

**UTVECKLING AV DET PRODUKTIVA
LEXIKONET I BERÄTTELSE SKRIVNA
AV FINSKA GYMNASISTER**

Marja Virtanen

Pro gradu – avhandling i svenska
Jyväskylä universitet
Institutionen för språk
Hösten 2015

Humanistinen tiedekunta	Kielten laitos
Tekijä: Marja Virtanen	
Työn nimi: Utveckling av det produktiva lexikonet i berättelser skrivna av finska gymnasister	
Oppiaine: Ruotsin kieli	Työn laji: Pro gradu – tutkielma
Aika: Lokakuu 2015	Sivumäärä: 83 + 1
<p>Tiivistelmä</p> <p>Sanastolla on keskeinen merkitys kielen oppimisessa ja osaamisessa. Varsinkin produktiivinen sanasto, jota kielenkäyttäjät pystyvät tuottaessaan hyödyntämään, on siinä erityisen tärkeässä roolissa sillä se mahdollistaa kommunikoinnin muiden kielenkäyttäjien kanssa. Tämän tutkielman tavoitteena oli selvittää, kuinka lukiolaisten produktiivinen sanavarasto kehittyi lukio-opintojen pakollisten ruotsin ja englannin kurssien aikana. Erityisenä mielenkiinnonkohteena olivat neljä avointa sanaluokkaa (substantiivit, verbit, adjektiivit, adverbit), joiden kehittymistä tarkasteltiin pääasiassa kvantitatiivisten tutkimusmenetelmien, kuten TTR-metodin, avulla. Laajemman kokonaiskuvan saamiseksi kvantitatiivista analyysia täydennettiin erillisillä, osin kvalitatiivisilla analyyseilla, joissa tutkittiin muun muassa adverbiaaleja, erisnimiä sekä kieltosanoja. Tutkimusaineisto koostui kaiken kaikkiaan 150 kirjoitetusta kertomuksesta, jotka oli kerätty Topling-tutkimusprojektin puitteissa kolmella eri mittauskerralla vuosina 2010–2013.</p> <p>Tutkimukseni tulokset osoittivat, että lukiolaisten produktiivinen sanasto kehittyi tietyiltä osin sekä ruotsissa että englannissa. Kehityksessä oli kielten välillä sekä eroja että yhtäläisyyksiä. Avointen sanaluokkien prosentuaaliset osuudet tekstien kokonaissanamäärästä kehittyivät samansuuntaisesti kummassakin kielessä. Toisaalta etenkin ruotsissa avointen sanaluokkien sanastosta tuli vaihtelevampaa opintojen aikana, kun taas englannissa niiden kehitys ei ollut yhtä suoraviivaista. Molemmissa kielissä erityisesti verbit kehittyivät melko vähän. Sen sijaan adverbiaaleja käytettiin selvästi monipuolisemmin kieliopintojen loppuvaiheessa kummassakin kielessä. Lisäksi tulokset osoittivat, että ruotsinkieliset kertomukset sisälsivät englanninkielisiä tarinoita huomattavasti enemmän sellaisia sanoja jotka jouduttiin jättämään varsinaisten analyyksien ulkopuolelle (mm. vieraskieliset sanat) ja täten analysoimaan omana erillisenä ryhmänä.</p> <p>Tutkimuksessa saadut tulokset tarjosivat pienen katsahduksen lukiolaisten sanaston kehitykseen. Tulevissa tutkimuksissa aihetta voisi tarkastella esimerkiksi eri tekstityyppien näkökulmasta.</p>	
Asiasanat: ord, lexikala ord, produktivt lexikon, ordklasser, svenska, engelska, gymnasium, TTR	
Säilytyspaikka: JYX	
Muita tietoja: Sivuainetutkielma englannin kieleen	

INNEHÅLL

1	INLEDNING	7
2	TEORETISK BAKGRUND.....	10
2.1	Ett ord och problematiken omkring dess definitioner.....	10
2.2	Lexikala ord och grammatiska ord.....	13
2.3	Öppna ordklasser och ordklassindelning i svenska och engelska.....	14
2.3.1	Substantiv	16
2.3.2	Adjektiv	16
2.3.3	Verb	17
2.3.4	Adverb	18
2.3.5	Adverbial	18
2.3.6	Ordklassernas andelar i skriftliga texter	19
2.4	Receptivt och produktivt lexikon	20
2.5	Produktiv ordanvändning	23
2.6	Ordinläring och ordinlärningsstrategier	24
2.6.1	Ordinläringprocess och utveckling av lexikonet	25
2.6.2	Ordinlärningsstrategier	27
2.7	Lexikonets storlek och problematiken med uppskattningar.....	29
2.8	Läroplanen (GLGY 2003) och lexikonet	30
2.9	Resultat i tidigare forskning	32
3	SYFTE, MATERIAL OCH METOD.....	34
3.1	Syfte	34
3.2	Material	35
3.2.1	Toppling-projektet och materialet i allmänhet	35
3.2.2	Materialinsamlingen	38
3.2.3	Informanterna	39
3.3	Metod	40
3.3.1	Bakgrunden: kvantitativa och kvalitativa forskningstraditioner.....	40
3.3.2	Analysens gång och analysmetoder	41
4	RESULTAT.....	46
4.1	Den kvantitativa analysen – svenska.....	46
4.1.1	Förändringar i ordklassernas procentuella andelar	46

4.1.2	Förändringar i ordvariation inom ordklasserna	48
4.2	De separata analyserna – svenska	50
4.2.1	Adverbial	50
4.2.2	Egennamn	51
4.2.3	Negationsord.....	52
4.2.4	Ord som utelämnades i andra analyser	53
4.3	Den kvantitativa analysen – engelska	56
4.3.1	Förändringar i ordklassernas procentuella andelar	56
4.3.2	Förändringar i ordvariation inom ordklasserna	58
4.4	De separata analyserna – engelska	60
4.4.1	Adverbial	60
4.4.2	Egennamn	61
4.4.3	Negationsord.....	62
4.4.4	Ord som utelämnades i andra analyser	63
4.5	Jämförelser mellan svenska och engelska.....	65
5	DISKUSSION	69
5.1	Sammandrag av resultaten	69
5.2	Validitet och reliabilitet.....	74
6	AVSLUTNING	78
	LITTERATUR.....	80
	BILAGA 1: BERÄTTELSEUPPGIFTEN.....	84

1 INLEDNING

Inläring av ord på ett främmande språk är en komplex process (Tornberg 2009: 130, 136), men det har visats att språkinlärare ändå tycker att den är en mycket betydande del av språkinläringen (se t.ex. Jokinen & Pitkäljärvi 2005). Enligt gymnasister kan ordinläringen vara språkkunskapernas allra viktigaste delområde eftersom ord behövs för all kommunikation (Jokinen & Pitkäljärvi 2005: 64). Lexikonet är därmed en essentiell del av språkkunskaper, eftersom det anknyter till språkkunskapernas alla delområden, såsom talande och skrivande, och möjliggör därigenom kommunikation med andra människor (jfr Enström 2004: 171, Nation 2001). Utan ett tillräckligt omfattande lexikon är det således omöjligt att uttrycka ens känslor och åsikter på ett lyckat sätt.

Utöver det tillräckliga lexikonet är det också viktigt att veta hur språkbrukare kan lära sig nya användbara ord och fraser som de sedan kan använda effektivt i sin egen produktion i varierande kontexter. Ordinläringen är egentligen en process som pågår hela livet; språkanvändare lär sig kontinuerligt såväl nya ord som nya betydelser hos ”gamla” ord även i modersmålet (Thornbury 2002: 1). Därför påpekar Enström (2004: 184) att språkanvändarnas lexikala kunskaper aldrig kan vara fullständigt färdiga. Av dessa olika anledningar är det centralt att vara medveten om hur nya ord lärs in och hur lexikonet utvecklas under språkinlärningsprocessen.

Syftet med denna magisteravhandling är att redogöra för hur finska gymnasisternas produktiva lexikon i de två obligatoriska språkämnen, svenska och engelska, utvecklas under språkstudierna i den treåriga finska gymnasieutbildningen. I stället för hela lexikonet har jag avgränsat analyseringen till fyra öppna ordklasser, det vill säga substantiv, adjektiv, verb och adverb, som också kan kallas lexikala ord. Jag betraktar temat ur tre olika synvinklar: å ena sidan ligger intresset på förändringar mellan ordklassernas användning i skriftliga texter och å andra sidan på förändringar inom ordklasserna. Dessutom kartlägger jag potentiella likheter och skillnader mellan de svenska och de engelska texterna. Som informanter har jag gymnasister från flera gymnasier i två olika områden i Finland, nämligen Österbotten och Mellersta Finland. Materialet omfattar sammanlagt 150 berättelser varav hälften (75 stycken) är skrivna på engelska och hälften (75) på

svenska. Texterna har insamlats inom ramar för Topling-projektet (forskningsprojektet *Inlärningsgångar i andraspråket*, se också avsnitt 3.2.1) år 2010-2013.

Undersökningen är mestadels kvantitativ: berättelserna analyseras genom att använda några relativt enkla räkneoperationer, såsom procenträkning, medelvärden och TTR (*type-token ratio*). TTR är en metod som kan utnyttjas för att ta reda på lexikal diversitet i texter: *token* står för antalet alla ord i en text och *type* uppger antalet olika ord i texten. Siffran (*ratio*) demonstrerar deras inbördes relation (*type / token*). (Milton 2009: 125-126). Numeriska resultat kompletteras med mer kvalitativa analyser och därför analyseras exempelvis adverbial, egennamn och negationsord separat. Dessutom excerpas belysande exempel ur materialet.

Föreliggande studie har tre centrala begrepp: ett ord, lexikala ord och ett produktivt lexikon. *Ett ord* definieras på basis av Malmgrens (1994: 12) förslag: ett ord omfattar såväl ordets grundform som dess olika böjningsformer. Med *lexikala ord* avses innehållsord, med andra ord substantiv, adjektiv, verb och adverb (Thornbury 2002: 3, Carter 2012: 23-24). Den tredje viktiga termen, *ett produktivt lexikon*, används för att hänvisa till ord som språkbrukare kan använda i sin egen muntliga eller skriftliga språkproduktion (Milton 2009: 13). Alla dessa tre centrala begrepp kommer att diskuteras och definieras noggrannare i kapitel 2.

Allmänt anser jag att det är viktigt att undersöka detta tema, eftersom det mig veterligen inte finns några tidigare studier om exakt samma tema och det finns därmed behov av preliminära studier inom forskningsfältet. Dessutom utgör de fyra öppna ordklasserna en central del av lexikonet och de har en viktig roll i kommunikation på grund av deras relativt stora informationsinnehåll (se avsnitt 2.2). Jag förväntar också att studien erbjuder ny och nyttig information om lexikala inlärningsgångar för språklärare samt språkforskare.

Härnäst ska jag presentera teoretiska utgångspunkter till denna studie och definiera dess centrala begrepp i mer detalj. I kapitel 2 kommer jag således att redogöra för till exempel egenskaper hos öppna ordklasser, kunskaper som behövs för att kunna använda ett ord produktivt och några viktiga aspekter med tanke på ordinlärning i allmänhet. Efter

den teoretiska bakgrunden går jag över till att beskriva syfte, material och metod i föreliggande undersökning: jag presenterar de exakta forskningsfrågorna, materialet, informanterna samt de valda metoderna i kapitel 3. Därefter rapporterar jag mina forskningsresultat. I kapitel 5 ska jag diskutera dessa resultat och de anknyts till studiens teoretiska referensram. Till sist sammanfattas studiens relevantaste synpunkter i kapitel 6.

2 TEORETISK BAKGRUND

I detta kapitel beskrivs ett urval teoretiska utgångspunkter som är viktiga i föreliggande undersökning. Först diskuterar jag olika förslag till potentiella definitioner av begreppet ord (avsnitt 2.1). Därefter ska jag redogöra för lexikala och grammatiska ord i avsnitt 2.2. Öppna ordklasser och deras typiska andelar i skriftlig kommunikation presenteras i avsnitt 2.3 och distinktionen mellan receptivt och produktivt lexikon betraktas i avsnitt 2.4. I detta kapitel ska jag även behandla produktiv ordanvändning (avsnitt 2.5), ordinlärning och utveckling av lexikon i allmänhet (avsnitt 2.6), lexikonets storlek (avsnitt 2.7) samt läroplanen i gymnasieutbildningen (avsnitt 2.8). Till slut ger jag en sammanfattande överblick av tidigare forskning om liknande teman i avsnitt 2.9.

2.1 Ett ord och problematiken omkring dess definitioner

Ett ord är ett av de tre mest centrala begreppen i föreliggande undersökning. Såsom till exempel Hultman (2003: 28) påpekar är det ett så vanligt begrepp i vardagen att det kan vara svårt att lägga märke till problematiken omkring termen och dess olika definitioner. Fast det kanske inte är viktigt att definiera begreppet *ord* i vardaglig språk användning, påverkar emellertid ens förståelse av ordet på ett centralt sätt hur man förstår hela lexikonet och dess struktur. Valet av en viss definition avgör exempelvis vilket ord som egentligen räknas som ett ord, vilket vidare påverkar hur totalantalet ord i texter räknas (jfr Hultman 2003: 30).

Det finns således ett relativt stort antal olika definitioner av ordet. En orsak till detta är att lingvister har försökt definiera termen redan under flera decennier, vilket har resulterat i varierande förslag till lämpliga definitioner. Trots att det finns många alternativa definitioner av begreppet, har det dock visat sig vara rätt så svårt att definiera det på ett entydigt och vetenskapligt hållbart sätt. (Hultman 2003: 28, Carter 2012). I detta avsnitt presenterar jag några av de vanligaste definitionerna av begreppet *ord* och problematiken omkring dem.

En av de mest kända definitionerna är troligen en ortografisk definition av ett ord (Malmgren 1994: 12, Carter 2012). Enligt den ortografiska definitionen kan ett ord identifieras i skrivna texter med hjälp av mellanrum: ordet är således ”en följd av bokstäver som står mellan två spatier” (Malmgren 1994: 12). Carter (2012: 20) konstaterar att denna definition i första hand syftar på skrivna texter eftersom ordgränser inte är lika tydliga i tal. Definitionen är dock inte fullständigt hållbar ens i skrift eftersom det finns många undantag från definitionen: enligt den innehåller exempelvis det engelska ordet ”will not” två ord och ordet ”cannot” endast ett ord (ibid.). Möjligheten att skriva ett och samma ord på flera olika sätt är också problematiskt med tanke på denna definition: till exempel det engelska ordet ”postbox” kan skrivas som ”post-box” eller ”post box”. Ett problem med definitionen är även faktumet att ordets olika böjningsformer betraktas som skilda ord: därigenom räknas exempelvis böjningsformer av det svenska ordet ”hund” (hunden, hundar, hundarna) som olika ord även om de har ungefärligen samma betydelse och syftar till en viss typ av referent i världen. Eftersom den ortografiska definitionen av ordet inte tar hänsyn till vare sig grammatiken eller betydelsen hos ord, kan den konstateras att vara ofullkomlig i viss mån. (Carter 2012: 20).

En annan välkänd definition av ordet är följande: “Ord är den minsta språkbit som ensam kan utgöra ett yttrande” (Hultman 2003: 29). Definitionen är ursprungligen formulerad av Bloomfield, och dess grundläggande utgångspunkt är stabiliteten hos ord. Det betyder att ordet kan vara som självständigt yttrande och kan användas till exempel som svar på en fråga. (Carter 2012: 21, Hultman 2003: 28) Problematiskt med denna formulering är dock dess väsentliga antagande om ordets stabilitet: om ordet kan stå ensamt, kan det inte vidare indelas, vilket exempelvis gör analyseringen av idiom utmanande. Dessutom påpekar Carter (2012: 21, se även Sjöström 2006: 76) att det finns en stor mängd ord som inte kan stå ensamt, såsom konjunktioner, artiklar och prepositioner.

Det finns också orddefinitioner som baserar sig på semantik. Ordet kan exempelvis definieras som språkets minsta meningsfulla enhet vilket möjliggör identifieringen av olika betydelser hos ett ord så länge som de är separata semantiska enheter (Carter 2012: 21). En nackdel med detta definitionsförsök är dock dess antagande om entydiga relationer mellan alla ord och deras betydelse. Det finns till exempel ord vars betydelse är singular, fast själva ordformen består av flera parter. I dessa fall kan ordets betydelse

inte analyseras separat utan att förlora ordets riktiga betydelse. Exempel på sådana svenska ord är "språklärare" och "uttalsundervisning" samt engelska ord "bus conductor" och "model railway". (ibid.).

Malmgren (1994: 13) påpekar att den potentiella homonymin hos ett ord dock är den faktor som brukar orsaka det allvarligaste problemet med orddefinitionerna. Enligt Sjöström (2006: 42) avses med homonymi två ord som uttalas eller stavas på samma sätt men som betraktas som olika ord. Ett exempel på ett sådant ord som har en identisk grundform men olika betydelser är ordet "fot", vilket kan hänvisa till antingen en kroppsdel eller ett mått (Sjöström 2006: 42). Såsom Carter (2012: 20) framför är den centrala frågan här om dessa ord ska behandlas som ett ord eller flera olika ord. Malmgren (1994: 13) föreslår att svaret på frågan är att betydelskillnader mellan ord avgör om det är frågan om samma ord eller inte.

Även flera andra faktorer kan göra formuleringen av en vetenskapligt hållbar definition av ordet svårare. Carter (2012) framför att exempel på sådana faktorer är ordbetoning och idiom (se också Malmgren 1994). Man har försökt definiera ordet som enhet med bara en betonad stavelse. Problematiskt med denna orddefinition är emellertid faktumet att alla ord inte har betoning: till exempel prepositioner, pronomen och konjunktioner betonas inte i vanligt språkbruk. (Carter 2012: 22). Utöver ordbetoningen orsakar också idiom problem, såsom Malmgren (1994: 12-13) påpekar, eftersom man inte kan dra slutsatser av deras helhetsbetydelse genom att analysera betydelser hos skilda komponenter. Därför uppger han att problem som idiom orsakar till orddefinitioner kan vara omöjliga att avgöra på något entydigt sätt.

Enligt Malmgren (1994: 12) är det oftast mest praktiskt att definiera ett ord så att det omfattar både ordets grundform och olika böjningsformer. Han framför vidare att grundformen är den form som vanligen står som uppslagsord i ordböcker och den anses därför vara som representant för ordet (så kallad *lemma*). Hultman (2003: 30) anger att det är frågan om lemma-nivån när man exempelvis vill redogöra för hur litet eller stort lexikon en författare har. Dessutom är definitionen högst praktisk: Carter (2012: 22-23) förklarar att den inkluderar ordets grammatiska varianter såsom verbböjning och möjliggör polysemi hos ord. Därigenom antas det att till exempel former "komma"

och ”kommer”, ”är” och ”var” samt ”hunden” och ”hundar” hör till samma ord (”komma”, ”vara”, ”hund”). Av dessa olika praktiska anledningar använder jag denna definition av ordet som en grundläggande utgångspunkt i min avhandling.

2.2 Lexikala ord och grammatiska ord

Enligt Thornbury (2002: 3) kan ord och ordklasser indelas i två huvudkategorier: *lexikala ord* (eng. *content words*) och *grammatiska ord* (eng. *grammatical / functional words*). *Lexikala ord* är så kallade *innehållsord*, vilket betyder i praktiken att ett relativt stort informationsinnehåll är typiskt för dem. Verb, substantiv, adjektiv och adverb används ofta som klassiska exempel på lexikala ord (Carter 2012: 23-24). *Grammatiska ord* kan däremot exemplifieras med hjälp av ordklasser såsom artiklar, prepositioner och konjunktioner, och deras viktigaste uppgift är att visa och sätta relationer mellan de mer innehållstunga lexikala orden (Sjöström 2006: 76, Hultman 2003: 37). I denna studie ligger intresset dock först och främst på de fyra ovannämnda lexikala ordklasserna.

Lexikala ord och grammatiska ord skiljer sig från varandra på några väsentliga sätt. Enligt Carter (2012: 24) är det vanligt att lexikala ord har många ordformer, såsom böjningsformer; grammatiska ord har däremot oftast bara en ordform som representerar ordet (*lemma*; se också avsnitt 2.1). Dessutom är en viktig skillnad mellan lexikala ord och grammatiska ord deras totalantal: exempelvis Carter (2012: 23-24) konstaterar att antalet grammatiska ord är relativt begränsat, medan antalet innehållsord är mycket sannolikt obegränsat. Han påpekar också att form och betydelse hos lexikala ord kan förändras under tiden; grammatiska ord är däremot mer stabila till sin natur och förändras inte så lätt. Detta är också en central orsak till varför begreppet *lexikala ord* ofta används för att referera till så kallade *öppna ordklasser*. På ett motsvarande sätt kan termen *grammatiska ord* kopplas till *slutna ordklasser* (Carter 2012: 24). Även Hultman (2003: 37) tar upp detta förhållande mellan dessa begrepp. Han förklarar att typiskt för öppna ordklasser är att de kontinuerligt kan få nya medlemmar eftersom nya ord behövs för nya företeelser. Han menar vidare att ord som hör till slutna ordklasser kan i stället anges i en form av en lista eftersom dessa ordklasser sällan eller aldrig får nya medlemmar.

I nästa avsnitt (avsnitt 2.3) diskuteras den traditionella ordklassindelningen lite noggrannare. Jag ska också sammanfatta de viktigaste egenskaperna hos de fyra öppna ordklasserna (eller: lexikala ord) i svenska och engelska. Dessutom kommer jag att beskriva hur andelen ordklasser vanligen distribueras i skriftliga texter.

2.3 Öppna ordklasser och ordklassindelning i svenska och engelska

Ett vanligt sätt att beskriva lexikonets struktur är att använda den traditionella ordklassindelningen som utgångspunkt. Eftersom ordklasser och lexikon tydligt anknyter till varandra, är det nödvändigt att definiera vad begreppet *lexikon* egentligen betyder. Till exempel Sjöström (2006: 73) ger en användbar definition till det: ”Lexikonet utgör den samling av språkliga symboler språkanvändaren har i sitt minne och vars enheter hon eller han kan känna igen och återkalla vid behov.” Han konstaterar att lexikonet består av relationer som bildar ett nätverk och ”där uppdelningen av ordförrådet i lexikala klasser, eller ordklasser, är den mest grundläggande.” Han (2006: 81-82) nämner att ord i svenska språket traditionellt indelas i sammanlagt nio ordklasser: verb, substantiv, adjektiv, adverb, prepositioner, pronomen, konjunktioner, räkneord och interjektioner. Samma ordklasser räknas också i engelskan, men man brukar också tillägga två övriga ordklasser, nämligen hjälpverb (*auxiliary verbs*) och bestämmare (*determiners*). Därmed finns det elva ordklasser i engelska språket. (Greenbaum 1996: 90, 92). Fokus i denna undersökning ligger dock inte på alla dessa ordklasser utan jag har avgränsat det till de fyra öppna ordklasserna: substantiv, adjektiv, verb och adverb. Dessa ordklasser presenteras noggrannare nedan.

Enligt Sjöström (2006: 79) baserar ordklassindelningen i svenska sig på tre olika slags kriterier: syntaktiska, morfologiska och semantiska. Samma kriterier gäller också engelskans ordklassindelning (Greenbaum 1996: 90). Även Hultman (2003: 36) har en likartad uppfattning: han påpekar att ord kan formulera ganska naturliga grupper, eftersom de ofta har liknande betydelser, böjningar och grammatiska funktioner i fraser. Därför menar han att sådana grupper kan användas väl som grund för att indela språkets ord i olika ordklasser. Sjöström (2006: 79) redovisar att syntaktiska kriterier handlar om hur ett ord kan kombineras med andra ord i fraser och satser. Greenbaum (1996: 95) tillägger att de således omfattar ordets grammatiska funktioner i relation till andra ord. Mor-

fologiska kriterier refererar däremot till ordets formvariation, det vill säga hur ord i en viss ordklass kan böjas. Exempel på detta är pluralformer hos substantiv samt verbböjning. (Sjöström 2006: 79, Greenbaum 1996: 94). Till slut gäller semantiska kriterier typiska betydelser, med andra ord generaliseringar, hos ord i en ordklass. Till exempel substantiv är vanligen namn på saker, människor och platser. (Sjöström 2006: 79, Greenbaum 1996: 94). Fast exempelvis Greenbaum (1996: 95) framför att de syntaktiska kriterierna är mest reliabla kriterierna för att fastställa ordklassindelningen, verkar det dock vara så att ingen av de tre ovan presenterade kriterierna är ensamt tillräckligt fullständig eller omfattande. Alla tre kriterierna behövs för en bra beskrivning av ordklassindelningen eftersom de kompletterar varandra (jfr Greenbaum 1996: 95).

Den traditionella ordklassindelningen är dock inte fullständigt oproblematiserad. En anledning till detta är, såsom Hultman (2003: 36) konstaterar, att de traditionella ordklasserna inkluderar undergrupper som vidare består av varierande ord. En annan orsak till problematiken omkring ordklassindelningen är faktumet att ord kan höra till flera olika ordklasser samtidigt (ibid.). Till exempel Thornbury (2002: 3) påpekar att en stor del av engelska ord kan klassificeras samtidigt till två eller flera ordklasser. Detsamma gäller också svenska ord. Ett illustrerande exempel på denna företeelse är det svenska ordet ”fjärran” vilket kan höra till fyra ordklasser: adjektiv, adverb, substantiv och prepositioner. (Hultman 2003: 36).

Trots att den traditionella ordklassindelningen har sina brister, har jag i min pro gradu-avhandling emellertid valt att betrakta ord ur dess synvinkel. Jag koncentrerar mig på ordklasserna i stället för ordgrupper (såsom fraser, idiom och kollokationer) eftersom jag är intresserad av hur användningen av de lexikala ordklasserna förändras i skriftliga texter under gymnasiet. Dessutom ligger intresset på språklig variation inom ordklasser (se en noggrannare beskrivning av syftet i avsnitt 3.1). Eftersom ord samtidigt kan höra till många olika ordklasser, avgörs varje analyserat ords huvudsakliga ordklass på basis av kontexten. Härnäst sammanfattar jag de viktigaste egenskaperna hos de fyra öppna ordklasserna i var sitt eget avsnitt (avsnitt 2.3.1 – 2.3.4). Beskrivningarna gäller både svenska och engelska, men även eventuella skillnader mellan språken redogörs vid behov.

2.3.1 Substantiv

Vanligen betecknar substantiv föremål, människor, djur och företeelser (Hultman 2003: 39). I svenskan har de grammatiskt genus, och de kan därigenom indelas i två grupper: utrum (*n*-ord, såsom ”en katt”) och neutrum (*t*-ord, såsom ”ett hus”) (Hultman 2003: 48). Även i engelskan kan substantiv klassificeras i två huvudgrupper på basis av användningen av de två obestämda artiklarna ”a” och ”an”. Substantiv har oftast både singular- och pluralform: i svenskan kan substantivens pluralformer grupperas i fem olika deklinationer medan pluralformerna i engelskan oftast anges med hjälp av suffixen –s (”a dog” – ”two dogs”). Dessutom brukar man göra distinktionen mellan artnamn och egennamn. (Greenbaum 1996: 96-97). Såsom till exempel Sjöström (2006: 119) påpekar, hänvisar artnamn till så kallade vanliga substantiv. De kan vara antingen räkningsbara eller icke-räkningsbara / kollektiva; denna distinktion motsvarar i viss mån distinktionen mellan konkreta (”ett hus”, ”en hund”) och abstrakta (”kärlek”) substantiv (Greenbaum 1996: 97-98). Egennamn är däremot namn på bland annat specifika människor, platser och institutioner och de har därmed unika referenser. Detta betyder också att egennamn är icke-räkningsbara. (Greenbaum 1996: 97). Enligt Hultman (2003: 39) är substantivens viktigaste uppgift att fungera som kärna i nominalfraser: ”hennes snälla hund” och ”their beautiful house”.

2.3.2 Adjektiv

Adjektiv uttrycker egenskaper hos huvudordet: Alexander (1988: 106) konstaterar att adjektiv exempelvis kan användas för att beskriva huvudordets kvalitet, storlek, form och färg. Sjöström (2006: 87) menar vidare att egenskaperna kan vara antingen graderbara, såsom ”lång” och ”kort”, eller icke-graderbara, såsom ”död” och ”levande”. Graderbara adjektiv brukar ha sammanlagt tre former: positiv, komparativ och superlativ. Dessa former möjliggör jämförelser mellan olika slags föremål. (Greenbaum 1996: 139). Hultman (2003: 40) påpekar att utöver varierande egenskaper kan adjektiv även beteckna tillstånd, såsom ”en ren tröja”. Adjektiv kan användas antingen attributivt, det vill säga inför substantiv (”ett vackert hus”, ”a beautiful house”), eller predikativt, vilket då vanligen betyder att de används i samband med vara- och bli-verb (”Huset är vackert”, ”The house is beautiful”) (Greenbaum 1996: 134, se också Alexander 1988).

2.3.3 Verb

Verb utgör en ordklass som består av ord med varierande betydelser: verb kan exempelvis beteckna aktivitet ("leka"), övergång ("blekna") och tillstånd ("leva"). Deras viktigaste funktion är att vara som huvudord i verbfraser och de ligger till grund för satsernas konstruktion. (Hultman 2003: 41). Sjöström (2006: 88-90) påpekar att verb ofta anses vara den mest centrala ordklassen eftersom de "ställer specifika krav på sin syntagmatiska omgivning". Han förklarar att verb kan indelas till exempel i dynamiska och statiska samt transitiva och intransitiva verb. Enligt honom uttrycker dynamiska verb en förändring (till exempel verb "försvinna" och "växa") medan statiska verb uttrycker en oföränderlig process ("vara" och "finnas"). Hultman (2003: 143) framför att skillnaden mellan transitiva och intransitiva verb är att de förstnämnda kompletteras i normala fall med objekt.

Verb kan syfta till antingen nutid (presens) eller dåtid (preteritum), vilket innebär att verb har två huvudsakliga tempuskategorier (Greenbaum 1996: 81, Hultman 2003: 152). Presens, perfektum och futurum används för att uttrycka nutid, medan preteritum, pluskvamperfektum och futurum preteritum uttrycker dåtid (Hultman 2003: 262). Med tanke på tempusformer är engelska verb lite mer komplexa till sin natur jämfört med svenska verb. Alexander (1988: 159) påpekar att båda tempuskategorierna i engelskan inkluderar två olika former: *simple* (till exempel "I write") och *progressive* (till exempel "I am writing"). En liknande indelning i två former finns inte i svenska språket utan den sistnämnda formen (*progressive*) uttrycks med hjälp av andra verbkonstruktioner, såsom "hålla på att göra något".

En annan viktig skillnad mellan svenskan och engelskan är också antalet böjningskategorier. Verb i svenskan kan nämligen indelas i fyra böjningskategorier (eller fem om oregelbundna verb utgör en egen grupp) medan verb i engelskan brukar man kategorisera på basis av böjningen i två huvudgrupper: regelbundna och oregelbundna verb (för en noggrannare deskription av engelskans verbböjning, se t.ex. Greenbaum 1996).

2.3.4 Adverb

Enligt Greenbaum (1996: 141) utgör adverb en relativt heterogen ordklass eftersom deras funktioner och platser i satser kan variera. Dessutom menar han att ordklassen består av några överlappande undergrupper. Adverb är nämligen en grupp av ord som används för att uttrycka mycket varierande betydelser: allmänt taget kan det konstateras att de betecknar olika typer av omständigheter (Hultman 2003: 168). De kan således uttrycka tid ("då"), rum ("här"), grad ("ganska"), sätt ("skickligt"), orsak ("slumpmässigt") och modalitet ("knappast") (Sjöström 2006: 93). Hultman (2003: 169) tillägger att de även kan beteckna logiska relationer ("alltså"), negation ("inte") och talarens attityd ("gärna"). I denna undersökning ska jag inkludera alla dessa typer av adverb när jag analyserar materialet och redogör för utvecklingen av adverb (se också kapitel 3 för en mer exakt beskrivning av materialet och analyseringen). Det enda undantaget är dock negationsord (det svenska ordet "inte" och det engelska ordet "not"): Greenbaum (1996: 93) poängterar att negationsord i engelskan inte hör till ordklassen adverb. Av denna anledning kan jag inte analysera dem som en del av ordklassen adverb utan de kommer att analyseras separat både i svenska och engelska.

Alexander (1988: 122) påpekar att adverb kan bestämma verb, adjektiv, övriga adverb, prepositionsfraser, substantiv och hela meningar. En del av adverb kan också böjas och kompareras: till exempel ordet "gärna" har komparativformen "hellre" och superlativformen "helst" (Hultman 2003: 42, 173). Ett närliggande begrepp till adverb är *adverbial* som kommer att betraktas lite närmare i nästa avsnitt.

2.3.5 Adverbial

Såsom det konstaterades i det föregående avsnittet (avsnitt 2.3.4) är begrepp *adverb* och *adverbial* nära varandra. Det finns dock åtminstone en väsentlig skillnad mellan dem: exempelvis Greenbaum (1996: 141-142) redogör för denna skillnad och framhäver att termen *adverb* är namnet på en ordklass medan *adverbial* refererar till en satsdel och utgör därmed inte någon ordklass. På grund av detta analyseras adverb som en del av den kvantitativa analysen i föreliggande studie medan adverbial betraktas separat med hjälp av mer kvalitativa analysmetoder (se också kapitel 3 för en mer exakt beskrivning av analysmetoder i studien).

Adverbial är oftast frivilliga satsdelar, vilket betyder att en mening kan kvarstå grammatiskt korrekt och begriplig även om man tar bort alla dess adverbial (Greenbaum 1996: 70). Det finns emellertid adverbial som är nästan obligatoriska satsdelar: till exempel Hultman (2003: 239) framför att "[v]issa verb mer eller mindre kräver adverbial (...) I sådana fall kan adverbialen inte gärna utelämnas, i varje fall inte utan att verbet får en annan betydelsenysans". Exempel på sådana fall är: "Kaisa reser till Sverige", "Hon kommer från Finland", "Hon vill gå genom skogen" (se också Hultman 2003: 239).

Adverbial kan alltså ha en viktig kommunikativ betydelse, eftersom de kan erbjuda nyttig information (Greenbaum 1996: 70). Enligt Hultman (2003: 238) kan adverbial förekomma i tämligen varierande former. Exempel på dem är prepositionsfraser ("på natten"), bisatser ("eftersom hon älskar sitt arbete"), *med*-fraser ("med pappret i hand", "med honom") och nominalfraser ("nästa år"). Adverbialen kan således uttrycka bland annat tid (tidsadverbial: "på morgonen") och rum (rumsadverbial: "i Finland"). (ibid.: 238-239). Eftersom adverb kan förmedla liknande betydelser, anses adverb och adverbial ofta som närliggande. Adverbialen inkluderar dock ofta flera ord, medan adverb är i många fall uttryck som består av bara ett ord (se exempel i avsnitt 2.3.4).

Eftersom det finns ett relativt stort antal olika adverbial, har intresset i denna studie begränsats till bara vissa typer av adverbial, det vill säga till adverbial som är mest relevanta med tanke på studiens syfte och omfattning. Därför ligger fokuset företrädesvis på följande typer av adverbial: adverbial som anger tid, rum, medel (t.ex. "resa med bil") och adverbial som är *med*-fraser (sv. "med mina föräldrar") respektive engelskans *with*-fraser (t.ex. "with my parents"). Däremot uteslöts hela bisatser från det analyserade materialet. (se också Hultman 2003: 238-9)

2.3.6 Ordklassernas andelar i skriftliga texter

Ordklassernas typiska andelar i skriftliga texter skiljer sig från varandra. Till exempel Melin och Lange (2000: 168) har sammanfattat ordklassernas vanliga andelar i olika skriftliga texttyper. På basis av detta sammandrag kan man konkludera att andel substantiv av löpord är störst i alla texttyper och deras antal varierar mellan 21,9 % (gym-

nasistprosa) och 27,6 % (nyheter). Ett stort antal substantiv gör en text nominal (Melin och Lange 2000: 15). Även Hultman (2003: 44) konstaterar att substantiv brukar vara den dominerande ordklassen: han skriver att deras andel i skriftliga texter vanligen är mellan 10 och 30 % av alla löpord, beroende på framställningssätt. Dessutom påpekar han att en hög andel substantiv är typisk för koncentrerat skriftspråk och en låg andel ”för talspråk och informellt, talspråksbetonat skriftspråk.”

Enligt Hultman (2003: 76) varierar andel adjektiv i allmänhet mellan 5 och 10 % av löpord i skrift. Melin och Lange (2000: 168) presenterar liknande andelar: det framgår av deras sammandrag att antal adjektiv är relativt stabilt i olika texttyper, eftersom det varierar mellan 5,9 % (nyheter) och 8,1 % (ledare). I gymnasistprosa är andel adjektiv 6,7 % (ibid.). Melin och Lange (2000: 15) uppger också att ”[a]djektiv kan ge texten ökad specifikation men bidrar inte till ökad täthet eller komplexitet.” Även andel adverb liknar andel adjektiv för den är också under 10 % av löporden i skriftliga texter. I gymnasistprosa är andel adverb 9,7 % vilken är lite större än andel adjektiv (6,7 %). (ibid.: 168).

Hultman (2003: 140) framför att ordklass verb vanligen omfattar approximativt 15-20 % av texternas löpord. Han konstaterar att ett stort antal verb (cirka 20 % eller mer) hänvisar till talspråk och ett lägre antal (cirka 15-16 %) till skriftspråk. Även Melin och Langes (2000: 168) sammanfattning visar likartade siffror till andel verb i skriftliga texter: deras andel varierar mellan 16,2 % (bruksprosa) och 18,9 % (gymnasistprosa). Dessutom framgår det också att andel verb av texternas löpord är näst störst (efter substantiv) i olika skriftliga texttyper. Om texten innehåller en hög andel verb, brukar texten kallas verbal (ibid.: 15).

2.4 Receptivt och produktivt lexikon

Utöver den ovan presenterade ordklassindelningen kan lexikonet beskrivas ur en helt annorlunda synvinkel och det kan därför indelas på åtminstone två övriga sätt. I stället för distinktionen mellan öppna och slutna ordklasser baserar dessa klassificeringar sig på uppfattningar om språkkunskapernas delfärdigheter och därigenom på kunskaper om ord som behövs eller förutsätts i olika språkanvändningssituationer. Inom lingvistik

har lexikonet nämligen traditionellt klassificerats i *ett produktivt* och *ett receptivt lexikon*. Med *ett produktivt lexikon* avses ord som inlärare själva kan använda i tal och skrift, medan med *ett receptivt lexikon* avses ord som inlärarna kan förstå (Milton 2009: 13). Utöver produktivt och receptivt lexikon har man också använt motsvarande begrepp *aktivt* respektive *passivt ordförråd* för att redogöra för denna klassificering av lexikonet. Exempelvis Nation (2001) diskuterar distinktionen och menar att receptivt eller passivt lexikon framför allt kan kopplas till lyssnande och läsning medan produktivt eller aktivt lexikon snarare refererar till skrivande och talande. Därför kan det konstateras att man brukar anknyta dessa två lexikon till olika delfärdigheter av språkkunskaper (se även Milton 2009: 13).

De två ovan presenterade klassificeringarna av lexikonet (passivt – aktivt respektive receptivt – produktivt lexikon) är dock inte fullständigt entydiga, eftersom dessa begrepp i sig inte är helt oproblematiska. Till exempel Enström och Holmegaard (1994: 175) uppger att begreppet *passiv* är vilseledande eftersom ”det verkligen inte är frågan om att vara passiv” och därför vore det bättre att använda termerna receptivt och produktivt lexikon. Även Tornberg (2009: 127) föreslår en alternativ term till passivt lexikon: hon konstaterar att i stället för passivt ordförråd borde man använda begreppet *ordförståelse* eftersom förståelsen egentligen har visats sig vara aktiv. Nation (2001: 24) har också en liknande uppfattning. Han menar att även termerna *receptiv* och *produktiv* är till en viss grad problematiska, eftersom receptiva språkfärdigheter är meningsskapande processer. Detta betyder att exempelvis läsning är en process som förutsätter produktion.

Melka (1997) diskuterar också problematiken omkring dessa begrepp och deras definitioner. Hon poängterar att det är nästan omöjligt att definiera det receptiva och det produktiva lexikonet på något entydigt sätt, vilket exempelvis beror på att det kan vara svårt att avgöra när språkbrukarens kunskap om ett ord är receptiv och när det är frågan om produktiv kunskap. Därför påpekar hon att det kan vara mer rimligt att betrakta relationen mellan det receptiva och det produktiva lexikonet som olika grader av kunskap, med andra ord som ett kontinuum, snarare än två helt annorlunda och separata kunskaper om ordet.

I denna avhandling använder jag emellertid begreppen produktivt och receptivt lexikon för tydlighetens skull: jag har intresserat mig för utvecklingen av det produktiva lexikonet i skriftliga texter som gymnasister har skrivit under sina språkstudier i gymnasiet. Eftersom jag har avgränsat min studie till skriftlig kommunikation, anser jag att det är motiverat att använda begreppet *produktivt lexikon* för att hänvisa till ord som informanterna verkligen använder i sina texter.

Det receptiva och det produktiva lexikonet skiljer sig från varandra på flera olika sätt. Enligt Enström (2004: 176) förutsätter förståelse och produktion olika slags kunskaper om ordet: hon påpekar att det är mer utmanande att använda ett ord i sin egen produktion än att kunna förstå det. Även Nation (2001: 24-25) menar att användning av det receptiva lexikonet är mindre krävande, eftersom det innebär att inlärare iakttar ett ord och därefter återkallar ordets betydelse från sitt minne. Det produktiva bruket av lexikonet kräver däremot att inlärarna kan både återkalla och producera en talad eller en skriven form av ordet (ibid.). Till exempel i receptiva språkanvändningstillfällen behöver inläraren troligen inte kunna ett ord fullkomligt utan en approximativ uppfattning om ordets betydelse är oftast tillräcklig, men för att kunna använda ett ord produktivt ska inläraren ha mera omfattande kunskaper om ordet (Enström 2004: 176). Dessa olika förutsättningar för den produktiva ordanvändningen diskuteras noggrannare i avhandlingens nästa avsnitt (avsnitt 2.5).

Det finns även några skillnader mellan det receptiva och det produktiva lexikonet angående deras storlek. Flera forskare (se t.ex. Thornbury 2002: 15, Milton 2009: 117, Carter 2012: 203) verkar vara eniga om att det receptiva lexikonet är större än det produktiva, vilket i praktiken betyder att språkanvändare förstår flera ord än de kan producera. Enligt Milton (2009: 117) antas det produktiva lexikonet omfatta ca 50-80 % av det receptiva lexikonet. Även Melka (1997: 92) har en liknande uppfattning: hon konstaterar att det receptiva lexikonets storlek uppskattas vara åtminstone dubbelt jämfört med det produktiva lexikonets storlek.

Utöver storleken av det produktiva och det receptiva lexikonet är också deras inbördes ordning en signifikant aspekt som antas skilja dem från varandra, eftersom förståelse vanligen föregår produktion (Thornbury 2002: 15). Enström (2004: 176-177) påpekar

dock att detta inte innebär att alla ord som inlärarna behärskar receptivt automatiskt blir en del av deras produktiva lexikon. Om inlärarna inte behöver använda vissa ord produktivt, så kvarstår de endast i deras receptiva lexikon (Enström och Holmegaard 1994: 175). Härnäst går jag över till att beskriva hur språkbrukare kan använda ord i sin egen produktion och vilka slags kunskaper om ordet det förutsätter.

2.5 Produktiv ordanvändning

Olika orddefinitioner (se avsnitt 2.1) och att kunna ett ord är relativt tätt anknutna till varandra. Dessutom är distinktionen mellan det receptiva och det produktiva lexikonet (se avsnitt 2.4) central om man vill förstå vilka slags kunskaper om ordet språkanvändare behöver i olika kontexter. I detta avsnitt ska jag således diskutera vad det egentligen betyder att kunna använda ett ord och vilka som är de mest väsentliga faktorer samt egenskaper hos själva ordet som språkanvändarna ska behärska. Intresset ligger alltså framför allt på den produktiva ordanvändningen.

Att kunna ett ord är en komplex process som omfattar olika dimensioner och varierande kunskaper om ordet. Om man betraktar processen på en mycket generell nivå, kan man konstatera att för att kunna ett ord ska språkanvändare behärska både ordets form och betydelse (Thornbury 2002: 15). Denna kunskap om ordet räcker emellertid relativt sällan och språkanvändare borde därför känna till även några andra aspekter av ordet: enligt Thornbury (2002: 15) ska språkbrukare veta till exempel ordets grammatiska funktioner (det vill säga ordklassen) och vilka kollokationer samt konnotationer ordet kan kopplas till. Dessutom ska de vara medvetna om såväl en talad och en skriven form av ordet samt dess frekvens (ibid.: 16).

Även Enström och Holmegaard (1994: 178) diskuterar olika delfärdigheter som utgör en basis för att kunna använda ett ord och som behövs i synnerhet i skriftlig kommunikation. De nämner att språkbrukare ska behärska en korrekt stavning och användning av ordets böjningsformer. Utöver dessa borde de vara medvetna om till exempel ordets syntaktiska egenskaper, såsom vilka prepositioner ska kopplas till vissa verb. De borde också känna till ordets bibetydelser vid sidan av dess huvudsakliga betydelse. Om ett ord har ett eller flera synonymer, skulle det vara idealt att språkanvändare skulle veta

olika betydelskillnader mellan dem. (Enström och Holmegaard 1994: 178, se också Enström 2004: 176).

Såsom det konstaterades redan i det föregående avsnittet (avsnitt 2.4) ställer ordets receptiva förståelse och produktiva användning annorlunda krav på språkbrukare, för de förutsätter olika slags kunskaper om ordet (Enström 2004: 176). Att producera ett ord ställer relativt höga krav i synnerhet på inlärarna eftersom de ska behärska många olika aspekter av ordet, såsom hur det stavas och i vilka sammanhang det kan användas. I situationer där inlärarna endast ska förstå ordet behöver de däremot inte kunna det lika fullständigt. (Enström 2004: 176). På basis av dessa synpunkter kan man dra två slutsatser: den första är att det är mer krävande att producera ett ord på ett korrekt sätt än att enbart förstå det rätt. Den andra slutsatsen är att en korrekt ordproduktion förutsätter ett större antal olika delfärdigheter och kunskaper om ordet jämfört med rätt ordförståelse.

Laufer (1997: 142) påpekar att i en ideal situation skulle språkanvändare känna till och behärska alla egenskaper hos ett ord. Hon framför att detta oftast gäller bara språkets utbildade förstaspråkstalare, medan mer typiskt för språkinlärare är att kunna ett ord delvis. Hon förklarar att språkinlärarna vanligen känner till och behärskar en del av egenskaper hos olika ord, men inte alla. Att kunna ett ord betyder inte, såsom Melka (1997: 87) påpekar, att språkanvändare vet ordet antingen fullständigt eller inte alls, utan det kan beskrivas som helhet som omfattar både produktiva och receptiva kunskaper om ordet. Även Thornbury (2002: 16) konstaterar att inte ens avancerade språkinlärare vet alla dimensioner av ord, vilket beror på att ordinlärning är progressiv och långvarig till sin natur. En viktig aspekt till diskussionen anges av Enström och Holmegaard (1994: 178): de poängterar att språkbrukare troligen inte ens behöver alla delfärdigheter och kunskaper för varje ord. Att känna till alla ord fullständigt och att kunna använda alla dessa ord produktivt är således inte nödvändiga för att uttrycka sig och att kommunicera med andra människor på ett lyckat sätt.

2.6 Ordinlärning och ordinlärningsstrategier

Utveckling av lexikonet och ordinlärningsstrategier kopplar i varandra, eftersom en effektiv användning av ordinlärningsstrategier kan påverka inlärares ordinlärning på ett

positivt sätt. I det föregående avsnittet (avsnitt 2.5) diskuterades vilka kunskaper om ett ord språkanvändare behöver för att kunna använda ett ord produktivt. I detta avsnitt har jag för avsikt att presentera hur lexikonet egentligen utvecklas och vilka som är de mest grundläggande faktorer som antagligen påverkar ordinlärningsprocessen (avsnitt 2.6.1). Utöver detta ska jag kortfattat beskriva olika ordinlärningsstrategier och hur språkinlära- ren kan utnyttja dem under språkinlärningsprocessen (avsnitt 2.6.2). Hela avsnittet skrivs huvudsakligen ur *andraspråkinlärares* (=L2-inlärare) synvinkel, men även skillnaderna mellan förstaspråkstalare och andraspråksinlärare tas upp vid behov.

2.6.1 Ordinlärningsprocess och utveckling av lexikonet

Forskare verkar vara eniga om att utvecklingen av lexikonet spelar en central roll i språkinläring (se t.ex. Thornbury 2002: 14), men det finns däremot inte allmänt samförstånd bland lingvister om hur språkbrukare egentligen tillägnar sig ord (Pavičić Takač 2008: 4). På grund av detta existerar det ett relativt stort antal olika teorier om ordinlärningsprocessen och utvecklingen av lexikonet (för en översikt om olika teorier, se t.ex. Pavičić Takač 2008). Även om det inte ännu finns en allmänt accepterad teori om tillägnandet av lexikonet (ibid.: 4), upprepas många faktorer i olika teorier. Dessa faktorer och deras inflytande på ordinläring diskuteras härnäst.

En viktig faktor som antas påverka ordinläringen är relationen mellan språkanvändaren och språket ifråga, det vill säga om det är frågan om språkbrukarens förstaspråk (eng. *native language*) eller andraspråk/främmande språk (eng. *second language* respektive *foreign language*). Ordinlärningsprocessen kan nämligen se lite annorlunda ut beroende på om det gäller förstaspråkstalare (L1-talare) eller andraspråksinlärare (L2-inlärare) av ett språk. Med begreppet *andraspråkinlärare* refererar jag här till språkinlärare som håller på att lära sig ett språk antingen som andraspråk eller som främmande språk. På ett liknande sätt omfattar termen *andraspråk* (L2) språk som språkbrukaren använder som andraspråk och som främmande språk. Enligt Ellis (1997) är en viktig skillnad mellan L1-talare och L2-inlärare att L2-inlära- rarna kan utnyttja sina existerande system som hjälpmedel när de håller på att lära sig nya ord på L2:

Unlike young children learning their native language, older second- and foreign-language learners have already developed rich conceptual and semantic systems

which are already linked to L1. In the first instance at least, the acquisition of L2 words usually involves a mapping of the new word form onto pre-existing conceptual meanings or onto L1 translation equivalents as approximations. (Ellis 1997: 133-134)

En fördel som L2-inlärare har jämfört med L1-talare är alltså att deras inre semantiska och konceptuella system är anknutna till deras L1. Därför antar Ellis (1997: 133-134) att tillägnandet av nya ord på L2 förutsätter åtminstone under ett inledande skede att inläraren sammanbinder nya L2 ordformer antingen till ordens approximativa L1-översättningar eller till deras existerande konceptuella betydelser. Unga L1-talare har däremot inte motsvarande system till reds. Även om det kan vara betydelsefullt att kunna sammanfoga L2-ord med motsvarande L1-ord i inledningsfasen, är det dock viktigt att minnas, såsom Pavičić Takač (2008: 16) påpekar, att ordinläringen också gäller flera andra aspekter av ord än bara semantiska (se också avsnitt 2.5).

Ellis' ovanstående framförande tyder också på en annan viktig komponent i utvecklingen av lexikonet, nämligen minnet. Även Tornberg (2009: 124) uppger att minnet spelar en viktig roll i ordinläringen: hon konstaterar att särskilt långtidsminnet är centralt, eftersom informationen ska kodas in i detta minnessystem om man har för avsikt att minnas den också efter en lång tid. Efter en lyckad inkodning är informationen lagrad i långtidsminnet tills man behöver den (ibid.). Pavičić Takač (2008: 10) påpekar ändå att inläring av nya ord inte är någon lineär process, utan inläraren kan exempelvis glömma några aspekter av ordet (se också diskussionen i avsnitt 2.5). Därför kan man inte anta att all information om ett ord som språkinläraren har kodat in i minnet kvarstår oföränderlig.

Det finns också ett urval hypoteser om hur ord egentligen är lagrade i språkanvändarens minne. Till exempel Hulstijn (2000, refererad i Pavičić Takač 2008: 14) diskuterar olika existerande hypoteser om hur ord som språkbrukaren känner till (både på L1 och på L2) kan vara lagrade i minnessystemet. Hulstijn framför att en möjlighet är att alla ord (L1 och L2 ord) befinner sig i ett och samma lager. Ett annat alternativ är att L1 och L2 ord är lagrade i sina egna, separata lager. Forskare har också övervägt möjligheten att alla besläktade ord skulle kunna vara lagrade i ett gemensamt lager, medan ord som är språkspecifika vore i skilda lager. (ibid.). Konsensus om vilken hypotes som stämmer verkar dock inte ännu finnas bland lingvister, men det som forskare tycks vara eniga om

är att ord inte är lagrade i språkanvändarnas minne på samma sätt som ord är organiserade i ordböcker (Thornbury 2002: 16). Det är alltså snarare frågan om ett nätverk mellan ord än en organiserad lista på olika ord, vilket betyder att minnet inte verkar lagra ord vare sig slumpmässigt eller i form av en lista. I själva verket ligger ord på ett högst organiserat sätt i minnet och det finns ett stort antal relationer mellan dem, vilket ibland kallas *mentalt lexikon*. (ibid.) När inläraren förstår denna inre struktur hos lexikonet, blir inläring av nya ord mycket lättare (Enström 2004: 181-182).

Den sista viktiga faktor som är betydande med tanke på ordinlärningsprocessen är kontext. Till exempel Nagy (1997: 64) framhåller kontextens betydande roll under processen och motiverar sitt argument ur två olika synvinklar. Å ena sidan är kontexten viktig, eftersom betydelsen hos ett ord oftast avgörs på basis av kontexten. Å andra sidan kan kontexter stödja utvidgandet av språkinlärnarnas lexikala kunskaper, eftersom de kan fungera som nyttig språklig input från vilken språkinlärnarna kan tillägna sig ett stort antal lexikalisk information. (ibid.) Kontextens roll i ordinläring diskuteras i mer detalj i samband med ordinlärningsstrategier i nästa avsnitt (avsnitt 2.6.2).

2.6.2 Ordinlärningsstrategier

Ordinläringen är alltså en kreativ men komplex process (Tornberg 2009: 130, 136), men olika slags ordinlärningsstrategier kan användas för att underlätta denna process. Strategierna kan sägas fungera som hjälpmedel som språkinlärare kan dra nytta av vid behov. Till exempel Thornbury (2002: 2) konstaterar att det är viktigt att inlärarna har strategier vilka de kan utnyttja när det finns brister i deras lexikon. Han menar vidare att strategierna kan hjälpa inlärarna till exempel att förstå betydelsen hos främmande ord och kända ord som används på ett sätt som inlärarna inte är vana vid.

Enström (2004) diskuterar två olika typer av ordinlärningsstrategier: *implicita* och *explicita* strategier. Till exempel läsning kan fungera som en implicit ordinlärningsmetod, men läsning ensam räcker ganska sällan att garantera en tillräcklig utveckling av lexikonet samt effektiv ordinläring. I stället förutsätter ordinläring inlärarens egen aktivitet, det vill säga att inläraren kan använda olika inlärningsstrategier och sin egen slutledningsförmåga. För närvarande är det dock oklart på vilka sätt och hur mycket inlä-

rarna utnyttjar ord som de har tillägnat sig via gissningsstrategier i sin egen produktion. (Enström 2004: 180-181) Explicita ordinlärningsmetoder kan däremot ge inlärare redskap för tänkandet som kan hjälpa dem när de möter nya ord. Sådana metoder kan till exempel hjälpa inlärarna avgöra några essentiella aspekter av ordet ifråga, såsom dess ordklass och kombinationsmöjligheter med andra ord. För att maximera inlärningsresultat borde inlärarna dock använda både implicita och explicita ordinlärningsstrategier. (ibid.:182-183)

Även användning av kontexten kan betraktas som en typ av ordinlärningsstrategi. Bland annat Tornberg (2009: 127) argumenterar för att det är viktigt att inlärare försöker utnyttja sina förkunskaper och kontexter för att förstå okända ord i en text snarare än att läraren först introducerar de viktigaste orden till dem. På detta sätt lär inlärarna sig att det mest centrala är att kunna läsa en text som helhet (ibid.: 129). Dessutom påpekar Tornberg (2009: 130, 136) att i språkundervisningen är det viktigt att analysera ord i kontexter eftersom det möjliggör ordens rätta användning i övriga kontexter. Hon menar därmed att inlärare har lärt sig ett ord när de korrekt kan tillämpa det i olika former i varierande kontexter (jfr avsnitt 2.5).

Det finns också några andra fördelar med att utnyttja kontexten under ordinlärningsprocessen. Pavičić Takač (2008: 17) framför att en stor mångfald av varierande kontexter är viktig om man vill lära sig nya ord på L2: hon uppger att ordinläring är möjlig när inläraren har fått tillräckligt begriplig input. Enström (2004: 180) har en liknande uppfattning och förtydligar att språkinlärare oftast ska möta nya ord flera gånger i annorlunda kontexter innan orden kan bli inlärd. Avslutningsvis kan det konkluderas att kontexten inte endast fungerar som hjälpmedel i slutledningen av ordens betydelse (det vill säga som ordinlärningsstrategi) utan den möjliggör hela ordinläringen. Ju mer avancerade språkinlärare är, desto mer kan de sannolikt dra nytta av ordinläringen i kontexten (ibid.: 181).

Det är också högst viktigt att inlärare lär sig nya ord även utanför skolan. Detta förutsätter i flesta fall att inlärare kan utnyttja ordinlärningsstrategier effektivt för att utöka sitt eget ordförråd. (Pavičić Takač 2008: 24) Att kunna utvidga lexikonet självständigt är

väsentlig eftersom ordinlärning är på gång kontinuerligt; man kan aldrig bli färdig med den (Enström 2004: 18).

2.7 Lexikonets storlek och problematiken med uppskattningar

Språkets lexikon förändras kontinuerligt, eftersom nya ord exempelvis behövs för att beskriva nya företeelser – med andra ord är de nödvändiga för att motsvara språkbrukarnas uttrycksbehov. Samtidigt kan onödiga uttryck falla ur vardagligt språkbruk. Även ”gamla” ord kan få nya betydelser. På basis av dessa förändringar kan man konkludera att lexikonet inte är någon oförändrad konstruktion. (Sjöström 2006: 75) Likartade förändringar gäller också individuella språkanvändarnas lexikon: nya ord lärs in medan några andra ord faller ur minnet.

Det har visat sig vara relativt utmanande att uppskatta lexikonets storlek, eftersom flera olika faktorer kan påverka resultaten, inte minst de ovan beskrivna förändringarna i lexikonet. En avgörande faktor är forskarens definition av ett ord (se också avsnitt 2.1): till exempel Greenbaum (1996: 430) framhäver att problematiskt med uppskattningar är hur man har definierat ett ord, det vill säga vilka ord som kategoriseras som olika ord och vilka ordformer som däremot hör till ett och samma ord. Angående resultaten är det också signifikant vilken slags kunskap om ord man är ute efter: om intresset ligger på receptiva kunskaper, kan resultaten vara mycket annorlunda än i samband med produktiva kunskaper. Till exempel Enström (2004: 177) konstaterar att det finns varierande uppskattningar om lexikonets storlek, och resultaten om det receptiva och det produktiva lexikonets storlek avviker mycket från varandra. Hon uppger att en orsak till detta är otydliga definitioner av begreppen receptivt respektive produktivt lexikon (se också diskussionen i avsnitt 2.4). En annan potentiell förklaring till skillnader mellan de två lexikonets storlek kan helt enkelt vara flera forskarens (Thornbury 2002: 15, Milton 2009: 117, Carter 2012: 203) antagande om att det receptiva lexikonet är större än det produktiva. Därför ska höga uppskattningar om lexikonets storlek tolkas att hänvisa till det receptiva lexikonet. Antalet ord som hör till språkanvändarens produktiva lexikon är vanligen mycket mindre. (Greenbaum 1996: 430)

Utöver ens definition av ett ord och distinktionen mellan det receptiva och det produktiva lexikonet finns det också några övriga faktorer som kan påverka resultaten om lexikonets storlek. Exempel på sådana faktorer är bland annat prov- och testtypen, kontextens roll i provet samt bedömning av test som mäter informanternas lexikaliska kunskaper. En belysande diskussion om dessa faktorer anges av Melka (1997).

Med tanke på ordförrådets storlek i svenskan, påpekar Enström (2004: 172) att resultaten hos infödda talare är varierande, och därför finns det inte ännu en allmänt accepterad uppskattning om dess storlek. Hon anger emellertid att om språkanvändaren behärskar ungefär 10 000 ord så kan det antas att hen kan svenska språkets allra viktigaste lexikon (ibid.: 173). När det gäller engelskan som modersmål, konstaterar Aitchison (1987: 6-7, refererad i Greenbaum 1996: 430) att det genomsnittliga antalet ord som en utbildad vuxen människa behärskar är åtminstone cirka 50 000 ord, men antalet kan vara till och med 250 000 ord. Barn som är elva år gammal behärskar åtminstone 10 000 ord medan det motsvarande antalet ord för barn som är 14 år gammal är 20 000 ord (ibid.: 596).

2.8 Läroplanen (GLGY 2003) och lexikonet

Läroplanen (GLGY 2003) är ett av de viktigaste dokument som styr gymnasiernas verksamhet i Finland och dess styrande verkan framförs tydligt också i gymnasielagen (629/1998). Meningen med Grunderna för läroplanen i gymnasieutbildning för ungdomar (GLGY 2003) är att definiera hur undervisning och fostran ska ordnas i gymnasier, och därför kan utbildningsanordnaren ”inte lämna grunderna för läroplanen obeaktade eller avvika från dem” (GLGY 2003: 4). Varje gymnasium formulerar också sin egen läroplan på basis av de gemensamma nationella grunderna. Gymnasisternas kunskaper baserar sig på deras föregående studier i den nioåriga grundläggande utbildningen (för vidare information, se GLGU 2004).

Språkstudier är en obligatorisk del av gymnasisternas studier i det finska gymnasiesystemet (Gymnasielagen 629/1998). Gymnasister ska studera åtminstone ett främmande språk (oftast engelska) och det andra inhemska språket (antingen svenska eller finska beroende på studentens modersmål) under gymnasiet. Som sagt, behandlar föreliggande studie gymnasisternas utveckling i både engelska och svenska. Informanterna i studien

har läst engelska som A-språk (lång lärokurs) vilket betyder att de har påbörjat sina studier i årskurserna 1-6 i den grundläggande utbildningen. I gymnasiet omfattar studierna i A-engelskan sammanlagt sex obligatoriska kurser och två fördjupade, valfria kurser. (GLGY 2003: 103-105). Informanterna har däremot studerat svenska som B1-språk (en obligatorisk medellång lärokurs) och detta innebär att deras svenskstudier har påbörjat lite senare i den grundläggande utbildningen, oftast i årskurs 7. De obligatoriska studierna i medellång svenska innehåller totalt fem kurser i gymnasiet. Utöver dem kan studenter välja två övriga kurser som är fördjupade. (ibid.: 96-98).

Med tanke på språkkunskapernas olika delfärdigheter, är det värt att notera att lexikon och ordinlärning relativt sällan nämns på något explicit eller implicit sätt i läroplanen. Detta gäller både svenska och engelska. I samband med kursbeskrivningar av båda språken konstateras dock att ett av målen är att gymnasister lär sig att kommunicera på ett sätt som är typiskt för språket i fråga och dess kulturområde (GLGY 2003: 94, 102). Fast detta inte uttrycker behovet av ett tillräckligt lexikon explicit, kan det ändå antas att lexikonet ligger till grund för de ovannämnda kommunikativa kunskaperna, eftersom det är svårt att kommunicera på ett målspråkligt sätt utan tillräckliga lexikala kunskaper. Ordinlärningen nämns även explicit några gånger i läroplanen: det uppges att det är viktigt att såväl det svenska som det engelska ordförrådet utvidgas under språkstudierna. Utöver detta är syftet att språkbruket blir exaktare och mångsidigare. (ibid.: 95, 103).

När det gäller i synnerhet svenska språket, nämns lexikonet och inlärningen av nya ord explicit endast i samband med de egentliga kursbeskrivningarna av de två första kurserna (RUB 1 och RUB 2, se GLGY 2003: 96-97), där det påpekas att syftet är att befästa ordförrådet (RUB 1) och att vidga det (RUB 2). Angående engelskan, nämns ordförrådet enbart i kursbeskrivningen av den första obligatoriska kursen där det anges att kursen ska befästa inlärarnas lexikon (ibid.: 103). På de fördjupade svenska och engelska kurserna är avsikten att utveckla språkkunskaperna mångsidigt, vilket antagligen gäller också utvecklingen av studenternas lexikon. (ibid. 97-98, 104-105). Några andra antydningar till lexikala kunskaper verkar inte finnas i läroplanen.

Avslutningsvis kan konstateras att ordinlärningen inte ofta nämns explicit eller implicit i den gällande läroplanen. Å ena sidan är det lite oroande eftersom det kan förmedla bil-

den av språkkunskaper där lexikonet anses vara något ”extra” (jfr Thornbury 2002: 14) eller där lexikonet inte behandlas likvärdigt i jämförelse med andra delkunskaper, såsom skrivande och talande. Utbildningsstyrelsen håller emellertid på att förnya läroplanen och man vet inte ännu hurdan roll lexikonet slutligen har i den blivande läroplanen, som kommer att träda i kraft senast 1.8.2016 (Utbildningsstyrelsen 2015). Å andra sidan skrivs läroplanerna ofta på en mycket allmän nivå och de syftar inte ens till att ge för detaljerade råd om undervisningens innehåll. På grund av detta är det också viktigt att minnas att det alltid beror på läraren hur mycket språkkunskapernas olika delområden övas.

2.9 Resultat i tidigare forskning

Lexikonet i allmänhet var ett relativt outforskat område inom språkvetenskap under flera decennier. På grund av detta ansågs dess betydelse länge som minimal i språkundervisning och språkinläring. (Carter 2012: 178-179) Även Thornbury (2002: 14) påpekar att vokabulär och ordinläring tidigare behandlades endast som språkundervisningens biprodukt. Under de senaste decennierna har forskare dock blivit mer intresserade av att undersöka lexikon (Carter 2012: 179, Milton 2009: 1-5). Enligt Thornbury (2002: 14) har lingvister nuförtiden intresserat sig för till exempel ordfrekvenser och några skilda aspekter av lexikonet, såsom kollokationer (se t.ex. Pomell 2014). Även språkinlärares ordinlärningsstrategier (se t.ex. Jokinen och Pitkäljärvi 2005) har varit av intresse. Däremot finns det mig veterligen inte några tidigare studier om temat för denna avhandling, det vill säga utvecklingen av det produktiva lexikonet i svenska och engelska under den finska gymnasieutbildningen.

Fast det inte finns studier om exakt samma tema, är det dock möjligt att redogöra för några resultat om ordklassernas inläring på basis av andra föregående studier. Exempelvis Jokinen & Pitkäljärvi (2005) undersökte gymnasisternas åsikter om ordinläring och vokabulärundervisning genom att genomföra en enkätundersökning. Resultaten i deras studie visade att gymnasister tycker att ordinläring har en mycket viktig roll i språkinläringen. Jokinen & Pitkäljärvi (2005) kartlade också ordklassens roll och betydelse i ordinläringen och var därmed intresserade av om gymnasister anser att ord inom någon ordklass är lättare eller svårare att lära sig än ord inom någon annan ord-

klass. De fann att flera gymnasister var av den åsikten att adjektiv och substantiv kändes lättare i jämförelse med verb, men en del av informanter ansåg att det är precis tvärtom att verb är lättare att lära sig (ibid.: 66). Detta innebär att svaren var ganska varierande. Därför konkluderade Jokinen & Pitkäjärvi (2005: 66) att själva ordklassen troligen inte har en avgörande roll utan några andra aspekter av ordet, såsom stavning och längd, kan orsaka större problem i ordinläringen.

Även Laufer (1997) tar upp ordklassens påverkan på inläring av ord. Hon framför att substantiv oftast är lättast att lära sig medan adverb är mest utmanande. Adjektiv och verb antas vara där emellan. (ibid.: 148) Milton (2009: 37) redogör för likartade resultat och konstaterar att substantiv verkar vara lättare att lära sig än verb; däremot är verb lättare än adjektiv. Dessa resultat skulle kunna tyda på att språkinlärare troligen också behärskar substantiv bättre än andra öppna ordklasser och kan därigenom använda mer varierande substantiv i sin egen produktion.

Ett stort antal olika ord i en text kan betyda att språkinlärarens lexikon är omfattande (Thornbury 2002: 136). Detta kan vara förknippat med inlärarens kunskapsnivå och begåvning: till exempel Milton (2009: 126) konstaterar att begåvade språkinlärare som har ett omfattande produktivt lexikon är mer sannolikt kapabla att producera ett större antal olika ord i skriftliga uppgifter än inlärare som har ett litet produktivt lexikon. Därigenom kan antalet olika ord eventuellt vara en signal om inlärarens (produktiva) kunskapsnivå i språket ifråga. Greenbaum (1996: 431) har en liknande uppfattning för han påpekar att "[t]he higher the ratio of types to tokens, the richer the vocabulary". Dessutom kan andelen innehållsord i en text (eng. *lexical density*) spegla inlärarens kunskapsnivå, eftersom skriftliga texter skrivna av skickliga inlärare oftast innehåller en hög andel innehållsord (Thornbury 2002: 136).

Härnäst ska jag gå över till att beskriva syfte, material och metoder i studien. Även informanterna och analyseringsprocessens gång presenteras i nästa kapitel (kapitel 3).

3 SYFTE, MATERIAL OCH METOD

I detta kapitel presenterar jag först syftet med studien och de exakta forskningsfrågorna (avsnitt 3.1). Därefter beskriver jag Topling-projektet och materialet, materialinsamlingens gång samt informanterna i avsnitt 3.2. Till sist redogör jag för analysmetoder, olika skeden i analyseringsprocessen och ord som inkluderas samt ord som utelämnas i analysen i föreliggande studie (avsnitt 3.3).

3.1 Syfte

Syftet med denna studie är att ta reda på utvecklingen av gymnasisternas produktiva lexikon i svenska och engelska under de obligatoriska språkstudierna i gymnasieutbildningen. Framför allt betraktar jag temat ur de fyra öppna ordklassernas synvinkel: jag har intresserat mig för förändringar mellan de öppna ordklassernas användning i elevtexter. Därför försöker jag exempelvis utreda om gymnasister använder ord som hör till en viss ordklass, såsom verb, mer i sina texter i början av deras språkstudier än i slutet av deras obligatoriska språkstudier. Jag har således för avsikt att betrakta utvecklingen av ordklassernas inbördes användning under gymnasiestudiernas gång. Därutöver ligger intresset även på förändringar inom ordklasserna, det vill säga ordvariation. Detta betyder i praktiken att jag försöker kartlägga inlärningsgångar i gymnasisternas lexikala kunskaper, med andra ord om de kan använda mer varierande lexikon i slutet än i början av språkstudierna. Slutligen syftar jag till att jämföra svenska och engelska med varandra och beskriva likheter och skillnader angående de ovan presenterade aspekterna.

Syftet med avhandlingen kan sammanfattas i följande forskningsfrågor:

1. Utvecklas gymnasisternas produktiva lexikon i svenska och engelska i skrivna texter under gymnasieutbildningen? I så fall hur?
 - Förändras användningen av de fyra öppna ordklasserna?
 - Blir ordvariationen större inom ordklasserna under gymnasiets gång?
2. Finns det likheter och skillnader mellan svenska och engelska i utvecklingen av det produktiva lexikonet?

I stället för hela lexikonet har jag således avgränsat undersökningen till lexikala ord (öppna ordklasser) och jag fokuserar därmed på verb, substantiv, adjektiv och adverb i informanternas berättelser. Jag har valt lexikala ord till utgångspunkten i avhandlingen, eftersom de är innehållstunga och helt enkelt mer varierande till sin natur än grammatiska ord (se också avsnitt 2.2). Därför var de ett naturligt val till utgångspunkten i studien. För att få en mer omfattande uppfattning om gymnasisternas produktiva lexikon, ska jag också studera bland annat utveckling av adverbial och egennamn i berättelserna (se avsnitt 3.3).

På basis av de teoretiska utgångspunkter som beskrevs i kapitel 2 är det möjligt att formulera några allmänna och preliminära hypoteser om studiens forskningsresultat. Framför allt förväntar jag att resultaten i studien visar att informanternas produktiva lexikon utvecklas och utvidgas i båda språken under gymnasiet, eftersom gymnasisterna troligen lär sig många nya ord på kurserna. Under gymnasiestudiernas gång kan detta resultera i större ordvariation inom ordklasserna i gymnasisternas egen produktion. Dessutom är det sannolikt att andelen verb i texterna är relativt stabil och stor vid alla undersökningstillfällen, eftersom verb utgör en essentiell ordklass i båda språken. Det är också troligt att substantiv och verb är de två största ordklasserna vid varje undersökningstillfälle, medan andel adverb och adjektiv är mindre, och de båda utgör under 10 procent av löporden (jfr Melin och Lange 2000, Hultman 2003; se också avsnitt 2.3.6). Till sist antar jag att de engelska texterna är lexikalt mer varierande än de svenska texterna, eftersom informanterna har studerat engelska längre än svenska (för vidare information om informanterna, se avsnitt 3.2.3).

3.2 Material

3.2.1 Toppling-projektet och materialet i allmänhet

Materialet i föreliggande undersökning har samlats in i forskningsprojektet *Inlärningsgångar i andraspråket*, som också kallas *Toppling*-projektet. Toppling var ett gemensamt projekt för Institutionen för språk och Centrum för tillämpad språkvetenskap vid Jyväskylä universitet och det genomfördes år 2010-2013. Finlands Akademi och Jyväskylä universitet finansierade projektet. Projektet syftade till att ”undersöka inläring av

skriftliga färdigheter i finska, engelska och svenska som andraspråk i det finska utbildningssystemet” (Topling 2010-2013). Topling är en fortsättning till arbete som började i ett tidigare forskningsprojekt *Cefling* (för mer information, se Cefling 2007-2009).

Materialet som insamlades inom ramar för Topling-projektet är mestadels longitudinellt, vilket möjliggör att betrakta till exempel utveckling av och potentiella förändringar i informanternas språkfärdigheter. Svenska data har samlats in från sammanlagt tre olika skolnivåer i Finland och det består således av texter som högstadieelever, gymnasister och vuxna språkinlärare har skrivit. Även det engelska materialet omfattar alla dessa skolnivåer men det kompletterades också med data från lågstadienivå. Allmänt taget kan det konstateras att Topling var ett omfattande forskningsprojekt och antalet data är mycket stort. I denna undersökning har jag dock valt att koncentrera mig på enbart gymnasisternas skriftliga texter och mitt material består därför av berättelser som informanterna skrev vid tre olika undersökningstillfällen under deras språkstudier i gymnasiet (se också avsnitt 3.2.2 och 3.2.3 för en noggrannare beskrivning av materialinsamlingen och informanterna).

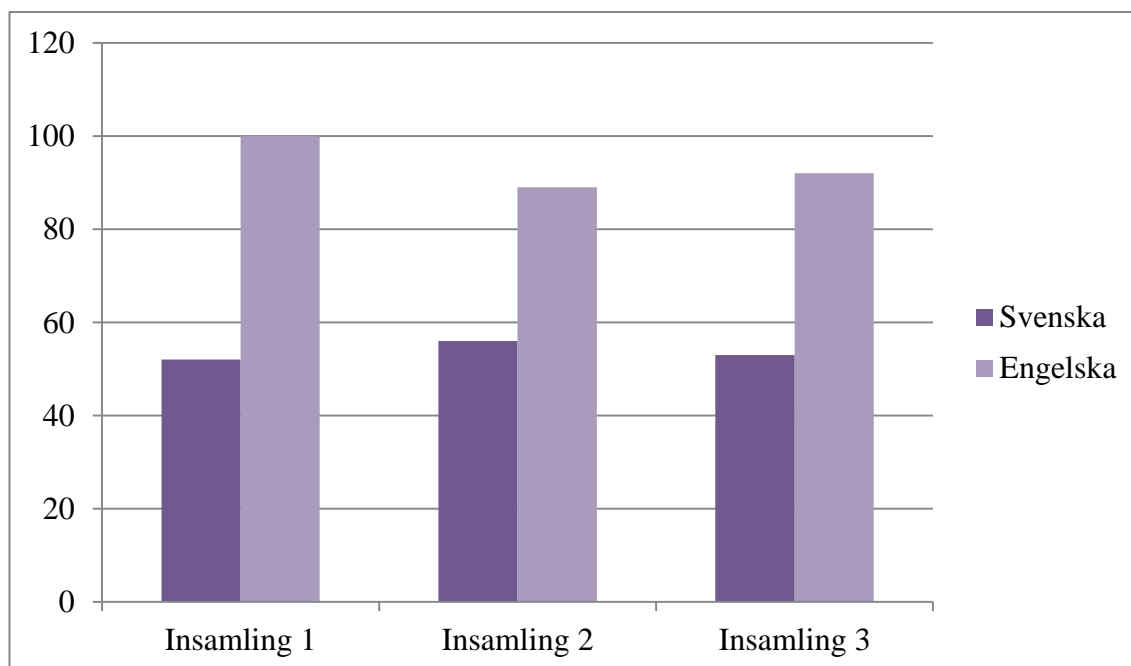
Materialet i föreliggande undersökning består av sammanlagt 150 berättelser varav 75 är från det svenska och 75 från det engelska Topling-materialet skrivna av finska gymnasister. Detta betyder att materialet omfattar 25 texter på båda språken från varje materialinsamlingstillfälle. Tabell 1 nedan visar totalantalet ord i det svenska och i det engelska materialet:

TABELL 1. Totalantalet ord i det svenska (N=75) och i det engelska (N=75) materialet.

Språket →	Svenska	Engelska
Insamlingsgången ↘		
Insamling 1	1298	2506
Insamling 2	1397	2216
Insamling 3	1323	2288
Totalt	4018	7010

Som framgår av Tabell 1, innehöll det engelska materialet ett större totalantal ord än det svenska materialet. Tabellen visar också att antalet ord varierade i viss mån mellan data-

insamlingsgångerna inom de svenska och de engelska texterna. Detta betyder också automatiskt att texternas genomsnittliga längd varierade mellan språken. Utöver detta fanns det även variation inom båda språken mellan olika materialinsamlingstillfällen. Figur 1 nedan illustrerar de svenska och de engelska texternas genomsnittliga längd vid alla tre materialinsamlingstillfällen:



FIGUR 1. Förändringar i texternas genomsnittliga längd vid alla tre materialinsamlingstillfällen under gymnasiet.

Som Figur 1 också visar är de engelska berättelserna mycket längre än de svenska berättelserna vid alla tre materialinsamlingstillfällen. En sannolik orsak till detta är att informanterna i denna studie har studerat engelska längre än svenska, vilket troligen påverkar deras egen textproduktion positivt. Skillnaden i textlängden mellan språken är som störst i samband med den första datainsamlingsgången då de engelska berättelserna (textlängden: ca 100 ord) var nästan dubbelt så långa som de motsvarande svenska berättelserna (textlängden: ca 52 ord). Skillnaden mellan språken blir lite jämnare tills det andra undersökningstillfället, då den genomsnittliga textlängden för de engelska texterna är ca 89 ord jämfört med de svenska texternas 56 ord. Vid den tredje datainsamlingsgången är skillnaden mellan språken lite större igen med tanke på textlängden (engelskan: 92 ord, svenskan: 53 ord).

Dessutom framgår det av Figur 1 att de svenska berättelsernas genomsnittliga längd kvarstår relativt stabil under gymnasiestudiernas gång (52 – 56 – 53 ord) vilket kan tyda på att informanternas språkkunskapers potentiella utveckling inte syns i en form av ett större ordantal i dessa berättelser. Textlängden i de motsvarande engelska berättelserna varierar i lite större utsträckning (100 – 89 - 92 ord). Det är anmärkningsvärt att de engelska texterna är längst redan vid den första datainsamlingsgången. En orsak till detta kan vara själva deltagandet i en undersökning: det har troligen varit nytt för gymnasisterna och de har därför satsat på sitt skrivande, vilket också syns i textlängden. Vid de två andra datainsamlingsgångerna har de redan varit bekanta med projektet och uppgifterna, vilket kan ha resulterat i lite kortare texter.

3.2.2 Materialinsamlingen

Materialet i denna avhandling omfattar således svenska och engelska berättelser som gymnasister skrev vid tre olika undersökningstillfällen under deras obligatoriska språkstudier i gymnasiet. Såsom det konstaterades redan i det föregående avsnittet (avsnitt 3.2.1), har materialet insamlats inom ramar för Topling-projektet, vilket innebär att jag själv inte har samlat in det utan jag har fått färdigt material.

Det svenska materialet insamlades när informanterna deltog i kurserna 1 eller 2, 3 och 5, och det engelska materialet samlades in på kurserna 1, 4 och 6. Det första undersökningstillfället i båda språken var på hösten 2010. Den andra gången då gymnasisterna bads att skriva texterna var antingen under vårtermin 2011 (en del av de svenska texterna) eller hösttermin 2011 (alla engelska texter och en del av de svenska texterna). Den sista datainsamlingsgången angående både det svenska och det engelska materialet var på våren 2012, då informanterna hade studerat i nästan två år på gymnasiet. Materialet är därmed autentiskt och longitudinellt, vilket betyder att det kan ge några inblickar i informanternas produktiva språkkunskaper i skriftlig kommunikation i början, i mitten och i slutet av deras obligatoriska språkstudier i gymnasieutbildningen.

Vid varje undersökningstillfälle skrev gymnasister sammanlagt fem skriftliga uppgifter. Dessa uppgiftstyper omfattade tre olika slags e-postmeddelanden till varierande mottagare (vän, lärare och nätbutik), en insändare och en berättelse. Informanterna fick inte

använda några hjälpmedel, såsom ordböcker, när de skrev texterna. Dessutom ställdes det ingen exakt minimi- eller maximilängd på texterna men gymnasisterna instruerades emellertid att skriva åtminstone några meningar i alla uppgifter.

I föreliggande avhandling ligger fokus på den femte uppgiftstypen, det vill säga berättelser. I anvisningarna till denna uppgift bads studenterna att beskriva någon minnesvärd erfarenhet eller händelse och berätta varför den är minnesrik (se Bilaga 1). I stället för de andra uppgiftstyperna bestämde jag mig för att analysera berättelserna i studien, eftersom jag antar att de troligen återspeglar informanternas lexikala kunskaper mer sanningsenligt än de övriga uppgifterna. Detta antagande kan motiveras genom att konstatera att instruktionerna till uppgiften var relativt fria, vilket vidare kan möjliggöra användning av ett mer varierat lexikon. Även teman för texterna kan vara relativt personliga och individuella, vilket sannolikt gör dem också lexikalt mer varierande och därigenom skiljer de på ett positivt sätt från de andra texttyperna. Av dessa olika anledningar var berättelserna det naturligaste valet till utgångspunkten i undersökningen.

3.2.3 Informanterna

Informanter i denna studie är finska gymnasister från flera gymnasier i Mellersta Finland och Österbotten. Särskilt Mellersta Finland är ett högst finskspråkigt område i Finland (Tilastokeskus 2015). Informanterna har i huvudsak läst engelska som A-språk (lång lärokurs) och svenska som B1-språk (medellång lärokurs) (se också avsnitt 2.8). Deltagandet i Topling-projektet var fullständigt frivilligt för gymnasisterna. Jag bestämde mig för att välja gymnasister som informanter i avhandlingen av många olika anledningar. Framför allt har de studerat såväl svenska som engelska längre än till exempel högstadieelever, vilket sannolikt innebär att deras texter är längre och även lexikalt mer varierande. Därigenom är det troligt att det finns mer att analysera i texterna. En annan anmärkningsvärd orsak till informantvalet är att det kan antas att gymnasister är mer vana vid att skriva hela texter än exempelvis högstadieelever som kan sägas vara i ett relativt tidigt skede i språkinlärningsprocessen. Den tredje viktiga anledning som påverkade informantvalet är brist på tidigare studier om detta tema (se också diskussionen i avsnitt 2.9). Gymnasister som informanter uppfyller således syftet med studien bäst.

I denna avhandling analyserar jag inte utvecklingen hos enskilda individer utan jag ska betrakta informanternas utveckling som grupp. På grund av detta har det ingen betydelse om samma informanter har skrivit texter både på svenska och på engelska; en del av informanterna har skrivit texterna på båda språken, medan en del av dem har deltagit i Topling-projektet bara inom ramarna för ett språk. Eftersom deltagandet i projektet var hela tiden frivilligt, är det möjligt att en och samma informant har deltagit i undersökningen antingen en, två eller tre gånger. Därför är det viktigt att påpeka att någon informant kan ha skrivit texterna vid alla tre undersökningstillfällen, men det är också lika möjligt att informanten har tagit del i undersökningen bara en gång. Detta spelar dock inte någon signifikant roll i föreliggande undersökning, eftersom intresset inte ligger på utvecklingen hos individer utan på utvecklingen hos hela informantgruppen under gymnasiestudiernas gång.

3.3 Metod

3.3.1 Bakgrunden: kvantitativa och kvalitativa forskningstraditioner

Denna studie är mestadels kvantitativ till sin natur, vilket följaktligen innebär att materialet analyseras med hjälp av kvantitativa analysmetoder. Den kvantitativa analysen kompletteras med separata analyser som är en kombination av kvalitativa och kvantitativa forskningstraditioner (se avsnitt 3.3.2 nedan). Utöver detta ska jag excerpiera ett urval belysande exempel ur materialet för att illustrera den numeriska informationen. Fast huvudvikten ligger på de kvantitativa analysmetoderna, innehåller undersökningen alltså även några kvalitativa drag eftersom jag försöker få en så omfattande uppfattning om informanternas produktiva lexikon som möjligt.

När man stöder sig på den kvantitativa forskningstraditionen i sin undersökning, vill man beskriva och analysera något fenomen med hjälp av numerisk information (Kalaja m.fl. 2011: 19). Även ett ganska stort antal informanter samt deras representativitet är essentiella för en kvantitativ studie för forskaren är oftast ute efter generaliserbara resultat. Figurer och tabeller är i en central roll när resultat rapporteras. (ibid.) Strävan efter objektivitet anknyts i synnerhet till denna typ av studier.

Den kvalitativa forskningstraditionen karaktäriseras av forskarens vilja att förstå ett fenomen och dess egenskaper via analysen. Analysen utgår från det egentliga materialet och forskarens egna kategoriseringar. (Kalaja m.fl. 2011: 20). I samband med resultatrapportering är det också vanligt att exciperera exempel ur ens data (ibid.). Enligt Hirsjärvi m.fl. (2003: 152) försöker man forska i det valda fenomenet så mångsidigt som möjligt om man väljer att utföra en kvalitativ undersökning.

Som sagt är min undersökning en kombination av båda forskningstraditionerna, men kvantitativa metoder dominerar under hela analysprocessens gång. Själva analysprocessen och sätten hur kvantitativa respektive kvalitativa analysmetoder har utnyttjats i analysen presenteras i nästa avsnitt (avsnitt 3.3.2). Enligt Kalaja m.fl. (2011: 20) är det nu förtiden motiverat att betrakta det valda forskningstemat ur flera olika synvinklar och genom att använda ett urval metoder, eftersom det kan ge mer information om företeelsen än användningen av bara en metod. Hirsjärvi m.fl. (2003: 125) menar också att de två traditionerna kompletterar varandra och är egentligen tätt förknippade, eftersom all mätning inkluderar både kvantitativa och kvalitativa drag. Forskaren behöver således inte nödvändigtvis göra ett absolut val mellan en kvantitativ och en kvalitativ studie.

3.3.2 Analysens gång och analysmetoder

Jag fick ett färdigt material som hade insamlats inom ramar för Topling-projektet. Allt material var i en elektronisk form vilket underlättade analysprocessen. Eftersom antalet data var relativt omfattande, kunde jag inte inkludera allt material i studien utan jag begränsade det till ett urval texter som sedan analyserades noggrant. I princip valde jag texterna slumpmässigt: jag valde varannan text ur materialet så att jag hade sammanlagt 25 texter från varje datainsamlingsgång. Det slutliga antalet texterna var därmed 75 per språk och på det sättet blev totalantalet texter 150.

Enligt Melin och Lange (2000: 15) är verb, substantiv och adjektiv de primära ordklasserna och de är därför mest centrala i de flesta textanalyserna. Dessa tre ordklasser kompletterades med ordklassen adverb i denna undersökning för att få studien att omfatta alla öppna ordklasserna. Analyseringen av materialet bestod av två delar: den

första delen var den kvantitativa analysen av de öppna ordklasserna och den andra delen var de separata analyserna av adverbial, egennamn, negationsord och ord som utelämnades från andra analyser, såsom finska ord. De fyra separata analyserna innehöll alltså såväl kvantitativa som kvalitativa drag. Denna tvådelade analysprocedur upprepades totalt sex gånger, eftersom materialet från varje materialinsamlingstillfälle betraktades individuellt.

Den första delen av analysprocessen, det vill säga den kvantitativa analysen av berättelserna, bestod av flera successiva steg. Först räknades texternas genomsnittliga längd genom att indela totalantalet ord med antalet berättelserna (se också Figur 1 i avsnitt 3.2.1). På nästa steg utelämnades alla grammatiska ord, såsom artiklar och prepositioner, så att endast lexikala ord kvarstod i materialet. Därefter tog jag reda på hur stor andel av alla ord hör till de lexikala respektive grammatiska ord, vilket skedde genom att dela antalet lexikala ord med totalantalet ord. Det kvarstående procenttalet visade automatiskt andelen grammatiska ord.

I följande fas ordnades lexikala ord i en alfabetisk ordning med hjälp av datorprogrammet Words funktioner: därigenom fick jag en färdig, alfabetisk lista på alla lexikala ord i materialet. Efter detta grupperade jag ord på basis av ordklassen (verb, substantiv, osv.) och räknade totalantalet ord (så kallade *tokens*) för varje ordklass. Jag betraktade således en ordklass åt gången. Ordklassernas procentuella andelar räknades genom att dela antalet alla ord i en ordklass med totalantalet ord (till exempel: antalet alla verb indelat med antalet alla ord). Det var viktigt att räkna de olika ordklassernas procentuella andelar eftersom på det sättet var det möjligt att kartlägga potentiella förändringar i ordklassernas användning i texterna under gymnasiet. Därefter räknade jag antalet *olika* ord (så kallade *types*) för varje ordklass. Till sist delade jag antalet olika ord (till exempel antalet alla olika verb) med antalet alla ord (till exempel antalet alla verb) inom varje ordklass och därigenom fick jag ett eget TTR-värde (för mer information om TTR, se nästa avsnitt) för alla fyra ordklasser. Detta steg var det sista i den kvantitativa analysdelen av materialet.

TTR (*type-token ratio*) är alltså en metod som kan användas för att redogöra för lexikal diversitet (eng. *lexical diversity*) i texter. Där står *token* för antalet alla ord i en text och

type uppger antalet olika ord i texten. Siffran (*ratio*) demonstrerar deras inbördes relation (*types* indelat med *tokens*). Ju närmare resultaten är 1, desto mer varierat är ordförrådet och desto bättre är resultaten. Den grundläggande hypotesen bakom TTR är således att ju bättre språkinläraren behärskar språket ifråga, desto större är antalet olika ord i hans texter, vilket vidare resulterar i högre TTR-värden. (Milton 2009: 125-126) TTR-metoden används vanligen på textnivån men i föreliggande studie utnyttjades den för att ta reda på ordvariation inom ordklasserna. Genom att använda TTR-metoden i studien var det möjligt att redovisa hur ordvariationen inom ordklasserna utvecklas under informanternas språkstudier. En annan orsak till varför jag bestämde mig för att utnyttja TTR är att den är tämligen enkel att använda. Även Milton (2009: 126-127) redovisar ett par andra fördelar med TTR: enligt honom kan TTR ge riktningsgivande resultat om storleken av språkinlärarens produktiva kunskaper. Utöver detta är den objektiv (*ibid.*). Nackdelar med TTR-metoden kommer att diskuteras i samband med metodkritik i kapitel 5.

Efter den kvantitativa analysen genomförde jag separata analyser av materialet. De separata analyserna sammanfogade drag från kvalitativa och kvantitativa forskningstraditioner och fokuserades till adverbial, negationsord, egennamn och ord som utelämnades från andra analyser (se också Tabeller 2 och 3 nedan). I praktiken utfördes de separata analyserna genom att plocka alla analyserade uttryck (adverbial, negationsord, och så vidare) ur materialet. Därefter räknade jag totalantalet uttryck för varje grupp (till exempel antalet negationsord). Till sist analyserade jag ord och uttryck inom grupperna mer kvalitativt, vilket exempelvis skedde genom att gruppera uttryck till mindre undergrupper som kommer att presenteras i samband med resultat i kapitel 4.

Jag ansåg att det var viktigt att komplettera den kvantitativa analysen med separata, mer kvalitativa analyser. En naturlig orsak till detta var strävan efter att få en så omfattande allmänbild om informanternas produktiva lexikon som möjligt. Å ena sidan finns det också vissa skillnader mellan svenska och engelska, vilket kunde ha påverkat slutresultaten: till exempel negationsord kategoriseras till adverb i svenskan, medan engelskans negationsord inte hör till denna ordklass (Greenbaum 1996: 93). På grund av detta var en separat analys av negationsord nödvändig. Å andra sidan kan behovet av separata analyser motiveras på basis faktumet att det finns ord som inte kan klassificeras till nå-

gon av de existerande ordklasserna: till exempel Greenbaum (1999: 93) påpekar att sådana ord är engelskans negationsord ”not” samt ordet ”there” när det används som ett existentiellt uttryck. Allt som har inkluderats i den kvantitativa analysen och i de separata analyserna har sammanfattats i Tabell 2 nedan.

TABELL 2. En sammanfattning av ord och uttryck som betraktas i den kvantitativa analysen och i de separata analyserna.

Den kvantitativa analysen:	De separata analyserna:
<ul style="list-style-type: none"> ➤ Adverb (exklusive negationsord: eng. “not”, sv. “inte”) ➤ Substantiv (inklusive artnamn, exklusive egennamn) ➤ Verb (exklusive alla hjälpverb) ➤ Adjektiv och deras komparativformer (t.ex. ”bra”, ”bättre”, ”bäst”) 	<ul style="list-style-type: none"> ➤ Adverbial (med vissa undantag, se avsnitt 2.3.5) ➤ Egennamn (namn på människor, platser, band, osv.) ➤ Negationsord ➤ Ord som inte inkluderades i vare sig den kvantitativa analysen eller i andra separata analyser (se Tabell 3)

Gemensamma principer tillämpades för både den kvantitativa analysen och de separata analyserna för att avgöra vilka ord som exkluderas och vilka som inkluderas i analyserna. För att börja med fel, var *begripligheten* det viktigaste kriteriet: till exempel små stavfel (t.ex. eng. ”complytely” i stället för ”completely”), felaktiga böjningsformer (eng. ”striked” i stället för ”struck”) och små semantiska fel (eng. ”ice cubes” i stället för ”hailstones”) resulterade inte i utelämnandet av ord så länge som dessa ord var förståeliga och deras betydelser samt ordklasser var tydliga. Däremot exkluderades alla sådana ord och uttryck vars betydelser var obegripliga. Utöver detta utelämnades också ord och uttryck som inte var ord på målspråket: exempel på dem är bland annat finska och tyska ord i svenska berättelser. I Tabell 3 nedan presenteras ord och uttryck som utelämnades från den kvantitativa analysen och de separata analyserna av adverbial, egennamn och negationsord. Även dessa utelämnade ord och uttryck betraktades med hjälp av kvantitativa och kvalitativa metoder (det vill säga de behandlades på samma

sätt som adverbial, negationsord och egennamn), och resultaten kommer att beskrivas i avhandlingens resultatdel (se avsnitt 4.2.4 och 4.4.4).

TABELL 3. Ord och uttryck som utelämnades från den kvantitativa analysen och de separata analyserna av adverbial, negationsord och egennamn.

Ord och uttryck som har exkluderats:

- ord från främmande språk (t.ex. ”Hon är sexton **jahre alt**” [interferens från tyska])
- ord och uttryck vars betydelse och/eller ordklass var oförståelig (t.ex. ”Det är så roligt **gor orsaken** med dem”)
- påhittade ord, dvs. ord som inlärares själv har uppfunnit (t.ex. ”Vi har **messadet** mycket” [i stället för exempelvis ”skickat textmeddelanden”])
- felaktiga ordval som avsevärt förhindrar begripligheten (t.ex. ”Jag **gjorde** på sjukhuset” [i stället för ”gick”], ”bara min fötter blev **vett** [i stället för ”blöta” eller ”sura”])

En annan viktig princip i analyseringen var ordens ordklasstillhörighet. Såsom det diskuterades i avsnitt 2.3, är det självklart att ord kan höra till flera olika ordklasser samtidigt. I denna studie avgjorde kontexten ordets huvudsakliga ordklass i analysen. På det sättet var det möjligt att inkludera alla ord bara en gång i studiens kvantitativa analysdel. Däremot innehöll de separata analyserna automatiskt även ett antal ord som redan hade analyserats som en del av den kvantitativa analysdelen. Till exempel meningen ”Vi bodde i hotel” analyserades på följande sätt: bodde = verb, hotel = substantiv, i hotel = adverbial. Såsom detta exempel visar, analyserade jag exempelvis artnamn (”hotel”) i studiens kvantitativa analysdel (med andra ord som en del av ordklassen substantiv). Eftersom ordet ”hotel” också är en del av ett adverbial (”i hotel”), inkluderades det även i den separata analysen av detta adverbial.

I nästa kapitel (kapitel 4) ska jag gå över till att presentera studiens resultat.

4 RESULTAT

Forskningsresultaten i studien redovisas och förklaras i detta kapitel. Material som inkluderas i analysen består alltså av 150 berättelser och de har insamlats vid tre olika materialinsamlingstillfällen. Jag har analyserat berättelserna med hjälp av både kvantitativa och kvalitativa analysmetoder. I samband med procenttal anges resultat huvudsakligen med en decimals noggrannhet, medan TTR-värden presenteras med två decimalers noggrannhet. I början av kapitlet ska jag redogöra för resultat för svenska berättelser: först beskriver jag resultat som baserar sig på den kvantitativa analysen (avsnitt 4.1) och därefter presenterar jag resultat från de separata analyserna (avsnitt 4.2). Efter dessa rapporteras resultat för engelska berättelser på ett liknande sätt (avsnitt 4.3 – 4.4). Avslutningsvis jämförs resultaten mellan de två språken i avsnitt 4.5.

4.1 Den kvantitativa analysen – svenska

4.1.1 Förändringar i ordklassernas procentuella andelar

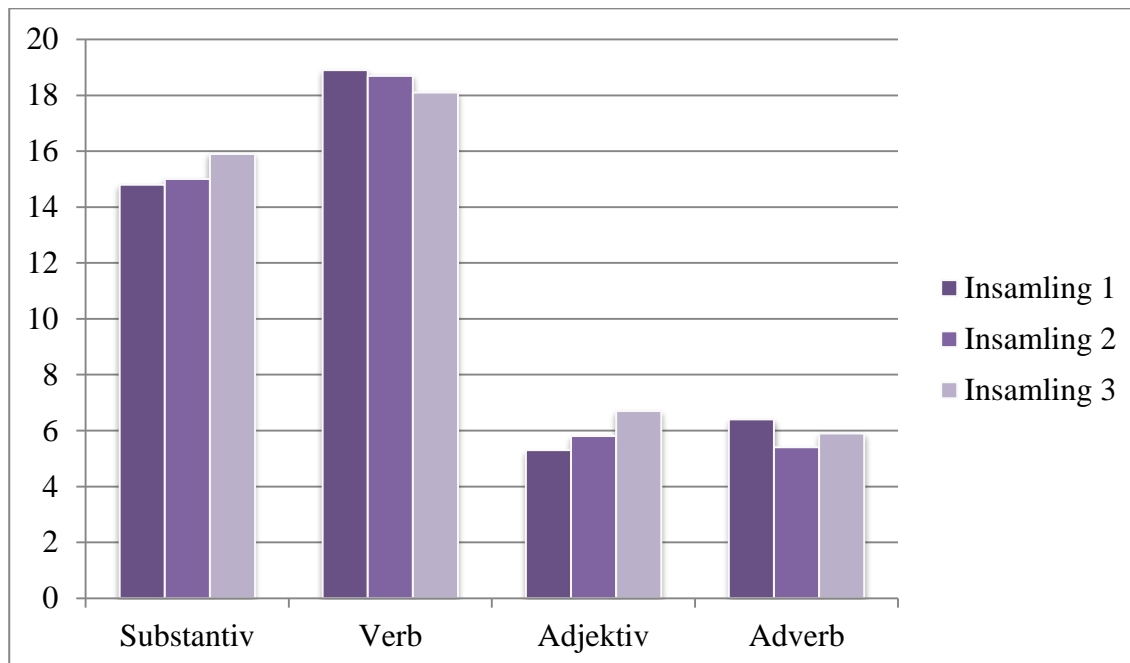
Vid den första datainsamlingsgången var informanterna på antingen sin första eller andra obligatoriska kurs i svenska språket i gymnasiet. Av alla ord i deras berättelser hörde 45,4 % till lexikala ord och 54,6 % till grammatiska ord, vilket tyder på att en ganska klar andel av texternas ord var grammatiska ord. Med tanke på de fyra öppna ordklassernas procentuella andelar kan det konstateras att deras andelar relativt tydligt fördelade sig i informanternas berättelser. Den största öppna ordklassen var verb eftersom 18,9 % av alla ord hörde till ordklassen verb vid det första undersökningstillfället. Substantiv utgjorde den näst största ordklassen för 14,8 % av orden representerade denna ordklass. Dessutom framgick det av resultaten att de två minsta ordklasserna var adverb (6,4 %) och adjektiv (5,3 %) vars procentuella andelar dock liknade varandra i relativt stor utsträckning. På basis av forskningsresultaten kan det därmed konkluderas att skillnaden mellan de två minsta ordklasserna (adverb och adjektiv) var ganska liten, medan skillnaden mellan de två minsta och de två största (verb, substantiv) ordklasserna var tämligen tydlig.

Det andra undersökningstillfället var när informanterna deltog i sin tredje kurs i svenska. Då var andelen lexikala ord 44,8 % och andelen grammatiska ord 55,2 %. Andelen lexikala ord var därmed lite lägre jämfört med deras andel i berättelserna vid den första datainsamlingsgången (45,4 %), medan andelen grammatiska ord var lite högre (55,2 %) i jämförelse med det första undersökningstillfället (54,6 %). Verb utgjorde fortfarande den största öppna ordklassen (18,7 %) och substantiv var den näst största ordklassen (15,0 %) i berättelser som informanterna skrev vid den andra datainsamlingsgången. Andelen verb minskade dock 0,2 procentenheter från det första undersökningstillfället (från 18,9 % till 18,7 %), medan andelen substantiv ökade i samma avseende, det vill säga 0,2 procentenheter (från 14,8 % till 15,0 %). Adjektiv och adverb kvarstod som de två minsta öppna ordklasserna: adjektivens procentuella andel var 5,8 % och adverbens andel var 5,4 %. Detta tyder på att skillnaden mellan dessa två ordklasser blev ännu mindre vid den andra datainsamlingsgången. Orsaken till detta är att andelen adverb minskade en procentenhet (från 6,4 % till 5,4 %) och samtidigt ökade andelen adjektiv 0,5 procentenhet (från 5,3 % till 5,8 %).

Vid det tredje och sista undersökningstillfället var informanterna på den femte kursen i svenskan. Kursen är den sista obligatoriska svenskkursen under gymnasieutbildningen. Då innehöll berättelserna 46,6 % lexikala ord och 53,4 % grammatiska ord, vilket innebär att andelen lexikala ord var som störst vid den sista datainsamlingsgången. Resultaten betyder också automatiskt att andelen grammatiska ord var som lägst då: deras andel var 54,6 % vid den första och 55,2 % vid den andra datainsamlingsgången. I de sista berättelser som informanterna skrev var verb den största öppna ordklassen: de omfattade 18,1 % av alla orden i texterna. Efter verb var substantiv den näst största ordklassen (15,9 %). Adjektiv och adverb var fortfarande de två minsta ordklasserna: andelen adjektiv var 6,7 % och andelen adverb var 5,9 % vid den sista materialinsamlingsgången.

Sammanfattningsvis kan konstateras att ordklassernas procentuella andelar förändrade under svenskstudiernas gång i gymnasiet. För att börja med verb kan det konkluderas att verb var den enda ordklass vars användning (det vill säga procentuella andelar) i texterna minskade under gymnasiets gång. Däremot ökade användningen av substantiv och adjektiv kontinuerligt under språkstudierna, medan användningen av adverb i texterna

vid de tre undersökningstillfällena var mer varierande i jämfört med de andra ordklasserna. Figur 2 nedan illustrerar dessa förändringar.



FIGUR 2. Förändringar i de fyra öppna ordklassernas procentuella andelar i berättelser (N=75) skrivna på svenska vid tre olika undersökningstillfällen under de obligatoriska språkstudierna i den finska gymnasieutbildningen.

4.1.2 Förändringar i ordvariation inom ordklasserna

Utöver förändringar i ordklassernas procentuella andelar var syftet med studien också att kartlägga hur ordvariation inom de fyra öppna ordklasserna utvecklas under språkstudiernas gång. I berättelser som informanterna skrev vid det första undersökningstillfället fick ordklassen substantiv det högsta TTR-värdet (0,48), medan det lägsta TTR-värdet räknades för ordklassen verb (0,22). TTR-värdena för adjektiv och adverb var mycket nära varandra (adjektiv: 0,38 respektive adverb: 0,37). Alla TTR-värden i berättelserna var därmed relativt låga vid det första undersökningstillfället, eftersom alla värden var under 0,50. Av resultaten att döma kan man konkludera att lexikonet inom ordklasserna varierade i ganska låg grad i de första analyserade berättelserna.

TTR-värden för alla ordklasserna förbättrades mellan de två första undersökningstillfällena. Vid den andra materialinsamlingsgången var TTR-värdet för ordklassen substantiv fortfarande det högsta (0,52) och det var lite högre än i samband med de första

berättelserna (0,48). Även om det högsta TTR-värdet räknades igen för substantiv, utvecklades ordvariationen inom adjektiv emellertid mest mellan de två första undersökningstillfällena: TTR-värdet för ordklassen adjektiv ökade från 0,38 till 0,47. Utöver detta visade resultaten även att TTR-värdet för adjektiv var det näst största efter substantiv. Det lägsta TTR-värdet (0,24) fick ordklassen verb och det var också den ordklass som utvecklade sig minst mellan den första och den andra datainsamlingen. Vid den andra datainsamlingsgången var TTR-värdet för adverb 0,41, vilket var lite högre än ordklassens första TTR-värde (0,37).

Ordklassernas TTR-värden förbättrade sig delvis också mellan det andra och det tredje undersökningstillfället. Det högsta TTR-värdet vid den sista materialinsamlingsgången var 0,52 och det räknades igen för substantiv. Det är ändå anmärkningsvärt att ordvariationen inte längre utvecklades för precis samma TTR-värde räknades redan för texter som var skrivna på det föregående undersökningstillfället. Adjektiv fick fortsättningsvis det näst högsta TTR-värdet (0,48), men det var bara lite högre än det föregående värdet (0,47). På ett motsvarande sätt utvecklade TTR-värdet för verb: det sista TTR-värdet för verb var 0,25 medan dess föregående värde var 0,24. Den enda ordklass vars TTR-värde sjönk mellan den andra och den tredje insamlingsgången var adverb: TTR-värdet sjönk från 0,41 till 0,37.

Sammanfattningsvis kan konstateras att alla fyra ordklasserna utvecklade sig i viss mån under gymnasiet, eftersom ordvariationen inom ordklasserna i huvudsak blev större. TTR-värdena kvarstod ändå relativt låga (huvudsakligen under 0,50) under gymnasiet. Ordklassen adjektiv utvecklades mest medan utvecklingen av de tre övriga ordklasserna var mer måttfull. Resultaten och utvecklingen av ordvariationen inom ordklasserna har sammanfattats i Tabell 4 nedan.

TABELL 4. Ett sammandrag av ordklassernas TTR-värden i svenska berättelser (N=75) skrivna vid tre olika materialinsamlingsgångar under gymnasiestudierna.

Insamlingsgången → Ordklassen ↘	Insamling 1	Insamling 2	Insamling 3
Substantiv	0,48	0,52	0,52
Verb	0,22	0,24	0,25
Adjektiv	0,38	0,47	0,48
Adverb	0,37	0,41	0,37

4.2 De separata analyserna – svenska

4.2.1 Adverbial

Adverbial analyserades alltså separat i denna studie (för en noggrannare information om analyseringens gång, se avsnitt 3.3.2, och för en beskrivning av adverbial, se avsnitt 2.3.5). I berättelser utgjorde adverbial naturligt tre huvudgrupper: 1.) rumsadverbial, 2.) tidsadverbial, och 3.) övriga adverbial, varav den sistnämnda exempelvis omfattade *med*-fraser och enstaka exempel på några andra adverbial än rums- eller tidsadverbial. Berättelser som hade insamlats vid det första undersökningstillfället innehöll sammanlagt 98 adverbial, vilket betyder att varje informant i genomsnitt använde fyra adverbial i sin text. Majoriteten av adverbialen var rumsadverbial (55 stycken, 56,1 %), och exempel på informanternas adverbialval inom denna grupp var ”i köket”, ”i Helsingfors” och ”i min fasters gamla hus”. Den näst största gruppen var tidsadverbial: 32 adverbial (32,7 %) hörde till denna grupp, och de kan exemplifieras med följande uttryck: ”på lördag”, ”på morgonen” och ”efter konsert”. Den minsta gruppen var därmed övriga adverbial med 11 uttryck (11,2 %), såsom ”med min familj”.

Berättelser skrivna vid den andra datainsamlingsgången innehöll lite färre adverbial än de första berättelserna för totalantalet adverbial minskade från 98 till 85. Det fanns därmed ungefär tre adverbial per text. Resultatet är intressant eftersom texternas genomsnittliga längd ändå var relativt stabil mellan datapunkterna 1 och 2 (se Figur 1 i avsnitt 3.2.1). Utöver förändringen i adverbialens totalantal var en annan intressant skillnad mellan de första och de andra berättelserna att skillnaden i mängden av rumsadverbial

och tidsadverbial blev mycket jämnare: totalantalet för båda adverbialtyperna var nämligen precis detsamma (37 stycken, 43,5 %) i berättelser som var skrivna vid det andra undersökningstillfället. Antalet övriga adverbial var 11 (12,9 %) då.

Totalantalet adverbial ökade igen mellan de två sista datainsamlingsgångerna, för berättelserna från det sista undersökningstillfället omfattade 98 adverbial, vilket är exakt samma antal som i de första berättelserna. Varje berättelse inkluderade i genomsnitt fyra adverbial, och den största adverbialgruppen var rumsadverbial (43 stycken, 43,9 %) vid det sista undersökningstillfället. Nästan lika mycket användes tidsadverbial (39 stycken, 39,8 %) medan en klar minoritet hörde till övriga adverbial (16 stycken, 16,3 %).

Sammanfattningsvis kan konstateras att totalantalet adverbial kvarstod relativt stabilt i berättelser som informanterna skrev under deras gymnasiestudier. En viktig förändring i användningen av adverbial var att de kvantitativa skillnaderna mellan olika adverbialgrupper, i synnerhet mellan tids- och rumsadverbial, blev jämnare i berättelser mot slutet av gymnasiet. Detta kan delvis förklaras med att totalantalet rumsadverbial minskade i texterna även om de kvarstod den största adverbialgruppen vid alla undersökningstillfällen.

4.2.2 Egennamn

Utöver adverbial analyserades också egennamn och förändringar i deras användning separat. Analysen omfattade sådana egennamn som var skilda enheter (till exempel namn på platser, människor och husdjur) och som således inte var en del av ett adverbial. Till exempel i meningen ”Helsingfors är vackert” skulle ordet ”Helsingfors” analyseras som egennamn, medan i meningen ”Våren är vacker i Helsingfors” skulle det analyseras som en del av adverbialet ”i Helsingfors”, eftersom jag inte ville inkludera samma egennamn två gånger i de separata analyserna (det vill säga som en del av både analyseringen av adverbial och egennamn).

Analysen av egennamn resulterade i fyra mindre undergrupper: 1.) namn på personer, 2.) namn på platser, 3.) namn på band och sångare, och 4.) övriga egennamn (till exempel namn på husdjur och händelser). Undergrupperna bildades eftersom egennamn tillhö-

rande till dessa fyra grupper förekom oftast i materialet. Berättelser från det första undersökningstillfället innehöll sammanlagt 27 egennamn, varav åtta (29,6 %) var namn på personer (till exempel "Kalle") och fyra (14,8 %) var namn på platser ("Helsingfors"). Den tredje undergruppen var störst vid denna materialinsamlingsgång med elva (40,7 %) namn på band och sångare ("Metallica"). Den fjärde undergruppen innehöll fyra egennamn (14,8 %) varav alla var namn på husdjur.

Berättelser från det andra undersökningstillfället innehöll lite flera egennamn, för deras totalantal var 32. Majoriteten av dem (14 stycken, ca 44 %) var namn på platser, vilket också innebär att undergruppens totalantal ökade i viss utsträckning i jämförelse med de första berättelserna (från fyra till 14). De två näst största undergrupperna var övriga (tio stycken, ca 31 %) och namn på personer (åtta stycken, 25 %). Jämfört med de första berättelserna utgjorde den tredje undergruppen (namn på band och sångare) den minsta gruppen vid den andra datainsamlingsgången, eftersom det inte fanns några exempel på dem i berättelserna.

Berättelser som hade insamlats vid den sista datainsamlingsgången inkluderade minst egennamn; deras totalantal var bara 11. Den största undergruppen var namn på platser (5 stycken, 45,5 %) också vid detta undersökningstillfälle, och den näst största gruppen var övriga (3 stycken, 27,3%). Dessutom fanns det två exempel på namn på personer och ett namn på band. Undergruppernas inbördes fördelning liknade därigenom den i berättelserna från den andra materialinsamlingsgången.

Allmänt taget kan det konstateras att totalantalet egennamn ganska drastiskt minskade efter de två första datainsamlingsgångerna. När det gäller undergrupperna, minskade i synnerhet andelen namn på band och sångare (undergrupp nr 3) mot slutet av informanternas gymnasiestudier.

4.2.3 Negationsord

Den separata analysen av svenska negationsord koncentrerade sig således på det svenska negationsordet "inte". Totalantalet negationsord i informanternas berättelser kvarstod mycket stabilt under gymnasiestudier, för deras antal förändrades nästan inte

alls mellan de tre materialinsamlingsgångerna. Vid det första undersökningstillfället var totalantalet negationsord 17, vilket utgjorde 1,3 % av totalantalet ord i berättelserna. Mängden negationsord kvarstod densamma mellan de två första datainsamlingsgångerna, eftersom deras antal var 17 (1,2 %) i berättelser skrivna vid det andra undersökningstillfället. Däremot sjönk totalantalet negationsord lite mellan den andra och den tredje materialinsamlingsgången: i berättelser som informanterna skrev vid den sista datainsamlingsgången var deras totalantal 15 (1,1 %). Å ena sidan minskade användningen av negationsord alltså lite både numerärt och procentuellt, men å andra sidan var förändringarna inte anmärkningsvärda. Därför kan konkluderas att inga signifikanta förändringar skedde under gymnasiet i användningen av negationsord i berättelserna.

4.2.4 Ord som utelämnades i andra analyser

Ord som helt uteslöts från den kvantitativa analysen och de separata analyserna av adverbial, egennamn samt negationsord analyserades som en egen grupp. Ord kategoriserades till denna grupp på basis av de kriterier som presenterades i avsnitt 3.3.2 (se särskilt Tabell 2). Dessa exkluderade ord och uttryck klassificerades till tre huvudgrupper under analyseringen: 1.) ord från främmande språk, 2.) ord och uttryck vars betydelse och/eller ordklass var oförståelig, och 3.) ord som utelämnades på grund av någon annan orsak. Inom den tredje gruppen, det vill säga ”ord som utelämnades på grund av någon annan orsak”, var anledningar till varför ord uteslöts från andra analyser mycket varierande: till exempel ord som inläraren själv hade hittat på samt felaktiga ordval som försvårade meningens eller hela textens begriplighet klassificerades till denna grupp. Tabell 5 sammanfattar resultat för huvudgrupperna vid alla tre datainsamlingsgångerna:

Tabell 5. Ett sammandrag av utelämnade ord i svenska berättelser (N=75) vid de tre materialinsamlingsgångerna.

Insamlingsgången → Huvudgruppen ↘	Insamling 1	Insamling 2	Insamling 3
1.) ord från främmande språk	13 (54,2 %)	33 (76,7 %)	11 (48 %)
2.) ord och uttryck vars betydelse och/eller ordklass var oförståelig	11 (45,8 %)	4 (9,3 %)	5 (22 %)
3.) ord som utelämnades på grund av någon annan orsak	0 (0 %)	6 (14,0 %)	7 (30 %)
Totalantalet utelämnade ord:	24 (100 %)	43 (100 %)	23 (100 %)

Som Tabell 5 visar utelämnades sammanlagt 24 ord eller uttryck från analyseringen av de berättelser som insamlades vid det första undersökningstillfället. Då innehöll berättelserna 13 ord från främmande språk (54,2 %), och den största delen av dem var från engelskan: till exempel orden ”icehockey”, ”guitar”, ”thing” och ”kept” förekom i materialet. Totalt 9 ord av dessa 13 exkluderade ord var substantiv, och resten var adjektiv och verb. I materialet fanns även 11 ord vars betydelse och/eller ordklass var obegriplig (45,8 %, huvudgrupp nr 2): exempel på dem är ”(...) men jag inte tycka att göra till Helsingfors” och ”Jag kunde inte spara, för jag kröss”. Däremot fanns det inga ord i berättelser som exkluderades på grund av någon annan orsak (den tredje huvudgruppen) vid den första datainsamlingsgången.

Det framgår också av Tabell 5 att resultat för den andra materialinsamlingsgången avviker relativt mycket från de två andra insamlingsgångerna. Då utelämnades sammanlagt 43 ord och uttryck från materialet. Majoriteten av dem var ord från främmande språk (totalt 33 stycken, 76,7 %), medan de två övriga huvudgrupperna utgjorde mindre delar av exkluderade ord: det fanns fyra ord (9,3 %) som utelämnades eftersom deras betydelse och/eller ordklass var oförståelig och sex ord (14,0 %) som utelämnades på basis av någon annan orsak. Den största skillnaden i resultaten mellan den första och den andra datainsamlingsgången var antalet ord från främmande språk: deras totalantal var alltså 13 i berättelser som insamlades vid det första undersökningstillfället, men vid det andra undersökningstillfället fanns det 33 sådana ord. Den största delen av dem (26 av 33) var engelska ord, såsom ”volleyball”, ”songs” och ”bake”, men även tyska ord (3 av

33) och finska ord (4 av 33) användes i berättelserna. 21 av orden från främmande språk var substantiv och resten av dem representerade övriga öppna ordklasser.

Tabell 5 visar också att totalantalet utelämnade ord var 23 för berättelser från det tredje undersökningstillfället. Nästan hälften av dem (11 stycken, 48 %) var ord från främmande språk, antingen från engelskan (till exempel: "(...) det var så **horrible**") eller tyskan ("min bästa **tag** var (...)"). Utöver detta fanns det sammanlagt fem 5 ord (22 %) vars betydelse och/eller ordklass var omöjlig att avgöra (till exempel: "Det är så roligt **gor orsaken** med dem") och sju ord (30 %) vilka exkluderades på grund av någon annan orsak, såsom felaktiga ordval ("Jag **gjorde** på sjukhuset"). Resultat för berättelser som samlades in vid det sista undersökningstillfället liknade därmed i viss utsträckning resultat för de första berättelserna.

I allmänhet kan konstateras att resultaten för alla materialinsamlingsgångar tydde på att i genomsnitt ett ord per text utelämnades. Vid varje undersökningstillfälle var ord från främmande språk den vanligaste orsak som ledde till att utesluta ett ord i berättelser. I de flesta fallen hade informanter lånat dessa ord från engelskan. Resultaten kan innebära att inlärnarna möjligen ville kompensera eventuella brister i sina svenskkunskaper genom att använda andra språk, såsom engelska, som de hade studerat längre. Resultaten är alltså intressanta ur ett inlärningsperspektiv, eftersom utelämnade ord och uttryck i sig kan spegla några aspekter av inlärarspråket. Dessutom framgick av resultaten att andelen utelämnade ord var som störst vid det andra undersökningstillfället, vilket kan bero på att informanterna höll på processa ny information. Undersökningstillfället var nämligen i mitten av deras språkstudier, och man därför anta att mängden av ny information kan vara stor då: å ena sidan har de troligen redan lärt sig mycket på föregående kurser, men å andra sidan finns det fortfarande information som de borde tillägna sig under gymnasiet. Att informanter håller på att processa information kan därmed synas i en form av ett större antal ord från främmande språk i berättelserna.

4.3 Den kvantitativa analysen – engelska

4.3.1 Förändringar i ordklassernas procentuella andelar

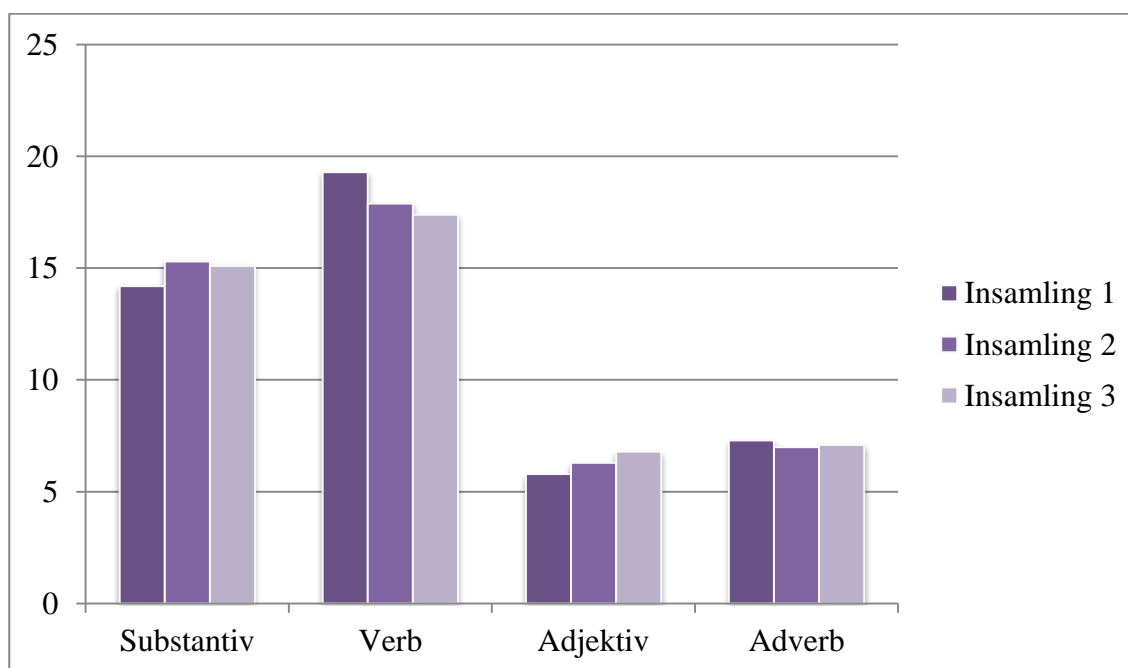
Vid den första datainsamlingsgången var informanterna på den första obligatoriska kursen i engelskan i gymnasiet. Då innehöll deras berättelser cirka 46,7 % lexikala ord och 53,3 % grammatiska ord, vilket betyder att en relativt stor andel av orden var grammatiska. Angående ordklassernas procentuella andelar visade resultaten att skillnaden mellan den största ordklassen (verb, 19,3 %) och den minsta ordklassen (adjektiv, 5,8 %) var ganska stor. Den näst största ordklassen efter verb var substantiv (14,2 %). Till sist kan det konstateras att andelen adverb (7,3 %) var relativt nära andelen adjektiv (5,8 %). Dessa två ordklasser utgjorde därigenom de två minsta ordklasserna i berättelserna som informanterna skrev vid den första materialinsamlingsgången.

Informanterna i studien deltog i sin fjärde obligatoriska kurs i engelskan när de skrev berättelserna för andra gången. Vid det andra undersökningstillfället innehöll berättelserna 46,6 % lexikala ord och 53,4 % grammatiska ord. Dessa andelar var därmed nästan identiska med andelarna i de första berättelserna, då andelen lexikala ord var 46,7 % och andelen grammatiska ord var 53,3 %. De olika ordklassernas procentuella andelar förändrades däremot i viss grad mellan den första och den andra datainsamlingsgången. Den största ordklassen var fortsättningsvis verb (17,9 %) men ordklassens procentuella andel av alla ord sjönk ungefärligen 1,4 procentenheter från det första undersökningstillfället då deras andel var 19,3 %. Substantiv utgjorde igen den näst största ordklassen, och deras procentuella andel ökade från 14,2 % till 15,3 % vilket innebär cirka en procentenhets ökning till föregående andel. Slutligen visade resultaten att 6,3 % av alla orden representerade ordklassen adjektiv, medan 7,0 % av texternas ord hörde till ordklassen adverb. Andelen adjektiv ökade 0,5 procentenheter medan andelen adverb minskade 0,3 procentenheter från det första undersökningstillfället. Av resultaten att döma kvarstod dessa två ordklasser som de två minsta ordklasserna även på den andra datainsamlingsgången.

För tredje gången skrev informanterna berättelser när de var på sin sista obligatoriska kurs i engelska i gymnasiet. Berättelser som insamlades vid detta sista undersökningstillfälle inkluderade 46,4 % lexikala ord och 53,6 % grammatiska ord. Andelen lexikala

ord och andelen grammatiska ord liknade därigenom i mycket stor utsträckning andelarna vid de två första datainsamlingsgångerna (lexikala ord: 46,7 % och 46,6 %, grammatiska ord: 53,3 % och 53,4 %). Under gymnasiet minskade andelen lexikala ord sammanlagt 0,3 procentenheter medan andelen grammatiska ord ökade i samma avseende. Förändringar var emellertid mycket små och därför kan det konstateras att andelen lexikala ord och andelen grammatiska ord var mycket stabil under hela gymnasiet. När det gäller ordklassernas procentuella andelar, hade deras andelar varit relativt oföränderliga jämfört med den andra datainsamlingsgången. Ordklassen verb var fortfarande den största öppna ordklassen (17,4 %), och adjektiv (6,8 %) utgjorde den minsta ordklassen. Det är dock märkningsvärt att andelen verb igen hade sjunkit medan adjektivens andel höll på att öka. Den näst största ordklassen var substantiv (15,1 %): dess procentuella andel sjönk 0,2 procentenheter från det andra undersökningstillfället (15,3 %). Andelen adverb ökade däremot lite (från 7,0 % till 7,1 %), och dess andel i de sista berättelserna var alltså 7,1 %

Sammanfattningsvis kan konstateras att ordklassernas procentuella andelar i berättelserna förändrades under gymnasiet, men förändringarna var ändå relativt små. Den enda ordklass vars procentuella andel kontinuerligt ökade var adjektiv medan verb utgjorde den enda ordklass vars användning i förhållande till andra ordklasser minskade vid varje undersökningstillfälle. Användningen av de två övriga ordklasserna, substantiv och adverb, gick inte på lika tydligt sätt utan deras procentuella andelar varierade lite mer. Figur 3 nedan tydliggör förändringar i ordklassernas andelar vid alla tre undersökningstillfällena.



FIGUR 3. Förändringar i de fyra öppna ordklassernas procentuella andelar i berättelser (N=75) skrivna på engelska vid tre olika undersökningstillfällen under språkstudierna i gymnasiet.

4.3.2 Förändringar i ordvariation inom ordklasserna

Vid det första undersökningstillfället skilde ordklassernas inre ordvariationer ordklasserna från varandra relativt tydligt. Det fanns en ordklass vars TTR-värde var högre än genomsnittet (det vill säga högre än 0,50) och den ordklassen var adjektiv. För adjektiv räknades följaktligen det högsta TTR-värdet (0,51), men substantivens TTR-värde var nästan lika högt (0,50). TTR-värdet för adverb var 0,32, och det lägsta TTR-värdet fick ordklassen verb (0,28). Skillnaderna i ordvariationen mellan ordklasserna (i synnerhet mellan det lägsta TTR-värdet och de två högsta TTR-värdena) var alltså relativt stora i berättelser som informanterna skrev i början av deras språkstudier.

TTR-värden förändrades i viss mån mellan den första och den andra materialinsamlingsgången. Den största skillnaden mellan dessa två insamlingsgångar var att bara en ordklass, nämligen adverb, fick ett högre TTR-värde på det andra undersökningstillfället (0,39) än på det första undersökningstillfället (0,32). Med tanke på de övriga ordklasserna var det så att TTR-värden antingen sjönk eller kvarstod. Det högsta TTR-värdet var 0,49 och det räknades för ordklassen adjektiv. Detta värde var lite lägre än värdet för de första berättelserna (0,51). Det näst största TTR-värdet fick substantiv (0,45),

men också dess TTR-värde sjönk lite (från 0,50 till 0,45). Det lägsta TTR-värdet räknades igen för verb, och det var exakt detsamma som värdet för berättelserna från det första undersökningstillfället (0,28).

I berättelser från den sista datainsamlingsgången varierade ordklassernas TTR-värden igen på lite annorlunda sätt i jämförelse med de två föregående undersökningstillfällena. Det allra högsta TTR-värdet (0,54) i hela studien (inklusive det svenska materialet) räknades vid denna insamlingsgång, och detta värde fick ordklassen adjektiv. Den näst största ordklassen var fortsättningsvis substantiv. Dess TTR-värde var 0,48, vilket var lite högre än dess föregående värde (0,45) men lägre än ordklassens första värde (0,50). TTR-värdet för adverb var 0,35. Det lägsta TTR-värdet räknades fortfarande för verb (0,28). Värdet kvarstod vid alla tre undersökningstillfällena.

Resultaten och utvecklingen av ordvariationen i berättelserna har sammanfattats i Tabell 6 nedan. Såsom tabellen och resultaten visar, var utvecklingen av ordklassernas inre ordvariation inte fullständigt linjär utan ganska varierande. Dessutom verkade det vara så att utvecklingen avbröts mellan de två första datainsamlingsgångerna, vilket exempelvis framgick av några lite sämre TTR-värden för berättelserna från det andra undersökningstillfället. Vid den sista insamlingsgången var en del av värdena igen lite högre.

TABELL 6. Sammandrag av ordklassernas TTR-värden i engelska berättelser (N=75) skrivna vid tre olika materialinsamlingsgångar under gymnasiestudier.

Insamlingsgången →	Insamling 1	Insamling 2	Insamling 3
Ordklassen ↘			
Substantiv	0,50	0,45	0,48
Verb	0,28	0,28	0,28
Adjektiv	0,51	0,49	0,54
Adverb	0,32	0,39	0,35

4.4 De separata analyserna – engelska

4.4.1 Adverbial

Engelska adverbial analyserades på ett motsvarande sätt som de svenska adverbialen (jfr avsnitt 4.2.1), och de klassificerades därmed till samma huvudgrupper som var 1.) rumsadverbial, 2.) tidsadverbial, och 3.) övriga adverbial. Engelska berättelser från det första undersökningstillfället inkluderade sammanlagt 116 adverbial, vilket i praktiken betyder att varje berättelse ungefärligen innehöll fem adverbial. En ganska klar majoritet av adverbialen var rumsadverbial, för 53 adverbial (45,7 %) hörde till denna grupp. Exempel var följande uttryck: ”in Florida”, ”in our bed”, och ”in a field”. Dessutom fanns det 41 exempel på tidsadverbial (35,3 %) i berättelserna, och de utgjorde därigenom den näst största adverbialgruppen vid denna datainsamlingsgång. Till exempel följande uttryck användes i texterna: ”in the mornings and in the evenings”, ”in last spring”, och ”a few years ago”. Resten av adverbialen representerade gruppen *övriga*: det fanns 22 exempel (19,0 %) i texterna och kan exemplifieras med följande typer av uttryck: ”with my friend and my cousin”, ”because of heat”, och ”without them”.

Totalantalet adverbial förändrade lite mellan de två första materialinsamlingsgångerna för deras mängd ökade från 116 till 134. Förändringen i mängden av *övriga* adverbial förklarar detta: deras totalantal ökade nämligen från 22 till 40 och gruppen utgjorde den näst största adverbialgruppen (29,9 %) vid den andra datainsamlingsgången. Rumsadverbialen utgjorde fortfarande den största adverbialgruppen (56 stycken, 41,8 %) medan tidsadverbialen var den minsta gruppen (38 stycken, 28,4 %) vid detta undersökningstillfälle. Rums- och tidsadverbialens andelar liknade i stor utsträckning deras andelar i berättelser som informanterna skrev vid det första undersökningstillfället.

Berättelser från den sista materialinsamlingsgången inkluderade igen lite flera adverbial i jämförelse med de två föregående undersökningstillfällena för deras totalantal ökade till 142. I genomsnitt fanns det sex adverbial per text. I de sista berättelserna utgjorde tidsadverbial en knapp majoritet av alla adverbial, och deras totalantal var 58 stycken (40,8 %). Den näst största gruppen var rumsadverbial med 54 exempel (38,0 %), men skillnaden mellan tids- och rumsadverbialens antal var dock relativt liten. Resultaten innebär också att deras andelar blev mycket jämnare jämfört med de motsvarande ande-

larna i de två första berättelserna. Den minsta adverbialgruppen utgjordes av övriga adverbial (30 stycken, 21,1 %).

Sammanfattningsvis kan konstateras att adverbialens totalantal kontinuerligt ökade i texterna under gymnasiets gång och mellan de tre undersökningstillfällena. Utöver detta blev skillnaden i mängden tids- och rumsadverbial mycket jämnare i texterna mot slutet av de obligatoriska språkstudierna. Detta kan förklaras med att antalet rumsadverbial kvarstod relativt stabilt vid alla tre undersökningstillfällena medan antalet tidsadverbial betydligt ökade efter skrivandet av de två första berättelserna.

4.4.2 Egennamn

Analyseringen av engelska egennamn baserade sig på samma principer som användes redan i samband med egennamn i svenska berättelser. Därmed klassificerades de engelska egennamnen i fyra likartade undergrupper: 1.) namn på personer, 2.) namn på platser, 3.) namn på band och sångare, och 4.) övriga. Berättelser från det första undersökningstillfället innehöll sammanlagt 24 egennamn. En knapp majoritet av dem (åtta stycken, 33,3 %) var namn på personer (undergrupp nr 1), såsom "Kaisa" och "Kirsi". Undergrupperna 3 (namn på band och sångare, till exempel "Slayer") och 4 (övriga, till exempel "Alba" som är ett namn på ett hotell) utgjorde lika stora grupper med sju exempel (29,2 %) per grupp. Namn på platser (undergrupp nr 2, såsom "Finland") var den minsta undergruppen i de första berättelserna för endast två egennamn (8,3 %) hörde till den.

Texter som gymnasister hade skrivit vid det andra undersökningstillfället omfattade lite färre egennamn än de första texterna, för deras totalantal minskade från 24 till 18. Namn på personer var fortfarande den största gruppen (sju stycken, 38,9 %), men precis samma antal egennamn hörde till den fjärde undergruppen, det vill säga till gruppen övriga. De två minsta undergrupperna var däremot namn på platser och namn på band och sångare, och båda grupperna representerades av enbart två egennamn (11,1 %). Jämfört med resultaten för de första berättelserna var den största skillnaden mellan dessa två undersökningstillfällen att antalet namn på band och sångare sjönk ganska mycket (från sju till två stycken) vilket vidare också förklarar att totalantalet egennamn minskade från

24 till 18. Annars liknade resultaten i relativt stor utsträckning den första datainsamlingsgångens resultat.

Berättelser skrivna vid det sista undersökningstillfället inkluderade igen ett mindre antal egennamn i jämförelse med de två andra materialinsamlingsgångerna, eftersom de tredje berättelserna innehöll sammanlagt bara 13 egennamn. En klar majoritet av dem var namn på platser (åtta stycken, 61,5 %) och den näst största gruppen var övriga med sitt fyra exempel (30,8 %). Utöver detta fanns det bara ett exempel på namn på band och sångare (7,7 %). Däremot förekom det inga namn på personer i de sista texterna.

För att sammanfatta kan konstateras att trenden angående egennamnens totalantal kontinuerligt var nedåtgående. I synnerhet minskade namn på personer samt namn på band och sångare mellan olika undersökningstillfällen. Den enda egennamnsgrupp vars totalantal ökade var namn på platser, medan andelen övriga egennamn kvarstod relativt stabil. Å andra sidan var skillnaderna mellan olika undergrupper ändå mycket små. Därför kan det konkluderas att förändringarna i totalantalet egennamn var mer betydande än förändringarna mellan och inom undergrupperna.

4.4.3 Negationsord

Den separata analysen av engelska negationsord fokuserades till negationsordet ”not” och dess förkortade variant ”n’t”, som kan användas till exempel i strukturer ”don’t” och ”haven’t”. Totalantalet negationsord i berättelser varierade i viss mån mellan de tre datainsamlingsgångerna. Det högsta antalet negationsord räknades för berättelser som informanterna skrev vid det första undersökningstillfället: då var deras mängd 30, vilket utgjorde sammanlagt 1,2 % av totalantalet ord i informanternas berättelser. Den största förändringen i användningen av negationsorden skedde mellan de två första undersökningstillfällena för antalet negationsord minskade relativt mycket då: deras antal var 20 (0,9 %) i berättelser som var skrivna vid den andra datainsamlingsgången. Texterna från den första insamlingsgången innehöll i genomsnitt mer än ett negationsord (ca 1,2 negationsord per text), men de andra texterna inkluderade nödvändigtvis inte ens ett negationsord (ca 0,8 negationsord per text). Berättelser som insamlades vid det sista undersökningstillfället omfattade 24 negationsord (1,0 %), vilket betyder att deras totalantal

ökade jämfört med berättelser från det andra undersökningstillfället. Fast totalantalet negationsord varierade i viss utsträckning vid olika tidpunkter under gymnasiet, var variationen emellertid tämligen liten, och den enda märkningsvärda förändringen i berättelser skedde mellan den första och den andra datainsamlingsgången.

4.4.4 Ord som utelämnades i andra analyser

Ord som exkluderades i engelska berättelser analyserades på samma sätt som motsvarande ord från de svenska berättelserna (jfr avsnitt 4.2.4). Därför kategoriserades de exkluderade orden och uttrycken i samma tre huvudgrupper under analyseringen av det engelska materialet: 1.) ord från främmande språk, 2.) ord och uttryck vars betydelse och/eller ordklass var oförståelig, och 3.) ord som utelämnades på grund av någon annan orsak. Tabell 7 nedan sammanfattar resultat för varje huvudgrupp vid alla tre undersökningstillfällena.

Tabell 7. Ett sammandrag av utelämnade ord i engelska berättelser (N=75) vid de tre materialinsamlingsgångerna.

Insamlingsgången →	Insamling 1	Insamling 2	Insamling 3
Huvudgruppen ↘			
1.) ord från främmande språk	1 (16,7 %)	6 (66,7 %)	5 (62,5 %)
2.) ord och uttryck vars betydelse och/eller ordklass var oförståelig	4 (66,7 %)	0 (0 %)	3 (37,5 %)
3.) ord som utelämnades på grund av någon annan orsak	1 (16,7 %)	3 (33,3 %)	0 (0 %)
Totalantalet utelämnade ord:	6 (100 %)	9 (100 %)	8 (100 %)

Som framgår av Tabell 7, utslöts endast sex ord från det engelska material som hade insamlats vid det första undersökningstillfället. I dessa berättelser fanns det ett ord från främmande språk, nämligen ordet ”kirurgian” (sv. ”kirurg”, eng. ”surgeon”) som troligen hade fått intryck av det finska ordet ”kirurgi”. Dessutom var det sammanlagt fyra sådana ord vars ordklass och/eller betydelse var omöjlig att bestämma (den andra hu-

vudgruppen), vilket kan belysas med hjälp av ett följande exempel: ”I trawelled with my friend and my cousin to the festival **spot**”. Bara ett ord utelämnades på grund av någon annan orsak (huvudgrupp nr 3): ordet var ”learning” och det användes i meningen ”She was **learning** mine same times”. Ordet exkluderades eftersom ordet ”teaching” skulle ha varit det naturligaste ordvalet på basis av kontexten. Utöver detta förhindrade ordet ”learning” i viss mån korrekt förståelse av den ifrågavarande meningen.

Dessutom visar Tabell 7 att ett lite större antal ord (sammanlagt nio stycken) utelämnades i berättelser som gymnasisterna hade skrivit vid det andra undersökningstillfället. Majoriten av dessa ord (sex av nio exkluderade ord) hörde till den första huvudgruppen, det vill säga de var ord från främmande språk. Ett exempel är följande mening: ”I regretted my **päätös** to not to go to choirs (...)” (”päätös” = eng. ”decision”, sv. ”ett beslut”). Fem av dessa ord hade lånats in från finska språket. Berättelser skrivna vid den andra datainsamlingsgången innehöll inga ord vars betydelse eller ordklass var obegriplig (huvudgrupp nr 2), men det fanns sammanlagt tre ord som blev uteslutna på grund av någon annan orsak (huvudgrupp nr 3), vilket var en aning högre andel än i samband med de första berättelserna.

Till sist visar Tabell 7 också att totalantalet utelämnade ord kvarstod nästan detsamma mellan de två sista datainsamlingsgångerna, eftersom antalet exkluderade ord var åtta i de sista berättelserna (i jämförelse med nio i de andra berättelserna). Fem av dessa åtta ord representerade ord från främmande språk, varav av fyra var finska ord (till exempel ordet ”opisto”, vars betydelse är ungefärligen ”institut” på svenska och ”institute” på engelska). Ett av de fem orden var ett svenskt ord ”centrum” (jfr eng. ”center”, ”[city] centre”). Dessutom fanns det tre ord som exkluderades eftersom deras betydelse och/eller ordklass inte var möjliga att avgöra men däremot fanns det inga ord som skulle ha utelämnats på basis av någon annan anledning.

Sammanfattningsvis kan det konkluderas att totalantalet uteslutna ord var relativt litet vid alla datainsamlingsgångerna, och dessutom var det också mycket stabilt i berättelserna mellan de tre undersökningstillfällena. Den vanligaste orsaken till utelämnandet av ett ord var att ordet hade lånats in från ett främmande språk, vilket oftast var finskan.

4.5 Jämförelser mellan svenska och engelska

För att börja med förändringar i ordklassernas procentuella andelar kan det konstateras att andelarna förändrades i viss grad i båda språken mellan undersökningstillfällena, men förändringarna var dock huvudsakligen tämligen små. Användningen av verb minskade i både svenska och engelska texter. Däremot använde gymnasister kontinuerligt flera adjektiv i sina berättelser. Utöver detta användes adverb i svenska och engelska berättelser på ett likartat sätt: förändringar i användningen av adverb var lite mer varierande i texterna än förändringar i användningen av de andra lexikala orden. Den enda väsentliga skillnaden mellan svenska och engelska berättelser var förknippad med substantiv: i svenska berättelser ökade andelen substantiv jämnt under gymnasiet, medan i engelska berättelser var förändringarna mer varierande och visade inte en lika tydlig utveckling.

När det gäller utvecklingen av ordvariation inom ordklasserna, framgick av resultaten att ordklassernas TTR-värden förändrades mellan undersökningstillfällena. Gemensamt för både svenska och engelska berättelser var att den största ordvariationen visades av adjektiv och substantiv, varav av den förstnämnda hade de största TTR-värdena i de engelska texterna och den sistnämnda i de svenska texterna. Ordklassen verb fick de lägsta TTR-värdena i båda språken vid varje materialinsamlingstillfälle. Resultaten mellan språken skiljde sig från varandra med tanke på en allmän utveckling av ordklassernas TTR-värden: i svenska berättelser verkade värdena bli huvudsakligen högre under gymnasiet, medan i engelska berättelser var utvecklingen lite mer varierande till sin natur. Tabell 8 nedan sammanfattar TTR-värdena för både svenska och engelska texter vid alla tre undersökningstillfällena.

TABELL 8. Sammandrag av utvecklingen av de öppna ordklassernas TTR-värden i svenskan och engelskan under språkstudierna i gymnasiet.

Datainsamlingsgången och ordklassernas TTR- värden	Svenskan	Engelskan
Insamling 1		
Substantiv	0,48	0,50
Verb	0,22	0,28
Adjektiv	0,38	0,51
Adverb	0,37	0,32
Insamling 2		
Substantiv	0,52	0,45
Verb	0,24	0,28
Adjektiv	0,47	0,49
Adverb	0,41	0,39
Insamling 3		
Substantiv	0,52	0,48
Verb	0,25	0,28
Adjektiv	0,48	0,54
Adverb	0,37	0,35

De första separata analyserna koncentrerade sig på texternas adverbial, som naturligt bildade tre huvudgrupper: rumsadverbial, tidsadverbial och övriga adverbial. I allmänhet kan konstateras att gymnasister använde flera adverbial när de skrev på engelska än när de skrev på svenska, och situationen var densamma vid varje undersökningstillfälle. Dessutom kvarstod totalantalet adverbial relativt stabilt i svenska berättelser mellan olika undersökningstillfällen, för antalet varierade mellan 85 (den andra materialinsamlingsgången) och 98 (den första och den tredje materialinsamlingsgången). Detta kan möjligen förklaras med att även totalantalet ord var stabilt vid alla datainsamlingsgångar (se Figur 1 i avsnitt 3.2.1). I jämförelse med engelskan visade resultaten att totalantalet adverbial jämnt ökade i informanternas texter: de första berättelserna innehöll 116 adverbial, de andra berättelserna 134 adverbial och de sista berättelserna 142 adverbial. Ett gemensamt drag för de svenska och de engelska texterna var en liknande användning av olika adverbialgrupper vid det första undersökningstillfället då rumsadverbial utgjorde den största gruppen, tidsadverbial den näst största gruppen och övriga adverbial den minsta gruppen. Efter detta blev de kvantitativa skillnaderna mellan olika adverbial-

algrupper jämnare i båda språken, i synnerhet mellan tids- och rumsadverbial (se också avsnitt 4.2.1 och 4.4.1).

Även egennamn bildade naturligt fyra mindre undergrupper i berättelserna (namn på personer, namn på platser, namn på band och sångare samt namn på övriga egennamn). Allmänt taget fanns det inga stora skillnader mellan språken i totalantalet egennamn: i engelskan varierade deras totalantal mellan 13-24 och svenska berättelser innehöll 11-32 egennamn beroende på materialinsamlingsgången. Det märkningsvärda var dock att egennamnens totalantal sjönk i svenska texter mellan den andra och den tredje datainsamlingsgången (från 32 till 11), medan i engelska texter minskade det hela tiden, det vill säga direkt efter det första undersökningstillfället (24 – 18 – 13). Även skillnaden mellan språken i egennamnens totalantal var som störst i berättelser från det andra undersökningstillfället (svenska texter: 32 egennamn jfr engelska texter: 18 egennamn). Å andra sidan visade analysen ändå några likheter mellan språken eftersom antalet namn på personer samt namn på band och sångare sjönk i båda språken mellan det första och det sista undersökningstillfället.

Resultaten angående negationsord visade att gymnasister generellt använde flera negationsord i engelska berättelser än i svenska berättelser vid varje undersökningstillfälle. Användningen av engelska negationsord var dock ganska varierande, för deras totalantal varierande beroende på datainsamlingsgången (30 – 20 – 24) vilket vidare tyder på att förändringarna inte var lineära. Däremot skedde det inte några betydande förändringar i användningen av negationsord i svenska berättelser utan deras totalantal kvarstod mycket stabilt under gymnasiet (17 – 17 – 15).

Studiens sista separata analys fokuserade ord som utelämnades i andra analyser, och dessa ord och uttryck grupperades i tre mindre undergrupper under analyseringens gång: ord från främmande språk, ord och uttryck vars betydelse och/eller ordklass var oförståelig, samt ord som utelämnades på grund av någon annan orsak. I allmänhet visade resultaten att svenska berättelser innehöll ett högre antal ord som måste utelämnas på varje undersökningstillfälle än engelska berättelser. När det gäller svenskan, var den vanligaste orsaken till utelämnandet av ett ord att ordet hade lånats från ett främmande språk, vilket i flesta fall var engelskan. Det fanns dock några enstaka exempel på tyska

och finska ord. Med tanke på engelskan tydde resultaten på att antalet uteslutna ord var relativt litet på alla tre datainsamlingsgångerna (6 – 9 – 8), och deras antal var stabilt i berättelserna mellan de tre undersökningstillfällena. Fast de exkluderade orden representerades ganska jämnt i alla undergrupper, var den vanligaste orsaken till utelämnandet av ett ord samma som i svenskan, det vill säga att ordet hade lånats in från ett främmande språk. Oftast var dessa ord från finskan.

Härnäst ska jag gå över till att diskutera mina forskningsresultat i mer detalj. Resultaten betraktas i synnerhet ur mina forskningsfrågors synvinkel (se avsnitt 3.1) och de anknyts också till studiens teoretiska utgångspunkter som presenterades i kapitel 2.

5 DISKUSSION

I detta kapitel ska jag sammanfatta mina forskningsresultat och diskutera dem i förhållande till både forskningsfrågor som ställdes i kapitel 3 och avhandlingens teoretiska bakgrund. Även hypotesernas riktighet kommer att beskrivas. Till slut ska jag ta ställning till validitet och reliabilitet i föreliggande studie och diskutera metodkritiken (avsnitt 5.2).

5.1 Sammandrag av resultaten

Min första forskningsfråga var: ”Utvecklas gymnasisternas produktiva lexikon i svenska och engelska i skrivna texter under gymnasieutbildningen? I så fall hur?” I samband med denna egentliga huvudfråga ställde jag också två mindre delfrågor som var: ”Förändras användningen av de fyra öppna ordklasserna?” och ”Blir ordvariationen större inom ordklasserna under gymnasietets gång?” För att börja med svenskan, visade resultaten att användningen av de lexikala orden förändrades under gymnasietets gång i informanternas berättelser. Verb och substantiv var de två största ordklasserna vid alla tre undersökningstillfällen och adjektiv samt adverb utgjorde de två minsta ordklasserna. Resultaten stämmer därmed i stor utsträckning med resultat i tidigare forskning (jfr Melin och Lange 2000, Hultman 2003). Särskilt adjektivens och adverbens procentuella andelar var mest samstämmiga med tidigare studier, eftersom deras andelar var under 10 % vid varje datainsamlingsgång, och till exempel adjektivens procentuella andel i de sista svenska berättelserna var precis desamma (6,7 %) som deras typiska andel är i gymnasistprosa (se Melin och Lange 2000: 168). Å andra sidan har tidigare studier visat att substantiv oftast är den dominerande ordklassen och verb utgör den näst största ordklassen efter den (Melin och Lange 2000, Hultman 2003). I denna studie var deras inbördes ordning motsatt, och substantivens andelar var lägre (14,8 % - 15,9 %) än deras typiska andelar i gymnasistprosa (21,9 %, se Melin och Lange 2000: 168). Resultaten visade också att både substantivens och adjektivens procentuella andelar kontinuerligt ökade medan verbens procentuella andel sjönk lite. Adverbens andelar varierade däremot mest och ingen lineär utveckling i deras användning kunde upptäckas.

Även ordvariationen blev större inom vissa ordklasser under svenskstudierna. Substantiv fick de högsta TTR-värdena vid alla datainsamlingsgångerna, men ordvariationen inom ordklassen adjektiv utvecklades mest: TTR-värden ökade nämligen från 0,38 till 0,48. Däremot visade resultaten att informanterna inte använde mer varierande verb och adverb i sina berättelser, eftersom dessa ordklassernas TTR-värden var nästan identiska hela tiden. Enligt min åsikt kan resultaten ge stöd till antagandet om att substantiv är lättast och adverb mest utmanande att lära sig (Laufer 1997: 148, Milton 2009: 37), vilket vidare betyder att det kan finnas skillnader mellan ordklasserna med tanke på hur lätta eller svåra de är att tillägna sig. Å andra sidan kan resultaten tolkas så att även om ordvariationen inom vissa ordklasser blev bättre under gymnasiet, var förändringarna emellertid relativt små. Utöver detta kvarstod TTR-värdena huvudsakligen under det genomsnittliga (0,50), vilket tyder på att lexikonet inte är särskilt varierande. Sammanfattningsvis kan konstateras att flera undersökningar om inläringen av ordklasserna borde genomföras för att få mer information om deras utveckling och potentiella skillnader mellan dem.

Med tanke på de separata analyserna av det svenska materialet, framgick av resultaten att totalantalet adverbial var relativt stabilt under gymnasiet. Berättelser från det första undersökningstillfället innehöll mest rumsadverbial, men skillnader i mängden av rums- och tidsadverbial blev mindre redan vid det andra materialinsamlingstillfället. Därför kan det konkluderas att användningen av adverbialen blev mångsidigare mot slutet av de obligatoriska svenskstudierna. Utöver detta visade resultaten att egennamnens totalantal sjönk i berättelser efter den andra datainsamlingsgången, medan antalet negationsord var nästan detsamma i texterna hela tiden. När det gäller ord som utelämnades från andra analyser, tydde resultaten på att den vanligaste orsaken till utelämnandet av ett ord var att det var från ett främmande språk, som i flesta fall var engelskan.

För att sammanfatta resultaten kan konstateras att gymnasisternas produktiva lexikon i svenska delvis utvecklades i berättelser under gymnasieutbildningen. Positiv utveckling syntes exempelvis i att gymnasister var mer kapabla att använda mer varierade adjektiv i sina sista texter än i sina första texter. Utöver detta var substantivens TTR-värden två gånger över det genomsnittliga vilket också var en positiv signal. Dessutom blev gym-

nasisternas adverbialbruk mer varierande mot slutet av svenskstudierna, eftersom de kvantitativa skillnaderna mellan rums- och tidsadverbial jämnades ut betydligt.

När det gäller det engelska materialet, visade resultaten att de lexikala ordens procentuella andelar i berättelserna förändrade under gymnasiet, men förändringarna var dock inte särskilt stora. De två största ordklasserna i berättelserna utgjordes av verb och substantiv vid varje datainsamlingsgång, medan adjektiv och adverb var de två minsta ordklasserna. Verbens procentuella andel minskade ändå kontinuerligt; däremot var adjektiv den enda ordklass vars procentuella andel ökade hela tiden. Användningen av substantiv och adverb var inte lika lineärt utan de var lite mer varierande i gymnasisternas berättelser.

Med tanke på ordklassernas TTR-värden, framgick av resultaten att utvecklingen av ordvariationen var relativt varierande och visade inte någon betydande positiv utveckling. Det verkade nämligen vara så att utvecklingen delvis avslutades mellan de två första undersökningstillfällena, eftersom TTR-värden för två av de fyra öppna ordklasserna (adjektiv och substantiv) var sämre vid den andra datainsamlingsgången än vid den första gången. Utöver detta var TTR-värdet för ordklassen verb exakt detsamma (0,28) i alla tre berättelserna. Vid den sista insamlingsgången var adjektivens och substantivens TTR-värden lite högre, vilket innebär att ordklasserna blev lite mångsidigare efter det andra undersökningstillfället. Fast resultaten inte visade någon lineär positiv utveckling för de lexikala ordens ordvariation, är det dock anmärkningsvärt att till exempel adjektivens TTR-värde var två gånger över det genomsnittliga (0,51 i de första berättelserna och 0,54 i de sista berättelserna). Även TTR-värdet för substantiv var två gånger nära det genomsnittliga: i samband med de första berättelserna (0,50) och de sista berättelserna (0,48).

Angående de separata analyserna av det engelska materialet, framgick av resultaten att adverbialens totalantal kontinuerligt ökade under gymnasiet. Antalet rumsadverbial kvarstod relativt stabilt medan antalet tidsadverbial blev betydligt större efter det andra undersökningstillfället. Detta betyder att de kvantitativa skillnaderna mellan rums- och tidsadverbial blev jämnare, vilket vidare kan tyda på att användningen av adverbialen blev mångsidigare i berättelserna mot slutet av engelskstudier. Utöver detta

minskade totalantalet egennamn kontinuerligt i gymnasisternas texter, medan mängden negationsord var mer varierande till sin natur. Till sist kan konstateras att mycket få ord utelämnades från analyserna; totalantal exkluderade ord var under tio stycken vid varje undersökningstillfälle. Majoriteten av dessa ord hade lånats in från ett annat språk, vilket oftast var finskan.

Sammanfattningsvis kan konstateras att gymnasisternas produktiva lexikon i engelska språket utvecklades i viss utsträckning. Följande resultat kan antas vara signaler på att informanternas produktiva lexikon utvecklades på ett positivt sätt: två gånger var TTR-värden för adjektiv högre än det genomsnittliga, och även adverbialens användning blev mångsidigare. Dessutom var antalet utelämnade ord mycket litet under hela studien. Å andra sidan var en stor del av förändringarna relativt små, och några ordklasser, såsom verb, verkade inte alls utvecklas, vilket vidare tyder på att hela det produktiva lexikonet inte förändrade på samma sätt eller utvecklades i samma takt. I själva verket var utvecklingen i de flesta fall varierande, det vill säga inte tydligt antingen positivt eller negativt.

Forskningsresultaten i båda språken stämmer mestadels med mina preliminära hypoteser (se avsnitt 3.1), eftersom resultaten i denna studie visade att informanternas produktiva lexikon delvis utvecklades under gymnasiet. Jag förväntade dock att de positiva förändringarna skulle ha varit lite större och mer synliga och att de skulle ha gällt alla ordklasser; nu blev till exempel ordvariationen större inom bara en del av ordklasserna. Jag antog också att andelen verb var relativt stabil och stor vid alla undersökningstillfällena, eftersom verb utgör en central ordklass i båda språken. Resultaten bekräftade antagandet även om deras procentuella andel minskade under gymnasiets gång. Dessutom var min hypotes om att substantiv och verb skulle vara de två största ordklasserna i båda språken vid varje undersökningstillfälle helt korrekt. Adverb och adjektiv utgjorde de minsta ordklasserna, och deras procentuella andelar var under 10 % av löporden. Även detta resultat stämde med mitt preliminära antagande samt med tidigare forskning.

Min andra forskningsfråga var: ”Finns det likheter och skillnader mellan svenska och engelska i utvecklingen av det produktiva lexikonet?”. Av forskningsresultaten att döma fanns det både likheter och skillnader mellan språken. Gemensamt för språken var att

ordklassernas procentuella andelar i huvudsak utvecklades på ett liknande sätt: till exempel verbens andelar minskade i både svenska och engelska berättelser under gymnasietets gång. Utöver detta fanns det också likheter angående ordvariationen: substantiv och adjektiv fick de högsta TTR-värdena i båda språken, medan ordklassen verb utvecklades minst. Å andra sidan skiljde sig språken från varandra med tanke på ordvariationen: TTR-värdena i svenskan blev mestadels större, medan i engelskan illustrerade utvecklingen av ordklassernas TTR-värden inte en lika rätlinjig positiv utveckling utan resultaten varierade i större utsträckning. Därmed visade resultaten att min preliminära hypotes inte stämde, eftersom jag antog att de engelska texterna skulle vara lexikalt mer varierande än de svenska texterna.

Informanterna använde också flera negationsord och adverbial i de engelska berättelserna än i de svenska berättelserna, vilket skiljde sig språken från varandra. Å andra sidan utjämnades exempelvis skillnader i totalantalet rums- och tidsadverbial mycket i båda språken under gymnasiet. Med tanke på egennamn, kan det konstateras att språkens resultat liknade varandra ganska mycket, eftersom det inte fanns några stora skillnader mellan dem i totalantalet egennamn. Dessutom sjönk totalantalet egennamn kontinuerligt i både svenska och engelska. När det gäller antalet utelämnade ord, visade resultaten att svenska berättelser innehöll flera sådana ord som exkluderades i andra analyser. I båda språken var den vanligaste orsaken till utelämnandet av ett ord dock detsamma, det vill säga att ordet hade lånats in från ett främmande språk. I svenskan var detta språk oftast engelskan, medan i engelska texter var det i de flesta fallen finskan.

I allmänhet anser jag att resultaten ger anledning till att diskutera i synnerhet inläring av de öppna ordklasserna. Resultaten visade att alla ordklassernas utveckling inte var positiv under gymnasiet: till exempel ordklassen verb kvarstod relativt oföränderlig i berättelserna, vilket är lite oroväckande. Verb utgör dock en central ordklass i såväl svenska som engelska, och därför skulle det vara viktigt att gymnasister kan utvidga även denna ordklass under sina språkstudier i gymnasiet. Läroplanen och lämpliga undervisningsmaterial kunde vara sådana alternativ som skulle stödja också inläringen av denna ordklass. Såsom det diskuterades tidigare i avhandlingen (avsnitt 2.8), nämns ordinläringen explicit ganska sällan i den nuvarande läroplanen. Det nämns i läroplanen att det är viktigt att både det svenska och det engelska ordförrådet utvidgas under

språkstudierna, och syftet är att språkbruket blir exaktare och mångsidigare (GLGY 2003: 95, 103). Eftersom alla ordklassernas TTR-värden i informanternas berättelser inte blev mångsidigare, ger resultaten anledning till att undra om läroplanens syfte blir uppfyllt. Utöver läroplanen vore det också väsentligt att ta hänsyn till inläring av ordinlärningsstrategier även på gymnasienivå, eftersom de kan hjälpa inlärnarna när det finns brister i deras lexikon (Thornbury 2002: 2). Med hjälp av olika ordinlärningsstrategier kunde det vara möjligt att förbättra både receptiv och produktiv behärskning av ordklasserna i olika språk.

5.2 Validitet och reliabilitet

Undersökningens *validitet* och *reliabilitet* brukar diskuteras framför allt i samband med kvantitativa studier, eftersom termerna ursprungligen har använts inom den kvantitativa forskningstraditionen (Hirsjärvi m.fl. 2003: 214). Hirsjärvi m.fl. (2003: 213) framför att med hjälp av *validitet* värderas hur väl en forskningsmetod eller en mätare faktiskt mäter det som den syftade till att mäta. Enligt Kalaja m.fl. (2011: 21) är det vanligt att göra en distinktion mellan inre och yttre validitet. De menar att *inre validitet* refererar till faktorer på basis av vilka en forskare kan bedöma hur väl undersökningen i sig är planerad och genomförd. Exempel på sådana faktorer är bland annat valet av informanter, systematisk analysering av data och rapportering av forskningsresultat. Med *yttre validitet* värderas däremot möjligheter till att generalisera ens forskningsresultat. Resultatens generaliserbarhet är delvis beroende på antalet data. (Kalaja m.fl. 2011: 21).

Begreppet *reliabilitet* anknyter däremot till studiens tillförlitlighet: när studien är trovärdig, ger den inte slumpmässiga resultat. I praktiken betyder detta att resultaten är desamma om undersökningen upprepas. (Hirsjärvi m.fl. 2003: 213). Såsom Kalaja m.fl. (2011: 21) påpekar, hör dessa två kriterier, det vill säga reliabilitet och validitet, tätt ihop eftersom en studie inte kan sägas vara tillförlitlig förrän man har visat att den fyller krav för inre validitet. Studien ska också först ha tillräcklig inre validitet innan yttre validitet kan värderas (*ibid.*).

Föreliggande studie var således mestadels kvantitativ till sin natur, men kvantitativa metoder, såsom TTR, kompletterades med hjälp av separata, lite mer kvalitativa ana-

lyser. Enligt Milton (2009: 126) är TTR-metoden dock inte helt oproblematiske: han påpekar att en av TTR-metodens nackdelar är att den är sensitiv till textlängd. Såsom det konstaterades redan i avsnitt 3.2.1, var de engelska berättelserna avsevärt längre än de svenska berättelserna vid varje datainsamlingsgång, vilket exempelvis kan ha påverkat jämförelser mellan språkens TTR-värden på ett negativt sätt. Detta kunde ha förbättrats genom att välja sådana texter ur materialet som är ungefärligen lika långa men å andra skulldetta ha minskat textvalens slumpmässighet. Dessutom skulldet ha varit relativt utmanande att hitta ett tillräckligt stort antal texter som är av samma längd. Däremot varierade textlängden *inom språken* inte så mycket mellan olika undersökningstillfällen, och därför är det ganska sannolikt att jämförelser inom språken är lite mer pålitliga än jämförelser mellan språken.

I undersökningen använde jag alltså inga egentliga statistiska metoder, utan den kvantitativa analysen baserade sig på enkla räkneoperationer (såsom procenträkning) och TTR. På grund av detta är studiens kvantitativa resultat proportionella och måste tolkas i förhållande till varandra: när andelen någon ordklass minskar, betyder det automatiskt att andelen någon annan ordklass ökar. Detta innebär att språkbrukaren använder den sistnämnda ordklassen förhållandevis mer i sin produktion än tidigare. Genom att använda statistiska analysmetoder skulldet ha varit möjligt att undersöka om förändringarna i proportioner egentligen är statistiskt signifikanta eller inte.

I denna avhandling studerade jag gymnasisternas utveckling som grupp, det vill säga jag var inte intresserad av enskilda individers utveckling. I praktiken betydde detta att informanterna inte nödvändigtvis var desamma vid olika undersökningstillfällen. Om informanterna hade varit desamma och skrivit en berättelse vid alla tre materialinsamlingsgångar, skulldet ha varit möjligt att undersöka skribenternas individuella inlärningsgångar och utveckling mer i detalj. I detta fall skulldinformanternas och texternas totalantal ändå ha varit mycket små, för en minoritet av Topling-projektets deltagare skrev texter vid varje tre undersökningstillfällen, och detta skulldförmodligen ha minskat resultatens generaliserbarhet i föreliggande studie. Därför anser jag att det var motiverat att betrakta informanternas utveckling på gruppnivå i stället för individnivå.

Det är också väsentligt att påpeka att fokus i denna studie har varit på endast en texttyp, det vill säga berättelser som handlade om någon minnesvärd erfarenhet eller händelse i informanternas liv. På grund av detta speglar resultaten gymnasisternas lexikala kunskaper bara inom denna genre och kan inte generaliseras till andra texttyper. Materialet var emellertid ganska omfattande och inkluderade sammanlagt 150 berättelser: därför kan man anta att det reflekterar förändringar i det produktiva lexikonet inom denna genre på ett ganska lyckat sätt. Det är ändå möjligt att resultaten skulle ha varit annorlunda om studien hade innehållit flera olika texttyper.

Intresset i föreliggande undersökning har varit på användningen av de fyra öppna ordklasserna, vilket betyder att jag inte har betraktat utvecklingen av hela det produktiva lexikonet och alla ordklasserna. Utöver detta analyserade jag endast skriftligt språk: om studien hade inkluderat både muntliga och skriftliga data, skulle resultaten ha varit mångsidigare, eftersom det är möjligt att gymnasister använder och behärskar lexikonet på ett lite olika sätt i muntliga och skriftliga språkbrukssituationer. Å andra sidan var det ett medvetet val att koncentrera sig på skriftligt språk. Enligt min åsikt är det ändå viktigt att framföra att resultaten inte reflekterar gymnasisternas hela lexikala repertoar utan en begränsad del av den.

Till sist är det viktigt att påpeka att berättelsen var den sista uppgiften vid varje undersökningstillfälle, vilket vidare kan ge anledning till att undra hur mycket informanterna egentligen orkade satsa på uppgiften, eftersom de hade skrivit fyra texter före den. Denna synvinkel kan få stöd av faktumet att berättelsernas genomsnittliga längd kvarstod relativt stabil under gymnasiet (se Figur 1 i avsnitt 3.2.1), vilket var lite överraskande, eftersom det skulle kunna ha varit sannolikt att berättelserna skulle ha blivit längre mot slutet av studierna, i samma takt med informanternas utvecklade språkkunskaper. Dessutom var deltagandet i Topling-projektet frivilligt och texter som gymnasister skrev inom projektets ramar hade ingen påverkan till exempel på deras kursbetyg. Av dessa anledningar är det motiverat att ta upp informanternas motivation till skrivandet av Topling-uppgifterna. Genom att komplettera de valda kvantitativa forskningsmetoderna med mer kvalitativa metoder, såsom intervjuer, skulle det ha varit möjligt att få en mer omfattande helhetsbild om gymnasisternas lexikala inlärningsgångar och eventuellt också om deras motivation.

Sammanfattningsvis anser jag att studien fyller kravet för validitet. De valda forskningsmetoderna fungerade väl, och jag fick svar på alla mina forskningsfrågor. Jag planerade undersökningen noggrant på förhand och till exempel kriterier för vilka ord som inkluderas och vilka ord som exkluderades definierades tydligt före analyseringen av materialet. I kapitel 4 syftade jag till att beskriva resultaten så detaljerat som möjligt. Dessutom var antalet data relativt stort, och texterna valdes slumpmässigt. Även om antalet materialet var tämligen stort, är resultaten inte helt generaliserbara, eftersom generaliserbarheten skulle ha förutsatt ett ännu större antal texter och troligen också texter som representerar olika genrer. Med tanke på reliabilitet, anser jag att undersökningen är tillförlitlig, eftersom resultaten sannolikt vore desamma om studien upprepas. Det är ändå alltid möjligt att kategorisera materialet på flera olika sätt vilket naturligtvis påverkar resultaten. Man skulle kunna ha indelat till exempel adverbial i mer än tre huvudgrupper, och resultaten skulle därmed ha sett olikt ut.

Nästa kapitel är det sista i föreliggande studie. Det syftar till att sammanfatta hela undersökningen och ge några konkreta förslag till vidare forskning.

6 AVSLUTNING

Syftet med denna studie var att undersöka hur gymnasisternas produktiva lexikon i svenskan och engelskan utvecklas under de obligatoriska språkstudierna i den finska gymnasieutbildningen. Huvudintresset låg på de öppna ordklassernas utveckling i skriftliga berättelser. Jag betraktade temat ur tre olika perspektiv: jag tog reda på förändringar i de öppna ordklassernas procentuella andelar i berättelser och utveckling av ordvariationen inom dessa ordklasser. Dessutom redogjorde jag för likheter och skillnader mellan svenska och engelska. Data bestod av 150 berättelser varav hälften var från det svenska Topling-materialet och hälften från det engelska Topling-materialet. Materialet hade ursprungligen insamlats i Topling-projektet år 2010-2013.

Undersökningen var huvudsakligen kvantitativ till sin natur men de numeriska resultaten kompletterades med separata, mer kvalitativa analyser av bland annat adverbial, egennamn och negationsord. Resultaten i studien visade att gymnasisternas produktiva lexikon i svenska och engelska utvecklades i viss grad i berättelserna under gymnasieutbildningen. Till exempel ordklassernas procentuella andelar utvecklades mestadels på samma sätt i båda språk. Dessutom blev gymnasisternas adverbialbruk mångsidigare mot slutet av språkstudierna, eftersom de kvantitativa skillnaderna mellan rums- och tidsadverbial utjämnades. Ordklassernas inre ordvariation utvecklades i synnerhet i svenska texter, eftersom gymnasister generellt var mer kapabla att använda mer varierande lexikala ord i sina tredje berättelser än i sina första berättelser. Å andra sidan var en stor del av förändringarna relativt små i både svenska och engelska, och några ordklasser, såsom verb, verkade inte alls utvecklas, vilket kan tyda på att hela det produktiva lexikonet inte utvecklades lika mycket. Dessutom inkluderade de engelska berättelserna färre sådana ord som utelämnades i analyser än de svenska texterna. Den vanligaste anledningen till att ett ord exkluderades var att ordet var från ett främmande språk. Avslutningsvis kan det således konstateras att informanternas produktiva lexikon utvecklades i viss mån mellan olika undersökningstillfällen i svenska och engelska, men det fanns såväl likheter som skillnader mellan språken.

Lexikonet och ordinläringen är viktiga med tanke på hela språkinlärningsprocessen samt kommunikation i allmänhet. Om man tänker sig en vanlig vardaglig situation där man kommunicerar med andra människor, är det troligen viktigare att man har ett tillräckligt stort (produktivt) ordförråd än att man har till exempel perfekta grammatiska kunskaper men ett mycket begränsat (produktivt) lexikon. Med hjälp av ord är det möjligt att förmedla ens tankar och åsikter och några små grammatiska brister förhindrar sällan kommunikationen. Nuförtiden uppskattas lexikaliska kunskaper mer än tidigare, och exempelvis Thornbury (2002: 14) framför att lexikonet inte längre anses vara något extra utan det har en viktig betydelse i sig. Lexikonet är nämligen förknippat med språkkunskapernas alla delområden. Såsom Viberg (2004: 197) påpekar, är ordinläringen dock en process som pågår hela livet även på modersmålet: det finns alltid något nytt att lära sig.

Det finns fortfarande mycket att forska i inom detta forskningsfält. I framtiden skulle det vara intressant och viktigt att undersöka ordklassernas samt det produktiva lexikonets utveckling i större utsträckning: man kunde till exempel ta reda på hur användningen av de slutna ordklasserna utvecklas i skriftliga texter under språkinlärningsprocessen. Det vore också intressant att jämföra olika ordklassernas procentuella andelar mellan olika texttyper och mellan muntliga och skriftliga texter. Det finns alltså flera möjligheter till lämpliga teman för framtida studier. Jag önskar att min undersökning har lyckats att ge några preliminära inblickar i området och kanske även inspirerat några andra forskare till vidare studier om likartade teman.

LITTERATUR

- Aitchison, J. 1987. *Words in the Mind: An Introduction to the Mental Lexicon*. Oxford: Basil Blackwell.
- Alexander, L. G. 1988. *Longman English Grammar*. London & New York: Longman.
- Carter, R. 2012. *Vocabulary: Applied linguistic perspectives*. Abingdon: Routledge.
- Cefling. 2007-2009. *Linguistic Basis of the Common European Framework for L2 English and L2 Finnish*.
<https://www.jyu.fi/hum/laitokset/kielet/tutkimus/hankkeet/paattyneet-hankkeet/cefling/en> (Hämtad: 22.2.2015)
- Ellis, N.C. 1997. Vocabulary acquisition: word structure, collocation, word-class, and meaning. I: Schmitt, N., och McCarthy, M. (red.): *Vocabulary: Description, Acquisition and Pedagogy*. Cambridge: Cambridge University Press. 122-139.
- Enström, I., & Holmegaard, M. 1994. Ordförråd och ordinläring. I: Cerú, E. (red.): *Svenska som andraspråk. Mera om språket & inläringen. Lärarbok 2*. Arlöv: Berlings. 164-191.
- Enström, I. 2004. Ordförråd och ordinläring – med särskilt fokus på avancerade inlärare. I: Hyltenstam, K., och Lindberg, I. (red.): *Svenska som andraspråk – i forskning, undervisning och samhälle*. Lund: Studentlitteratur AB. 171-195.
- GLGU 2004 = Grunderna för läroplanen för den grundläggande utbildningen. Utbildningsstyrelsen. <http://www02.oph.fi/svenska/ops/grundskola/LPgrundl.pdf> (Hämtad: 26.6.2015)
- GLGY 2003 = Grunderna för läroplanen i gymnasieutbildning för ungdomar. Utbildningsstyrelsen. <http://www02.oph.fi/svenska/ops/gymnasiet/gymnlpg.pdf> (Hämtad: 26.6.2015)
- Greenbaum, S. 1996. *The Oxford English Grammar*. Oxford: Oxford University Press.

Gymnasielag (629/1998).

<https://www.finlex.fi/sv/laki/ajantasa/1998/19980629?search%5Btype%5D=pika&search%5Bpika%5D=1998%2F629> (Hämtad: 26.6.2015)

Hirsjärvi, S., Remes, P., och Sajavaara, P. 2003. *Tutki ja kirjoita*. Helsingfors: Kustannusosakeyhtiö Tammi. 6. – 9. upplagan.

Hulstijn, J. 2000. Mnemonic methods in foreign language vocabulary learning: Theoretical considerations and pedagogical implications. I: Coady, J., och Huckin, T. (red.): *Second Language Vocabulary Acquisition*. Cambridge: Cambridge University Press. 203-224.

Hultman, T. G. 2003. *Svenska Akademiens språklära*. Stockholm: Svenska Akademien.

Jokinen, M., & Pitkäljärvi, T. 2005. *Gymnasisternas åsikter om ordinlärning*. Pro gradu -avhandling. Jyväskylä: Institutionen för språk, svenska språket vid Jyväskylä universitet.

https://jyx.jyu.fi/dspace/bitstream/handle/123456789/11381/URN_NBN_fi_jyu-2005172.pdf?sequence=1 (Hämtad: 1.10.2014)

Kalaja, P., Alanen, R., och Dufva, H. 2011. Minustako tutkija? Johdattelua tutkimuksen tekoon. I: Kalaja, P., Alanen, R., och Dufva, H. (red.): *Kieltä tutkimassa. Tutkielman laatijan opas*. Helsingfors: Finn Lectura. 8-32.

Laufer, B. 1997. What's in a word that makes it hard or easy: some intralexical factors that affect the learning of words. I: Schmitt, N., och McCarthy, M. (red.): *Vocabulary: Description, Acquisition and Pedagogy*. Cambridge: Cambridge University Press. 140-155.

Malmgren, S. 1994. *Svensk lexikologi: Ord, ordbildning, ordböcker och orddatabaser*. Lund: Studentlitteratur.

Melin, L., & Lange, S. 2000. *Att analysera text. Stilanalys med exempel*. Lund: Studentlitteratur. Tredje upplagan.

- Melka, F. 1997. Receptive vs. productive aspects of vocabulary. I: Schmitt, N., och McCarthy, M. (red.): *Vocabulary: Description, Acquisition and Pedagogy*. Cambridge: Cambridge University Press. 84-102.
- Milton, J. 2009. *Measuring second language vocabulary acquisition*. Bristol; Buffalo: Multilingual Matters.
- Nagy, W. 1997. On the role of context in first- and second-language vocabulary learning. I: Schmitt, N., och McCarthy, M. (red.): *Vocabulary: Description, Acquisition and Pedagogy*. Cambridge: Cambridge University Press. 64-83.
- Nation, I.S.P. 2001. *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Pavičić Takač, V. 2008. *Vocabulary Learning Strategies and Foreign Language Acquisition*. Clevedon: Multilingual Matters.
- Pomell, M. 2014. *Kollokationer i texter skrivna av L2-inlärare av svenska*. Pro gradu - avhandling. Jyväskylä: Institutionen för språk, svenska språket vid Jyväskylä universitet.
<https://jyx.jyu.fi/dspace/bitstream/handle/123456789/43676/URN%3aNBN%3afi%3ajyu-201406111992.pdf?sequence=1> (Hämtad: 1.10.2014)
- Sjöström, S. 2006. *Språkets struktur, betydelse och användning*. Stockholm: Norstedts Akademiska Förlag.
- Thornbury, S. 2002. *How to teach vocabulary*. Harlow: Longman.
- Tilastokeskus. 2015. *Väestö kielen mukaan sekä ulkomaan kansalaisten määrä ja maapinta-ala alueittain 1980 – 2014*.
[http://pxweb2.stat.fi/Dialog/Varvalagg.asp?ma=060%5Fvaerak%5Ftau%5F107&ti=V%EAest%F6+kielen+mukaan+sek%E4+ulkomaan+kansalaisten+m%E4%E4r%E4+ja+maa%2Dpinta%2Dala+alueittain+1980+%2D+2014&path=../Database/StatFin/vrm/vaerak/&lang=3&xu=&yp=&nr=1&aggfile\(1\)=%2D+Maakunnat+2015&prevagg=NNN&mapname=&multilang=fi&aggdir1](http://pxweb2.stat.fi/Dialog/Varvalagg.asp?ma=060%5Fvaerak%5Ftau%5F107&ti=V%EAest%F6+kielen+mukaan+sek%E4+ulkomaan+kansalaisten+m%E4%E4r%E4+ja+maa%2Dpinta%2Dala+alueittain+1980+%2D+2014&path=../Database/StatFin/vrm/vaerak/&lang=3&xu=&yp=&nr=1&aggfile(1)=%2D+Maakunnat+2015&prevagg=NNN&mapname=&multilang=fi&aggdir1) (Hämtad: 23.6.2015)

Topling. 2010-2013. *Topling – Inlärningsgångar i andraspråket.*

<https://www.jyu.fi/hum/laitokset/kielet/tutkimus/hankkeet/topling/se> (Hämtad: 15.11.2014)

Tornberg, U. 2009. *Språkdiraktik*. Malmö: Gleerups Utbildning AB. Fjärde upplagan.

Utbildningsstyrelsen. 2015. *Uppdatering av grunderna för gymnasiets läroplan.*

http://www.oph.fi/lagar_och_anvisningar/laroplans-och_examensgrunder/gymnasiet/glp2016 (Hämtad: 27.6.2015)

Viberg, Å. 2004. Lexikal utveckling i ett andraspråk. I: Hyltenstam, K., och Lindberg, I. (red.): *Svenska som andraspråk – i forskning, undervisning och samhälle*. Lund: Studentlitteratur AB. 197-220.

