser som ligger i linje med de resultat som Ellis et al. (2008) presenterar. Med andra ord uppnår grupp L2/sen låga resultat för de flesta uttryck med relativt låg frekvens i korpusen; bland de uttryck, för vilka någon eller båda undergrupperna i grupp L2/sen når resultat nära eller lika med noll, förekommer *klättra på väggarna* 5 gånger (”*klättr.** [1 eller 2 ord emellan] på väggarna”) och *hålla modet/humöret uppe* sammanlagt 8 gånger (en gång på söksträngen ”*humöret uppe*” och 7 gånger på ”*modet uppe*”) på 19.4 miljoner ord. *Få något/någon på halsen* förekommer däremot 28 gånger (”*på halsen*”) och *sätta käppar i hjulen*, som tidigare nämnts, 30 gånger. Dessa frekvenser är alltså något högre, men fortfarande betydligt lägre än frekvensen för uttrycket *på egen hand*, där grupp L2/sen når sina högsta resultat. Frekvensen framstår dock inte som en påverkande faktor i samband med alla uttryck. Uttrycket *ta första steget*, där de sena andraspråksinlärarna når sina näst högsta resultat, förekommer endast 27 gånger i PAROLE (på söksträng ”*första steget*”). Man kan här alltså som mest hävda att uttryckets frekvens kan ses som en av flera möjliga faktorer som kan tänkas påverka elevernas produktiva behärskning av de konventionaliserade uttrycken i B-delen.

6.4.3. Lucktest utanför kontext

Som redan nämnts i avsnitt 6.4, når samliga grupper sitt lägsta delresultat i C-delen, där uppgiften består i att komplettera ett ofullständigt konventionaliserat uttryck utan kontext. Frågan som ska besvaras i detta avsnitt är i vilken mån resultaten för de enskilda uttrycken varierar inom grupperna och om det finns vissa uttryck som är speciellt svåra för en viss grupp. Ser man på tabell 8.2 som helhet, kan man konstatera att

variationen bland resultaten för de enskilda uttrycken inom varje grupp är relativt stor. Även när det gäller signifikanta resultatskillnader mellan startåldersgrupperna, finns det en del variation. Ett allmänt mönster som kan utläsas är dock att grupp L1 når de högsta resultaten för samtliga uttryck, följt av grupp L2/tidig. Grupp L2/sen når i samtliga fall de lägsta resultaten. Resultatskillnaderna mellan grupperna är dock inte alltid statistiskt signifikanta.

Vad gäller uttrycket *ta saken i egna händer*, som återfinns överst i tabell 8.2, når 3 av 4 grupper relativt höga resultat (mellan 78 % och 97 % av maximala antalet poäng). Endast de senare inlärarna i grupp L2/sen skiljer sig signifikant från övriga grupper med 47 % (p-värde <0.0001). Längre ner i tabellen blir de signifikanta skillnaderna mellan startåldersgrupperna fler. Resultatet för uttrycket *av hela mitt hjärta* är 92 % för grupp L1, medan grupp L2/tidig når ett signifikant lägre resultat än förstaspråksleverna med 71 %. Det finns ingen signifikant skillnad mellan grupp L2/tidig (72 %) och gruppen med startålder ≤10 i grupp L2/sen (61 %). Däremot når de senare inlärarna i grupp L2/sen det signifikant lägsta resultatet för uttrycket *av hela mitt hjärta* (42 %, p-värde <0.0001). I samband med nästa uttryck i tabell 8.2, *gå bakom ryggen på någon*, skiljer sig resultatet för båda undergrupperna i grupp L2/sen (50 % respektive 42 %) signifikant från resultatet för grupp L2/tidigt (71 %). Den senare når i sin tur ett signifikant lägre resultat än grupp L1, som når 96 % av maximal poäng för detta uttryck. Även om det alltså finns en del variation när det gäller resultaten för de enskilda uttrycken i C-delen, är det i samband med nästan hälften av uttrycken (7 av 15) endast förstaspråkslevernas resultat som skiljer sig signifikant från resultaten för samtliga andraspråksgrupper. Det gäller uttrycken *som gräddde på moset*, *rinna ut i sanden*, *ha ett finger med i spelet*, *med ngt i högsta hugg*, *på rak arm*, *få upp ögonen för något* (p-värden <0.0001) och *hålla någon i schack* (p-värde = 0.0002). Informanterna med startålder >10 i grupp L2/sen når nollresultat för de senare 6 av dessa 7 uttryck. Uttrycket *hålla någon i schack* är dessutom det uttryck som genererar lägst poäng av samtliga uttryck i enkäten och är med andra ord det uttryck som informantgruppen som helhet verka behärska sämst av dem som testats i undersökningen. Ett annat uttryck som sticker ut är *livet leker för någon*, där både grupp L1 och jämförelsegruppen endast når 27 % av maximalt resultat. Här finns heller ingen signifikant skillnad mellan någon

TABELL 8.2: Resultat på lucktestet i del C per uttryck (i fallande ordning)

	SUMMA alla informanter	GRUPP L1	GRUPP L2/tidig (startålder, ≤ 6)	GRUPP L2/ sen (startålder > 6)		JK	
				startålder ≤ 10	startålder > 10		
				totalt	totalt		
antal elever	186	74	75	18	19	37	26
UTTRYCK							
ta saken i egna händer	158 (85%)	72 (97%)	63 (84%)	14 (78%)	9 (47%)	23 (62%)	26 (100%)
av hela mitt hjärta	141 (76%)	68 (92%)	54 (72%)	11 (61%)	8 (42%)	19 (51%)	23 (88%)
gå bakom ryggen på ngn	141 (76%)	71 (96%)	53 (71%)	9 (50%)	9 (42%)	17 (46%)	21 (81%)
på löpande band	135 (73%)	65 (88%)	56 (75%)	11 (61%)	3 (16%)	14 (38%)	24 (92%)
ingen dans på rosor	106 (57%)	62 (83%)	36 (48%)	6 (33%)	2 (10%)	8 (22%)	25 (96%)
som gräddde på moset	81 (43%)	57 (76%)	19 (25%)	3 (16%)	2 (10%)	5 (13%)	24 (92%)
tid är pengar	77 (41%)	43 (57%)	21 (28%)	5 (28%)	8 (42%)	13 (35%)	19 (73%)
rinna ut i sanden	73 (39%)	58 (78%)	15 (20%)	0 (0%)	0 (0%)	0 (0%)	23 (88%)
ha ett finger med i spelet	57 (30%)	43 (57%)	12 (16%)	2 (11%)	0 (0%)	2 (5%)	19 (73%)
inte för fem öre	54 (29%)	40 (54%)	13 (17%)	1 (5%)	0 (0%)	1 (3%)	18 (69%)
med ngt i högsta hugg	41 (22%)	35 (47%)	5 (7%)	1 (5%)	0 (0%)	1 (3%)	19 (73%)
på rak arm	39 (21%)	32 (43%)	7 (9%)	0 (0%)	0 (0%)	0 (0%)	13 (50%)
livet leker för ngn	40 (21%)	20 (27%)	18 (24%)	2 (11%)	0 (0%)	2 (5%)	7 (27%)
få upp ögonen för ngt	37 (20%)	28 (38%)	9 (12%)	0 (0%)	0 (0%)	0 (0%)	15 (58%)
hålla ngn i schack	20 (11%)	17 (23%)	2 (3%)	1 (5%)	0 (0%)	1 (3%)	13 (50%)

av grupperna. Även detta uttryck hör till de tio uttryck i enkäten som har fått de lägsta poängen (se vidare avsnitt 6.6 om den kvalitativa analysen) men skillnaden i förhållande till *hålla någon i schack* är alltså att jämförelsegruppens resultat inte skiljer sig signifikant från L1-gruppens resultat.

Frågan om vilka faktorer som påverkar resultatet för de enskilda uttrycken är inte heller enkel att besvara för denna del av enkäten. Frekvensuppgifter i PAROLE och Google för de högst och lägst placerade uttrycken ger inga entydiga svar. Uttrycket *ta saken i egna händer*, som har fått högst poäng i C-delen, förekommer 25 gånger i PAROLE (söksträng: "saken i egna händer") och ca 550 000 gånger på Google ("saken i egna händer"). Uttrycket *hålla någon i schack*, som har fått lägst poäng i hela materialet, förekommer 56 gånger i PAROLE och får ca 46 000 träffar på Google ("i schack", Språkbanken 2010a, hålla "i schack" Google 2010a). De låga resultaten för de två ovannämnda uttrycken kan tänkas relateras till kompositionalitet och genomskinlighet; *ta saken i egna händer* är semantiskt mer genomskinligt och i viss mån även mer kompositionellt än uttrycket *hålla någon i schack*. Ordet *schack* är dessutom polysemt och kan tolkas bokstavligt som 'schackspel', om informanten inte identifierar den ofullständiga frasen (_____ *någon i schack*) som ges i enkäten som just det konventionaliserade uttrycket *hålla någon i schack*. Det senare tas även upp i samband med den kvalitativa analysen i avsnitt 6.6 nedan.

Det finns sammanlagt 3 uttryck som utmärker sig genom att hela grupp L2/sen når nollresultat i samband med dem. Uttrycken är *rinna ut i sanden*, *på rak arm* och *få upp ögonen för något*. Alla tre tillhör alltså de uttryck för vilka det finns en signifikant resultatskillnad mellan första- och andraspråkslever. När det gäller *rinna ut i sanden*, är skillnaden mellan grupp L1 (78 %) och andraspråksgrupperna L2/tidig och L2/sen (20 % respektive 0 %) tydlig. Jämförelsegruppen når 88 % av maximalt resultat för detta uttryck. Uttrycket *rinna ut i sanden* verkar alltså vara betydligt svårare för andraspråksleverna än för eleverna med svenska som modersmål. De andra två uttrycken får däremot relativt låga resultat även bland grupp L1 och jämförelsegruppen; *på rak arm* får 43 % respektive 50 % och *få upp ögonen för något* får 38 % respektive 58 % av maximalt resultat. Ser man på frekvensen för uttrycket *rinna ut i sanden*, hittar man 23 förekomster i PAROLE (söksträng: "ut i sanden") och

drygt 1,1 miljon träffar på Google (på samma söksträng). *På rak arm* förekommer endast 11 gånger i PAROLE ("*på rak arm*"). Däremot får man drygt 3,5 miljoner träffar på Google på söksträngen "*på rak arm*". Uttrycket *få upp ögonen för något* förekommer 67 gånger i PAROLE (söksträng: "*upp ögonen för*") och ca 930 000 gånger på Google (samma söksträng) (Språkbanken 2010a, Google 2010a). Frekvensuppgifterna är alltså även i detta fall svårtolkade. Inget av uttrycken är så lågfrekvent i referensmaterialen att man, mot bakgrund av Ellis et al. (2008), skulle kunna tolka det som att frekvensen påverkar andraspråkelevernas resultat för dessa uttryck. När det gäller uttrycken *på rak arm* och *få upp ögonen för något* är frekvenserna – i alla fall på Google – relativt höga men resultaten för alla informantgrupperna relativt låga. Här lär alltså andra faktorer ligga till grund för resultaten. *På rak arm* kan ses som icke-kompositionellt och ogenomskinlig medan *få upp ögonen för något* är semantiskt förhållandevis mer genomskinligt och till viss del mer kompositionellt. En annan faktor som skulle kunna ha betydelse är informanternas ålder. Man kan tänka sig att dessa uttryck helt enkelt inte är speciellt frekventa i den aktuella åldersgruppens språkbruk.

Sammanfattningsvis framgår det av de resultat som redovisats i detta avsnitt att det finns en signifikant skillnad i resultat mellan första- och andraspråkstalare för nästan hälften av uttrycken i C-delen. Man kan tolka detta som att förstaspråkstalare överlag har betydligt lättare för uppgiften att komplettera ett konventionaliserat uttryck utan hjälp av kontexten, vilket naturligtvis inte är ett oväntat resultat. Resultaten för de enskilda uttrycken varierar dock rätt mycket, även för de elever som har svenska som modersmål.

6.4.4. Kollokationer

I den sista delen av enkäten testas elevernas behärskning av sammanlagt 20 olika kollokationer. Av tabell 8.3 framgår att skillnaderna mellan grupperna inte är lika stora i denna del av materialet som i C-delen. Intressant är att skillnaden mellan grupp L1 och grupp L2/tidig i 16 av 20 fall inte är statistisk signifikant. Här skiljer sig alltså resultatet från den detaljerade analysen från det sammanlagda resultatet för hela del D (se avsnitt 6.3), där skillnaden mellan grupp L1 och grupp L2/tidig är den enda som är statistiskt signifikant i en jämförelse mellan alla informant-

TABELL 8.3: Resultat på matchningstestet i del D per uttryck (i fallande ordning)

	SUMMA alla informanter	GRUPP L1	GRUPP L2/tidig (startålder ≤ 6)	GRUPP L2/ sen (startålder > 6)			JK
				startålder ≤ 10		totalt	
				18	19		
antal elever	186	74	75	18	19	37	26
UTTRYCK							
allvarlig sjukdom	157 (84%)	69 (93%)	59 (79%)	15 (83%)	14 (76%)	29 (78%)	25 (96%)
vara/komma till nytta	149 (80%)	67 (91%)	58 (77%)	13 (72%)	11 (57%)	24 (65%)	24 (92%)
kasta en blick	147 (79%)	68 (92%)	54 (72%)	14 (78%)	11 (58%)	25 (67%)	25 (96%)
svårt sjuk	137 (74%)	62 (84%)	55 (73%)	12 (66%)	8 (48%)	20 (54%)	24 (92%)
få i uppgift	131 (70%)	65 (88%)	47 (63%)	10 (55%)	9 (48%)	19 (51%)	22 (85%)
olika slags	130 (70%)	61 (82%)	51 (68%)	8 (44%)	10 (48%)	18 (49%)	22 (85%)
djupt anderlag	127 (68%)	49 (65%)	52 (69%)	13 (72%)	13 (67%)	26 (70%)	22 (85%)
full fart	127 (68%)	61 (81%)	48 (64%)	11 (61%)	7 (43%)	18 (49%)	20 (77%)
evigt tacksam	123 (66%)	64 (86%)	41 (55%)	12 (66%)	6 (38%)	18 (49%)	25 (96%)
hamna i missbruk	123 (66%)	55 (74%)	46 (61%)	12 (66%)	10 (57%)	22 (59%)	24 (92%)
knappast troligt	119 (64%)	64 (86%)	40 (53%)	9 (50%)	7 (43%)	16 (43%)	23 (88%)
stor betydelse	116 (62%)	56 (76%)	43 (57%)	8 (44%)	9 (52%)	17 (46%)	25 (96%)
ge ett exempel	124 (66%)	57 (76%)	41 (55%)	14 (78%)	12 (67%)	26 (70%)	19 (73%)
bryta ett beroende	106 (57%)	51 (69%)	38 (51%)	8 (44%)	9 (52%)	17 (46%)	19 (73%)
politiskt aktiv	96 (52%)	50 (68%)	32 (43%)	9 (50%)	5 (33%)	14 (38%)	22 (85%)
bilda en union	97 (52%)	43 (58%)	38 (51%)	8 (44%)	8 (43%)	16 (43%)	19 (73%)
gå i arv	82 (44%)	44 (59%)	28 (37%)	6 (33%)	4 (21%)	10 (27%)	18 (69%)
dra en suck	71 (38%)	38 (51%)	21 (28%)	6 (33%)	6 (38%)	12 (32%)	20 (77%)
ligga till grund	60 (32%)	33 (45%)	18 (24%)	6 (33%)	3 (14%)	9 (24%)	10 (38%)
lyckligt ovetande	53 (28%)	36 (48%)	12 (16%)	2 (11%)	3 (14%)	5 (13%)	24 (92%)

grupperns resultat. Däremot finns det även i denna del av analysen en tydlig tendens, nämligen att grupp L1 alltid når högre resultat än samtliga andraspråksgrupper (se tabell 8.3). Skillnaden mellan förstaspråkseleverna och grupp L2/tidig är statistiskt signifikant för uttrycken få i uppgift, evigt tacksam, knappast troligt och lyckligt ovetande (p-värden <0.0001). Ett annat intressant mönster som framgår av tabell 8.3 är att skillnaden mellan grupp L2/tidig och de tidigare inlärarna i grupp L2/sen endast är signifikant i samband med kollokationen *olika slags* (p-värde = 0.0023). Ser man närmare på dessa två gruppers resultat i tabell 8.3, kan man konstatera att de tidigare inlärarna i grupp L2/sen dessutom når högre resultat än grupp L1 i vissa fall, t.ex. när det gäller *evigt tacksam* och *ge ett exempel*. Dessa skillnader är dock, som sagt, inte signifikanta.

En tendens som är konstant över samtliga deltester är att gruppen med startålder >10 når lägst resultat i jämförelse med övriga startåldersgrupper (samt jämförelsegruppen). I den detaljerade analysen av kollokationstestet skiljer sig resultaten för de sena inlärarna i L2/sen dock endast i 7 av 20 fall signifikant från resultaten för de övriga startåldersgrupperna. Detta gäller uttrycken *vara/komma till nytta* (p-värde = 0.0061), *kasta en blick* (p-värde = 0.0015), och *svårt sjuk* (p-värde = 0.0021), överst i tabell 8.3, samt *full fart* (p-värde = 0.0012) och *evigt tacksam* (p-värde <0.0001) i mitten av tabellen, och *politiskt aktiv* (p-värde = 0.0013) samt *gå i arv* (p-värde = 0.0046), som återfinns i nedre halvan av tabell 8.3. En skillnad i förhållande till resultaten för de enskilda uttrycken i delarna B och C är att det, i likhet med del A, inte finns några nollresultat för någon av kollokationerna.

I nästan hälften av fallen (9 av 20) finns det inga signifikanta skillnader i resultat mellan grupperna över huvud taget. Så likartade resultat förekommer inte i lika stor utsträckning i de övriga deltesten. Detta gäller också kollokationen *allvarlig sjukdom* som får den högsta poängen sammanlagt och därmed hamnar högst upp i tabell 8.3. Alla grupper når relativt höga resultat för detta uttryck, mellan 76 % och 93 %. När det gäller det uttryck som ligger näst sist i tabellen, *ligga till grund*, finns det heller inga signifikanta resultatskillnader mellan startåldersgrupperna. Något som är anmärkningsvärt med detta resultat är att jämförelsegruppen, vars resultat i övrigt ligger mellan 69 % och 96 %, endast når 38 % av maximalt resultat för denna kollokation.

Orsakerna till den relativt stora resultatskillnaden för dessa två kollokationer kan man återigen bara spekulera om. Ser man på uttryckens frekvens i PAROLE och på Google, kan man konstatera att den inte framstår som en avgörande faktor, då frekvensen för det näst sist placerade uttrycket *ligga till grund* är betydligt högre (193 på ”*ligga till grund*”³ i PAROLE, drygt 380 000 miljoner på *ligga till grund* på Google) än den för förstplacerade *allvarlig sjukdom*, som endast förekommer 7 gånger i PAROLE och ca 45 000 gånger på Google (”*allvarlig sjukdom*”, Språkbanken 2010a, ”*allvarlig sjukdom*”, Google 2010a). Det sist placerade uttrycket *lyckligt ovetande* får ungefär lika låga frekvenser som *allvarlig sjukdom* i referensmaterialen (4 i PAROLE och ca 28 000 på Google, söksträng: ”*lyckligt ovetande*”). Kompositionalitet kan däremot förmodas påverka resultatet, speciellt i fråga om *ligga till grund* som kan sägas vara både semantiskt betydligt mindre genomskinligt och mindre kompositionellt än kollokationen *allvarlig sjukdom*. Detta gäller till viss del även *lyckligt ovetande*, även om skillnaden inte är lika tydlig i det fallet. Betydelsen hos *ligga till grund* kan dessutom sägas vara tämligen abstrakt.

Utgår man, som i avsnitt 4.4, ifrån van Liers (1995) hypotes om skillnaden mellan första- och andraspråksbrukare, när det gäller att bilda de mest idiomatiska kollokationerna av ett antal givna ord, är resultatet av den detaljerade analysen som beskrivits här något oväntat. Signifikanta skillnader mellan grupp L1 och någon av andraspråksgrupperna föreligger endast i drygt hälften (11 av 20) av fallen. I 7 av dessa handlar det dessutom enbart om en signifikant skillnad mellan grupp L1 och andraspråksleverna med den högsta startåldern bland informanterna. Signifikanta skillnader mellan de olika startåldersgrupperna bland andraspråksleverna är få, vilket ligger i linje med den hypotes som redovisas i avsnitt 4.4. Även om det för många av uttrycken alltså inte finns några signifikanta skillnader mellan grupp L1 och andraspråksgrupperna, visar tabell 8.3 en relativt tydlig tendens till att grupp L1 konstant

³ I detta fall har jag sökt på samtliga verbformer för att undvika träffar på ”lägga till grund”. Sökuttrycket har dessutom trunckerats för att även få träff på former med ett eller två ord emellan verbet och prepositionen. Denna trunkering gav sammanlagt 9 (av 193) träffar.

presterar bättre för de enskilda uttrycken än de elever som har ett annat modersmål än svenska.

Det kan också konstateras att det, i likhet med de detaljerade resultaten för de övriga deltesten, finns få signifikanta resultatskillnader mellan jämförelsegruppen och grupp L1 i den detaljerade analysen av kollokationstestet; det är endast i samband med uttrycken *dra en suck* (p-värde = 0.0263), *stor betydelse* (p-värde = 0.0234) och *lyckligt ovetande* (p-värde <0.0001) som jämförelsegruppen når ett signifikant högre resultat (77 %, 96 % respektive 92 %) än grupp L1 (51 %, 76 % respektive 48 %). Även på detta sätt skiljer sig alltså resultatet av den detaljerade analysen från analysen av de samlade delresultaten för del D där jämförelsegruppen når ett signifikant högre resultat än grupp L1 i alla fyra delarna av materialet. Som i de övriga delarna kan man dock även i analysen av de enskilda uttrycken i D se en tendens till att jämförelsegruppen når konstant högre resultat än grupp L1 (förutom när det gäller resultaten för *ligga till grund*).

Slutligen kan noteras att de mer lågfrekventa kollokationer som tagits med i undersökningen hamnar relativt långt ned i tabellen, om än inte längst ned. Kollokationen *dra en suck* ligger med 71 poäng bland de sista tre uttrycken i tabellen, *bilda en union* och *bryta ett beroende* ligger med 97 respektive 106 poäng något högre upp. *Hamna i missbruk* ligger med 123 högst upp av de fyra lågfrekventa uttrycken. Den låga frekvensen i referensmaterialet framstår alltså knappast som den avgörande faktorn för resultaten i samband med de aktuella uttrycken – även om man kan tänka sig att den har viss betydelse – då det finns mer frekventa uttryck som fått lägre poäng.

6.5. Validitet

Som nämndes i avsnitt 5.5 har testmetodens validitet testats med hjälp av Pearsins korrelationskoefficienten. I detta syfte har korrelationer mellan resultaten för respektive uttryck och det totala resultatet samt delresultaten (A–D) för hela informantgruppen beräknats. Dessutom har medelvärden för resultaten på hela enkäten mellan de informanter som behärskar, och de som inte behärskar några av de svåraste respektive lättaste uttrycken (dvs. de uttryck som få respektive många informanter

har kompletterat på det förväntade sättet), jämförts varpå skillnaderna har testats för signifikans. Båda dessa analyser ger statistiskt sett samma p-värden och tjänar alltså till att synliggöra vissa statistiska samband; om resultatet för uttryck x korrelerar med hela informantgruppens resultat dels på hela enkäten, dels på de olika deltesterna, innebär det att många av informanterna som har nått höga resultat totalt har rätt på den enskilda uppgiften ifråga. Likaså, om de informanter som har kompletterat ett svårt uttryck på det förväntade sättet sammantaget når ett signifikant högre resultat på hela enkäten samt på de olika deltesterna än den mindre gruppen av informanter som avviker från det förväntade svaret, visar det på att det finns ett samband mellan informanternas positiva resultat för det enskilda uttrycket och deras högre resultat totalt. Samma gäller för de elever som inte har kompletterat ett lätt uttryck på det förväntade sättet; om deras resultat är signifikant lägre än resultaten för den större elevgruppen som besvarat uppgiften på det förväntade sättet, tyder det på att det finns ett samband mellan det avvikande svaret på den enskilda uppgiften och lågt resultat på hela enkäten respektive de olika deltesterna.

De samband som beskrivs ovan visar sig vara starkt signifikanta i det stora flertalet av fallen (i de flesta av dessa fall är r-värdet >0.25 och p-värdet <0.0001). Det uttrycket som sticker ut i detta sammanhang är *få något på halsen* där korrelationerna mellan resultatet för uttrycket och det totala resultatet samt delresultaten är svaga (r-värde totalt = 0.05833) och det inte kan påvisas en signifikant skillnad mellan medelvärdet av totalt resultat för de 126 informanter som behärskar uttrycket och medelvärdet för de 61 informanterna som inte gör det (p-värde totalt = 0.4278, jfr även tabell 8). Överlag kan dock konstateras att det finns en genomgående tendens resultaten emellan och att testuppgifterna verkar uppfylla syftet att diskriminera mellan de informanter som behärskar många av de konventionaliserade uttrycken ifråga och de som inte gör det.

6.6. Kvalitativ analys

Hittills har i första hand kvantitativa resultatskillnader mellan informantgrupperna diskuterats. I detta avsnitt kommer en del av materialet

(se avsnitt 5.5) beskrivas kvalitativt. En av de frågor som ska besvaras i detta sammanhang är vilka typer av avvikelser från de konventionaliserade konstruktionerna i fråga som finns i materialet. Efter analys av några av enkäterna med lägst resultat i varje grupp – sammanlagt 45 – har samtliga avvikande svar klassificerats på ett av följande sätt:

- *inget svar* – när informanten inte har kompletterat uttrycket i fråga
- *vanlig kontamination* – när kontaminationen i fråga förekommer hos flera (minst 5) informanter
- *idiosynkratisk kontamination* – när kontaminationen endast förekommer hos enskilda (högst 4) informanter
- *icke-idiomatiskt* – när informanten fyllt i ett ord som är grammatiskt och semantiskt möjligt, men gör att frasen ifråga avviker från det förväntade konventionaliserade uttrycket
- *ogrammatiskt* – när ordet som informanten kompletterat med orsakar ett grammatiskt eller semantiskt problem i den aktuella kontexten

Alla dessa kategorier är inte lika relevanta för alla de tio uttrycken. Skillnaden mellan *vanliga* och *idiosynkratiska kontaminationer* görs då det visat sig finnas kontaminationer i materialet som är vanligt förekommande. Gränsen för vad som kan anses som vanligt har satts vid 5 förekomster. ”Vanlig” är i detta fall i hög grad att betrakta som ett relativt begrepp och den valda gränsen kan naturligtvis diskuteras. Vid mindre än 5 förekomster bör dock ett uttryck knappast kunnas anses som vanligt i det här aktuella materialet. Därtill kommer att man i vissa fall kan mistänka ett samarbete mellan två eller möjligtvis flera elever, något som har lett till identiska lösningar av uppgiften ifråga. Det är alltså heller inte rimligt att se alla kontaminationer som förekommer mer än en gång som en egen kategori. Dessutom bör man komma ihåg att det alltid finns en spekulativ aspekt i att klassificera en fraseologisk enhet som en kontamination av två eller flera olika konventionaliserade uttryck. Man kan knappast bevisa att informanten har associerat till andra befintliga uttryck och därvid blandat ihop flera enheter i samma fras. Man kan dock argumentera för att sannolikheten att man av ren slump väljer en eller flera komponenter från ett visst uttryck och blandar dem med

komponenter från ett annat konventionaliserat uttryck, som uppvisar likheter i form och/eller betydelse, inte borde vara så stor.

Kategorin *icke-idiomatisk* är bred och inte helt lätt att avgränsa. Konnotationer kan sägas vara specialfall av icke-idiomatisk avvikelse. Mer problematisk är däremot gränsdragningen mellan icke-idiomatiska och grammatiska avvikelser, särskilt i samband med abstrakta och i relativt hög grad lexikaliserade konventionaliserade uttryck som t.ex. *hålla någon i schack*. Här är det i många fall inte självklart om användningen av ett annat verb än *hålla* (som i samband med detta uttryck kan betraktas som tämligen betydelsestomt) bör ses som icke-idiomatisk eller grammatisk. De ovannämnda kategorierna diskuteras närmare i följande avsnitt.

Av de 45 informanterna vars enkäter har analyserats kvalitativt ingår 14 i grupp L1, 15 i grupp L2/tidig och 16 i grupp L2/sen.

6.6.1. De svåraste uttrycken

De tio konventionaliserade uttryck som har fått lägst poäng i materialet, och därmed kan sägas visat sig vara svårast för informanterna i denna undersökning, presenteras nedan i stigande ordning, efter det resultat som informantgruppen som helhet har uppnått för uttrycket:

- *hålla någon i schack* (20/186)
- *sätta käppar i hjulen* (36/186)
- *få upp ögonen för något* (37/186)
- *på rak arm* (39/186)
- *livet leker för någon* (40/186)
- *med något i högsta hugg* (41/186)
- *klättra på väggarna* (43/186)
- *få något/någon på halsen* (50/186)
- *lyckligt ovetande* (53/186)
- *inte för fem öre* (54/186)

Bortsett från *lyckligt ovetande* (del D) återfinns alltså samtliga uttryck i delarna B (4 av tio) och C (6 av tio) i enkäten. A-delen är inte representerad bland de uttryck som fått lägst resultat. Detta kan sägas återspegla det mönster som framgår av den kvantitativa analysen.

Genom att titta närmare på de avvikande svaren för de ovannämnda uttrycken kan man tänkas komma något närmare en förklaring till, varför just dessa uttryck vållar störst problem för de eleverna som deltagit i undersökningen.

6.6.1.1. Hålla någon i schack

TABELL 9: Antal olika typer av avvikelser i samband med uttrycket *hålla någon i schack* per informantgrupp

	Inget svar	Vanliga kontaminationer	Idio-synkratiska kontaminationer	Icke-idiomatiska svar	Ogrammatiska svar	Förväntade svar
L1 (15)	1	–	–	9	5	–
L2/tidig (14)	9	–	–	3	2	–
L2/sen (16)	10	–	–	1	5	–
Totalt	20	–	–	13	12	–

Hålla någon i schack är det uttryck som har fått lägst poäng i materialet som helhet (20 av 186 möjliga). Bland de 45 informanterna som inkluderats i den kvalitativa delen av analysen har 20 avstått från att komplettera fragmentet _____ *någon i schack*. 19 av dessa är andraspråkselever (9 i grupp L2/tidig, 10 i grupp L2/sen). Svartsbortfallet är bland det högsta för de uttryck som diskuteras här.

Det förekommer inga kontaminationer i samband med uttrycket *hålla någon i schack* i den delen av materialet som analyserats kvalitativt. Däremot finns det 13 kompletteringar som har klassificerats som icke-idiomatiska. Dessa fördelar sig samtliga på exemplen *slå någon i schack* (L2/tidig, L2/sen), *sätta någon i schack* (L1), *ta någon i schack* och *dra någon i schack*. Alla dessa exempel, förutom *dra någon i schack*, förekommer flera gånger. Ett exempel som sticker ut något i sammanhanget är användningen av verbet *slå*. Det förekommer sammanlagt 5 gånger i denna del av materialet, vilket inte är helt oväntat, med tanke på att frasen *slå någon i schack* är såväl grammatiskt som semantiskt-pragmatiskt

fullt möjlig att använda, även om det inte kan betraktas som ett konventionaliserat uttryck i den här avsedda bemärkelsen.

Endast 4 av de icke-idiomatiska svaren återfinns bland andraspråkseleverna (3i grupp L2/tidig och 1 i grupp L2/sen). Eleverna i grupp L1 står alltså för 9 avvikande svar i denna kategori. Man kan tolka de hittills nämnda resultaten som att de andraspråkelever som inte behärskar uttrycket snarare avstått från att besvara uppgiften, medan fler elever i grupp L1 har försökt att komplettera uttrycket, i många fall dock utifrån en annan betydelse än den som avses i enkäten.

12 kompletteringar har klassats som ogrammatiska. 5 av dessa svar förekommer i grupp L1. Förstaspråkseleverna har producerat formerna **spela någon i schack* (2 gånger), **explodera någon i schack*, **göra någon i schack* och **som någon i schack*. I grupp L2/tidig förekommer 2 ogrammatiska svar, nämligen **sätt någon i schack* och **spela någon i schack*. I grupp L2/sen förekommer 5 ogrammatiska svar, nämligen **hela någon i schack*, **är någon i schack* (2 gånger) och **vinna/vann någon i schack*.

Ingen av de 45 informanterna har kompletterat det ofullständiga uttrycket på det förväntade sättet.

En faktor som verkar vålla problem för både första- och andraspråkseleverna är polysemin i samband med substantivet *schack*. Valet av verb som *vinna*, *slå* eller *spela*, istället för det konventionaliserade *hålla*, kan antas bero på att informanterna associerar till 'schackspel'. I dessa fall ligger problemet alltså på det receptiva planet; då eleverna utgår ifrån en annan betydelse av konstituenten *schack*, avviker följaktligen även de ord som de associerar med denna konstituent från den efterlysta formen (jfr Nation 2001, se avsnitt 3.1). Den ogrammatiska formen **spela någon i schack* (istället för *spela mot någon*) skulle dessutom kunna ses som influerad av engelskan, där formen *to play someone at chess* är grammatiskt möjlig. Det förekommer också exempel där man kan förmoda ett avsiktligt nonsenssvar (förmodligen i brist på kunskap om det förväntade svaret). Hit kan t.ex. användningen av verbet *explodera* räknas.

6.6.1.2. Sätta käppar i hjulen

TABELL 9.1: Antal olika typer av avvikelser i samband med uttrycket *sätta käppar i hjulen* per informantgrupp

	Inget svar	Vanliga kontaminationer	Idio-synkratiska kontaminationer	Icke-idiomatiska svar	Ogrammatiska svar	Förväntade svar
L1 (15)	3	–	1	8	2	1
L2/tidig (14)	10	–	1	2	1	–
L2/sen (16)	5	–	3	3	5	–
Totalt	18	–	5	13	8	1

Sammanlagt 20 av de 45 informanterna har avstått från att komplettera fragmentet *sätta _____ i hjulen*. Även i detta fall har flertalet noll-svar, dvs. 19 av dem, lämnats av andraspråkseleverna (9 i grupp L2/tidig och 10 bland de sena inlärarna).

Det finns sammanlagt 5 svar som har klassats som kontaminationer i detta material. I dessa fall har informanterna kompletterat det ofullständiga uttrycket med *stopp* (L2/sen), *eld* (L2/tidig och L1), *spiken* (L2/sen). När det gäller de två förstnämnda lösningsförslagen kan man förmoda att kontaminationer med de konventionaliserade uttrycken *sätta stopp för något*, *sätta eld på något* kan vara en förklaring. Här är alltså verbet *sätta* den gemensamma nämnare som verkar orsaka kontaminationerna. Användningen av substantivet *spiken* skulle kunna bero på en kontamination med *slår (sista) spiken i kistan för något*, som i viss mån även förekommer i varianten *sätta spiken i kistan för något* (se Google 2010b, Språkbanken 2010b).

Det finns 13 svar i materialet som har klassats som icke-idiomatiska. 5 av dessa har producerats av andraspråkselever (2 i grupp L2/tidig och 3 i grupp L2/sen) och 8 av elever i grupp L1. Bl.a. har ordformerna *sig*, *pinnen*, *pinnar*, *bromsen* och *in* använts. Dessutom har formerna *käppen* (L1) och *käpparna* (L2/sen) används. Även dessa svar har klassats som icke-idiomatiska. Detta kan dock diskuteras; Sköldbberg (2004) har t.ex. visat att idiomat *sätta käppar i hjulen* varierar en hel del i språkbruket.

Hennes material uppvisar dock endast variation i numerus av lexemen *käpp* och *hjul*. Formerna som förekommer här är alltså *sätta en käpp i hjulen/hjulet* och *sätta käppar i hjulet/hjulen*. De bestämda formerna av substantivet *käpp* (*käppen, käpparna*) använts däremot inte (Sköldberg 2004:207). I anslutning till Sköldbergs resultat och för att inte vålla ytterligare gränsdragningsproblem i samband med analysen, har jag alltså valt att räkna användningen av bestämd form som avvikelser från den konventionaliserade formen (enligt Svenkt språkbruk 2003).

8 av svaren i samband med *sätta käppar i hjulen* har klassats som ogrammatiska. 2 av dessa förekommer i grupp L1, 1 i grupp L2/tidig och 5 i grupp L2/sen. Det handlar om användningen av formerna **snur* och **tjappen* (grupp L1); partiklarna *dit* (L2/tidig), *upp*, *ned* och *fast* (L2/sen) samt substantivet *pinne* och verbet *snurra* (grupp L2/sen). *Pinne* kan sägas ligga ganska nära det efterlysta ordet i betydelse, men är inte en idiomatisk lösning på uppgiften. Avgörande för kategoriseringen är att användningen av denna ordform utan obestämd artikel bryter mot svenskan grammatiska regler. Även **sätta snurra i hjulen* är oförenligt med svenskans grammatik, oavsett om man betraktar *snurra* som ett substantiv eller ett verb. **Sätta tjappen i hjulen* skulle kunna tänkas vara ett resultat av stavfel och därmed ett försök att komplettera det ofullständiga uttrycket med ordformen *käppen*, vilket alltså i så fall snarare skulle vara ett exempel på ett icke-idiomatiskt svar. Eftersom ordformen dock i så fall skulle innehålla flera stavfel samtidigt, ligger det närmare till hands att tolka svaret som att informanten inte behärskar den korrekta formen och därför har producerat ett ogrammatiskt svar. Intressant nog förekommer just detta fall bland förstaspråkseleverna i grupp L1⁴.

Endast 1 av 45 informanter har svarat på det förväntade sättet. Informanten ingår i grupp L1. Uttrycket verkar alltså vålla problem för både första- och andraspråkselever. Svartsbortfallet är relativt stort, inte minst bland andraspråkseleverna. Man kan konstatera att det även förekommer en hel del avvikelser från den konventionaliserade formen

⁴ Här bör dock nämnas att informanten ifråga är speciell på det sättet att hon har angett svenska som enda modersmål och som det enda språk hon talar, men samtidigt skriver att hon inte är född i Sverige och att hon börjat lära sig svenska vid ettårsåldern.

bland förstaspråkseleverna. Det förekommer också ett antal svar som visar att informanten i fråga har partiell kunskap om uttrycket. Härtill kan svar som *sätta pinnar/pinnen i hjulen* och kanske även *sätta tjappen i hjulen räknas*. När det gäller svar som *sätta kätten/käpparna i hjulen* måste man utgå ifrån att informanterna i fråga känner till uttrycket, även om de använder en från den ”mest konventionaliserade” något avvikande form. Det senare är, som antytts tidigare, inte förvånansvärt med tanke på den variation som förekommer i samband med *sätta käppar i hjulen* (jfr Sköldberg 2004).

6.6.1.3. Få upp ögonen för något

TABELL 9.2: Antal olika typer av avvikelser i samband med uttrycket *få upp ögonen för något* per informantgrupp

	Inget svar	Vanliga kontaminationer	Idio-synkratiska kontaminationer	Icke-idio-matiska svar	Ogram-matiska svar	För-väntade svar
L1 (15)	8	–	4	1	1	1
L2/tidig (14)	10	–	2	1	1	–
L2/sen (16)	10	–	1	4	1	–
Totalt	28	–	7	6	3	1

Det förekommer sammanlagt 28 nollsvaret i samband med uttrycket *få upp ögonen för något*. Detta är det största svarsbortfallet bland de tio uttrycken som tas upp i den kvalitativa analysen. Även i detta fall står andraspråksgруппerna för större delen av svarsbortfallet (tio i grupp L2/tidig och 10 i L2/sen mot 8 i L1).

7 kompletteringar av *få upp _____ för något* har klassificerats som kontaminationer. 4 av dessa finns i grupp L1, nämligen exemplen *få upp huvudet för något*, *få upp hoppet för något*, *få upp modet för något*, *få upp hakan för något* samt *få upp den för något*. Det första exemplet kan ses som en kontamination med uttrycket *hålla huvudet högt*, som förekommer i A-delen av enkäten, vilket skulle kunna ha bidragit till att informanten blandat in en komponent från detta konventionaliserade ut-

tryck. Bakom det andra exemplet kan man förmoda en kontamination med uttrycken *ge upp hoppet* eller *hålla hoppet uppe*. Det tredje exemplet har analyserats som kontamination av *få upp ögonen för något* och det konventionaliserade uttrycket *hålla modet uppe*, som också förekommer i enkäten (del B). *Få upp hakan för något* kan ses som en kontamination med frasen *upp med hakan!* När det gäller det sistnämnda exemplet, finns det visserligen ingen kontext som kan indikera vad pronomenet *den* syftar på, vilket gör följande tolkning om möjligt ännu mer spekulativ än de övriga kontaminationsanalyserna. Då informanten är en 16-17-årig gymnasieelev, är det dock inte helt orimligt att förmoda en kontamination med frasen *få upp den* (som ofta används negerat i vardagligt språk, där *inte få upp den* betyder 'ha erektionsproblem'). Vid denna typ av svar ligger det nära till hands att utgå ifrån en medveten avvikelse från den konventionaliserade formen. Huruvida informanten behärskar det konventionaliserade uttrycket i fråga eller inte kan man i sådana fall bara spekulera om.

I grupp L2/tidig förekommer 2 svar som har klassats som kontaminationer. Eleverna har använt ordformerna *saken* och *humöret* för att komplettera det ofullständiga uttrycket ifråga. Formerna *få upp saken för något* och *få upp humöret för något* skulle kunna ses som kontaminationer med uttrycken *ta upp saken (med någon)* respektive *hålla humöret uppe*. Det senare uttrycket förekommer i B-delen av enkäten.

Ett av svaren i grupp L2/sen har klassats som kontamination. Även i detta fall har eleven använt substantivet *hakan* som komplettering.

6 av 45 svar i samband med uttrycket *få upp ögonen för något* har klassats som icke-idiomatiska. Ett av dessa förekommer i grupp L1. Frasen som eleven har producerat är *få upp *u*en för något* (som har tolkats som '*få upp [kuken] för något*'). Exemplet liknar det tidigare nämnda exemplet *få upp den för något*, som har räknats som kontaminerad form, då *få upp den*, i alla fall i vardagligt språk, kan ses som konventionaliserat uttryck (jfr Google 2010a). Även i fallet *få upp kuken för något* kan man förmoda en medvetet producerad avvikelse från den konventionaliserade formen ifråga. I grupp L2/tidig har 1 informant använt pronomenet *det* för att komplettera det ofullständiga uttrycket. De övriga 4 icke-idiomatiska svaren återfinns hos informanterna i grupp L2/sen. Här förekommer formerna *få upp glöd för något*, *få upp någon för något* (2 gånger) och *få upp det för något*.

3 av de 45 svaren i föreliggande del av materialet har klassificerats som ogrammatiska. En av informanterna i grupp L1 har kompletterat fragmentet med verbet *stå*, vilket resulterar i formen **få upp stå för något*. En alternativ kategorisering skulle i detta fall kunna vara den som kontamination av *få upp ögonen för något* med uttrycket *stå för något*.

I de övriga 2 fallen har informanterna använt reflexivt personligt pronomen (L2/tidigt och L2/sen), vilket resulterar i **få upp sig/mig för något*. De exempel som redovisats här tyder på att flertalet informanter har fokuserat på konstituenten *upp* vid försöket att hitta rätt komplettering till det aktuella uttrycksfragmentet. Det sista ledet i frasen, *för något*, verkar däremot inte ha beaktats i särskilt hög utsträckning i en del av exemplen.

Endast 1 av 45 informanter har kompletterat fragmentet *få upp _____ för något* på det förväntade sättet. Även i detta fall ingår informanten i grupp L1.

6.6.1.4. På rak arm

TABELL 9.3: Antal olika typer av avvikelser i samband med uttrycket på *rak arm* per informantgrupp

	Inget svar	Vanliga kontaminationer	Idio-synkratiska kontaminationer	Icke-idiomatiska svar	Ogrammatiska svar	Förväntade svar
L1 (15)	1	10	–	–	2	2
L2/tidigt (14)	6	4	1	2	1	–
L2/sen (16)	9	4	–	–	3	–
Totalt	16	18	1	2	6	2

16 av 45 informanter har inte kompletterat fragmentet *på rak _____*. I grupp L1 förekommer endast ett nollsvar. Andraspråkeleverna i grupp L2/tidigt står för 6, och de i grupp L2/en för 9 av de övriga nollsvaren.

Den kontaminerade formen *på rak sak* förekommer sammanlagt 9 gånger i denna del av materialet (5 i grupp L1, 2 i grupp L2/tidigt samt 2 i grupp L2/sen). Det ligger nära till hands att en kontamination av de

konventionaliserade uttrycken *på rak arm* och *rakt på sak* ligger till grund för denna avvikelse, då dessa uttryck dels har två lexikala komponenter gemensamma, dels till viss del ligger nära varandra i betydelsen. En annan vanlig kontamination, som förekommer 6 gånger (5 i grupp L1, 1 i grupp L2/tidig) är den mellan *på rak arm* och uttrycket *på egen hand* (det senare förekommer också i B-delen av enkäten), som resulterar i formen *på rak hand*. I detta fall kan det även tänkas spela roll att substantivet *hand* ligger nära *arm* på ett semantiskt plan. Här kan man, med Hyltenstams terminologi, tala om en approximation (eng. *close approximation*) till målformen (Hyltenstam 1992, se avsnitt 4.4.1). En kontaminerad form som förekommer 5 gånger och därmed också klassats som vanlig, är *på rak väg* (2 gånger i L1, en gång i L2/tidig, en gång i L2/sen). I detta fall förmodas en kontamination med uttrycket *raka vägen* ligga till grund för avvikelsen. I samband med uttrycket *på rak arm* förekommer alltså tre olika typer av vanliga kontaminationer, vilket är ovanligt i det kvalitativt analyserade materialet. En komplettering som har räknats som idiosynkratisk kontamination är *på rak sikt* (L2/tidig). Denna form kan ses som en kontamination mellan *på rak arm* och *på sikt*.

Det förekommer 2 icke-idiomatiska svar i de 45 enkäter som inkluderats i denna del av analysen, båda i grupp L2/tidig. Exempelen är *på rak rygg* och *på rak sida*.

Sammanlagt 6 av svaren har klassats som ogrammatiska. I grupp L1 har substantivet *kurva* använts av 2 informanter (som förmodas att ha samarbetat kring svaren) som komplettering. Formen **på rak kurva* har klassats som grammatiskt inkorrekt då substantivet *kurva* vanligtvis inte kan föregås av prepositionen *på*. I grupp L2/tidig förekommer den inkongruenta formen **på rak tal*. I grupp L2/sen har bl.a. formen **på rak spår* bildats. Även här är den valda kompletteringen alltså inkongruent med den givna formen *rak*. I detta fall kan man förmoda en kontamination med uttrycket *gå raka spåret*. I L2/sen förekommer också formerna **på rak saken* och **på rak fram*. Bakom det senare exemplet kan man förmoda en kontamination med adverbialt *rakt fram*. Vid klassificeringen har dock den ogrammatiska formen fokuserats (jfr analysen av *få upp stå för ngt* i 6.6.1.4).

2 informanter, som båda ingår i grupp L1, har svarat på det förväntade sättet. Som för de uttryck som diskuterats tidigare, är det alltså en-

dast ett mindre antal förstaspråkslever som klarar av att bilda det konventionaliserade uttrycket ifråga.

6.6.1.5. Livet leker för någon

TABELL 9.4: Antal olika typer av avvikelser i samband med uttrycket *livet leker för någon* per informantgrupp

	Inget svar	Vanliga kontaminationer	Idiosynkratiska kontaminationer	Icke-idiomatiska svar	Ogrammatiska svar	Förväntade svar
L1 (15)	4	–	–	3	4	4
L2/tidig (14)	7	–	–	2	4	1
L2/sen (16)	9	–	1	1	5	–
Totalt	20	–	1	6	13	5

När det gäller uttrycket *livet leker för någon*, har 20 av de 45 informanterna avstått från att komplettera det givna fragmentet *livet _____ för någon* (4 i grupp L1, 7 i L2/tidig och 9 i L2/sen).

Det förekommer endast 1 (idiosynkratisk) kontamination i samband med det aktuella uttrycket, nämligen formen *livet faller för någon* (L2/sen), som kan förmodas vara en sammanblandning av *livet leker för någon* och uttrycket *falla för någon/något*.

Däremot finns det 6 svar i denna del av materialet som klassats som icke-idiomatiska (3 i grupp L1, 2 i L2/tidig och 1 i L2/sen). I grupp L1 har informanterna bildat formerna *livet klarnar/exploderar/föds för någon*. Alla dessa kompletteringar är grammatiskt möjliga. När det gäller användningen av verbet *explodera*, kan man däremot diskutera om det är möjligt på det semantiska planet. Man skulle dock kunna tänka sig en slags överförd betydelse av *explodera* och formen har därmed inte klassats som ogrammatisk. Liknande gäller för användningen av *föds*. I grupp L2/tidig har formerna *livet rycker* respektive *är viktigt för någon* bildats. I det förstnämnda fallet är det något oklart vilken betydelse av ordet *rycker* informanten avser. Även här lär det handla om en överförd användning av verbet. Det andra exemplet är grammatiskt och seman-

tiskt fullt möjligt, men däremot inte konventionaliserat. Det förekommer även ett icke-idiomatiskt exempel i grupp L2/sen, nämligen frasen *livet kämpar för någon*. I detta fall skulle man kunna argumentera för en klassificering som ogrammatisk, då verbet *kämpa* vanligtvis kräver animat subjekt. Utgår man från en användning av *kämpa* i överförd betydelse, är exemplet dock inte mindre grammatiskt än det efterlysta konventionaliserade uttrycket *livet leker för någon*, då även verbet *leka* i bokstavlig betydelse kräver ett animat subjekt. Exemplet kan däremot inte anses som idiomatiskt i föreliggande form (jfr användningen av *explodera* och *födast* ovan).

Sammanlagt 13 av 45 svar i samband med uttrycket *livet leker för någon* har klassificerats som ogrammatiska. 4 av dessa har producerats av elever i grupp L1. Formerna som bildats här är **livet är för någon* (3 gånger) och **livet för för någon*. I grupp L2/tidig förekommer 4 ogrammatiska svar. Eleverna har bildat fraserna **livet är för någon* (2 gånger), **livet viktigt för någon* och **livet hör för någon*. De sammanlagt 5 ogrammatiska exempel som producerats i grupp L2/sen är samliga förekomster av **livet är för någon*. Användningen av verbet *vara* är alltså relativt vanlig. Även här resulterar kompletteringen i ett frasfragment och anses därför ogrammatisk.

5 av informanterna har besvarat uppgiften på det förväntade sättet genom att använda verbet *leka*, 4 i grupp L1 och 1 i grupp L2/tidig. *Livet leker för någon* tillhör alltså de relativt få konventionaliserade uttrycken i den kvalitativa analysen, som någon av andraspråkseleverna har bildat på det förväntade sättet.

6.6.1.6. Med något i högsta hugg

TABELL 9.5: Antal olika typer av avvikelser i samband med uttrycket *med något i högsta hugg* per informantgrupp

	Inget svar	Vanliga kontaminationer	Idiosynkratiska kontaminationer	Icke-idiomatiska svar	Ogrammatiska svar	Förväntade svar
L1 (15)	3	3	1	1	5	2
L2/tidig (14)	8	4	1	–	1	–
L2/sen (16)	7	4	2	1	2	–
Totalt	18	11	4	2	8	2

När det gäller kompletteringen av fragmentet *med något i högsta _____* har 18 av 45 informanter avstått från att besvara uppgiften. I grupp L1 finns 3 icke-kompletterade fragment, samt 8 i grupp L2/tidig och 7 i L2/sen.

Vidare finns det en typ av vanlig kontamination som förekommer sammanlagt 11 gånger i denna del av materialet, formen *med något i högsta grad*, som här antas vara resultatet av en kontamination av *med något i högsta hugg* och det konventionaliserade uttrycket *i (allra) högsta grad*. Denna kontaminerade form har bildats av 3 informanter i grupp L1, 4 i grupp L2/tidig och 4 i grupp L2/sen. Dessutom förekommer 4 idiosynkratiska kontaminationer (en i grupp L1, 1 i grupp L2/tidig och 2 i grupp L2/sen). I grupp L1 återfinns formen *med något i högsta beredskap*, som kan vara en kontamination med uttrycket *varalligga i högsta beredskap*. I detta fall kan man notera att det finns viss semantisk likhet mellan de kontaminerade formerna. I grupp L2/tidig har formen *med något i högsta domstolen* bildats, bakom vilken man kan förmoda en kontamination med den konventionaliserade nominalfrasen *högsta domstolen*. De två kontaminerade formerna i grupp L2/sen är *med något i högsta tid* och *med något i högsta kvalitet*. Här förmodas att kontaminationer med uttrycken *det är hög tid* och *vara av högsta kvalitet* ligger till grund för valet av respektive ord.

2 av svaren har klassificerats som icke-idiomatiska (ett i grupp L1 och ett i grupp L2/sen). 1 av eleverna i grupp L1 har använt substantivet *höjd* för att komplettera fragmentet ifråga. I grupp L2/sen har den icke-idiomatiska frasen *med något i högsta skolan* bildats.

8 av 45 svar i samband med uttrycket *med något i högsta hugg* har klassats som ogrammatiska. I grupp L1 har dessa kompletteringar resulterat i formerna **med något i högsta bomb*, **med något i högsta sker*, **med något i högsta hönset* och **med något i högsta strå* (2 gånger, informanterna antas ha samarbetat). När det gäller användningen av ordformen *hönset* skulle man kunna argumentera för en klassificering som kontamination med den konventionaliserade nominalfrasen *högsta hönset*. Eftersom formen är nonsensartad och inte heller grammatiskt välformad har den dock klassats som ogrammatisk (jfr ovan). I grupp L2/tidig har ett svar klassats som ogrammatiskt, nämligen formen **med något i högsta betyg*. Denna förekommer även i grupp L2/sen, tillsammans med den ogrammatiska frasen *med *med något i högsta skolan*.

Det finns endast 2 bland de 45 informanterna som har använt substantivet *hugg* och därmed bildat det konventionaliserade frasen *med något i högsta hugg*. Båda dessa informanter ingår i grupp L1.

6.6.1.7. Klättra på väggarna

TABELL 9.6: Antal olika typer av avvikelser i samband med uttrycket *klättra på väggarna* per informantgrupp

	Inget svar	Vanliga kontaminationer	Idiosynkratiska kontaminationer	Icke-idiomatiska svar	Ogrammatiska svar	Förväntade svar
L1 (15)	2	–	–	6	5	2
L2/tidig (14)	9	–	–	2	2	1
L2/sen (16)	9	–	–	5	2	–
Totalt	20	–	–	13	9	3

I likhet med de uttryck som diskuterats tidigare är den vanligaste ”avvikelsen” i samband med uttrycken *klättra på väggarna* ett uteblivet svar;

i 20 av 45 fall har informanterna alltså inte kompletterat detta uttryck alls. 18 av dessa 20 icke-svar förekommer i andraspråksgrupperna (9 i grupp L2/tidig, 9 i grupp L2/sen). Detta verkar alltså vara ett uttryck som vållar större problem för andraspråksleverna än för elever med svenska som förstaspråk.

13 av de 50 svaren i samband med *klättra på väggarna* har klassificerats som icke-idiomatiska (6 i grupp L1, 2 i grupp L2/tidig och 5 i grupp L2/sen). Några exempel är formerna *hänga, glo* och *äta* (L1), *stirra, slå* (L2/tidig) samt *måla* och *rita* (L2/sen) *på väggarna*.

Det förekommer inga kontaminationer i samband med *klättra på väggarna* i det kvalitativt analyserade materialet. Däremot finns 2 avvikelser som har klassats som ogrammatiska, som kan förmodas vara resultatet av något som liknar en kontamination. Två informanter i grupp L1 (som troligtvis har samarbetat) har kompletterat uttrycket med egennamnet *Karlsson*. Man kan tänka sig att det är Astrid Lindgrens barnboksfigur *Karlsson på taket* som ligger till grund för denna avvikelse, vilket innebär att svaret kan klassas som ett slags kontamination. *Karlsson på taket* är visserligen ett egennamn och skiljer sig därför från de konventionaliserade uttryck som är föremål för denna undersökning. Ordförbindelsen kan dock definitivt betraktas som konventionaliserad. Då svaret resulterar i nonsens-meningen

Det är mer underhållande än att se människor isolerade i ett hus *som *Karlsson på väggarna* för att de är uttråkade

har svaren dock klassats som ogrammatiska. I detta fall ligger det nära till hands att förmoda ett avsiktligt nonsens-svar. Om orsaken till detta sätt att svara är att eleverna inte behärskar det efterlysta konventionaliserade uttrycket, är däremot omöjligt att bedöma. Andra exempel på ogrammatiska svar i samband med uttrycket *klättra på väggarna* (sammanslagt 9) är ...*att se människor isolerade i ett hus *som att/hål* (L2/tidig), **klister/tavlor/bläddrar på väggarna* (L1). Användningen av verbet *bläddra* har klassats som ogrammatiskt då den resulterar i ett brott mot svenskans selektionsregler i den aktuella kontexten. Utgår man från ordets betydelse, kan valet verka helt obegripligt. Däremot finns en viss fonologisk och grafisk likhet med det efterlysta verbet *klättra*, vilket kan vara en möjlig förklaring till denna avvikelse.

Det finns 3 bland de 45 informanterna som har kompletterat uttrycket *klättra på väggarna* på det förväntade, konventionaliserade sättet. 2 av dem finns i grupp L1 och 1 i grupp L2/tidig.

6.6.1.8. Få någon/något på halsen

TABELL 9.7: Antal olika typer av avvikelser i samband med uttrycket *få någon/något på halsen* per informantgrupp

	Inget svar	Vanliga kontaminationer	Idiosynkratiska kontaminationer	Icke-idiosynkratiska svar	Ogrammatiska svar	Förväntade svar
L1 (15)	4	–	4	2	–	5
L2/tidig (14)	6	–	1	4	3	–
L2/sen (16)	8	–	3	–	5	–
Totalt	18	–	8	6	8	5

Även när det gäller uttrycket *få någon/något på halsen* är svarsbortfallet relativt högt. 18 av 45 informanter har inte kompletterat fragmentet *få en riktig pratkvavn på _____*. 14 av dessa är andraspråkselever (6 i grupp L2/tidig, 8 i L2/sen).

Det förekommer sammanlagt 8 idiosynkratiska kontaminationer i samband med *få någon/något på halsen* (4 i grupp L1, 1 i grupp L2/tidig och 3 i grupp L2/sen). Ett exempel från grupp L1 är *få en riktig pratkvavn på nacken*, som kan ses som kontamination med *ha (några år) på nacken*. Dessutom har ordformerna *köpet* och *hjärnan* använts i grupp L1. I dessa fall ligger det, på grund av uttryckens likartade struktur, nära till hands att utgå ifrån kontaminationer med uttrycken *få något på köpet* och *få något [t.ex. en låt] på hjärnan*. En av informanterna i grupp L1 har endast fyllt i bokstaven *g* för att komplettera uttrycksfragmentet ifråga. Detta skulle kunna ses som en möjlig kontamination av *få något på halsen* och uttrycket *ha något på g[ång]*. Vidare finns det 1 idiosynkratisk kontamination som förekommer sammanlagt 3 gånger i de båda andraspråksgруппerna. 2 elever i grupp L2/sen och 1 i grupp L2/tidig har använt ordformen *käft* för att komplettera det aktuella ut-

trycket. Här är alltså en kontamination med *få på käften* en möjlig förklaring. Dessutom förekommer 1 kontamination med uttrycket *få någon/något på kroken* i grupp L2/sen.

6 av svaren i samband med *få någon/något på halsen* har klassificerats som icke-idiomatiska. I grupp L1 har ordformerna *axlarna* och *axeln*, och i grupp L2/tidig *munnen*, *sig*, och *mig* (2 gånger) använts för att komplettera uttrycksfragmentet i fråga.

Ogrammatiska lösningar av den aktuella uppgiften förekommer endast bland andraspråkseleverna (sammanlagt åtta). Exempel på denna typ av avvikelse är svar som **få något på allmänt/att* (grupp L2/tidig) och **få något på människor/tungal kommentar* (grupp L2/sen).

5 informanter, som samtliga ingår i grupp L1, har bildat det konventionaliserade uttrycket på det förväntade sättet. Det förekommer alltså inga förväntade svar bland andraspråkseleverna i föreliggande del av materialet.

6.6.1.9. Lyckligt ovetande

TABELL 9.8: Antal olika typer av avvikelser i samband med uttrycket *lyckligt ovetande* per informantgrupp

	Inget svar	Vanliga kontaminationer	Idiosynkratiska kontaminationer	Icke-idiomatiska svar	Ogrammatiska svar	Förväntade svar
L1 (15)	3	–	–	7	1	4
L2/tidig (14)	8	–	–	3	3	–
L2/sen (16)	6	–	–	5	4	1
Totalt	17	–	–	15	8	5

Uttrycket *lyckligt ovetande* är den enda kollokation som hamnar bland de tio uttryck som fått lägst poäng i hela materialet. Sammanlagt 17 av 45 informanter har avstått från att komplettera fragmentet _____ *ovetande*. 3 av dessa nollsvar förekommer i grupp L1, 8 i grupp L2/tidig och 6 i grupp L2/sen.

Som tidigare nämnts består uppgiften i att para ihop sammanlagt 5 adjektiv med ett av 8 givna adverb, som i detta fall är *evigt*, *exakt*, *hemligt*, *knappast*, *lyckligt*, *negativt*, *politiskt* och *svårt*. Adjektiven, bortsett från *ovetande*, är *troligt*, *sjuk*, *aktiv* och *tacksam* (se även appendix 1). Det är följaktligen adverbena *negativt*, *exakt* och *hemligt* som inte bör användas alls i denna del av uppgiften.

Beskrivningen av olika typer av avvikelser är något vansklig i samband med tvåordskollokationer, då vissa av kategorierna som använts hittills är föga eller inte alls relevanta med tanke på hur uppgiften har ställts. Detta gäller i första hand kontaminationer. I många fall är det dock heller inte rimligt att klassificera avvikelser från kollokationen i fråga som ogrammatiska; när det gäller tvåordsuttryck av liknade struktur är det i de flesta fall en fråga om idiomacitet och språklig konvention när det gäller vilka ord som anses passa ihop. Med detta sagt, finns det dock några fall i materialet, där den komponent som informanten har valt är så oförenlig med adjektivet i fråga på ett semantiskt plan, att svaren klassats som ogrammatiska.

Resonemanget innebär att de flesta avvikelserna i samband med kollokationen *lyckligt ovetande*, bortsett från nollsvaren, har klassats som icke-idiomatiska. Av de sammanlagt 15 avvikelserna i denna kategori förekommer 7 i grupp L1, 3 i grupp L2/tidig och 5 i grupp L2/sen. Informanterna i grupp L1 har bildat formerna *svårt/politiskt/hemligt/evigt/knappast ovetande* (det sistnämnda exemplet förekommer 2 gånger). I grupp L2/tidigt förekommer *hemligt* och *politiskt* (2 gånger) *ovetande*. I grupp L2/sen har adverbena *knappast*, *politiskt* och *hemligt* (2 gånger) används för att komplettera den ofullständiga kollokationen.

De svar som innehåller adverbena *exakt* eller *negativt* har klassats som ogrammatiska då dessa adverb enligt svenskans semantiska regler och språkliga konventioner inte ingår i kollokationspotentialen för adjektivet *ovetande*. Dessa avvikelser förekommer sammanlagt 8 gånger. I grupp L1 är det bara *exakt* som använts som ogrammatisk komplettering. Informanterna i grupp L2/tidig har använt *exakt* en gång och *negativt* 2 gånger. I grupp L2/sen förekommer formen **negativt ovetande* 3 gånger och **exakt ovetande* en gång.

5 av 45 informanter har bildat den konventionaliserade formen *lyckligt ovetande* (4 i grupp L1 och 1 i grupp L2/sen). Det gör uttrycket till ett av dem som genererat flest förväntade svar bland de nio uttryck som

diskuterats hittills, vilket bekräftar det allmänna mönstret; ser man på hela materialet, har den aktuella kollokationen fått näst högsta antal poäng (53/186) bland de tio uttrycken som tas upp i den kvalitativa analysen.

6.6.1.10. Inte för fem öre

TABELL 9.9: Antal olika typer av avvikelser i samband med uttrycket *inte för fem öre* per informantgrupp

	Inget svar	Vanliga kontaminationer	Idio-synkratiska kontaminationer	Ickeidiomatiska svar	Ogrammatiska svar	Förväntade svar
				<i>ett femtio</i>		
L1 (15)	1	–	–	2	2	4
L2/tidig (14)	5	–	–	6	1	1
L2/sen (16)	5	–	–	2	6	3
Totalt	11	–	–	10	9	8

När det gäller uttrycket *inte för fem öre* har 11 av 45 informanter avstått från att besvara uppgiften. 1 av dessa ingår i grupp L1, 5 i grupp L2/tidig och 5 i grupp L2/sen.

Det förekommer inga kontaminationer i samband med detta uttryck. Däremot förekommer ett antal avvikelser som klassats som icke-idiomatiska och som alla är av samma typ. Denna typ av avvikelse består i användningen av ett annat räkneord än *fem* och förekommer hela 19 gånger i föreliggande del av materialet (4 i grupp L1, 7 i L2/tidig och 8 i L2/sen). De räkneord som används är *ett* (2 gånger i L1, 6 i L2/tidig, 2 i L2/sen) och *femtio* (2 gånger i L1, ett i L2/tidig, 6 i L2/sen). De avvikande räkneorden är den enda typen av icke-idiomatiska kompletteringar som förekommer i samband med uttrycket *inte för fem öre*. Man kan naturligtvis inte med säkerhet uttala sig om varför det är just *ett* eller *femtio öre* som används. Att man väljer en ännu mindre valör än *fem* är, med tanke på frasens betydelse, inte orimligt. Uttrycket används för att förstärka en negering, t.ex. för att betona att man 'inte alls tror på något, inte ens så lite som för det låga värdet av fem öre'. Om man endast

behärskar uttrycket partiellt, dvs. känner till dess betydelse men inte den konventionaliserade formen, kan det ligga nära till hands att ta till det lägsta värdet man kan tänka sig i öre. Här kan det dock noteras att en alternativ förklaring skulle kunna vara en kontamination med uttrycken *inte ett ruttet/rött öre*. En faktor som kan ligga till grund för användningen av *femtio* istället för *fem* är att *fem* och *femtio* är ganska lika till formen. En annan förklaring skulle kunna vara att informantens tankegång går via den bokstavliga betydelsen av konstruktionen *x öre*; då det nu-förtiden är femtio öre som är det lägsta myntet, är det logiskt att använda det i den aktuella kontexten.

Sammanlagt 8 av svaren i samband med *inte för fem öre* har klassats som ogrammatiska (4 i grupp L1, ett i L2/tidig och 3 i L2/sen). I grupp L1 har formerna **inte för korv öre*, **inte för ett korv öre*, **inte för mycket öre* och **inte för rätta öre* bildats. De två första exemplen kan förmodas beror på en kontamination med uttrycket *inte ett korvöre* (även här har dock klassificeringen baserats på den ogrammatiska formen). Exemplet **inte för en öre* i grupp L2/tidig avviker från den konventionaliserade formen både med avseende på idiomacitet och genuskongruens. I grupp L2/sen förekommer fraserna **inte för enda öre*, **inte för ända öre* och **inte för en ända öre*, som samtliga kan antas vara ett försök att bilda frasen *inte för ett enda öre*.

7 av 45 informanter har använt räkneordet *fem* och därmed kompletterat fragmentet *inte för _____ öre* på det förväntade sättet. 6 av dessa ingår i grupp L1 och 1 i grupp L2/tidig. Även här visar sig alltså samma tendens som i materialet som helhet; bland de tio uttrycken som ingår i den kvalitativa analysen har *inte för fem öre* fått högst poäng (54/186), sett till hela informantgruppen. Detta är alltså också fallet när det gäller de 45 enkäter som analyserats här.

6.6.2. De vanligaste avvikelserna

TABELL 10: Antal olika typer av avvikelser för samtliga tio uttryck per informantgrupp

	Inget svar	Kontaminationer			Ickeidiomatiska svar	Ogrammatiska svar	Förväntade svar
		vanliga	idiosynkratiska	alla			
L1 (15)	30	13	10	26	41	29	27
L2/tidig (14)	78	8	6	14	27	19	3
L2/sen (16)	78	8	10	18	29	35	1
Totalt	186	29	26	55	97	83	31

Som framgår av tabell 10, är den klart vanligaste ”avvikelsen” från de förväntade svaren i enkäten, att informanten har avstått från att komplettera uttrycksfragmentet i fråga. Det totala svarsbortfallet bland de 45 enkäterna som har analyserats är 186 av 450. Kontaminationer är betydligt mindre vanliga. Det finns sammanlagt 31 exempel på det som rubricerats som vanliga kontaminationer och 27 idiosynkratiska kontaminationer. De båda kontaminationstyperna är alltså ungefär lika vanliga i materialet (se tabell 10). De icke-idiomatiska svaren är, med sammanlagt 97 exempel, relativt vanligt förekommande. De ogrammatiska svaren hamnar med 81 förekomster på tredje plats. Att det finns relativt få exempel (sammanlagt 31) på de förväntade konventionaliserade uttrycken i detta material är givet, då lågt antal poäng har varit urvalskriteriet för de uttryck som analyserats här.

Jämför man antalet olika avvikelser i de tre informantgrupperna, visar sig några intressanta tendenser. I båda andraspråkgrupperna är svarsbortfallet betydligt högre än i grupp L1. Skillnaderna är däremot inte lika tydliga för övriga kategorier av avvikelser. Intressant är att kontaminationerna i det material som analyserats här verkar vara ungefär lika vanliga (t.o.m. något vanligare) bland förstaspråkseleverna som bland andraspråkseleverna, oavsett grupp. Detta ligger i linje med Hyltenstams tidigare refererade resultat (1992, jfr även Prentice & Sköldberg und. utg.). Som nämnts ovan är de icke-idiomatiska svaren de näst van-

ligaste efter nollsvaren. Antalet icke-idiomatiska svar i grupp L1 (41) är högre än i båda andraspråksgrupperna (26 i L2/tidig och 28 i L2/sen). När det gäller ogrammatiska svar är dessa, intressant nog, ungefär lika många i grupp L1 och L2/sen (29 respektive 35) men relativt få (19) i grupp L2/tidig.

I fråga om förväntade svar är resultatet tydligt och kan sägas ligga helt i linje med resultaten i den kvantitativa analysen (se tabellerna 8 – 8.3); nästan samtliga dessa svar (27 av 31) har producerats av eleverna i grupp L1. I grupp L2/tidigt förekommer endast tre, och i grupp L2/sen ett svar där eleverna har kompletterat uttrycksfragmentet på det förväntade sättet.

För att bekräfta de mönster som framträder i tabell 10 har ytterligare 10 enkäter av dem med lägsta resultat analyserats på samma sätt som de första 45. Tabell 10.1 visar att mönstret är i stort sett detsamma efter denna utökning av materialet.

TABELL 10.1: Antal olika typer av avvikelser för samtliga tio uttryck per informantgrupp – utökat material.

	Inget svar	Kontaminationer	Ickeidiomatiska svar	Ogrammatiska svar	Förväntade svar
L1 (19)	43	29	49	38	30
L2/tidig (16)	81	21	35	20	3
L2/sen (20)	100	23	36	41	1
Totalt	224	73	120	99	34

Resultaten av den kvalitativa analysen visar alltså på några intressanta aspekter med relevans för frågan varför vissa uttryck får låga resultat i undersökningen, såväl i hela materialet som i vissa av informantgrupperna. Det kan bl.a. konstateras att de låga resultaten hos de sena andraspråksinlärarna till stor del verkar bero på att eleverna helt enkelt har avstått från att komplettera de ofullständiga uttrycken i fråga. När det gäller förstaspråks eleverna verkar avvikelserna från den konventionaliserade formen i många fall vara en fråga om idiomacitet. Intressant nog förekommer det dock också relativt många svar som klassats som ogrammatiska i grupp L1, ungefär lika många som i grupp L2/sen och betyd-

ligt fler än i L2/tidig. I det sammanhanget måste det relativt låga antalet ogrammatiska lösningsförslag i den sistnämnda gruppen betraktas som ett något oväntat resultat, som dock ska ses i relation till det stora antalet nollsvår.

Överlag kan man säga att avvikelserna i grupp L1 är relativt jämnt fördelade över de olika kategorierna, till skillnad från andraspråksgrupperna, där kategorin *inget svar* är mycket mer markant.

6.6.3. Enkäterna med de högsta resultaten

En kvalitativ analys av enkäterna med högst resultat ger naturligtvis inte lika stort utslag som analysen av de enkäterna med lägst resultat i varje grupp. Men även om det inte finns lika många avvikelser att analysera i dessa enkäter, är det intressant att se i vilken mån mönstret i fråga om avvikelser från de konventionaliserade formerna bland de elever som klarar av större delen av uppgifterna skiljer sig ifrån mönstret som beskrivits i föregående avsnitt. Därför har enkäterna med högst resultat, dvs. minst 45 poäng, också analyserats kvalitativt (se avsnitt 5.5). 19 av de 20 informanterna som har nått 45 poäng eller mer ingår (inte helt oväntat, med tanke på de resultat som presenterats hittills) i grupp L1. Dessutom finns det 1 informant i grupp L2/tidig som har nått 46 poäng (jfr avsnitt 6.2.1).

Tabell 11 visar hur avvikelserna från de konventionaliserade formerna i samband med de tio svåraste uttrycken fördelar sig på de olika avvikelsetyper som nämnts i föregående avsnitt. Det framgår av tabellen att samtliga 10 uttryck som har visat sig svårast för informanterna i materialet som helhet även genererar avvikelser bland de elever som har nått höga resultat i undersökningen. Det uttryck som visat sig generera flest avvikande svar i materialet, med andra ord det uttryck som verkar vara svårast av samtliga konventionaliserade uttryck i materialet, nämligen *hålla någon i schack*, genererar också flest avvikande svar bland eleverna med högst resultat (se tabell 11). Endast 7 av 20 informanter har fått poäng för uttrycket. Förutom detta uttryck är det uttrycket *livet leker för någon* som får betydligt lägre poäng än de övriga bland de tio svåraste uttrycken. Med sammanlagt 8 poäng i de 20 enkäterna med högst resultat är *livet leker för någon* alltså ett av två uttryck med ett resultat under 14 poäng. I de 45 enkäterna med lägst resultat får uttrycket sammanlagt

5 poäng, vilket innebär ett av de högsta resultaten bland de tio analyserade uttrycken i den delen av materialet.

För de övriga 8 uttrycken ligger resultaten i de 20 enkäterna med högst resultat på en ganska jämn nivå (mellan 14 och 17 av 20 poäng). Detta mönster skiljer sig alltså något ifrån analysen av de 45 enkäterna med låga resultat, där uttrycken får resultat som är spridda över alla siffror mellan 0 och 5 poäng.

En annan tydlig skillnad jämfört med resultaten för den ursprungliga kvalitativa analysen är att nollsvaren är betydligt färre i enkäterna med högst resultat. 6 av de 10 uttrycken har inte fått några nollsvaret alls. Bland dessa finns även *hålla någon i schack*, som i övrigt har genererat de flesta avvikelserna i den aktuella delen av materialet. Än en gång är det uttrycket *livet leker för någon* som sticker ut något, då uttrycket med 6 nollsvaret har genererat ett större antal sådana avvikelser än de övriga uttrycken. I det avseendet liknar resultatet det för den kvalitativa analysen av enkäterna med lågt resultat, där *livet leker* har genererat 20 nollsvaret.

En skillnad i jämförelse med den första kvalitativa analysen är att de högstpresterande informanterna producerar få ogrammatiska svar. Av de sammanlagt 5 avvikelser som har klassats som ogrammatiska (jämfört med 81 ogrammatiska svar i den första kvalitativa analysen) förekommer 4 för uttrycket *livet leker för någon* och 1 för *med något i högsta hugg*.

Även när det gäller kontaminationer skiljer sig resultaten något från den första kvalitativa analysen; i de 45 enkäterna med lågt resultat förekommer sammanlagt 58 kontaminationer i samband med de 10 uttryck som har inkluderats i analysen. Dessa är jämfördelade på vanliga och idiosynkratiska kontaminationer. I de 20 enkäterna med högst resultat förekommer endast 9 kontaminationer. De kontaminationer som har visat sig vara vanligt förekommande i den första kvalitativa analysen verkar inte lika vanliga bland de högstpresterande informanterna. Endast 2 av dessa kontaminationer förekommer i den aktuella delen av materialet, nämligen kontaminationen av *med något i högsta hugg* och *i högsta grad* samt den av *på rak arm med på egen hand*.

Icke-idiomatiska svar är med 31 förekomster den klart vanligaste kategorin av avvikelser som producerats av de 20 informanterna med högst resultat överlag. Här finns ytterligare en skillnad jämfört med resultaten för den kvalitativa analysen av enkäterna med lågt resultat, där nollsvaren är vanligast. Dessutom kan det konstateras att uttrycket *hål-*

TABELL II: Typer av avvikelser i de 20 enkäterna med högst resultat

	Hålla i schack	Käppar i hjulen	Få upp ögonen	På rak arm	Livet leker	Högsta hugg	Klättra på väggarna	Få ngt/ngn på halsen	Lyckligt ovetande	Inre för fem öre	Totalt
avvikelser totalt	13	4	5	4	12	5	5	3	3	6	60
förväntade svar	7	16	15	16	8	15	15	17	17	14	140
inget svar	-	-	3	-	6	1	1	-	-	-	11
ogrammatiska svar	-	-	-	-	4	1	-	-	-	-	5
kontaminationer	-	1	1	-	2	-	-	2	-	2	8
vanliga kontaminationer	-	-	-	4	-	2	-	-	-	-	6
icke-idiomatiska svar	13	3	1	-	-	1	4	1	3	4	31

la någon i schack genererar betydligt fler icke-idiomatiska svar bland de högpressterande informanterna (13) än de övriga uttrycken, där antalet ligger mellan 0 och 5 förekomster av den aktuella typen av avvikelse. I likhet med de kvalitativa resultaten för enkäterna med lågt resultat är dock orsaken till avvikelsen även hos de högpressterande informanterna i många fall en bokstavlig tolkning av komponenten *schack* (som i 'schackspel').

I anslutning till detta kan man konstatera att skillnaden mellan resultaten av den första kvalitativa analysen och de avvikelser som de 20 högstpressterande informanterna producerar i första hand är kvantitativ. De avvikelser som de informanter som nått högst resultat ändå producerar, i samband med de tio aktuella uttrycken, är överlag av liknande karaktär som de avvikelser som produceras av informanterna med lågt resultat. I många fall används även samma ord för att komplettera uttrycksfragmenten i enkäten. Detta gäller bl.a. för de avvikande kompletteringar i samband med uttrycket *inte för fem öre*, där den vanligaste avvikelsen, även i denna del av materialet, är användningen av räkneordet *ett* istället för *fem*. Andra avvikelser som förekommer hos både låg- och högpressterande informanter är t.ex. *slå/sätta någon i schack*, *sätta stopp/pinnen/käppen i hjulen*, *stirrar på väggarna*, **livet är för någon*, *knappast ovetande* och, som tidigare nämnts, de kontaminerade formerna *på rak hand* och *med något i högsta grad*.

7. Sammanfattning och slutsatser

I föreliggande arbete har behärsknigen av ett antal svenska konventionaliserade uttryck bland ungdomar i flerspråkiga miljöer undersökts med hjälp av en enkätundersökning som innefattat både kvantitativa och kvalitativa analyser. En central fråga för undersökningen har varit huruvida behärsknigen av de aktuella uttrycken skiljer sig mellan första- och andraspråksinlärare av svenska, samt mellan andraspråksinlärare med olika startåldrar för inläringen av svenska.

Ytterligare frågeställningar för undersökningen har varit om det finns skillnader mellan olika typer av uttryck vad gäller informanternas behärskning av dessa och på vilket sätt informanterna avviker från de aktuella konventionaliserade former som efterlyses i enkäten. I de följande avsnitten sammanfattas och tolkas undersökningens viktigaste resultat.

7.1. Kvantitativ analys

Som framgår av den kvantitativa delen av detta arbete, finns det en rad signifikanta skillnader mellan startåldergrupperna, vilket tyder på att språklig bakgrund, och inte minst startåldern för inläring av ett andraspråk, är en relevant faktor i samband med behärsknigen av konventionaliserade uttryck. Resultaten som presenterats här kan sägas ligga i linje med tidigare studier som har visat att konventionaliserat språk är

en av de stora stötestenarna på väg mot en inföddlik språkbehärskning (jfr bl.a. Wray 2002, Abrahamsson & Hyltenstam 2009). I denna undersökning bekräftas detta genom signifikanta resultatskillnader mellan grupp L1 och grupp L2/tidig i varje deltest. Skillnaderna tyder på att även andraspråkselever på avancerad nivå med en relativt tidig start för inläringen har större svårigheter än förstaspråkselever när de ställs inför uppgiften att tolka och producera konventionaliserade uttryck på det sätt som krävs i föreliggande studie.

De signifikanta skillnaderna mellan grupp L1 och jämförelseklassen tyder på att klassmiljöernas heterogenitet i fråga om elevernas språkliga bakgrund kan ha viss betydelse när det gäller elevernas resultat i denna undersökning. Här måste man dock understryka att dessa resultatskillnader med hög sannolikhet också kan bero på en rad olika faktorer. I detta sammanhang kan skillnader i elevernas studiemotivation och/eller socioekonomiska faktorer (se avsnitt 5.1) vara minst lika avgörande för testresultaten som en språkligt homogen respektive heterogen skol- och klassrumsmiljö. Även om det alltså inte kan fastställas att resultatskillnaderna mellan förstaspråkseleverna och jämförelseklassen enbart beror på skillnaderna i klassernas språkliga sammansättning, kvarstår det faktum att den språkligt homogena jämförelseklassen når signifikant högre resultat än de förstaspråkselever som studerar i flerspråkiga klassmiljöer i alla deltest förutom del A.

7.1.1. Skillnader mellan deltesten

Formaten för de fyra deltesten är olika, vilket gör det något vanskligt att jämföra dem. Med andra ord är det naturligtvis inte helt överraskande att eleverna presterar bättre på ett flervals- och ett matchningstest än på ett lucktest, där de inte har givna lösningar att välja emellan. Man kan alltså inte dra slutsatsen att de konventionaliserade uttrycken i delarna A och D är lättare att förstå respektive bilda än dem i del B och C. Det innebär däremot inte att de signifikanta skillnaderna deltesterna emellan är ointressanta; dessa resultat säger något om svårighetsgraden i olika testuppgifter som i sin tur kan kopplas till olika aspekter av lexikal kunskap och behärskningen av konventionaliserade uttryck. Det finns två tydliga mönster när det gäller jämförelsen av deltesternas resultat inom de olika informantgrupperna; resultaten blir lägre med stigande

startålder för andraspråksinlärning i samtliga delar. Dessutom når samtliga grupper sina bästa resultat i A-delen och sina sämsta resultat i B-delen (se tabeller 7 och 7.1). Uppgifterna i del A – som verkar vålla minst problem för samtliga informantgrupper – kräver i första hand receptiv, och inte nödvändigtvis full behärskning av de konventionaliserade uttrycken i fråga (jfr Golden 2005). Även i D-delen kommer informanterna långt med sina receptiva kunskaper om kollokationerna, då konstituenterna är givna. Då orden ska matchas ihop, krävs det däremot även viss produktiv kunskap i fråga om de aktuella ordens kollokationspotential (Nation 2001). I lucktesten (B och C) är informanterna däremot tvungna att i högre grad använda produktiva kunskaper om de aktuella formerna, då det inte finns färdiga lösningar att välja emellan. Detta innebär att resultaten i dessa test säger något om det produktiva ordförrådets storlek och djup, vilket återspeglas i det höga svarsbortfallet i den sena startåldersgruppen i dessa delar av enkäten. När det gäller del C, måste informanterna behärska uttryckens form som helhet, då det inte är möjligt att sluta sig till det förväntade svaret med hjälp av kontexten eller utifrån de övriga konstituenternas betydelser. Detta visar sig, inte helt oväntat, vara svårast för samliga informantgrupper.

Förhållandet mellan de olika delresultaten är inte bara intressant med tanke på uppgifternas svårighetsgrad och olika aspekter av lexikal och fraseologisk kunskap, utan även med tanke på undersökningsmetodens reliabilitet. Det enhetliga mönster som beskrivits för delresultaten ovan är en indikation på kontinuitet och koherens mellan enkätens olika delar (jfr även avsnitt 6.4).

7.2. De enskilda uttrycken

Även när det gäller resultaten för de enskilda uttrycken i enkäten visar undersökningen på en rad intressanta skillnader, både mellan informantgrupperna och de olika uttrycken. Det kan dock konstateras att det är betydligt svårare att beskriva några tydliga mönster för sambandet mellan ett visst uttrycks egenskaper och det resultat det genererar i undersökningen. Det är med andra ord svårt att dra några generella slutsatser om möjliga orsaker till att vissa uttryck behärskas av fler informanter än andra. Faktorerna genomskinlighet och kompositionalitet

däremot tänkas ligga till grund för ett mer allmänt mönster som framgår av tabellerna 8 – 8.3. De uttryck som kan beskrivas som mest genomskinliga och kompositionella får i allmänhet högre resultat än de minst genomskinliga och minst kompositionella. Denna skillnad kan illustreras med uppgiftsparen *sätta färg på något* och *prata i nattmössan* (A), *ta första steget* och *sätta käppar i hjulen* (B), *ta saken i egna händer* och *hålla någon i schack* (C) samt *allvarlig sjukdom* och *ligga till grund*. Ett tydligt undantag från denna regel är *dra alla över en kam*, som definitivt kan betraktas som ett av de minst kompositionella och relativt ogenomskinliga i materialet, men ändå ligger på tredje plats i tabell 8.1. Liknande kan sägas för *komma till kritan*, som också genererar ett relativt högt resultat medan ett mer genomskinligt uttryck som *hålla modet/humöret uppe* i samma tabell ger ett relativt lågt resultat. Som nämnts i avsnitt 2.2 är bedömningen av ett uttrycks genomskinlighet emellertid inte oproblematiske, då den helt bygger på, i föreliggande fall författarens, intuition. Det finns inget objektiva sätt att bedöma hur genomskinligt ett uttryck är för någon som aldrig har mött det förut och begreppet i sig och bedömningarna som jag gör i samband med denna undersökning är att betrakta som relativa. Det finns alltså ingen skarp gräns mellan genomskinliga och ogenomskinliga uttryck och bedömningen handlar bl.a. om hur invecklad metaforiken inom ett uttryck är och om faktorer som betydelsestyngd eller polysemi hos de enskilda konstituenterna (jfr Svanlund 2009).

Det finns ytterligare möjliga påverkansfaktorer att diskutera i detta sammanhang, som t.ex. frekvens och i viss mån även sociala faktorer som informanternas ålder. När det gäller åldersfaktorn blir resonemangen dock ganska spekulativa, eftersom det inte har gjorts någon jämförelse med andra åldersgrupper. Inte heller frekvensdiskussionen är oproblematiske mot bakgrund av de problem som diskuteras i samband med frekvenssökningar på Google. Som tidigare nämnts är Googlesökningarna i första hand att betrakta som en referenspunkt och som komplement till sökningarna i PAROLE, som inte alltid räcker till för att ge en bild av hur vanliga relativt lågfrekventa språkliga fenomen av den typ som behandlas här faktiskt är i det allmänna språkbruket. Även om webben alltså inte lämpar sig som språkkorpus i vanlig mening med tanke på det ständigt föränderliga materialet (Moon 2008), måste man ändå diskutera i vilken mån man kan bortse från språkbruket på webben i un-

dersökningar som denna. Det handlar trots allt om en språklig källa som används alltmer av allmänheten – både produktivt och receptivt – och som tillhandahåller ett stort, aktuellt och varierat språkligt material. Sammanfattningsvis kan man konstatera att det finns vissa tendenser som pekar i samma riktning som resultaten i Ellis, Simpson-Vlach & Maynard (2008), men att det utifrån denna analys inte kan dras några entydiga slutsatser om frekvensens roll för informanternas behärskning av de aktuella konventionaliserade uttrycken.

7.3. Den kvalitativa analysen

Den kvalitativa beskrivningen utifrån olika typer av avvikelser från konventionaliserade konstruktioner har visat sig något vanskelig, då skillnaderna mellan de olika typerna i vissa fall är svåravgränsade. Detta gäller särskilt kategorierna *icke-idiomatisk* och *ogrammatisk*. En annan vanskelig kategori är kontaminationerna. Problemet här är att bedömningen alltid blir mer eller mindre spekulativ. Icke desto mindre visar de kvalitativa resultat som presenterats här en rad intressanta tendenser bl.a. när det gäller vilka strategier informanterna använder för att lösa uppgifterna och hur de associerar mellan olika lexikala enheter.

Att förstaspråkseleverna står för de flesta förväntade svaren visades redan i den kvantitativa analysen. I den kvalitativa analysen av det reducerade materialet framträder däremot en del tydliga och intressanta tendenser vad gäller sambandet mellan typ av avvikelser och informanternas språkliga bakgrund. Ett viktigt resultat är att andraspråkseleverna i större uträkning än eleverna med svenska som modersmål väljer att avstå från att lösa uppgifterna. Det verkar som att förstaspråkselever, när de inte behärskar ett uttryck, är mer benägna att pröva olika typer av lösningar med hjälp av olika strategier. Detta är naturligtvis inget förvånande resultat; förstaspråkseleverna kan antas vara betydligt tryggare i sin språkanvändning än i synnerhet eleverna med sen startålder för inläring. De sena andraspråksinlärarna kan heller inte förväntas ha kommit ikapp eleverna i grupp L1, bl.a. när det gäller det ordförråd de har att laborera med. Med anknytning till Bybee (1998, 2006) finns det här alltså skäl att anta att de helt enkelt inte har utvecklat representationer för orden och uttrycken ifråga, inte heller sådana som kan associeras till

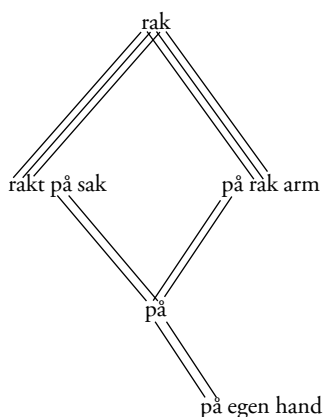
relaterade uttryck, vilket kan tänkas i högre grad vara fallet bland första-språkseleverna, som har fler och rikare språkliga erfarenheter när det gäller svenska och därmed en större uppsättning mentala representationer att välja bland (jfr Ekberg 1999).

I vissa fall kan den kvalitativa analysen ge en bättre bild av varför uttrycket i fråga har fått låga resultat i den kvantitativa analysen. Ett exempel är *inte för fem öre*, där det visar sig finnas en osäkerhet kring räkneordet, som dock inte verkar väljas helt godtyckligt. Det verkar snarare finnas ett tydligt mönster i fråga om vilka avvikande räkneord som kommer ifråga, nämligen den lägsta tänkbara valören i öre (*ett*) eller den närvarande lägsta valören för mynt (*femtio*). När det gäller användningen av *ett* skulle man dock, som tidigare nämnts, också kunna analysera avvikelserna som en kontamination av uttrycket *inte för fem öre* med andra konventionaliserade former som *inte ett ruttet öre* eller *inte ett rött öre*.

7.3.1. Lexikala komponenter, kontext och mental representation

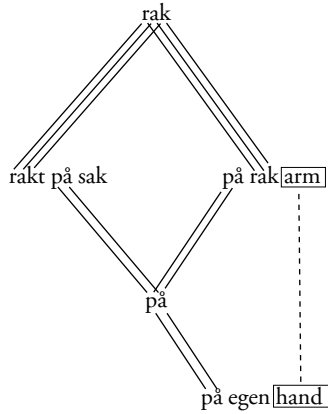
En annan aspekt som blir tydlig i den kvalitativa analysen av de 45 enkäterna med lågt resultat är att informanterna i många fall tar fasta på en av konstituenterna i uttrycksfragmentet och utifrån denna letar efter ord att komplettera med. Exempel på sådana konstituenten är partikeln *upp* i *få upp ögonen för något*, *högsta* i *med något i högsta hugg* och *sätta* i *sätta käppar i hjulen*, som verkar väcka associationer till en rad olika konventionaliserade konstruktioner, som i vissa fall bara har en konstituent gemensam med det efterlysta uttrycket (t.ex. *upp med hakan!*, *ta upp ett spår*, *ta upp saken*; *högsta hönset*, *högsta betyg*, *i högsta grad*; *sätta eld på något*, *sätta stopp för något* eller *sätta/slå spiken i kistan för något*). Detta är en tendens som visserligen i större utsträckning gäller för de lågpresterande informanterna, men som märks även bland de med de 20 högsta resultaten i undersökningen (se bl.a. kontaminationen av *med något i högsta hugg* och *i högsta grad*). Även om man utifrån föreliggande undersökning inte kan dra några generella slutsatser, är det intressant att fundera på vad detta kan säga om hur konventionaliserade flerordsuttryck lagras i det mentala lexikonet. Det verkar som att även enstaka lexikala

komponenter som är gemensamma för vissa uttryck kan ge upphov till kontamination av två uttryck som innehåller just denna komponent. Detta skulle kunna tyda på att olika konventionaliserade uttryck är sammanlänkade genom lexikala komponenter som är gemensamma för dem och därmed ge visst stöd för den lexikala modell som presenteras av Bybee (1998, 2006). Enligt Bybees modell skulle nätverket mellan olika uttryck med delvis samma komponenter utifrån föreliggande material kunna beskrivas som i figur 4.



FIGUR 4: Nätverk mellan de kontaminerade uttrycken och deras lexikala komponenter (efter Bybee 1998, 2006)

Samtliga slutsatser om den exakta karaktären hos dessa relationer i det mentala lexikonet är naturligtvis spekulativa. Man skulle även kunna hävda att gemensamma semantiska komponenter i olika uttryck med liknande struktur kan ge upphov till kontaminationer. Som nämnts i avsnitt 6.5.1.4, skulle man, med tanke på den kontaminerade formen *på rak hand*, även kunna förmoda att det finns en länk mellan komponenterna *arm* och *hand*, som tillhör samma semantiska fält. Enligt en sådan analys skulle relationerna mellan uttryck och lexikala komponenter kunna illustreras som i figur 4.1.



FIGUR 4.1: Nätverk mellan olika konventionaliserade uttryck och deras lexikala och semantiska komponenter

Det kan vidare noteras att det varierar kraftigt i vilken utsträckning informanter som nått lägsta resultat i sin grupp tar hänsyn till den övriga kontexten, dvs. till frasens övriga konstituenten och/eller den omgivande meningen. Detta gäller både form och betydelse och utgör i många fall skillnaden mellan icke-idiomatiska och ogrammatiska lösningsförslag. Det senare kan illustreras genom att jämföra formen *med något i högsta beredskap* med **med något i högsta hönset*, *på rak sak* med **på rak saken* eller **på rak fram* samt satserna *...att se människor isolerade i ett hus som stirrar på väggarna för att de är uttråkade* och *...att se människor isolerade i ett hus *som Karlsson på väggarna för att de är uttråkade*. När det gäller informanterna med högst resultat överlag, kan det konstateras att de avvikelser som produceras i mindre utsträckning är ogrammatiska och att dessa informanter också producerar färre kontaminationer, vilket kan vara ett tecken på att de i större uträkning tar hänsyn till kontexten (i de fallen där det finns en sådan) vid kompletteringen av uttrycksfragmenten i enkäten. Detta är naturligtvis en tentativ slutsats och det finns, som tidigare påpekats, en rad olika faktorer som kan orsaka de resultatskillnader mellan hög- och lågpresterande informanter som påtalats här.

Det förekommer också ett relativt stort antal avvikande exempel i materialet som liknar den efterlysta konventionaliserade formen på olika sätt. Dessa likheter kan ligga på ett semantiskt plan (*sätta pinnen i hjulen, få upp blicken för något*), på ett fonologiskt (*sätta tjappen i hjulen, bläddra på väggarna*) och/eller grafiskt plan (*med något i högsta högg*). Dessa typer av avvikelser tyder på att informanterna kan ha partiell kunskap om det konventionaliserade uttryck som efterlyses, dvs. de känner till minst en av uttryckets konstituenten och har en ungefärlig uppfattning om vad den betyder, hur den låter eller hur den ser ut i skrift. Detta kan även vara en möjlig förklaring till att två uttryck som liknar varandra i ett eller flera avseenden kontamineras. Ett sådant exempel är den tidigare nämnda formen *med något i högsta beredskap*, där de kontaminerade uttrycken *med något i högst hugg* och *varalligga i högsta beredskap* till viss del liknar varandra, både semantiskt och med tanke på de ingående lexikala komponenterna. De kvalitativa resultat som sammanfattas ovan tyder alltså på att informanterna använder sig av en rad olika strategier för att komma fram till en lösning när de inte behärskar ett uttryck till fullo.

I båda analysdelarna har informanterna beskrivits som grupp. Även om det framträder tydliga mönster i analysen, finns det naturligtvis även en del individuell variation i materialet som helhet såväl som inom startåldersgrupperna. Resultaten i grupp L2/sen varierar mellan 2 och 39 av 55 möjliga poäng. I grupp L1 finns en spridning mellan 8 och 53 (se avsnitt 6.2.1). I den kvalitativa analysen framträder också en del individuella, oväntade lösningsförslag. Överlag finns det dock tydliga tendenser och förvånansvärt många identiska, eller i alla fall likartade svar bland dem som på olika sätt avviker från den förväntade, konventionaliserade formen, även om man tar med i beräkningen att det i vissa fall troligtvis har förekommit samarbete mellan informanterna.

8. Avslutande diskussion

Studien som presenterats här bekräftar en del tidigare forskningsresultat angående skillnader i behärskning av konventionaliserade uttryck mellan första- och andraspråksinlärare (jfr bl.a. Ekberg 1997, 1998, Hyltenstam 1992 och Abrahamsson & Hyltenstam 2009). Undersökningen visar signifikanta skillnader mellan första- och andraspråkselever i behärskning av konventionaliserade uttryck och bekräftar dessutom på ett tydligt sätt att startålder för inläring är en avgörande faktor när det gäller behärsningen av konventionaliserade uttryck på ett andraspråk. Utöver detta bidrar studien till forskningen kring konventionaliserat språk med en undersökning av både receptiv och produktiv behärskning av olika typer av konventionaliserade uttryck bland elever med varierande språklig bakgrund samt en genomgång av enskilda konventionaliserade uttryck med hjälp av både kvantitativa och kvalitativa analyser. Arbetet kan på det sättet bidra till ökad kunskap om vilka svårigheter särskilt andraspråkselever ställs inför i samband med olika typer av konventionaliserade uttryck. Studien kan dessutom ge en bättre bild av vilka mönster som finns i avvikelser från konventionaliserade strukturer i informantgrupper med olika startålder för inläringen av svenska och olika prestationsnivåer på tester liknande de som har används här. De kvalitativa resultaten kan dessutom ge vissa indikationer på hur både första- och andraspråksinlärare associerar mellan lexikala enheter som har gemensamma komponenter eller som liknar varandra på andra sätt.

Det finns naturligtvis andra faktorer än de som redovisats i denna studie som kan tänkas påverka förståelsen och produktionen av kon-

ventionaliserade uttryck på ett andraspråk. Bl.a. hade en analys av sambandet mellan olika modersmål och olika typer av avvikelser från de konventionaliserade formerna på svenska varit intressant. En sådan undersökning hade kunnat ge svar på frågan vilken betydelse transfer från modersmålet (se Ringbom 2006) eller andra språk har för andraspråksanvändares resultat i den typen av tester som genomförts inom ramen för detta arbete. Det hade emellertid bl.a. krävts ett mer systematiskt informanturval med avseende på modersmål samt en hel del kunskap om konventionaliserade uttryck i de aktuella språken. Syftet med denna undersökning har dock inte varit att studera skillnader mellan olika språkgrupper, utan att undersöka skillnader, dels mellan första- och andraspråkselever i olika språkliga miljöer, dels mellan andraspråkselever med olika startålder för inläringen.

En annan möjlig analys av föreliggande material – som inte rymts inom ramen för detta arbete men som skulle vara särskilt intressant i anknytning till tidigare forskning av Ellis, Simpson-Vlach & Maynard (2008) – skulle vara en undersökning av eventuella samband mellan elevernas behärskning av konventionaliserade uttryck och uttryckens frekvens och Mutual Information i en större korpus. Detta skulle bl.a. kunna ge en bättre bild av orsakerna till en del av de resultatskillnader som framträder i denna undersökning och på så sätt vara ytterligare ett bidrag till forskningen om inläringen av konventionaliserade språkliga mönster. Dessutom skulle en sådan undersökning kunna vara en utmärkt utgångspunkt för fortsatt forskning, inte minst som underlag för utarbetandet av en fraslista för svenska i t.ex. akademiska sammanhang (jfr Ellis, Simpson-Vlach & Maynard 2008).

Ett metodiskt val som kan diskuteras är formen för uttrycksfragmenten i de delarna av enkäten där fragmenten ska kompletteras utan kontext. De former som har valts kan sägas motsvara ordboksformer, vilket inte är helt oproblemiskt då dessa sällan används i naturlig produktion. Detta gäller bl.a. användningen av *något* och *någon*; det är inte särskilt sannolikt att eleverna har mött formerna *med något i högsta hugg*, *livet leker för någon* eller *hålla någon i schack* många gånger i allmänt språkbruk. Användningen av dessa former kan alltså bidra till ytterligare dekontextualisering av de efterlysta konventionaliserade uttrycken. Det är däremot svårt att tänka sig vad man skulle kunna använda istället utan att ge en kontext till fragmenten ifråga.

Det bör också poängteras att de metoder som används i denna undersökning har sina klara begränsningar när det gäller att utröna i vilken utsträckning informanterna behärskar ett uttryck eller inte (jfr Henriksen 1999). Det är naturligtvis också svårt att dra några säkra slutsatser om vad enstaka elevers resultat kan tänkas bero på. Bl.a. är det för försöksledaren omöjligt att veta i vilken mån informanterna har varit motiverade att lösa uppgifterna på bästa sätt. Vissa svar som framträder i den kvalitativa analysen tyder på att informanterna ibland medvetet gett ett svar som skiljer sig från det efterlysta svaret. Detta kan man emellertid inte med säkerhet veta utan att t.ex. intervjua eleverna. Samma osäkerhet gäller för de fall där informanterna har avstått från att lösa uppgiften; om nollsvaren i det individuella fallet beror på bristande motivation, bristande självförtroende eller bristande kunskap om uttrycket ifråga, kan här inte med säkerhet fastställas. Det är heller inte möjligt att veta om informanterna kan ha haft partiell kunskap om uttrycket, men ändå varit så pass osäkra att de valt att avstå ifrån ett lösningsförslag. Det finns naturligtvis metoder som möjliggör en djupare analys av olika grader av behärskning av lexikala enheter. Ett exempel är så kallade *think-a-loud-protocols*, där informanterna ges möjlighet att högt för sig själv resonera kring ett ord eller en fras, medan de spelas in (jfr Spöttl & McCarthy 2004). Ett annat alternativ är att använda en blandning av flera olika metoder, t.ex. test av olika slag i kombination med intervjuer (jfr Ellis, Simpson-Vlach & Maynard 2008). Dessa metoder är dock ofta tidskrävande både för försöksledaren och deltagarna, vilket gör det svårt att genomföra en sådan undersökning med en större informantgrupp, i synnerhet när det handlar om ungdomar som kan vara svåra att motivera för uppgiften. En enkät har fördelen att den är relativt lätt att distribuera till ett större antal informanter och genererar ett material som är relativt lätthanterligt. Även om denna undersökning inte kan tillhandahålla en djupare analys av elevernas behärskning av olika typer av konventionaliserade uttryck, visar den att det finns tydliga och signifikanta skillnader i elevers behärskning av olika typer av konventionaliserade uttryck och att dessa skillnader kan kopplas till faktorerna språklig bakgrund och startålder för andraspråksinlärning (även om dessa naturligtvis inte är de enda bidragande faktorerna). Genom en kvalitativ beskrivning av olika typer av avvikelser från de efterlysta konventionaliserade formerna erhålls dessutom en bättre (om än inte heltäckande) bild av

tänkbara förklaringar till de låga testresultaten för vissa uttryck och till hur olika typer av avvikelser kan tänkas uppstå. Resultaten visar dessutom att det finns vissa samband mellan språklig bakgrund och olika typer av avvikelser i materialet. Undersökningen kan härigenom höja medvetenheten om olika typer av svårigheter som i synnerhet elever med flerspråkig bakgrund kan ställas inför i samband med tolkningen och användningen av konventionaliserade uttryck i svenskan. Resultaten som presenterats här kan också tjäna som underlag för fortsatt forskning, t.ex. vid en djupare analys av vissa uttryck som på olika sätt visat sig intressanta i denna undersökning.

Referenser

- Abrahamsson, Niclas & Hyltenstam, Kenneth (2004). Mognadsbegränsningar och den kritiska perioden för andraspråksinläring. I: Hyltenstam, Kenneth & Lindberg, Inger (red.), *Svenska som andraspråk – i forskning undervisning och samhälle*. Lund: Studentlitteratur. S. 221–258
- Abrahamsson, Niclas & Hyltenstam, Kenneth (2009). Age of Onset and Nativelikeness in a Second Language: Listener Perception Versus Linguistic Scrutiny. *Language Learning*, 59/2: 249–306.
- Allén, Sture (1975). *Nusvensk frekvensordbok 3. Ordförbindelser/Frequency Dictionary of Present-Day Swedish 3. Collocations*. (Data Linguistica 10.) Stockholm: Almqvist & Wiksell.
- Bartsch, Sabine (2004). *Structural and Functional Properties of Collocations in English*. Tübingen: Gunter Narr Verlag.
- Bialystok, Ellen (2001). *Bilingualism in Development. Language Literacy & Cognition*. Cambridge: Cambridge University Press.
- Biber, Douglas & Conrad, Susan (1999). Lexical bundles in conversation and academic prose. I: Hasselgard, Hilde & Oksefjell, Signe (eds.), *Out of Corpora: Studies in Honor of Stig Johansson*. Amsterdam: Rodopi. S. 181–189.
- Birdsong, David (1999). *Second Language Acquisition and the Critical Period Hypothesis*. Mahwah, N.J., London : Erlbaum
- Bolinger, Dwight (1975). *Aspects of Language*. (2nd edition.) New York: Harcourt Brace Jovanovich.

- Brown, Roger (1973). *A First Language*. Cambridge, Massachusetts: Harvard University Press.
- Butler, Yuko G. (2000). The age effect in second language acquisition: Is it too late to acquire native-level competence in a second language after the age of seven? I: Oshima-Takane, Y., Shirai, Y. & Sirai, H. (eds.), *Studies in Language Science 1*. Tokyo: The Japanese Society for Language Science. S. 159–169.
- Bybee, Joan (1998). The Emergent Lexicon. I: Gruber, M. Catherine, Higgins, Derrick, Olson, Kenneth S. & Wysocki, Tamra (eds.), *CLS 34: The Panels*. Chicago: Chicago Linguistic Society. S. 421–439.
- Bybee, Joan (2006). From Usage to Grammar: The Mind's Response to Repetition. *Language* 82/4: 711–733.
- Church, Kenneth W. & Hanks, Patrick (1990). Word association norms, mutual information, and lexicography. *Computational Linguistics* 16/1: 22–29.
- Coleman, J.S. (1966). *Equality of Educational Opportunity*. Washington, D.C.: US Government Printing Office.
- Costello, Fintan J. & Keane, Mark T. (2000): Efficient creativity: constraint-guided conceptual combination. *Cognitive Science* 24/2: 299–349
- Coulmas, Florian (1979). On the sociolinguistic relevance of routine formulae. *Journal of Pragmatics* 3: 239–266.
- Cowie, Anthony P. (1988). Stable and creative aspects of vocabulary use. I: Carter, Ronald & McCarthy, Michael. *Vocabulary and language teaching*. London: Longman. S. 126–139.
- Cruttenden, Alan (1981). Item-Learning and System-Learning. *Journal of Psycholinguistic Research*, 10/1: 79–88.
- Crystal, David (1997). *The Cambridge encyclopedia of language*. Cambridge: Cambridge University Press.
- DeKeyser, Robert M. & Sokalski, Karl J. (1996). The differential roles of comprehension and production practice, *Language Learning* 46: 613–642.
- Dijkstra, Ton & van Heuven Walter J. B. (1998). The BIA model and bilingual word recognition. I: Grainger, Jonathan & Jacobs, Arthur M. (eds.), *Localist connectionist approaches to human cognition*. Mahwah, NJ: Erlbaum. S. 189–225.

- Ekberg, Lena (1989). *Gå till anfall och falla i sömn. En strukturell och funktionell beskrivning av abstrakta övergångsfraser*. Lund: Lund University Press.
- Ekberg, Lena (1997). Diskursiva och syntaktiska mönster i återberättelser hos invandrarbarn i Rosengård. I: Håkansson, Gisela et al. (utg.), *Svenskans beskrivning 22*. Lund: Lund University Press. S. 97–110.
- Ekberg, Lena (1998). Regeltillämpning kontra lexikonkunskap i svenskan hos invandrarbarn i Rosengård. I: Møller, Janus, Quist, Pia, Holmen, Anne & Norrman Jørgensen, Jens (utg.), *Nordiske sprog som andetsprog. Københavnerstudier i tosprogethed 30*. København: Københavns universitet. S. 247–263.
- Ekberg, Lena (1999). Användningen av komplexa predikat hos invandrarbarn i Rosengård. I: Andersson, Lars-Gunnar, Lundqvist, Aina, Norén, Kerstin & Rogström, Lena (utg.), *Svenskans beskrivning 23. Förhandlingar vid Tjugotredje sammankomsten för svenskans beskrivning*. Lund : Lund University Press. S. 86–95.
- Ekberg, Lena (2004): Grammatik och lexikon i svenska som andraspråk på nästan infödd nivå. I: Hyltenstam, Kenneth & Lindberg, Inger (red.), *Svenska som andraspråk – i forskning, undervisning och samhälle*. Lund: Studentlitteratur. S. 259–276.
- Ellis, Nick C. (1995). The psychology of foreign language vocabulary acquisition: Implications for CALL. *Computer Assisted Language Learning* 8: 103–128.
- Ellis, Nick C. & Beaton, Alan (1993). Factors affecting foreign language vocabulary: imagery keyword mediators and phonological short-term memory. *Quarterly Journal of Experimental Psychology* 46: 533–558.
- Ellis, Nick. C., Simpson-Vlach, Rita & Maynard, Carson (2008). Formulaic Language in Native and Second Language Speakers: Psycholinguistics, Corpuslinguistics and TESOL. *TESOL Quarterly* 42/3: 375–396.
- Ellis, Rod, (1995). Modified oral input and the acquisition of word meanings. *Applied Linguistics* 16: 409–441.
- Enström, Ingegerd (1996). *Klara Verba. Andraspråksinlärares verb-användning i svenskan*. (Nordistica Gothoburgensia 18.) Göteborg: Acta Universitatis Gothoburgensis.

- Fernando, Chitra & Flavell, Roger (1981). *On idiom: Critical views and perspectives*. (Exeter Linguistic Studies 5.) Exeter: University of Exeter.
- Firth, John R. (1957). *Papers in Linguistics 1934-1951*. London: Oxford University Press.
- Fleischer, Wolfgang (1997). *Phraseologie der deutschen Gegenwartssprache*. (2., durchgesehene und ergänzte Auflage.) Tübingen: Niemeyer.
- Gläser, Rosemarie (1986). *Phraseologie der englischen Sprache*. Tübingen: Niemeyer.
- Gläser, Rosemarie (1988). The grading of idiomaticity as a presupposition for a taxonomy of idioms. I: Hüllen, Werner & Schulze, Rainier (eds.), *Understanding the lexicon. Meaning, sense and world knowledge in lexical semantics*. Tübingen: Niemeyer. S. 264–279.
- Golden, Anne (2005). *Å gripe poenget. Forståelse av metaforiske uttrykk fra lærebøker i samfunnskunnskap hos minoritets elever i ungdomsskolen*. Oslo: Unipub forlag.
- Granger, Sylviane (1993). The International Corpus of Learner English. I: Aarts, Jan, de Haan, Pieter & Oostdijk, Nelleke, *English Language Corpora: Design, Analysis and Exploitation*. Amsterdam: Rodopi. S. 57–71.
- Granger, Sylviane (1996). Learner English around the world. I: Greenbaum, Sidney (ed.), *Comparing English World-Wide*. Oxford: Clarendon Press. S. 13–24.
- Granger, Sylviane (1998). Prefabricated Patterns in Advanced EFL Writing: Collocations and Formulae. I: Cowie, Anthony P. (ed.), *Phraseology, Theory, Analysis and Applications*. Oxford: Oxford University Press. S. 145–160.
- Granger, Sylviane (2003). The International Corpus of Learner English: a new resource for foreign language learning and teaching and second language acquisition research. *TESOL Quarterly* 37/3: 538–546.
- Granger, Sylviane, Hung, Joseph & Petch Tyson, Stephanie (eds.) (2002). *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*. Amsterdam: John Benjamins.
- Greenberg, Joseph (1966). *Language universals: With special reference to feature hierarchies*. The Hague: Mouton.

- Große [Prentice], Julia (2008). Fraseologiska enheter på flerspråkig grund. En jämförelse mellan verklighet och fiktion. *Nordand. Nordisk tidskrift för andrespråksforskning* 3/2: 103–132.
- Gustafsson, J.-E. (1997). Measurement Characteristics of the IEA Reading Literacy Scales for 9- and 10-Year-Olds at Country and Individual Levels. *Journal of Educational Measurement* 34/3: 233–251.
- Halliday, M. A. K., McIntosh, Angus & Strevens, Peter D. (1964). *The Linguistic Sciences and Language Teaching*. London: Longmans.
- Halliday, M. A. K. (1966). Lexis as a linguistic level. I: Bazell, Charles E., Catford, John C., Halliday, M. A. K. & Robins, Robert H. (eds.), *In Memory of J.R. Firth*. London: Longmans. S. 148–162.
- Hausmann, Franz Josef (1985). Kollokationen im deutschen Wörterbuch. Ein Beitrag zur Theori des lexikographischen Beispiels. I: Bergenholtz, Henning & Mugdan, Joachim (Hrsg.), *Lexikographie und Grammatik. Akten des Essener Kolloquiums zur Grammatik im Wörterbuch 28.-30.6. 1984*. Tübingen: Niemeyer. S. 118–129.
- Henriksen, Birgit, (1999). Three Dimensions of Vocabulary Development. *Studies in Second Language Acquisition* 21: 303–317.
- Holmegaard, Margareta (2007). Långa ord – en svårighet för flerspråkiga studerande? I: Lindberg, Inger & Johansson Kokkinakis, Sofie, *OrdiL. En korpusbaserad kartläggning av ordförrådet i läromedel för grundskolans senare år*. Rapporter om svenska som andraspråk (ROSA) 8. Göteborg: Institutet för svenska som andraspråk, Göteborgs universitet. S. 135 –170.
- Hopper, Paul, J. & Thompson, Sandra A. (1980). Transitivity in grammar and didcourse. *Language* 56: 251–99.
- Hyltenstam, Kenneth (1992). Non-native features of Near-native speakers: On the Ultimate Attainment of Childhood L2 Learners. I: Harris, Richard (ed.), *Cognitive processing in Bilinguals*. Amsterdam; London, New York, Tokyo: North Holland. S. 351–368.
- Hyltenstam, Kenneth & Abrahamsson, Niclas, (2000). Who can become a native speaker in a second language? All or none? On the maturational constraints controversy in second language acquisition. *Studia Linguistica* 54/2: 150–166.
- Kjellmer, Göran (1994). *A Dictionary of English Collocations. Based on the Brown Corpus. I-III*. Oxford: Clarendon Press.

- Kokkinakis, Dimitrios (2006). Collection, Encoding and Linguistic Processing of a Swedish Medical Corpus – The MEDLEX Experience. *Proceedings of the 5th Language Resources and Evaluation Conference (LREC), May 2006*, Genoa, Italy.
- Körner, Svante & Wahlgren, Lars (2005). *Statistiska metoder* (2. uppl.). Lund: Studentlitteratur.
- Lamb, Sidney M. (2000). Neuro-cognitive structure in the interplay of language and thought. I: Pütz, Martin & Verspoor, Marjolijn, *Explorations in linguistic relativity*. Amsterdam: John Benjamins. S. 173–196.
- Langacker, Ronald (2000). A dynamic usage-based model. I: Barlow, Michael & Kemmer, Suzanne (eds.), *Usage Based Models of Language*. Stanford, California: CSLI Publications. S. 1–64.
- Langlotz, Andreas (2006). *Idiomatic Creativity. A cognitive-linguistic model of idiom representation and idiom-variation in English*. Amsterdam & Philadelphia: John Benjamins.
- Lenneberg, Eric (1967). *Biological Foundations of Language*. New York: Wiley & Sons.
- Lindberg, Inger 2004. Språkanvändning – kreativt skapande eller förutsägbara mönster? I: Olofsson, Mikael (red.), *Symposium 2003, Arena Andraspråk*. Stockholm: HLS Förlag. S. 61–81.
- Lindberg, Inger, (2007): Multilingual Education. A Swedish Perspective. I: Carlson, Marie, Rabo, Annika & Gök, Fatma (eds.), *Education in 'Multicultural' Societies. Turkish and Swedish Perspectives*. (Swedish Research Institute in Istanbul Transactions, vol. 18.) Istanbul: Svenska Forskningsinstitutet i Istanbul. S. 71–90.
- Lindberg, Inger & Johansson Kokkinakis, Sofie (2007). *OrdiL. En korpusbaserad kartläggning av ordförrådet i läromedel för grundskolans senare år*. Rapporter om svenska som andraspråk (ROSA) 8. Göteborg: Institutet för svenska som andraspråk, Göteborgs universitet.
- Locke, J., (1997). A theory of linguistic development. *Brain and Language* 58: 265–326.
- Lorentzen, Henrik (2008): Korpusbaseret lemma-eksklusion eller ”hvor gik ordene hen da de gik ud?”. I: Jóhannesson, Kristinn, Landquist, Hans, Lundqvist, Aina, Rogström, Lena, Sköldberg, Emma & Wallgren Hemlin, Barbro (utg.), *Nog ordat? Festskrift till Sven-Göran Malmgren*. (Meijerbergs Arkiv för svensk ordforskning 34.) Göteborg. S. 259–267.

- Luthman, Hans (2002). *Svenska idiom. 3500 vardagsuttryck*. Lund: Folkuniversitetets förlag.
- Makkai, Adam (1972). *Idiom structure in English*. The Hague & Paris: Mouton.
- Malmgren, Sven-Göran (2003). *Begå eller ta självmord? Om svenska kollokationer och deras förändringsbenägenhet*. *Språk och stil* 13: 123–168.
- McCarthy, Michael, (1998). *Spoken language and applied linguistics*. Cambridge: Cambridge University Press.
- Mel'čuk, Igor, (1998). Collocations and lexical Funktionen. I: Cowie, Anthony, P. (ed.), *Phraseology, Theory, Analysis and Applications*. Oxford: Oxford University Press. S. 23–53.
- Moon, Rosamund (1998). *Fixed Expressions and Idioms in English. A Corpus-Based Approach*. Oxford: Oxford University Press.
- Moon, Rosamund, (2008). Conventionalized *as*-similes in English. A problem case*. *International Journal of Corpus Linguistics* 13/1: 3–37.
- Mühlenbock, Katharina (2009). Readable, legible or plain words - Presentation of an easy-to-read Swedish corpus. Multilingualism, Proceedings of the 23rd Scandinavian Conference of Linguistics, *Studia Linguistica Upsaliensia* 8 s. 325–327.
- Mühlenbock & Johansson Kokkinakis (2009). LIX 68 revisited. An extended readability measure. Konferensbidrag på Corpus Linguistics Conference (CL) 2009. University of Liverpool, UK20-23 July 2009.
- Munro, Miles & Mann, Virginia (2005). Age of immersion as a predictor of foreign accent. *Applied Psycholinguistics* 26: 311–341.
- Nagy, William E., Herman, Patricia A. & Anderson, Richard C. (1985). Learning words from Context. *Reading Research Quarterly* 20: 233–253.
- Nation, Paul (2001). *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.
- Nist, Sherrie L. & Olejnik, Stephen (1995). The role of context and dictionary definitions on varying levels of word knowledge. *Reading Research Quarterly* 30: 172–193.
- O'Keefe, Anne, McCarthy, Michael & Carter, Ronald (2007). *From Corpus to Classroom: language use and language teaching*. Cambridge. Cambridge University Press.

- Pallier, C., Dehaene, J., Poline, J.-B., LeBihan, D., Argentié, A.-M., Dupoux, E. & Mehler, J. (2003). Brain Imaging of Language Plasticity in Adopted Adults: Can a Second Language Replace the First? *Cerebral Cortex* 2003 13: 155–161.
- Palmer, Harold E. (1933). *Second Interim Report on English Collocations*. Tokyo: Kaitakusha.
- Partee, Barbara H. (1984). Compositionality. I: Landman, Fred & Veltman, Frank (eds.), *Varieties of Formal Semantics. Proceedings of the fourth Amsterdam Colloquium, September 1982*. Dordrecht: Foris. S. 281–311.
- Paul, Peter V., Stallman, A. C. & O'Rourke, Joseph P. (1990). *Using three test formats to assess good and poor readers' word knowledge*, Technical Report No. 509 of the Center for the Study of Reading. Urbana-Champaign: University of Illinois.
- Pawley, Andrew & Syder, Frances H. (1983). Two puzzles for linguistic theory: nativelike selection and nativelike fluency. I: Richards, Jack & Schmidt, Richard (eds.), *Language and Communication*. London: Longman S. 191–221.
- Peters, Ann M. (1983). *Units of Language Acquisition*. Cambridge. Cambridge University Press.
- Prentice, Julia & Emma Sköldbäck (und. utg.) *Klättra på väggarna eller bara vara ett med soffan? Om figurativa ordförbindelser bland ungdomar i flerspråkiga skolmiljöer*. Språk och stil 20.
- Ringbom, Håkan (2006). On L1 Transfer in L2 Comprehension and L2 Production. *Language Learning* 42/1: 85–112.
- Schmitt, Nobert (2008). Instructed second language vocabulary learning. *Language Teaching Research* 12/3: 329–363
- Sinclair, John (1966). Beginning the Study of Lexis. I: Bazell, Charles E., Catford, John C., Halliday, M. A. K. & Robins, Robert H. (eds.), *In Memory of J.R. Firth*. London: Longmans. S. 410–430.
- Sinclair, John (1987). Collocation: A progress report. I: Steele, Ross & Treadgold, Terry (eds.), *Language Topics. Essays in honour of Michael Halliday*. Amsterdam: John Benjamins. S. 319–331.
- Skehan, Peter (1998). *A Cognitive Approach to Language Learning*. Oxford: Oxford University Press.
- Skolverket (2004). *Elever med utländsk bakgrund. Rapport till regeringen, 1 oktober 2004*. (Dnr 75 – 2004:545.) Stockholm: Skolverket.

- Skolverket (2005). *Elever med utländsk bakgrund. En sammanfattande bild*. Stockholm: Skolverket.
- Skolverket (2008). Studieresultat i gymnasieskolan – en statistisk beskrivning av ofullständiga gymnasiestudier. (Skolverkets aktuella analyser.) Stockholm: Skolverket.
- Skolverket (2009). *Vad påverkar resultaten i svensk grundskola? Kunskapsöversikt om betydelsen av olika faktorer*. Stockholm.
- Sköldberg, Emma (2004). *Korten på bordet. Innehålls- och uttrycksmässig variation hos svenska idiom*. (Meijerbergs Arkiv för svensk ordforskning 31.) Göteborg: Göteborgs universitet.
- Sköldberg, Emma & Toporowska Gronostaj, Maria (2008). Modell för beskrivning av kollokationer i ett medicinskt lexikon (MedLex). I: Svavarsdóttir, Ásta, Kvaran, Guðrun, Ingólfsson, Gunnlaugur & Jónsson, Jón Hilmar (red.), *Nordiske Studier i Leksikografi 9. Rapport fra konference om leksikografi i Norden, Akureyri 22. – 26. maj 2007*. Reykjavík: Nordisk Forening for Leksikografi. S. 433–446.
- Spöttl, Carol and Michael McCarthy (2004). Comparing knowledge of formulaic sequences across L1, L2, L3, and L4. I: Schmitt, Norbert (ed.), *Formulaic Sequences*. Amsterdam & Philadelphia: John Benjamins. S. 191–225.
- Stubbs, Michael (2001). *Words and Phrases. Corpus Studies of Lexical Semantics*. Oxford & Malden, Massachusetts: Blackwell.
- Svanlund, Jan (2001). *Metaforen som konvention. Graden av bildlighet i svenskans vikt- och tyngdmetaforer*. (Stockholm Studies in Scandinavian Philology. New series. 23.) Stockholm: Almqvist & Wiksell.
- Svanlund, Jan (2009). Lexikal etablering. *En korpusundersökning av hur nya sammansättningar konventionaliseras och får sin betydelse*. (Acta Universitatis Stockholmiensis. Stockholm Studies in Scandinavian Philology. New series. 52.) Stockholm. Stockholms universitet.
- Svensén, Bo (1987). *Handbok i lexikografi*. Stockholm: Norstedts ordbok.
- Thomas, Wayne & Collier, Virginia (2002). A National Study of School Effectiveness for Language Minority Students' Long-Term Academic Achievement. Center for Research on Education, Diversity & Excellence (CREDE).
- Van Lier, Leo (1995). *Introducing Language Awareness*. London: Penguin Books.

- Viberg, Åke (1993). Andraspråksinläring i olika åldrar. I: Cerú, Eva (red.), *Lärbok 2. Sevnska som andraspråk. Mera om språket och inläringen*. Stockholm: Natur och Kultur, Utbildningsradion. S. 13–83.
- Wahlgren, Lars (2005). *SPSS steg för steg*. Lund: Studentlitteratur.
- Wiktorsson, Maria (2003). *Learning Idiomaticity. A corpus-based study of idiomatic expressions in learners' written production*. Lund: Lund University.
- Wray, Alison, 2002: *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press.
- Wray, Alison (2008). *Formulaic Language: Pushing the Boundaries*. Oxford: Oxford University Press.
- Wray, Alison & Perkins, Michael R. (2000). The functions of formulaic language: an integrated model. *Language and Communication* 20: 1–28.
- Yang, Kajsa & Gustafsson, Jan - Eric (2003). Measuring Socioeconomic Status at Individual and Collective Levels. *Educational Research and Evaluation* 10/3: 259–288.
- Yorio, Carlos A. (1989). Idiomaticity as an indicator of second language proficiency. I: Hyltenstam, Kenneth & Obler, Loraine K. (eds.), *Bilingualism across the lifespan*. Cambridge: Cambridge University Press. S. 55–72.

Övriga källor

- Google (2010a). <www.google.se>, hämtat 11 januari 2010.
- Google (2010b). <www.google.se>, hämtat 6 februari 2010.
- Skolverket 2010, <<http://siris.skolverket.se>>, hämtat 20 juni 2010.
- Språkbanken (2010a). <<http://spraakbanken.gu.se/parole/>>, hämtat 11 januari 2010.
- Språkbanken (2010b). <<http://spraakbanken.gu.se/konk/>>, hämtat, 6 februari 2010.
- Svenskt språkbruk. Ordbok över konstruktioner och fraser (2003)*. Stockholm: Norstedts ordbok.

Appendix

**Först några korta frågor om dig och din språkanvändning.
Kryssa i de alternativ som beskriver dig!**

- Tjej
- Kille

Gymnasieprogram:

- Är född i Sverige och har svenskfödda föräldrar
- Är född i Sverige och har utlandsfödda föräldrar
- Är född i Sverige och har **en** utlandsfödd förälder
- Är inte född i Sverige. Kom till Sverige vid _____ års ålder
- Började lära mig svenska när jag var _____ år gammal

Vilket/vilka är ditt/dina modersmål? _____

Vilket/vilka språk talar ni hemma i familjen?

Vilket/vilka språk talar du mest med dina kompisar?

Har du under din skoltid läst svenska som andraspråk och i så fall när och hur länge?

Nej

Ja

När? _____

Hur många år? _____

Har du under din skoltid läst ämnet modersmål?

Nej

Ja

När? _____

Hur många år? _____

Vilket typ av svenskundervisning deltar du i just nu? svenska
 svenska som andraspråk
 båda
 vet ej

A

Vad tror du att följande uttryck betyder? Välj det alternativ som du tycker stämmer bäst. Markera bara ett alternativ!

Exempel:

ha händerna fulla

- vara rik
- ha mycket att göra
- inte kunna röra sig
- Inget av alternativen stämmer, uttrycket betyder följande:

1) prata i nattmössan

- prata i sömnen
- prata otydligt, mumla
- säga dumheter, inte veta vad man talar om
- Inget av alternativen stämmer, uttrycket betyder följande:

2) fastna för något

- inte kunna ta sig loss från något, sitta fast
- bli förtjust i något och bestämma sig för att t.ex. köpa det
- inte förstå något
- Inget av alternativen stämmer, uttrycket betyder följande:

3) något är/blir en ond cirkel

- när människor med onda avsikter bildar en grupp
- när en tråkig händelse leder till nästa
- när någon skadar sig gång på gång
- Inget av alternativen stämmer, uttrycket betyder följande:

4) något står en upp i halsen

- när man är tvungen att ta hand om något
- när man inte kan svälja något
- när man har fått nog, när man är mycket trött på något
- Inget av alternativen stämmer, uttrycket betyder följande:

5) hålla huvudet högt

- gå upprätt
- se ner på andra människor
- behålla sin stolthet
- Inget av alternativen stämmer, uttrycket betyder följande:

6) ha fria tyglar

- vara lösaktig
- få göra precis som man vill
- strunta i något
- Inget av alternativen stämmer, uttrycket betyder följande:

7) ha mage att göra något

- ha den fysiska kraften att göra något
- våga göra något som andra kan uppfatta som fräckt
- vara lagom tjock för att göra något
- Inget av alternativen stämmer, uttrycket betyder följande:

8) sätta färg på något

- måla om
- göra något trevligare, gladare eller mer intressant
- dela in, kategorisera något med hjälp av olika färger
- Inget av alternativen stämmer, uttrycket betyder följande:

9) stå i rampljuset

- vara i centrum för allmänhetens intresse
- spela teater
- vara bländad av ljus
- Inget av alternativen stämmer, uttrycket betyder följande:

10) innan man vet ordet av

- när man inte kommer på vad man ska säga
- när något händer så snabbt att man inte hinner reagera
- när man aldrig hört talas om något
- Inget av alternativen stämmer, uttrycket betyder följande:

B

Här kommer exempel på några vanliga uttryck och fraser. Fyll i det ord som du tycker saknas i de olika uttrycken.

Exempel:

I Expedition Robinson måste man jobba för att överleva. Det funkar inte att åka snälskjuts för då blir man utrustad direkt.

Samma sak gäller dokusåpan Farmen. Vill man vara kvar där så måste man dra sitt strå till stacken.

1) När jag misslyckas försöker jag alltid att *hålla* _____ *uppe* och inte deppa i onödan.

2) Det är år 2003 och alla borde vara jämlika.

Jag vill inte _____ *alla män över en kam* men de flesta vet nog vad jag pratar om.

3) Jag erkänner, om jag hade makten att ändra det förflutna eller mitt eller ditt öde, så skulle jag nog _____ *för frestelsen*.

4) På Fame Factory får deltagarna lära sig hur man agerar på scen och skriver egna låtar. Det är mer underhållande än att se människor isolerade i ett hus som _____ *på väggarna* för att de är uttråkade.

5) När jag sedan blev äldre och skulle gå i skolan, fick tiden helt plötsligt betydelse. Från att ha blivit körd av föräldrarna till dagis, skulle jag nu ta mig till skolan *på egen* _____ och vara där i tid.

6) Ibland låter man helt enkelt bli att prata med folk, av rädsla att *få en riktig pratkvart på* _____.

7) Gör din röst hörd! Få politikerna att förstå att det är vi, nästa generation, som ska styra samhället i framtiden, som blir lidande. Något måste hända och det är vi ungdomar som måste ta det *första* _____ !

8) Kan det vara så att vi skjuter upp saker i onödan och när det sedan verkligen _____ till kritan klagar vi på att vi har haft för lite tid för att åstadkomma ett bra resultat?

9) En tolkning av scenen från filmen ”Moderna tider”, där Charlie Chaplin med våld vrider tillbaka tiden, skulle kunna vara att Chaplin inte vill följa med i utvecklingen. Därför försöker han bokstavligt talat att sätta _____ i hjulen.

10) Ryck upp er alla latmaskar! *Ta er i* _____ och börja träna!

C

Fyll i det ord som fattas i uttrycken nedan. Fundera inte för länge på varje uttryck utan skriv det som du först kommer att tänka på.

Exempel:
gå ner i varv

- 1) tid är _____
- 2) ta saken i egna _____
- 3) som _____ på moset
- 4) av hela mitt _____
- 5) _____ någon i schack
- 6) livet _____ för någon
- 7) på rak _____
- 8) på löpande _____
- 9) ha ett _____ med i spelet

- 10) gå bakom _____ på någon
11) inte för _____ öre
12) ingen dans på _____
13) med något i högsta _____
14) rinna ut i _____
15) få upp _____ för något

D

Vilka ord passar bäst tillsammans? Välj det ord som du tycker passar bäst bland de olika alternativen. Tänk på att alla de alternativa orden inte behöver användas!

Exempel:

Fyll i det ord som saknas i följande uttryck. Välj bland följande alternativ: *bilda, dra, föra, hålla, leda, lägga, spela, ta*
dra en slutsats

Fyll i det ord som saknas i följande uttryck. Välj bland följande alternativ: *hamna, hoppa, komma, leva, ligga, sitta, stå*
leva i celibat

Fyll i det ord som saknas i följande uttryck. Välj bland följande alternativ: *djup, evigt, full, negativ, politisk, stor, svensk, typisk*
svensk medborgare

Fyll i det ord som saknas i följande uttryck. Välj bland följande alternativ: *djupt, evigt, exakt, fullt, mycket, stort, särskilt, typiskt*
exakt likadan

Fyll i det ord som saknas i följande uttryck. Välj bland följande alternativ: *bilda, bryta, dra, föra, ge, kasta, leda, lägga*

- 1) _____ en blick
- 2) _____ ett exempel
- 3) _____ ett beroende
- 4) _____ en union
- 5) _____ en suck

Fyll i det ord som saknas i följande uttryck. Välj bland följande alternativ:
få, gå, hamna, komma, ligga, sträva, stå, vara

- 6) _____ i uppgift
- 7) _____ till nytta
- 8) _____ i missbruk
- 9) _____ till grund
- 10) _____ i arv

Fyll i det ord som saknas i följande uttryck. Välj bland följande alternativ:
allvarlig, djup, full, lätt, mycket, olika, stark, stor

- 11) _____ slags
- 12) ett _____ andetag
- 13) _____ fart
- 14) en _____ sjukdom
- 15) _____ betydelse

Fyll i det ord som saknas i följande uttryck. Välj bland följande alternativ:
Välj bland: *evigt, exakt, hemligt, knappast, lyckligt, negativt, politiskt, svårt*

16) _____ troligt

17) _____ sjuk

18) _____ ovetande

19) _____ aktiv

20) _____ tacksam

