

Syntaktisk komplexitet och korrekthet i finska universitetsstudenters skriftliga inlärarvenska: en jämförelse mellan två färdighetsnivåer

Syntactic complexity and accuracy in Finnish university students' written L2 Swedish: a comparison between two proficiency levels

Mari Mäkilä

Doktorand i Nordiska språk, Institutionen för språk- och översättningsvetenskap, Åbo universitet

mari.makila@utu.fi

Sammanfattning

I denna studie undersöks hur syntaktisk komplexitet och korrekthet skiljer sig på två olika färdighetsnivåer i skriftlig inlärarvenska samt vilken relation som föreligger mellan komplexitet och korrekthet på dessa färdighetsnivåer. Enligt tidigare studier ökar både komplexiteten och korrektheten i syntaktiska strukturer i samband med ökade kunskaper i målspråket, men det finns också studier som har visat på motstridiga resultat (Wolfe-Quintero et al., 1998). Vidare är det oklart om relationen mellan dessa komponenter är konkurrerande (Skehan, 1998) eller stödjande (Robinson, 2001).

Materialet i denna studie består av 31 uppsatser skrivna av finska universitetsstudenter. Fokuset ligger på strukturernas yttre komplexitet och följande komplexitetsmått används i analysen: den genomsnittliga längden på T-enheter och satser, andelen bisatser och förekomsten av bisatser med satsadverbial och topikaliserade strukturer. Korrekthetsaspekten avgränsas till fel i ordföljden och mäts med andelen felfria satser, felfria bisatser och felfria bisatser med satsadverbial samt förekomsten av inversion jämfört med det totala antalet topikaliserade strukturer.

Resultaten tyder på att färdighetsnivån huvudsakligen påverkar korrekthetsaspekten medan komplexiteten också kan påverkas av andra faktorer, såsom inlärarens allmänna kännedom om strukturer i målspråket. Något överraskande visar resultaten inga statistiskt signifikanta korrelationer mellan komplexitet och korrekthet. Därmed visar resultaten inget stöd för vare sig konkurrerande eller stödjande samband mellan dessa komponenter, med eventuellt undantag för en tendens till konkurrerande relation mellan bisatser med satsadverbial och korrekthet på den högre färdighetsnivån och mellan topikaliserade strukturer och korrekthet på den lägre nivån.

Nyckelord

färdighet, inlärarvenska, komplexitet, korrekthet

Abstract

This study examines how syntactic complexity and accuracy differ at two proficiency levels in written L2 Swedish, and how complexity stands in relation to accuracy at these proficiency levels. Based on the evidence from earlier research, both complexity and accuracy increase when learners become more proficient in the target language (see *e.g.* Wolfe-Quintero, Inagaki & Kim, 1998). However, some studies have resulted in mixed results. In addition, it is unclear

whether the relationship between complexity and accuracy is competitive (Skehan, 1998) or supportive (Robinson, 2001).

The data consist of 31 essays written by Finnish university students. The focus is on surface-level complexity, and the following measures for complexity are used: mean length of T-unit and clause, dependent clause ratio, dependent clauses with a sentence adverbial, and topicalised structures per T-unit. Accuracy is limited to word order and measured with error-free clauses, dependent clauses, dependent clauses with a sentence adverbial, and the use of inversion after topicalised structures.

The results indicate that proficiency mainly affects accuracy, whereas complexity is also affected by other factors, such as learners' knowledge of linguistic structures in general. Somewhat surprisingly, no statistically significant correlations were found between complexity and accuracy. Therefore, the results do not show any support for either a competitive or a supportive relationship between these components, with potential exceptions for a tendency towards competitive relationship between dependent clauses with a sentence adverbial and accuracy at the higher proficiency level, and between topicalised structures and accuracy at the lower level.

Keywords

accuracy, complexity, L2 Swedish, proficiency

Inledning

När språkinlärare blir mer kompetenta i målspråket ökar deras förmåga att producera olika slags syntaktiska strukturer, och därmed kan deras produktion bli mer komplex (se t.ex. Wolfe-Quintero, Inagaki & Kim, 1998, s. 4, 69). Därtill kan språket bli mer korrekt i samband med ökade språkfärdigheter. En intressant fråga är om relationen mellan komplexitet och korrekthet är konkurrerande eller stödjande, och vilka faktorer som påverkar denna relation. Särskilt på lägre färdighetsnivåer kan inlärare ha begränsade resurser när de processar målspråket, vilket kan leda till en konkurrerande relation exempelvis mellan komplexitet och korrekthet (Skehan, 1998). Detta betyder att inlärare kan fästa vikt vid antingen komplexitet eller korrekthet i produktionen. Under vissa omständigheter kan inlärare dock dela sin uppmärksamhet mellan både komplexiteten och korrektheten (Robinson 2001, se också Kowal, 2016, s. 213). Förmågan att samtidigt fästa vikt vid båda dessa komponenter har förknippats med ökade språkfärdigheter (se t.ex. Spoelman & Verspoor, 2010).

Syftet med denna studie är att redovisa hur komplexiteten och korrektheten i syntaktiska strukturer skiljer sig på två olika färdighetsnivåer¹ i skriftlig L2-svenska². Därtill studeras sambandet mellan komplexitet och korrekthet på dessa färdighetsnivåer. Med dessa syften strävar jag efter att fylla ett antal luckor i tidigare forskning om syntaktisk komplexitet och relationen mellan komplexitet och korrekthet. Tidigare studier om såväl relationen mellan komplexitet och färdighetsnivå som mellan komplexitet och korrekthet har visat på delvis motstridiga resultat (se diskussionen i Housen, Kuiken & Vedder, 2012; Kuiken, Vedder, Housen & De Clercq, 2019, s. 163), varför det behövs mer forskning om ämnet. Vidare har målspråket varit engelska i merparten av dessa studier (se dock t.ex. Norrby & Håkansson, 2007), och eftersom tidigare studier har visat att graden av komplexitet varierar mellan olika

språk (t.ex. Bernardini & Granfeldt, 2019) är det värdefullt att studera komplexitet och sambandet mellan komplexitet och korrekthet också i andra målspråk, i detta fall svenska.

I föreliggande studie definieras komplexitet som en strukturell egenskap i språk som kan mätas med hjälp av längdbaserade mått och andelen underordnade satser. Vidare uppfattas strukturer som är svåra för inlärare och som lärs in sent som komplexa (De Clercq & Housen, 2017, s. 316–317). På basis av tidigare forskning (se t.ex. Rahkonen & Håkansson, 2008) har förekomsten av bisatser med satsadverbial och topikaliserade strukturer (dvs. påståendeformade strukturer som inleds med någon annan satsdel än subjekt, Teleman et al., 2001) valts som komplexitetsmått i den aktuella studien. Därmed operationaliseras syntaktisk komplexitet genom längden på T-enheter (dvs. enheter som består av en huvudsats och alla eventuella bisatser samt meningsfragment, se Hunt, 1965) och satslängden, användningen av underordning samt förekomsten av bisatser med satsadverbial och topikaliseringar i den aktuella studien. Granskningen av korrektheten är avgränsad till *syntaktisk* korrekthet, dvs. endast fel i ordföljden beaktas i analysen. Längden på syntaktiska enheter, förekomsten av bisatser, bisatser med satsadverbial och topikaliserade strukturer studeras i relation till ordföljdens (in)korrekthet vid granskningen av sambandet mellan syntaktisk komplexitet och korrekthet.

Artikeln består av fem avsnitt. I det följande avsnittet presenterar jag den teoretiska referensramen, några centrala begrepp och ett urval tidigare studier kring komplexitet och korrekthet. Därefter redogör jag för materialet, metoderna och målet i studien, och efter detta följer analysdelen. I det avslutande avsnittet sammanfattar och diskuterar jag resultaten.³

Teoretisk referensram och tidigare forskning

Den teoretiska referensramen för denna undersökning baserar sig på kognitiva teorier och på språkbruksbaserade teorier om språkinlärning (t.ex. Ellis, 2019; Wulff & Ellis, 2018). Inlärares kognitiva resurser har en central roll när inlärare processar och producerar målspråket. Dessa resurser kan påverkas av flera faktorer, exempelvis av färdighetsnivån. Inom ramen för denna studie betyder detta att det kan finnas skillnader i exempelvis vilket samband som föreligger mellan komplexitet och korrekthet i inlärares produktion på olika färdighetsnivåer (mer om detta i samband med diskussionen om relationen mellan komplexitet och korrekthet).

Vidare tillhör denna studie CAF-forskningstraditionen (förkortningen står för complexity, accuracy och fluency) där forskare har beskrivit inlärares performans med hjälp av komponenter i CAF-triaden, dvs. komplexitet, korrekthet och flyt (för en omfattande beskrivning av CAF-studier, se t.ex. Housen et al., 2012; Wolfe-Quintero et al., 1998). I denna forskningstradition hävdas att L2-språkfärdighet är mångdimensionell till sin natur, och att man med hjälp av CAF-komponenterna kan analysera inlärarespråkets holistiska utveckling och mäta inlärares språkfärdighet på dessa enskilda områden (Housen et al., 2012, s. 2). I den aktuella studien undersöks komplexitet och korrekthet samt sambandet mellan dessa komponenter hos inlärare av svenska.

Komplexitet och korrekthet i inlärarespråk

Alla CAF-komponenter är mångdimensionella och särskilt har komplexitet definierats på olika sätt i tidigare studier (se t.ex. Bulté & Housen, 2018, s. 148; Pallotti, 2015, s. 118–120, Wolfe-Quintero et al., 1998, s. 69, 101). Komplexitet kan delas in i absolut och relativ komplexitet som båda ingår i forskningsobjektet här. Med absolut komplexitet avses antalet element och sambanden mellan dessa element inom ett system, medan relativ komplexitet syftar till komplexitet i förhållande till inlärarespråket (De Clercq & Housen, 2017, s. 316–317). Det är viktigt att hålla isär dessa två typer av komplexitet (se diskussionen i Pallotti, 2015). Genom att undersöka både absolut och relativ komplexitet syftar jag till att avslöja intressanta skillnader mellan de färdighetsgrupper som undersöks här (se också De Clercq & Housen, 2017, s. 316).

Vidare kan komplexitet delas in i syntaktisk och lexikal komplexitet, och därtill förekommer syntaktisk komplexitet på olika språkliga nivåer, exempelvis på fras- och satsnivå (t.ex. Biber, Gray & Staples, 2016, s. 664; Lambert & Kormos, 2014). I denna studie avgränsas granskningen av komplexitet till strukturernas yttre uppbyggnad, varför syntaktisk komplexitet definieras som strukturella egenskaper som framkommer i långa och inbäddade enheter (absolut komplexitet). Med komplexa strukturer avses även strukturer som lärs in sent och som används först på högre färdighetsnivåer (relativ komplexitet), i detta fall bisatser med satsadverbial och topikaliserade strukturer (se t.ex. Rahkonen & Håkansson, 2008).

En förenklad definition av korrekthet är att inlärarespråket kan användas målspråksenligt (Housen et al., 2012, s. 3). Att operationalisera korrekthet är däremot inte så enkelt. Särskilt i samband med inlärarespråk producerat av informanter på lägre färdighetsnivåer kan avgränsningen vara utmanande, eftersom felet i produktionen kan vara mycket frekventa (Kowal, 2016, s. 46–47; Kuiken & Vedder, 2012, s. 146). Eftersom fokus i föreliggande studie ligger på syntaktiska egenskaper beaktas endast felet i ordföljden. Jag vill ändå påpeka att syftet inte är att lyfta fram felet i sig, utan att undersöka hur korrektheten interagerar med färdighetsnivån och komplexiteten i informanternas produktion.

Komplexitet och korrekthet på olika färdighetsnivåer

Enligt de s.k. CAF-hypoteserna börjar inlärarespråket producera mer felfritt och komplext språk när deras språkkunskaper blir bättre (Kuiken & Vedder, 2012, s. 155, 164; Wolfe-Quintero et al., 1998, s. 4). Det finns starka belägg för att korrektheten korrelerar med inlärarespråkets färdighetsnivå, men det finns också studier som visar det motsatta (se Wolfe-Quintero et al., 1998, s. 37–38). Tidigare studier uppvisar motstridiga resultat även i fråga om sambandet mellan komplexitet och färdighetsnivå. Komplexiteten korrelerar med färdighetsnivån exempelvis i Bulté och Housen (2018) samt i Berggreen och Sørland (2016) och för de flesta måtten i Kuiken och Vedder (2019), medan resultaten exempelvis i Kuiken & Vedder (2012) och i Granfeldt (2007) visar på motstridiga resultat om relationen mellan komplexitet och färdighetsnivå.

Variationen i resultaten kan bero på flera faktorer. Enligt Wolfe-Quintero et al. (1998, s. 35) kan begränsade kunskaper i målspråket manifesteras antingen som enkelt språk eller som målspråksavvikande strukturer (se också Skehan, 1998). Detta påverkar sambandet mellan

komplexitet och färdighetsnivå respektive korrekthet och färdighetsnivå. Övriga faktorer som särskilt påverkar sambandet mellan komplexitet och färdighetsnivå innefattar exempelvis olika målspråk (se t.ex. Kuiken & Vedder, 2012) och olika mått som används i studierna (se t.ex. Bulté & Housen, 2018). Det har framkommit att inlärare på lägre färdighetsnivåer kan uppvisa komplexitet på andra språkliga nivåer än inlärare på högre nivåer (Lambert & Kormos, 2014). Exempelvis kan inlärare först börja producera fler underordnade satser, men ökande variation i de producerade strukturerna kan leda till en lägre andel underordnade satser. På de mer avancerade nivåerna börjar inlärare uppvisa komplexitet särskilt på satsnivå. Vidare kan individuella skillnader i komplexiteten vara betydliga (Bulté & Housen, 2018; Norrby & Håkansson, 2007). I den aktuella studien ligger fokus dock på gruppnivå, och syftet är att redovisa hur komplexitet, korrekthet och sambandet mellan dessa skiljer sig på två olika färdighetsnivåer.

Komplexitet på olika färdighetsnivåer diskuteras även i den gemensamma europeiska referensramen för språk (CEFR 2007, se också Kowal, 2016, s. 35). Enligt CEFR (2007, s. 63) består inlärares texter av fristående enheter på nivå A1. På nivå A2 börjar inlärare sätta samman enheterna och producera längre meningar, men även på nivå B1 kan språket vara relativt enkelt. På nivå B2 är strukturerna mer komplexa, och för att nå nivåerna C1 och C2 måste inlärare kunna producera välstrukturerade och komplexa texter.

Gyllstad, Granfeldt, Bernardini och Källkvist (2014) granskar relationen mellan CEFR-nivå och komplexitet i skriftlig inlärarengelska, -franska och -italienska. Komplexiteten mäts med en genomsnittlig längd på T-enheterna, satslängd och underordnade satser per T-enhet. Enligt deras studie är nivå B1 en brytningspunkt mellan enkla och komplexa strukturer i CEFR-deskriptorerna (se Gyllstad et al., 2014, s. 5–6). Vidare visar resultaten i deras studie att det finns en linjär, positiv korrelation mellan CEFR-nivå och komplexitet. Den aktuella studien syftar till att öka kunskap om hur språkfärdigheten interagerar med komplexitet i skriftlig L2-svenska.

Relationen mellan komplexitet och korrekthet

Som framkommit spelar inlärares kognitiva resurser en central roll när inlärare använder sitt L2 (se t.ex. Ellis, 2019, se också Wolfe-Quintero et al., 1998, s.69). Det finns olika syn på hur inlärare kan utnyttja dessa resurser under olika omständigheter (se t.ex. Housen et al., 2012, s. 6). Enligt Limited Attentional Capacity-modellen (Skehan, 1998) är inlärares kognitiva resurser begränsade när de processar målspråket, och därför kan inlärare bara fästa vikt vid ett begränsat antal CAF-komponenter i produktionen. Med andra ord måste inlärare exempelvis prioritera korrektheten på komplexitetens bekostnad, eller vice versa (mer om interaktionen mellan komplexitet, korrekthet och flyt, se t.ex. Housen et al., 2012: 5–8; Kowal, 2016, s. 175–176, 182, 184, 213). Detta kallas trade-off. I Multiple Resources Attentional-modellen (Robinson, 2001) hävdas att inlärare har flera förråd av resurser som de kan utnyttja samtidigt. Uppgiftens kognitiva krav kan styra inlärares uppmärksamhet så att både komplexiteten och korrektheten befrämjas i produktionen (se diskussionen i Kormos & Trebits, 2012, s. 444–445; Kuiken & Vedder, 2012). Enligt dessa två kontrasterande modeller kan interaktionen mellan komplexitet och korrekthet (samt flyt som inte tillhör forskningsobjektet i denna studie)

därmed vara konkurrerande eller stödjande (se Housen et al., 2012, s. 6–7; Kowal, 2016, s. 139). Denna interaktion mellan CAF-komponenterna är dock inte en stabil företeelse (Housen et al., 2012, s. 5). Varje komponent är mångdimensionell (t.ex. komplexitet på olika språkliga nivåer), och alla dimensioner är i konstant interaktion med varandra.

Flera faktorer kan påverka sambandet mellan komplexitet och korrekthet (se t.ex. Kormos & Trebits, 2012, s. 441–447, se också diskussionen i Kuiken et al., 2019, s. 163). Utöver uppgiftens kognitiva krav styrs sambandet av interna faktorer, särskilt av språkfärdigheten (se t.ex. Spoelman & Verspoor, 2010; Gunnarsson, 2012). Som konstaterats kan det speciellt på lägre färdighetsnivåer vara utmanande att fästa uppmärksamhet vid både komplexiteten och korrektheten (Skehan, 1998), medan det på högre färdighetsnivåer kan vara möjligt att samtidigt iaktta båda komponenterna (Robinson, 2001). Exempelvis konstaterar Spoelman och Verspoor (2010, s. 550) att komplexitet och korrekthet verkar konkurrera med varandra i början av inläringen, medan relationen mellan komponenterna på högre färdighetsnivåer kan vara antingen konkurrerande eller stödjande (se också Kowal, 2016, s. 185–186, 214). I föreliggande studie undersöks vilka skillnader i sambandet mellan komplexitet och korrekthet som förekommer på två olika färdighetsnivåer i L2-svenska. Speciellt intressant är att redogöra för skillnader mellan olika mått eftersom tidigare forskning tyder på att det finns skillnader i komplexiteten på olika språkliga nivåer mellan olika färdighetsnivåer (t.ex. Lambert & Kormos, 2014).

Komplexa strukturer i L2-svenska

I den aktuella studien mäts komplexitet med hjälp av längdbaserade mått och mått som är baserade på underordning. Vidare används två specifika mått som granskar användningen av topikaliserade strukturer och bisatser med satsadverbial. Dessa specifika mått har valts på basis av tidigare studier (se t.ex. Rahkonen & Håkansson, 2008; Åberg, 2020) som tyder på att topikaliserade strukturer och bisatser med satsadverbial lärs in sent och är utmanande strukturer för inlärare av svenska (för studier om L2-norska, se Jensen, 2018). Inläringen av dessa syntaktiska strukturer beskrivs också i de s.k. PT-nivåerna (Processbarhetsteori⁴, Pienemann, 1998, se också t.ex. Norrby & Håkansson, 2007). I början producerar inlärare mest huvudsatser med SVX-ordföljd (t.ex. subjekt–verb–objekt) och eventuella försök att använda underordnade satser och topikaliseringar resulterar i målspråksavvikande strukturer. Först på senare färdighetsnivåer blir underordnade satser (särskilt med satsadverbial) och topikaliserade strukturer mer frekventa och målspråksenliga.

Enligt Bohnacker och Lindgren (2014; se också Håkansson & Nettelbladt, 1993, s. 133, 136) är 30–40 % av strukturerna i skriften L1-svenska topikaliserade. Enligt tidigare studier (se t.ex. Rahkonen & Håkansson, 2008) börjar svenskinlärare emellertid använda topikaliserade strukturer mer frekvent först på de högre färdighetsnivåerna. Därtill kan inlärare av svenska ha svårigheter med att lära sig använda inversion efter topikalisering (V2-regeln, *Igår spelade han fotboll*, se t.ex. Rahkonen & Håkansson, 2008; Åberg, 2020).

Underordnade satser är också en del av en naturlig och varierande användning av svenska. Wijers (2018, s. 103) konstaterar att andelen bisatser i hennes inlärmaterial relativt snabbt uppnår den motsvarande andelen i L1-materialet. Skillnaden mellan L1- och L2-

språkanvändare är mer uppenbar i fråga om olika bisatstyper (se Wijers, 2018). Kowal (2016, s. 86) påpekar däremot att informanterna i hennes studie börjar använda underordning först senare i inlärningsgången. Vidare kan andelen underordnade satser minska på de högre färdighetsnivåerna när inlärare börjar variera sina strukturer mer och använda exempelvis satsförkortningar i sina texter (t.ex. Kuiken & Vedder, 2012; Ortega, 2003, s. 514).

Bisatser med målspråksenlig ordföljd (med och utan satsadverbial) framkommer relativt sent i inlärarspråket oavsett inlärningskontext (Rahkonen & Håkansson, 2008, s. 149, 151). Rahkonen och Håkansson (2008, s. 153–155) påpekar att bisatser oftast förekommer utan satsadverbial i L1-svenska, och detta kan påverka inläringen av denna struktur. Enligt teorier om språkbruksbaserad inläring (se t.ex. Ellis, 2002; Wulff & Ellis, 2018) sker inläringen genom språkanvändning och frekvensen på strukturerna i inflödet spelar en central roll i inläringen. Begränsat inflöde kan förhindra inläringen av en viss struktur. Därför är det på sätt och vis naturligt att ordföljden i bisatser med satsadverbial vållar problem för inlärare av svenska.

Material, metod och frågeställningar

I detta avsnitt redogörs för materialet, metoderna och syftet i studien.

Material

Informanterna är finskspråkiga universitetsstudenter. Endast de informanter som har gett sitt samtycke till att delta i undersökningen ingår i studien. En del av informanterna studerar nordiska språk som huvudämne (n=14) och övriga informanter deltar i en kurs i svenska vid Centrum för språk och kommunikation⁵ (n=17). Kön fördelningen är relativt jämn med 16 kvinnliga informanter, 14 manliga och 1 som är ickebinär. Informanternas ålder varierar från 19 till 34 (medelvärde 22 år).

Materialet består av 31 uppsatser. Uppsatsuppgiften baserar sig på en bildserie och stödfrågor om drömyrken och yrkesval (se bilaga 1). Informanterna ombads beskriva vad som händer i bildserien (en deskriptiv del), berätta om sina egna erfarenheter (en narrativ del) och diskutera ämnet mer allmänt (en argumenterande del). Inga hjälpmedel var tillgängliga men informanterna hade möjlighet att använda ord från en stödordlista i sin produktion. Informanterna skrev uppsatserna på en lektion och hade ca 40 minuter på sig att planera och skriva sina texter (100–150 ord som mål).

Uppsatserna bedömdes av två erfarna, oberoende bedömare enligt en nivåskala som baserar sig på CEFR-skalor (se bilaga 2; CEFR, 2007, s. 63–64, 107, 109). I bedömningskriterierna betonas inlärarens lexikala repertoar och förmågan att förmedla information, medan de kriterier som har med komplexitet och korrekthet att göra har utelämnats ur kriterierna. På detta sätt minimeras risken för cirkelresonemang. Det är dock viktigt att påpeka att komplexiteten och korrektheten i strukturerna kan ha påverkat bedömningarna.

På basis av bedömningarna utformades två grupper enligt färdighetsnivå, en lägre nivå (FN1, motsvarar CEFR-nivåerna A1 och A2) och en högre nivå (FN2, CEFR-nivåerna B1, B2 och en text på C1⁶). På den lägre nivån (FN1) ligger 17 informanter (Språkcenterstudenterna) medan 14 informanter (huvudämnesstudenterna) ligger på den högre nivån (FN2). Materialet sammanfattas i tabell 1.

Tabell 1. Materialets storlek (ord)

Färdighetsnivå	Textlängd (ord)			
	medeltal	min	max	std.av.
FN1 (n=17)	160,1	67	279	61,23
FN2 (n=14)	303,7	185	492	93,84
Totalt (n=31)				

Som framgår av tabell 1 är texterna på FN2 avsevärt längre än på FN1, vilket bör beaktas när man tolkar resultaten (se diskussionen i Golden et al., 2017). Eftersom syftet är att analysera inlärares faktiska kunskap i att producera svenska, tas hela produktionen med i analysen och längden beaktas inte.

Metod

Komplexitet har operationaliserats på olika sätt i tidigare studier (se t.ex. Bulté & Housen, 2018, s. 151; Norris & Ortega, 2009; Ortega, 2003; Pallotti, 2015). I denna studie undersöker jag strukturernas yttre uppbyggnad, och därför använder jag komplexitetsmått som granskar enheternas yttre struktur (längd, underordning, topikalisering).

Tabell 2. Komplexitetsmått

Typen av mått	Mått (förkortning*)
absolut komplexitet	
genomsnittlig längd på enheter	ord per T-enhet (W/T) ord per sats (W/C)
underordning	bisatser per sats (DC/C)
relativ komplexitet	
specifika mått	bisatser med satsadverbial per T-enhet (DC+/T) topikaliserade strukturer per påståendeformad T-enhet (TOP/T(P))

*engelska förkortningar används för segmenteringsenheterna: W=word, T=(minimally) terminable unit, C=clause, DC=dependent clause, TOP=topicalised

I komplexitetsstudier är det ofta nödvändigt att använda flera mått (Norris & Ortega, 2009). Komplexitetsmåten i denna studie kombinerar absolut (dvs. antalet element och sambanden mellan dessa element) och relativ komplexitet (dvs. komplexitet i förhållande till inlärares) (se De Clercq & Housen, 2017, s. 316–317). Syftet är inte att likställa dessa två typer av komplexitet (se diskussionen i Pallotti, 2015) utan att studera vilka skillnader i komplexitet som förekommer mellan två färdighetsnivåer mätt med olika typer av mått.

Även om längdbaserade mått tillhör de oftast rekommenderade komplexitetsmåten (se t.ex. Wolfe-Quintero et al., 1998) har användningen av dessa mått också kritiserats eftersom de inte visar enhetens inre struktur, dvs. vad som är komplext i enheten (se t.ex. Biber et al., 2016). I denna studie använder jag ändå längdbaserade mått eftersom det är strukturernas yttre komplexitet som undersöks, och tidigare forskning har visat att dessa mått effektivt skiljer mellan olika färdighetsnivåer.

På högre färdighetsnivåer börjar inlärare variera sina strukturer mer än på lägre nivåer (se t.ex. Kowal, 2016, s. 85–86), och därför kan exempelvis andelen underordnade satser minska i inlärares produktion (Lambert & Kormos, 2014). För att få fram komplexiteten på högre färdighetsnivåer bör man använda mått som fungerar på satsnivån (Biber et al., 2016, s. 664; Norris & Ortega, 2009). Därför analyseras längden på finita satser (inkl. både ett subjekt och ett finit verb, se Teleman et al., 2001) i denna studie. Därtill används två specifika komplexitetsmått, dvs. förekomsten av topikaliserade strukturer och bisatser som innehåller ett satsadverbial. Dessa mått har jag valt eftersom de granskar strukturer som enligt tidigare studier först används på högre färdighetsnivåer och är svåra för inlärare av svenska, och därmed mäter dessa mått den relativa komplexiteten i informanternas produktion.

I denna studie avgränsas korrekthetsaspekten till ordföljden. Korrekthetsmåten specificeras i tabell 3.

Tabell 3. Korrekthetsmått

Typen av mått	Mått (förkortning*)
allmänt mått	felfria satser av totala antalet satser (EFC/C)
specifika mått	felfria bisatser av totala antalet bisatser (EFDC/DC)
	felfria bisatser med satsadverbial av totala antalet bisatser med satsadverbial (EFDC+/DC+)
	inversion av totala antalet topikaliseringar (EFTOP/TOP)

*engelska förkortningar: EF=error-free, C=clause, DC=dependent clause, TOP=topicalised

Det allmänna måttet (felfria satser av totala antalet satser, EFC/C) ger en allmän bild av korrektheten i materialet. Därtill används specifika mått: andelen felfria bisatser (separat för alla bisatser och bisatser med satsadverbial) och förekomsten av inversion jämfört med det totala antalet topikaliserade strukturer.

Analysen av materialet har jag inlett med att segmentera hela materialet i T-enheter, satser och bisatser. Vidare har jag excerperat alla belägg på topikaliseringar och bisatser med satsadverbial samt räknat det totala antalet ord per text. Därefter har jag analyserat korrektheten av strukturerna och räknat andelen felfria enheter.

Analysen utförs på gruppnivå. Den deskriptiva statistiken innefattar medelvärdet och medianen, minimi- och maximivärdet samt standardavvikelsen och interkvartilavståndet för varje mått. Den statistiska analysen inledde jag med att testa om materialet är normalt fördelat.

Eftersom hela materialet inte uppfyller kraven för normalfördelning använder jag icke-parametriska test. Mann-Whitney U-testet används för att granska sambandet mellan färdighetsnivå och komplexitetsmått samt färdighetsnivå och korrekthetsmått. Effektstorlek (r , Larson-Hall, 2016, s. 475) och konfidensintervall för effektstorleken (CI, konfidensgraden 95 %, Larson-Hall, 2016, s. 149) rapporteras för varje mått. Några exempel på skillnaden mellan färdighetsnivåerna illustrerar jag med hjälp av låddiagram. Spearmans rangkorrelation (ρ) används för att granska relationen mellan de fem olika komplexitetsmåten och korrekthetsmättet felfria satser av det totala antalet satser (EFC/C) på de två olika färdighetsnivåerna. Andelen felfria satser av det totala antalet satser används som allmänt korrekthetsmått i denna del av analysen eftersom detta mått ger en allmän bild av korrektheten i informanternas produktion. Några exempel på skillnader i sambandet mellan komplexitet och korrekthet illustrerar jag med spridningsdiagram. Risken för falska positiva resultat kontrolleras med hjälp av False Discovery Rate (FDR⁷)-metoden (se Benjamini & Hochberg, 1995; Larson-Hall, 2016, s. 287, 476).

Mål

Syftet med denna studie är att undersöka hur syntaktisk komplexitet och korrekthet samt sambandet mellan dessa skiljer sig på två olika färdighetsnivåer i skriften L2-svenska. Syntaktisk komplexitet studeras på en yttre nivå och operationaliseras genom flera mått eftersom inlärare på olika färdighetsnivåer uppvisar syntaktisk komplexitet på olika språkliga nivåer (se Norris & Ortega, 2009; Lambert & Kormos, 2014; Kuiken & Vedder, 2019). Som framkommit mäts komplexitet med hjälp av längdbaserade mått, mått som är baserade på underordning och två specifika mått i den aktuella studien. Dessa mått innefattar längden på T-enheter, satslängden, andelen underordnade satser samt förekomsten av bisatser med satsadverbial och topikaliserade strukturer (se tabell 2). Granskningen av korrekthet avgränsas till ordföljden och mäts med andelen felfria satser av totala antalet satser, felfria bisatser av totala antalet bisatser, felfria bisatser med satsadverbial av totala antalet bisatser med satsadverbial och andelen inversion per totala antalet topikaliserade strukturer (se tabell 3). Artikeln syftar till att besvara följande forskningsfrågor (FF):

FF1) I vilken mån skiljer sig a) syntaktisk komplexitet, operationaliserad som absolut och relativ komplexitet och utvärderad enligt de fem olika komplexitetsmåten, och b) korrekthet, operationaliserad som syntaktisk målspråksenlighet och utvärderad enligt de fyra olika korrekthetsmåten, på två olika färdighetsnivåer?

FF2) Vilket samband föreligger mellan syntaktisk komplexitet och korrekthet på två olika färdighetsnivåer?

Utgående från tidigare studier (t.ex. Bulté & Housen, 2018; Kuiken & Vedder, 2019; Wolfe-Quintero et al., 1998) antar jag att produktionen på den högre färdighetsnivån är mer komplex än på den lägre nivån, vilket i fråga om denna studie betyder längre T-enheter, längre satser, (möjligtvis något) mer frekvent underordning (med och utan satsadverbial), och mer frekvent topikalisering (FF1a). Därtill antar jag att produktionen på den högre nivån är mer målspråksenlig än på den lägre nivån (FF1b). På grund av ett flertal påverkande faktorer är det

dock sannolikt att det finns variation mellan olika mått (se också Bulté & Housen, 2018). Exempelvis tyder tidigare forskning på att inlärare på högre färdighetsnivåer uppvisar syntaktisk komplexitet särskilt på satsnivå (mätt med satslängden i denna studie) och att andelen underordnade satser kan minska på högre nivåer på grund av ökande variation i de strukturer som inlärare producerar (se t.ex. Norris & Ortega, 2009; Lambert & Kormos, 2014). Jag kan därmed anta att det finns skillnader mellan färdighetsnivåerna speciellt i fråga om satslängden medan skillnaden i andelen underordning kan antas vara relativt liten. Vidare antar jag att relationen mellan komplexitet och korrekthet på den lägre nivån är konkurrerande, medan produktionen på den högre nivån kan visa tecken på både en konkurrerande och en stödjande relation (FF2, se t.ex. Spoelman & Verspoor, 2010). Därmed antar jag att korrelationskoefficienterna på FN1 visar på negativa samband mellan komplexitetsmåten och det allmänna korrekthetsmättet, medan sambanden på FN2 kan vara både negativa och positiva. Igen är det troligt att det finns skillnader i sambandet mellan komplexitet och korrekthet mätt med olika mått.

Resultat

I detta avsnitt presenterar jag först resultaten för relationen mellan färdighetsnivå och komplexitet (FF1a). Efter detta granskar jag relationen mellan färdighetsnivå och korrekthet (FF1b) och till sist sambandet mellan komplexitet och korrekthet på de två olika färdighetsnivåerna (FF2).

Relationen mellan färdighetsnivå och komplexitet

Bland informanterna i urvalet visar produktionen på något högre komplexitet på den högre färdighetsnivån (FN2) än på den lägre nivån (FN1) (se tabell 4).

Tabell 4. Deskriptiv statistik över komplexitet

		mv	median	min	max	std.av.	IQR
W/T	FN1	9,22	9,20	6,49	12,36	1,60	2,32
	FN2	11,06	10,97	8,30	13,77	1,73	2,66
W/C	FN1	6,09	5,96	4,68	7,89	0,83	1,15
	FN2	6,86	6,54	5,59	8,62	0,99	1,70
DC/C	FN1	0,32	0,30	0,06	0,55	0,15	0,26
	FN2	0,36	0,39	0,21	0,48	0,08	0,12
DC+/T	FN1	0,08	0,05	0,00	0,29	0,09	0,12
	FN2	0,11	0,07	0,00	0,44	0,12	0,14
TOP/T(P)	FN1	0,41	0,40	0,11	0,86	0,21	0,37
	FN2	0,44	0,45	0,25	0,69	0,12	0,19

W/T=ord per T-enhet, W/C=ord per sats, DC/C=bisatser per sats, DC+/T=bisatser med satsadverbial per T-enhet, TOP/T(P)=topikaliserade strukturer per påståendeformade T-enhet

Som framgår av tabell 4 är produktionen på FN2 mer komplex än på FN1 särskilt i fråga om längden på T-enheterna (W/T) och satslängden (W/C). Tabellen visar också att skillnaden inte är så stor beträffande andelen bisatser (DC/C) och topikaliserade strukturer (TOP/T(P)). Vad

gäller bisatser med satsadverbial är det viktigt att påpeka att det är fråga om ett begränsat antal belägg (21 belägg på FN1 och 38 på FN2). På FN1 finns det sex informanter som inte alls har producerat bisatser med satsadverbial, och på FN2 är det motsvarande antalet informanter två.

Skillnaden i komplexitet mellan färdighetsnivåerna testades med Mann-Whitney U-testet. Resultaten presenteras i tabell 5.

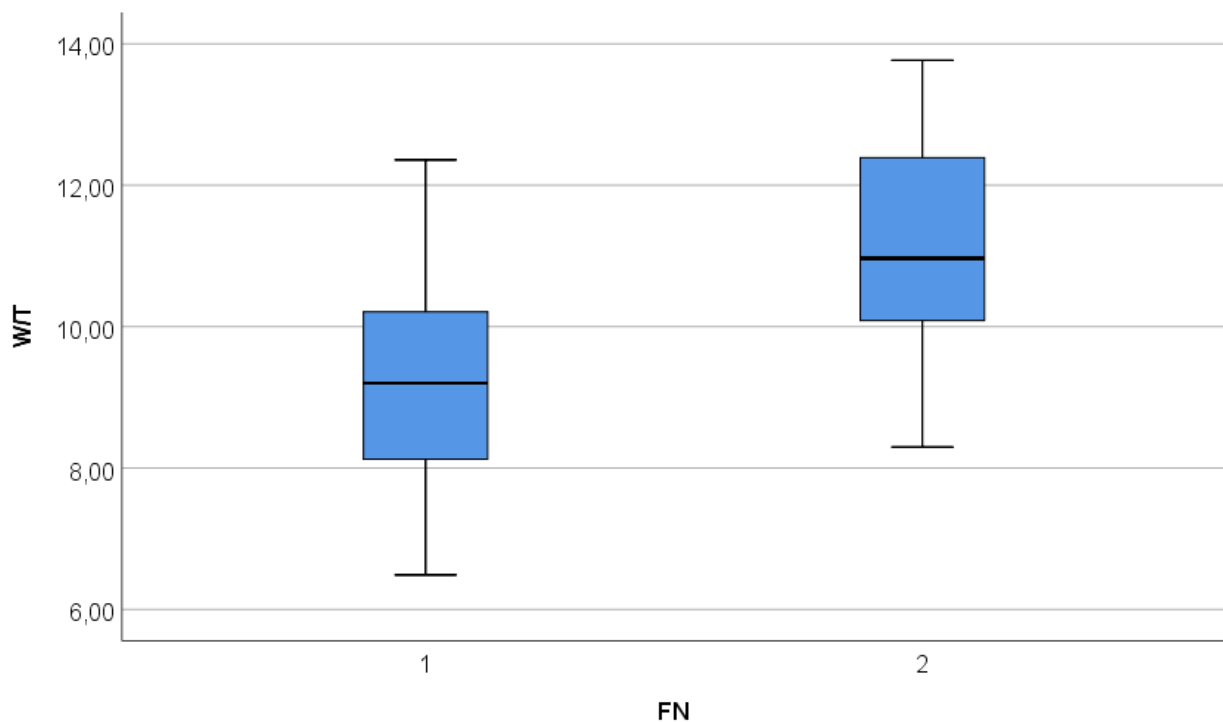
Tabell 5. Relationen mellan färdighetsnivå och komplexitetsmått

	W/T	W/C	DC/C	DC+/T	TOP/T(P)
p	0,005*	0,035	0,451	0,447	0,592
r	-0,60	-0,45	-0,16	-0,16	-0,11
CI	(-0,87, -0,26)	(-0,78, -0,05)	(-0,53, 0,21)	(-0,56, 0,27)	(-0,51, 0,24)

* statistiskt signifikant efter FDR-korrigerig

$$r = z/\sqrt{N}$$

Tabell 5 visar att skillnaden mellan färdighetsnivåerna är statistiskt signifikant beträffande längden på T-enheterna ($p = 0,005$) och att detta resultat i viss utsträckning är generaliserbart ($r = -0,60$, $CI = (-0,87, -0,26)$). Längden på T-enheterna på de två olika färdighetsnivåerna illustreras i figur 1.



Figur 1. Den genomsnittliga längden på T-enheter (W/T) på den lägre (FN1) och den högre (FN2) färdighetsnivån.

Figur 1 visar att skillnaden mellan färdighetsnivåerna är tydlig i fråga om längden på T-enheterna. T-enheterna på FN1 är genomsnittligt något under två ord kortare än på FN2.

De övriga måtten visar inga statistiskt signifikanta skillnader mellan färdighetsnivåerna (se tabell 5). Exempelvis är skillnaden mellan grupperna i fråga om andelen bisatser (DC/C) och topikaliserade strukturer (TOP/T(P)) liten (se tabell 4 och 5). En förklaring till detta kan vara att informanterna på den högre nivån producerar syntaktisk komplexitet på en annan lingvistisk nivå, exempelvis på satsnivå. En ytterligare möjlig orsak till avsaknaden av signifikanta skillnader mellan färdighetsnivåerna är att informanterna också på den lägre färdighetsnivån producerar relativt komplexa strukturer vad gäller användningen av bisatser och topikaliserade strukturer. I själva verket tillhör den informant som har använt mest underordning (0,55 DC/C) och topikalisering (0,86 TOP/T(P)) den lägre nivån. Detta illustreras i följande utdrag från produktionen av informant FN1.09 (exempel 1, topikaliserade strukturer understrukna).

- 1) När Sebastian var liten [DC C] drömde han [C] att han ska bli en astronaut. [DC C T] När han blir äldre [DC C] planerade han [C] att han skulle bli en berömd musiker. [DC C T] I gymnasiet drömde han [C] att han vill bli en läkare. [DC C T] På vuxen hans barnen hade samma problemar [C] som han hade [DC C] när han var liten. [DC C T] (FN1.09)

I detta utdrag har informant FN1.09 använt en topikaliserad struktur i varje T-enhet och varje T-enhet innehåller åtminstone en bisats.

Däremot kan relativt enkla strukturer förekomma i texter skrivna av informanter på FN2 (exempel 2, topikaliserade strukturer understrukna).

- 2) Barnet sitter i en låda och leker. [C T] Han vill bli astronaut. [C T] Han drömmer om en rymdraket. [C T] Några år senare vill han inte längre bli astronaut. [C T] Han har blivit äldre [C T] och nu är det musik [C] som intresserar honom. [DC C T] Efter ett par år börjar han fundera mer seriöst. [C T] Att bli läkare är hans nya dröm. [C T] (FN2.04)

I detta utdrag har informant FN2.04 använt relativt enkelt språk enligt definitionen i denna studie. Strukturerna är korta och utdraget innehåller endast en bisats.

Relationen mellan färdighetsnivå och korrekthet

Härnäst granskas relationen mellan korrekthet och färdighetsnivå. Korrekthetsaspekten avgränsas till syntaktisk korrekthet, dvs. att endast fel i ordföljden beaktas i analysen. I tabell 6 presenteras deskriptiv statistik över korrekthetsmått.

Tabell 6. Deskriptiv statistik över korrekthet

		mv	median	min	max	std.av.	IQR
EFC/C	FN1	0,73	0,74	0,56	1,00	0,12	0,17
	FN2	0,90	0,91	0,78	0,97	0,05	0,07
EFDC/DC	FN1	0,78	0,83	0,00	1,00	0,25	0,28
	FN2	0,83	0,86	0,69	0,94	0,09	0,16
EFDC+/DC+	FN1	0,03	0,00	0,00	0,33	0,10	0,00
	FN2	0,44	0,45	0,00	1,00	0,41	0,85
EFTOP/TOP	FN1	0,29	0,17	0,00	1,00	0,37	0,59
	FN2	0,89	0,91	0,63	1,00	0,12	0,15

EFC/C=felfria** satser av totala antalet satser, EFDC/DC=felfria** bisatser av totala antalet bisatser, EFDC+/DC+=felfria** bisatser med satsadverbial av totala antalet bisatser med satsadverbial, EFTOP/TOP=inversion av totala antalet topikaliseringar

**endast fel i ordföljd beaktas

Tabell 6 visar att informanterna på den högre färdighetsnivån (FN2) producerar fler målspråksenliga ordföljder än informanterna på den lägre nivån (FN1). Skillnaden är dock mycket liten i andelen felfria bisatser (EFDC/DC). Som framkommit är det viktigt att notera att bisatser med satsadverbial är lågfrekventa i materialet (21 på FN1 och 38 på FN2) och att det finns informanter som inte alls har producerat belägg på dessa strukturer (6 informanter på FN1 och 2 på FN2).

Skillnaden i korrektheten mellan färdighetsnivåerna testades med Mann-Whitney U-testet. Resultaten presenteras i tabell 7.

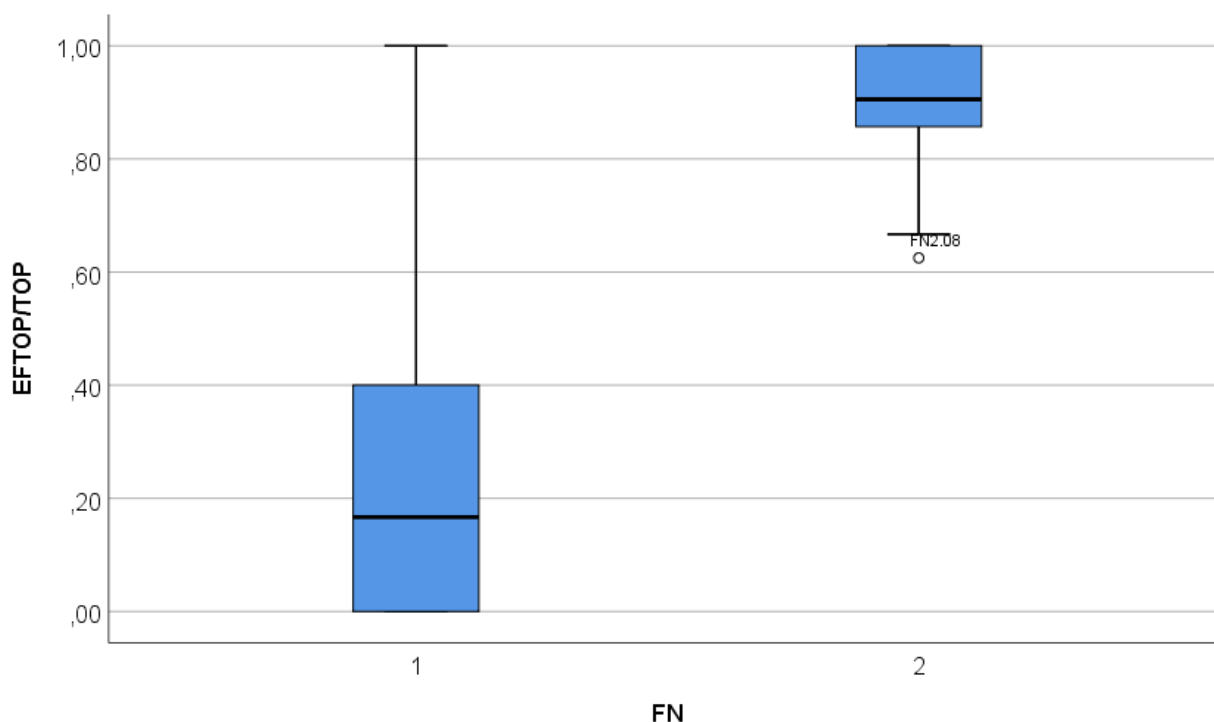
Tabell 7. Relationen mellan färdighetsnivå och korrekthetsmått

	EFC/C	EFDC/DC	EFDC+/DC+	EFTOP/TOP
p	0,000*	0,905	0,001*	0,000*
r	-0,80	-0,03	-0,59	-0,77
CI	(-1,00, -0,47)	(-0,40, 0,35)	(-0,84, -0,31)	(-0,98, -0,51)

* statistiskt signifikant efter FDR-korrigerig

$r = z/\sqrt{N}$

Som framgår av tabell 7 är skillnaden mellan färdighetsnivåerna statistiskt signifikant vad gäller andelen felfria satser (EFC/C, $p = 0,000$), andelen felfria bisatser med satsadverbial (EFDC+/DC+, $p = 0,001$) och andelen inversion av det totala antalet topikaliserade strukturer (EFTOP/TOP, $p = 0,000$). Effektstorleken av dessa resultat varierar från en måttlig (EFDC+/DC+, $r = -0,59$) till en relativt stor (EFC/C, $r = -0,80$ och EFTOP/TOP, $r = -0,77$) effekt. Särskilt i samband med andelen felfria satser och felfri inversion är resultaten generaliserbara. Skillnaden i andelen inversion av det totala antalet topikaliserade strukturer illustreras i figur 2.



Figur 2. Andelen inversion av det totala antalet topikaliserade strukturer (EFTOP/TOP) på den lägre (FN1) och den högre (FN2) färdighetsnivån.

Som framgår av figuren är skillnaden mellan färdighetsnivåerna tydlig i fråga om andelen inversion av det totala antalet topikaliserade strukturer. På FN2 producerar informanterna betydligt fler topikaliserade strukturer med inversion än informanterna på FN1.

Relationen mellan komplexitet och korrekthet

Härnäst presenteras vilket samband som föreligger mellan komplexitet och korrekthet på de två olika färdighetsnivåerna. Korrekthetsmättet felfria (dvs. syntaktiskt målspråksenliga) satser av det totala antalet satser (EFC/C) används som allmänt korrekthetsmått i denna del av analysen.

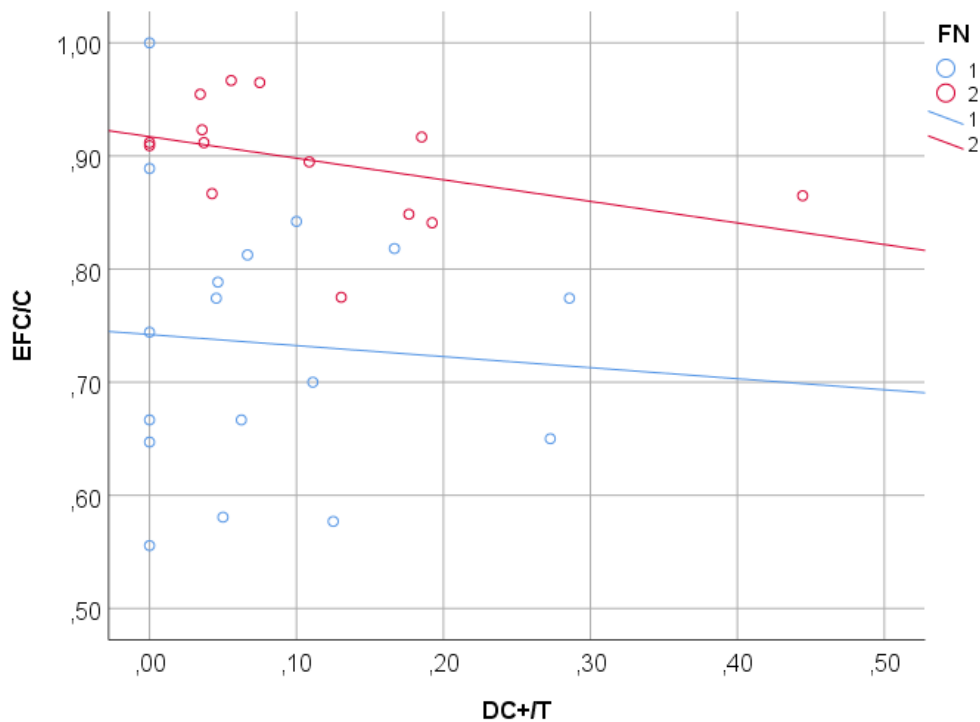
Tabell 8. Relationen mellan komplexitet och korrekthet på den lägre och på den högre färdighetsnivån

	W/T EFC/C		W/C EFC/C		DC/C EFC/C		DC+/T EFC/C		TOP/T(P) EFC/C	
	FN1	FN2	FN1	FN2	FN1	FN2	FN1	FN2	FN1	FN2
**rho	-0,12	-0,26	-0,29	0,06	0,12	-0,31	-0,01	-0,49	-0,51	-0,05
p	0,642	0,370	0,251	0,846	0,646	0,283	0,964	0,077	0,038	0,872
CI	(-0,63, 0,39)	(-0,98, 0,47)	(-0,88, 0,29)	(-0,65, 0,78)	(-0,45, 0,69)	(-0,83, 0,23)	(-0,62, 0,60)	(-0,95, -0,00)	(-0,81, -0,18)	(-0,79, 0,70)

** Spearmans rangkorrelation

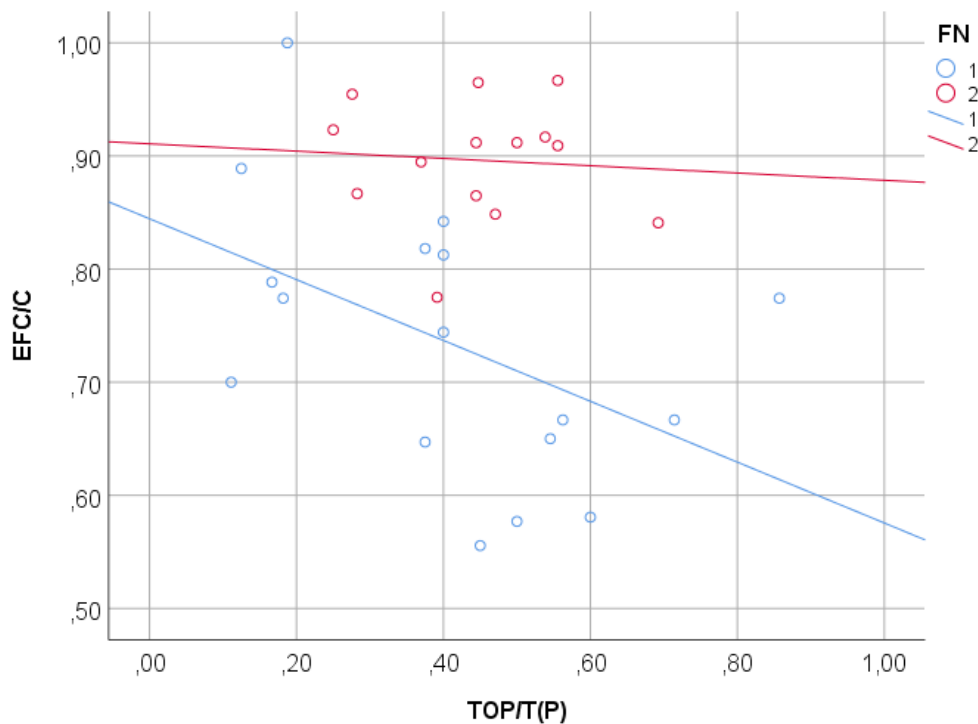
Som framgår av tabell 8 visar granskningen av sambandet mellan komplexitet och korrekthet inga eller mycket svaga korrelationer för största delen av måtten. Det finns inga statistiskt

signifikanta korrelationer efter FDR-korrigering. Resultaten visar dock vissa tendenser i detta urval vad gäller relationen mellan komplexitet och korrekthet. Dessa tendenser illustreras i figur 3 och 4.



Figur 3. Relationen mellan måttet andelen bisatser med satsadverbial per T-enhet (DC+/T) och det allmänna korrekthetsmättet (EFC/C) på den lägre (FN1=blå) och den högre (FN2=röd) färdighetsnivån.

Som framgår av figur 3 finns det i urvalet en tendens till trade-off mellan måttet bisatser med satsadverbial per T-enhet (DC+/T) och korrekthetsmättet (EFC/C) på FN2 (röd linje i figuren). Figuren visar att en större andel bisatser med satsadverbial kan leda till flera målspråksavvikande ordföljder på denna nivå, vilket syns som lägre värden på korrekthetsmättet.



Figur 4. Relationen mellan måttet andelen topikaliserade strukturer (TOP/T(P)) och det allmänna korrekthetsmåttet (EFC/C) på den lägre (FN1=blå) och den högre (FN2=röd) färdighetsnivån.

Figur 4 visar att det i urvalet finns tendenser till trade-off mellan andelen topikaliserade strukturer (TOP/T(P)) och korrekthetsmåttet (EFC/C) på FN1 (blå linje i figuren). Användningen av topikalisering verkar leda till en lägre andel målspråksenliga ordföljder på denna nivå.

Sammanfattningsvis kan det konstateras att resultaten i denna studie inte visar något stöd för vare sig konkurrerande eller stödande samband mellan komplexitet och korrekthet på gruppnivå (med eventuellt undantag för en konkurrerande relation mellan bisatser med satsadverbial och korrekthet på FN2 samt mellan topikaliserade strukturer och korrekthet på FN1).

Sammanfattande diskussion

Syftet med denna studie har varit att redogöra för sambandet mellan färdighetsnivå och syntaktisk komplexitet, mellan färdighetsnivå och syntaktisk korrekthet samt mellan komplexitet och korrekthet i skriftlig L2-svenska. Tyngdpunkten i studien ligger på strukturernas yttre uppbyggnad.

Enligt resultaten uppvisar informanterna i urvalet något högre komplexitet på den högre färdighetsnivån (FN2) än på den lägre nivån (FN1) (FF1a). Skillnaden är dock statistiskt signifikant endast i fråga om längden på T-enheterna ($p = 0,005$, $r = -0,60$, CI [-0,87; -0,26]).

Även satserna är längre på FN2 än på FN1, och detta är i linje med antagandet om att det finns skillnader i komplexiteten på satsnivå mellan färdighetsnivåerna. Skillnaden är ändå inte statistiskt signifikant efter FDR-korrigerig.

Skillnaden mellan färdighetsnivåerna är relativt liten exempelvis i samband med andelen bisatser (se också Wijers, 2018; men jfr Kowal, 2016). Detta kan bero på att informanterna på FN2 varierar sina strukturer mer och producerar komplexa strukturer på en annan språklig nivå (t.ex. på frasnivå) än informanterna på FN1 (t.ex. Lambert & Kormos, 2014). Resultatet bekräftar därmed antagandet om att skillnaden i förekomsten av bisatser är relativt liten mellan färdighetsnivåerna i denna studie. Därtill kan uppgiftstypen ha påverkat bisatsförekomsten, exempelvis så att den narrativa och den argumenterande delen kan ha medfört frekvent användning av bisatser. Utöver dessa orsaker kan avsaknaden av signifikanta skillnader mellan färdighetsnivåerna bero på den relativt höga komplexiteten på FN1. Detta avviker från resultaten exempelvis i Gyllstad et al. (2014, se också diskussionen i Gunnarsson, 2012). Informanterna i den aktuella studien kan uttrycka sig på ett komplext sätt på sitt modersmål och är bekanta med komplexa strukturer på flera språk (se *relative cognitive maturity*, Foster et al., 2000, s. 355). Detta kan leda till att somliga informanter på FN1 (t.ex. FN1.22) kan producera komplexa strukturer på svenska även om deras kommunikativa kompetens och ordförråd representerar lägre språkfärdighetsnivåer (jfr kriterierna i CEFR, 2007, s. 63). Exempelvis kan några informanter på den lägre nivån visa tecken på komplext språkbruk vad gäller användningen av topikaliserade strukturer (jfr t.ex. Åberg, 2020). Kännedom om språk och strukturer i allmänhet kan påverka komplexiteten på strukturernas yttre nivå i hög grad, till och med mer än själva färdigheten i målspråket (jfr t.ex. Gyllstad et al., 2014), och här kan även informanternas L1 påverka då finskan har relativt fri ordföljd (se t.ex. Åberg, 2020, s. 33, 34, 238).

Vidare visar resultaten att bisatser med satsadverbial är relativt lågfrekventa även på den högre färdighetsnivån. Bisatser med satsadverbial är överhuvudtaget sällsynta i svenska (Rahkonen & Håkansson, 2008, s. 153), och detta kan påverka förekomsten av dessa strukturer i inlärares produktion (om relationen mellan inflöde och utflöde, se t.ex. Ellis, 2002).

I motsats till Gyllstad et al. (2014) visar resultaten här att komplexiteten inte nödvändigtvis korrelerar med färdighetsnivån utan kan påverkas av flera andra faktorer (se också Kuiken & Vedder, 2019, s. 207). Detta förstärker uppfattningen om komplexitet som en komplex och mångdimensionell företeelse. Exempelvis tyder resultaten här på att det individuella sättet att producera språk kan påverka komplexiteten i produktionen. Å ena sidan finns det informanter på båda färdighetsnivåerna som producerar relativt enkla strukturer. Exempelvis använder informant FN2.04 enkelt språk jämfört med andra informanter på FN2. Å andra sidan finns det informanter på båda nivåerna som föredrar komplext språk. En del informanter på FN1 (t.ex. FN1.09) producerar mycket komplexa strukturer. Intressant är att de mest komplexa strukturerna i fråga om underordning och topikalisering förekommer i texter som tillhör den lägre nivån. Den avgörande faktorn verkar vara det individuella sättet att producera språk (se också Norrby & Håkansson, 2007; Kowal, 2016). Färdighetsnivån spelar då inte så stor roll utan några inlärare är mer riskvilliga med sin produktion (risktagare i Norrby & Håkansson, 2007) medan andra nöjer sig med enklare strukturer (återanvändare i Norrby & Håkansson,

2007). I framtida studier är det därför viktigt att studera komplexitet på individnivå. På grund av den mångdimensionella karaktären och de faktorer som påverkar komplexiteten (t.ex. L1, målspråk, texttyp, individuell variation och eventuell trade-off mellan CAF-komponenterna) verkar det vara utmanande att entydigt beskriva komplexiteten på en viss färdighetsnivå (se också diskussionen i Bulté & Housen, 2018), men dessa studier på gruppnivå erbjuder värdefull information om allmänna tendenser i inlärares produktion och lägger grund till studier på individnivå. Resultaten från studier på gruppnivå kan användas som jämförelsedata i analyser av individers produktion (se också Kowal, 2016, s. 208).

Produktionen på FN2 är mer målspråksenlig än på FN1 (FF1b). Exempelvis har informanterna på FN2 producerat en större andel topikaliserade strukturer med inversion ($p = 0,000$), men även på denna nivå producerar ett antal informanter målspråksavvikande strukturer. Svårigheterna med inversion efter topikalisering är väldokumenterade (t.ex. Håkansson & Nettelbladt, 1993; Åberg, 2020; Rahkonen & Håkansson, 2008, se också Jensen, 2018), och denna studie ger ytterligare belägg på dessa svårigheter i automatiseringen av den målspråksenliga ordföljden. Skillnaden mellan färdighetsnivåerna är statistiskt signifikant också i fråga om andelen felfria (dvs. syntaktiskt målspråksenliga) satser av det totala antalet satser ($p = 0,000$) och andelen felfria bisatser med satsadverbial av det totala antalet bisatser med satsadverbial ($p = 0,001$). Särskilt i samband med felfria satser ($r = -0,80$, $CI = (-1,00, -0,47)$) och andelen inversion ($r = -0,77$, $CI = (-0,98; -0,51)$) är resultaten generaliserbara.

På gruppnivå (dvs. på varken den lägre eller den högre färdighetsnivån) visar resultaten inget stöd för vare sig en konkurrerande eller en stödjande relation mellan komplexitet och korrekthet (FF2). Detta avviker från antagandet om att relationen mellan komplexitet och korrekthet är mer konkurrerande på den lägre än på den högre nivån. Det finns dock tendenser i urvalet till en något starkare trade-off-effekt mellan vissa aspekter på olika färdighetsnivåer. Resultaten tyder på att det finns tendenser i urvalet på FN2 till trade-off mellan bisatser med satsadverbial och korrekthet ($\rho = -0,49$, $p = 0,077$, $CI = (-0,95, -0,00)$). Däremot finns det tendenser i urvalet på FN1 till trade-off mellan topikaliserade strukturer och korrekthet ($\rho = -0,51$, $p = 0,038$, $CI = (-0,81, -0,18)$) på FN1. Att resultaten här inte ger något stöd för antingen Limited Attentional Capacity-modellen (Skehan, 1998) eller Multiple Resources Attentional-modellen (Robinson, 2001) utesluter dock inte möjligheten till en konkurrerande eller en stödjande relation mellan andra språkliga nivåer inom komplexitet och korrekthet (se Housen et al., 2012, s. 5), varför det behövs mer forskning om relationen mellan dessa komponenter på olika språkliga nivåer.

I tidigare studier (se t.ex. Kowal, 2016, s. 39–40, 204–205, 216) har det framkommit att inlärarespråket kan utvecklas på olika sätt beroende av inlärares ålder, utgångsnivå och erfarenhet av språk. Inlärare kan antingen börja med enkla strukturer som först blir målspråksenliga och efter detta mer komplexa (se t.ex. CEFRL, 2007, s. 111), eller inlärare kan redan ha en bred repertoar av strukturer som så småningom blir mer målspråksenliga (se t.ex. Housen et al., 2012, s. 7). Det förstnämnda fallet verkar gälla yngre inlärare och nybörjare, medan det senare fallet kan vara vanligare med äldre eller erfarna inlärare. Yngre inlärare brukar ha en snävare repertoar av strukturer medan vuxna, såsom informanterna i denna studie, redan kan uttrycka sig på ett komplext sätt på ett eller flera språk (se Norrby & Håkansson,

2007, s. 45). Därför kan språket på lägre färdighetsnivåer hos yngre inlärare vara enklare än språket på lägre färdighetsnivåer hos äldre inlärare (se Wolfe-Quintero et al., 1998, s. 98). Å andra sidan kan särskilt vuxna inlärare kontrollera sin produktion och prioritera korrekthet på komplexitetens bekostnad (se t.ex. Norrby & Håkansson, 2007, s. 45). Dessa skillnader i prioriteringen av komplexitet eller korrekthet kan ha påverkat resultaten i föreliggande studie. Som framkommit uppvisar vissa informanter på den lägre nivån relativt hög komplexitet i sin produktion, vilket minskar skillnaden mellan färdighetsnivåerna. Det är också möjligt att en del av informanterna på den högre nivån har prioriterat korrekthet och (medvetet eller omedvetet) producerat mindre komplexa strukturer, vilket igen kan minska skillnaden mellan grupperna (se också Kowal, 2016, s. 139).

Både CAF-hypoteserna (Wolfe et al., 1998) och CEFR (2007) presenterar en förenklad syn på exempelvis sambandet mellan komplexitet och färdighetsnivå. I samband med komplexitet kan kontexten påverka de egenskaper som förväntas av texten (jfr t.ex. ett e-postmeddelande och en akademisk text). Därtill iakttar de förenklade CAF-hypoteserna inte den eventuella trade-off-effekten mellan komponenterna. Enligt det syfte som presenteras i CEFR (2007, s. 1) är det möjligt att bedöma språkfärdigheten hos inlärare oberoende av mål språk och utbildningsnivå med hjälp av CEFR-nivåskalorna. Gyllstad et al. (2014) har emellertid visat att komplexiteten på B-nivån skiljer sig mellan olika mål språk (se också Bernardini & Granfeldt, 2019), och som diskuterats, kan komplexiteten påverkas av informanternas ålder, kännedom om språk och strukturer i allmänhet samt individuella sätt att producera språk. Därmed behövs det mer empiriska studier som iakttar denna mångfald av påverkande faktorer (se också Golden et al., 2017, s. 234).

Resultaten måste tolkas med försiktighet på grund av ett antal begränsande faktorer. För det första är informantgrupperna relativt små och de producerade texterna relativt korta (se diskussionen i Golden et al., 2017, s. 239–240). Med större material skulle man kunna få mer generaliserbara resultat. För det andra är texterna på FN2 längre än på FN1, och de olika delarna inom uppsatserna varierar i längd (se t.ex. Biber et al., 2016, s. 644; Bulté & Housen, 2018, s. 150; Golden et al., 2017, s. 237). Den uppgift som användes i materialinsamlingen för denna studie består av tre delar som motsvarar olika texttyper. Denna typ av uppgift erbjuder inlärare en möjlighet att använda en bred repertoar av syntaktiska strukturer, men tyvärr försvårar de olika delarna i uppgiften tolkningen av resultaten. Exempelvis kan argumenterande texter innehålla flera bisatser på grund av frekvent motivering. Den argumenterande delen i materialet är ofta kortare på den lägre nivån, och i några texter fattas denna del helt och hållet.

En ytterligare begränsande faktor är operationaliseringen av både komplexitet och korrekthet. Valet av mått påverkar resultaten, särskilt när man jämför komplexitet på olika färdighetsnivåer (se t.ex. Norris & Ortega, 2009; Gyllstad et al., 2014, s. 23). I fortsättningen är det därför viktigt att fördjupa analysen och undersöka exempelvis de satstyper som inlärare använder. Vidare är relationen mellan olika komplexitets- och korrekthetsmått i denna studie inte helt oproblematiserad på grund av de obligatoriska kontexter som användningen av satsadverbial i bisatser och topikaliserade strukturer (jfr t.ex. längdmåtten) skapar i produktionen.

I språkundervisning och -bedömning bör relationen mellan olika komponenter i inlärarespråket beaktas (se diskussionen i Kuiken & Vedder, 2012, s. 165). Som resultaten i denna studie tyder på är relationen inte alltid entydig. Det viktigaste är att utvecklingen inte likställs med felfrihet, eftersom olika komponenter i inlärarespråket utvecklas samtidigt, och ibland kan utvecklingen i en av komponenterna påverka övriga komponenter negativt (trade-off-effekt). I olika texter krävs det olika slags strukturer och språkinlärare borde uppmuntras att uttrycka sig på ett varierande sätt istället för endast försöka undvika fel.

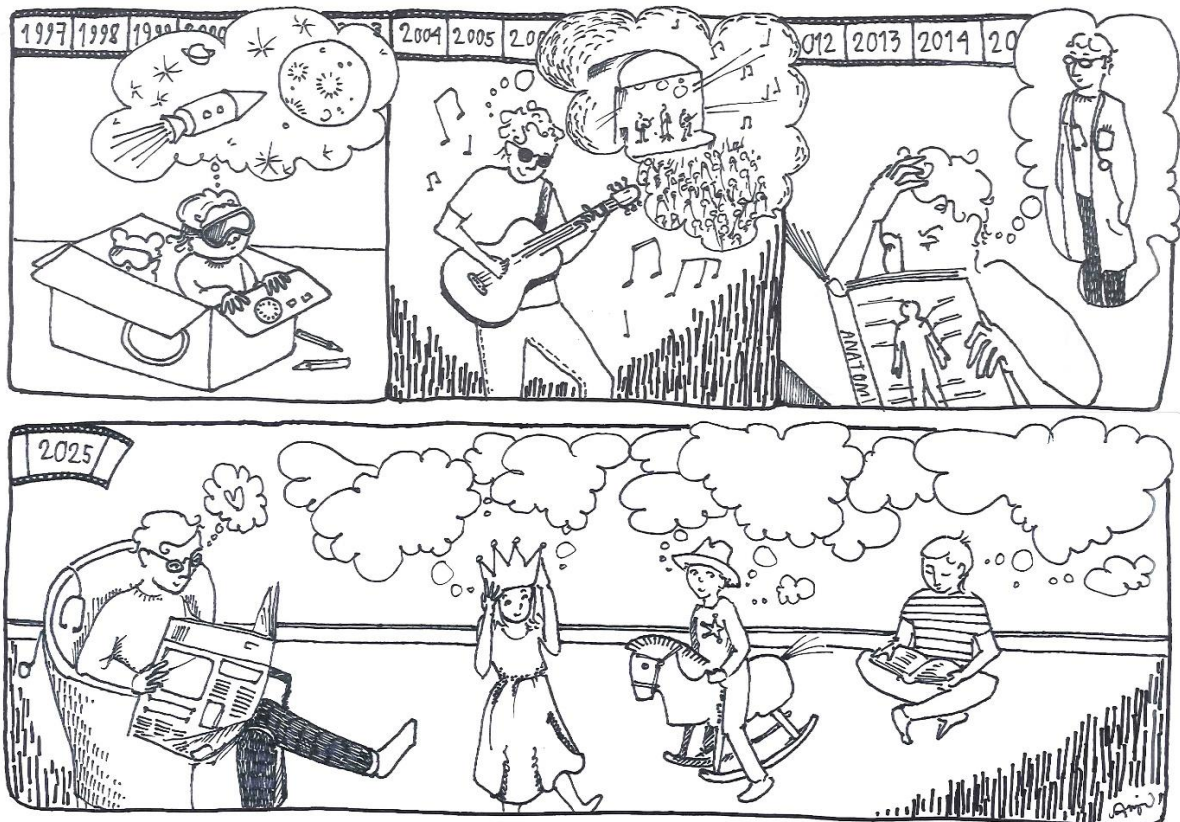
Litteratur

- Benjamini, Y. & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society*, 57(1), 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Berggreen, H. & Sørland, K. (2016). Syntaktisk kompleksitet i et skriftlig innlærerspråkmateriale. *NOA: norsk som andrespråk*, 32(1–2), 31–75.
- Bernardini, P. & Granfeldt, J. (2019). On Cross-Linguistic Variation and Measures of Linguistic Complexity in Learner Texts: Italian, French and English. *International Journal of Applied Linguistics*, 29, 211–232. <https://doi.org/10.1111/ijal.12257>
- Biber D., Gray B. & Staples S. (2016). Predicting Patterns of Grammatical Complexity Across Language Exam Task Types and Proficiency Levels. *Applied Linguistics*, 37(5), 639–668. <https://doi.org/10.1093/applin/amu059>
- Bohnacker, U. & Lindgren, J. (2014). Fundament, formellt subjekt och frekvens: Ordföljdsmonster i svenska, nederländska och hos vuxna inlärare av svenska. *Språk och stil*, 24, 33–71.
- Bulté, B. & Housen, A. (2018). Syntactic complexity in L2 writing: Individual pathways and emerging group trends. *International Journal of Applied Linguistics*, 28(1), 147–164. <https://doi.org/10.1111/ijal.12196>
- CEFR (2007). *Gemensam europeisk referensram för språk: lärande, undervisning och bedömning*. Stockholm: Skolverket.
- De Clercq, B. & Housen, A. (2017). A Cross-Linguistic Perspective on Syntactic Complexity in L2 Development: Syntactic Elaboration and Diversity. *The Modern Language Journal*, 101, 315–334. <https://doi.org/10.1111/modl.12396>
- Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, 24(2), 143–188. <https://doi.org/10.1017/S0272263102002024>
- Ellis, N. C. (2019). Essentials of a Theory of Language Cognition. *The Modern Language Journal*, 103, 39–60. <https://doi.org/10.1111/modl.12532>
- Golden, A., Kulbrandstad, L. A. & Tenfjord, K. (2017). Evaluation of texts in tests, or: Where is the dog buried? I A. Golden, S. Jarvis & K. Tenfjord (Red.), *Crosslinguistic Influence and Distinctive Patterns of Language Learning: Findings and Insights from a Learner Corpus* (s. 231–271). Bristol: Multilingual Matters. <https://doi.org/10.21832/9781783098774-011>

- Granfeldt, J. (2007). Speaking and Writing in L2 French: Exploring Effects on Fluency, Complexity and Accuracy. I S. van Daele, A. Housen, F. Kuiken, M. Pierrard & I. Vedder, (Red.), *Complexity, Accuracy and Fluency in Second Language Use, Learning & Teaching* (s. 87–98).
- Gunnarsson, C. (2012). The development of complexity, accuracy and fluency in the written production of L2 French. I A. Housen, F. Kuiken & I. Vedder (Red.), *Dimensions of L2 Performance and Proficiency. Complexity, Accuracy and Fluency in SLA* (s. 247–276). Amsterdam: John Benjamins. <https://doi.org/10.1075/llt.32.11gun>
- Gyllstad, H., Granfeldt, J., Bernardini, P. & Källkvist, M. (2014). Linguistic correlates to communicative proficiency levels of the CEFR: The case of syntactic complexity in written L2 English, L3 French and L4 Italian. *EUROSLA Yearbook*, 14, 1–30. <https://doi.org/10.1075/eurosla.14.01gyl>
- Håkansson, G. & Nettelbladt, U. (1993). Developmental sequences in L1 (Normal and Impaired) and L2 acquisition of Swedish syntax. *International Journal of Applied Linguistics*, 3(2), 131–157. <https://doi.org/10.1111/j.1473-4192.1993.tb00047.x>
- Housen, A., Kuiken, F. & Vedder, I. (2012). Complexity, accuracy and fluency: Definitions, measurement and research. I A. Housen, F. Kuiken & I. Vedder (Red.), *Dimensions of L2 Performance and Proficiency. Complexity, Accuracy and Fluency in SLA* (s.1–20). Amsterdam: John Benjamins. <https://doi.org/10.1075/llt.32.01hou>
- Hunt, K. (1965). *Grammatical Structures Written at Three Grade Levels*. Champaign, IL: National Council of Teachers of English.
- Jensen, B. U. (2018). Syntaks i norsk innlærerspråk: empiriske funn. I A.-K. H. Gujord & G. T. Randen (Red.), *Norsk som andrespråk – perspektiver på læring og utvikling* (s. 235–260). Oslo: Cappelen Damm Akademisk.
- Kormos, J. & Trebits, A. (2012). The Role of Task Complexity, Modality, and Aptitude in Narrative Task Performance. *Language Learning*, 62(2), 439–472. <https://doi.org/10.1111/j.1467-9922.2012.00695.x>
- Kowal, I. (2016). *The Dynamics of Complexity, Accuracy and Fluency in Second Language Development*. Kraków: Jagiellonian University Press. <https://doi.org/10.4467/k9474.72/16.16.6986>
- Kuiken, F. & Vedder, I. (2012). Syntactic Complexity, Lexical Variation and Accuracy as a Function of Task Complexity and Proficiency Level in L2 Writing and Speaking. I A. Housen, F. Kuiken & I. Vedder (Red.), *Dimensions of L2 Performance and Proficiency. Complexity, Accuracy and Fluency in SLA* (s. 143–170). Amsterdam: John Benjamins. <https://doi.org/10.1075/llt.32.07kui>
- Kuiken, F. & Vedder, I. (2019). Syntactic Complexity across Proficiency and Languages: L2 and L1 writing in Dutch, Italian and Spanish. *International Journal of Applied Linguistics*, 29, 192–210. <https://doi.org/10.1111/ijal.12256>
- Kuiken, F., Vedder, I., Housen, A. & De Clercq, B. (2019). Variation in Syntactic Complexity: Introduction. *International Journal of Applied Linguistics*, 29, 161–170. <https://doi.org/10.1111/ijal.12255>
- Lambert, G. & Kormos, J. (2014). Complexity, Accuracy, and Fluency in Task-based L2 Research: Toward More Developmentally Based Measures of Second Language Acquisition. *Applied Linguistics*, 35(5), 607–614. <https://doi.org/10.1093/applin/amu047>

- Larson-Hall, J. (2016). *A guide to doing statistics in second language research using SPSS and R* (2.uppl.). New York: Routledge.
- Norrby, C. & Håkansson, G. (2007). The Interaction of Complexity and Grammatical Processability: The case of Swedish as a foreign language. *International Review of Applied Linguistics in Language Teaching*, 45(1), 45–68. <https://doi.org/10.1515/iral.2007.002>
- Norris, J. M. & Ortega, L. (2009). Towards an Organic Approach to Investigating CAF in Instructed SLA: The case of complexity. *Applied Linguistics*, 30(4), 555–578. <https://doi.org/10.1093/applin/amp044>
- Ortega, L. (2003). Syntactic Complexity Measures and Their Relationship to L2 Proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics*, 24(4), 492–518. <https://doi.org/10.1093/applin/24.4.492>
- Pallotti, G. (2015). A Simple View of Linguistic Complexity. *Second Language Research*, 31(1), 117–134. <https://doi.org/10.1177/0267658314536435>
- Pienemann, M. (1998). *Language processing and second language development: processability theory*. Amsterdam: Benjamins.
- Rahkonen, M. & Håkansson, G. (2008). Production of written L2 Swedish – processability or input frequencies. I H-J. Kessler (Red.), *Processability Approaches to Second Language Development and Second Language Learning* (s. 135–161). Cambridge: Cambridge Scholars Publishing.
- Robinson, P. (2001). Task complexity, cognitive resources, and syllabus design: a triadic framework for examining task influences on SLA. I P. Robinson (Red.), *Cognition and Second Language Instruction* (s. 287–318). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139524780.012>
- Skehan, P. (1998). *A Cognitive Approach to Language Learning*. Oxford: Oxford University Press.
- Spoelman, M. & Verspoor, M. (2010). Dynamic Patterns in Development in Accuracy and Complexity: A longitudinal case study in the acquisition of Finnish. *Applied Linguistics*, 31(4), 532–553. <https://doi.org/10.1093/applin/amq001>
- Teleman, U., Hellberg, S. & Andersson, E. (2001). *Inledning till grammatiken*. Svenska Akademien: Stockholm.
- Wijers, M. (2018). The Role of Variation in L2 Syntactic Complexity: A Case Study on Subordinate Clauses in Swedish as a Foreign Language. *Nordic Journal of Linguistics*, 41(1), 75–116. <https://doi.org/10.1017/S0332586517000233>
- Wolfe-Quintero, K., Inagaki, S. & Kim, H.-Y. (1998). *Second Language Development in Writing: Measures of Fluency, Accuracy, and Complexity*. Honolulu: University of Hawaii Press.
- Wulff, S. & Ellis, N. C. (2018). Usage-based approaches to second language acquisition. I D. Miller, F. Bayram, J. Rothman & L. Serratrice (Red.), *Bilingual Cognition and Language. The state of the science across its subfields* (s. 37–56). Amsterdam: John Benjamins.
- Åberg, A.-M. (2020). Effekten av explicit undervisning på inläring av explicit och implicit kunskap om ordföljden i svenska. Doktorsavhandling. Turun yliopiston julkaisuja, Annales Universitatis Turkuensis 507 Humaniora: Åbo universitet.

Bilaga 1.



© Anja Karppinen

Ammatilliset haaveet (=Drömyrken)

1. Kerro, mitä kuvasarjassa tapahtuu. (=Berätta vad som händer i bildserien.)
2. Mikä on sinun unelmatyösi? Mistä haaveilit, kun olit lapsi? Missä näet itsesi kymmenen vuoden kuluttua? (=Vad är ditt drömyrke? Vad ville du bli när du var barn? Var befinner du dig om tio år?)
3. Pohdi, miten ja miksi ammatilliset haaveet muuttuvat. (=Diskutera hur och varför yrkesmässiga drömmar förändras.)

Pyri tuottamaan tekstiä mahdollisimman paljon, vähintään 100–150 sanaa. Ylärajaa ei ole.
(=Försök producera så mycket text som möjligt, åtminstone 100–150 ord. Ingen maximigräns.)

Tukisanoja (= Stödord):

aikuinen = vuxen

ammatti = ett yrke (-t, -n, -na)

astronautti = en astronaut (-en, -er, -erna)

avaruus = en rymd (-en, -er, -erna)

koulutus = en utbildning (-en, -ar, -arna)

kuuluja = berömd (-t, -da)

lapsi = ett barn (-et, -, -en)

lääkäri = en läkare (-n, -, -arna)

muusikko = en musiker (-n, -, -na)

opiskella = studera (-ar, -ade, -at)

palkka = en lön (-en, -er, -erna)

suunnitella = planera (-ar, -ade, -at)

tulevaisuus = framtid (-en)

unelmoida = drömma (-er, -de, -t)

Bilaga 2.

SKRIFTLIG PRODUKTION

Nivå	Deskriptorer
C2	Kan skriva tydliga och flytande texter med en effektiv stil, kan tillämpa en lämplig disposition som underlättar för läsaren att ta fram det väsentliga, behärskar en bred språklig repertoar (inbegripet vardagliga och idiomatiska uttryck) och kan använda den till att formulera tankar exakt, betona vissa saker och få fram nyanser
C1	Kan skriva tydliga, välstrukturerade texter och utförligt utveckla och ge stöd för ståndpunkter med hjälp av understödjande argument samt avrunda med en lämplig slutsats, har en bred språklig repertoar, kan omformulera (vid behov) och behöver inte begränsa det han/hon vill säga
B2	Kan skriva tydliga, detaljerade texter och beskrivningar där tankegången framgår tydligt i en sammanhängande text, kan argumentera genomtänkt och lyfta fram viktiga punkter, samt ge relevanta understödjande fakta, kan förklara fördelar och nackdelar med olika ställningstaganden, har ett stort ordförråd när det gäller de flesta allmänna ämnen och kan variera formuleringarna för att undvika upprepningar
B1	Kan skriva sammanhängande texter genom att sätta samman en rad element i en logisk följd, kan beskriva erfarenheter och känslor, kan skriva en detaljerad beskrivning av bildserien, har ett tillräckligt stort ordförråd för att skriva om de flesta vardagliga ämnen (t.ex. familj, hobbyer, intressen, arbete och aktuella händelser)
A2	Kan skriva en beskrivning i form av separata sekvenser, kan skriva om saker i vardagsmiljön (t.ex. människor, platser och studieerfarenheter), har ett tillräckligt stort ordförråd (också utanför stödordlistan) för att beskriva enkla situationer
A1	Kan skriva fristående meningar om sig själv och vad han/hon gör, har ett mycket grundläggande förråd av ord som rör särskilda konkreta situationer (t.ex. i bildserien), beskriver bildserien (mestadels) med hjälp av stödorden
under A1	

(CEFR, 2007, s. 63–64, 107, 109)

¹ Färdighetsnivå kan definieras på olika sätt i olika sammanhang. I denna studie definieras informanternas färdighetsnivå enligt de texter som de har producerat för föreliggande studie. Den lägre färdighetsnivån motsvarar CEFR-nivåerna A1 och A2, och den högre nivån motsvarar nivåerna B1 och B2 (+ en text på nivå C1).

² I denna studie används L2 som paraplyterm som täcker både andraspråks- och främmandespråksinläring.

³ Jag vill tacka Svenska Litteratursällskapet i Finland för finansiering som möjliggjorde materialinsamlingen för denna studie och Jouko Katajisto för all hjälp med statistiken. Därtill vill jag tacka de anonyma granskarna för deras värdefulla kommentarer.

⁴ Ursprungligen användes denna teori i analysen av muntlig L2-produktion men senare har den tillämpats också i studier av skriftlig L2-produktion.

⁵ Centrum för språk och kommunikation ordnar de obligatoriska kurserna i svenska som enligt lagen ingår i universitetsexamina.

⁶ Komplexiteten i denna text motsvarar komplexiteten i texter producerade av de övriga informanterna på den högre nivån, och därför tas också denna text med i analysen.

⁷ Vid upprepade jämförelser är det viktigt att korrigera p-värden eftersom risken för falska positiva resultat ökar. FDR avviker något från andra metoder (t.ex. Bonferroni) som används för att kontrollera risken för falska resultat och accepterar en högre risk för falska positiva resultat (se t.ex. Benjamini & Hochberg, 1995, s. 291, 298; Larson-Hall, 2016, s. 287, för en omfattande beskrivning av metoden, se Benjamini & Hochberg, 1995).