SYNTACTIC COMPLEXITY IN L2 SWEDISH

The Effects of Language Proficiency, Mode of Production, and Individual Variation

Mari Mäkilä

SYNTACTIC COMPLEXITY IN L2 SWEDISH

The Effects of Language Proficiency, Mode of Production, and Individual Variation

Mari Mäkilä

University of Turku

Faculty of Humanities
School of Languages and Translation Studies
Department of Scandinavian Languages
Doctoral programme in Languages and Translation Studies (Utuling)

Supervised by

Professor Emerita Sinikka Lahtinen
University of Turku
Turku, Finland

Professor Pekka Lintunen
University of Turku
Turku, Finland

Reviewed by

Professor Emerita Gisela Håkansson
Lund University
Lund, Sweden

Professor Iwona Kowal
Jagiellonian University
Kraków, Poland

Opponent

Professor Iwona Kowal
Jagiellonian University,
Kraków, Poland

ABSTRACT

This dissertation examines syntactic complexity in learner Swedish. Syntactic
complexity is explored through both absolute and relative complexity. Absolute
complexity refers to linguistic, structural complexity, and relative complexity can be
seen as complexity in relation to the learner. Thus, syntactic complexity is defined
as a structural property of learner language that can be observed in the varied use of
long and embedded units of language and in the use of demanding structures.
Syntactic complexity is inherently multifaceted and influenced by many factors. The
present study focuses on the effects of language proficiency, mode of production,
and individual variation. The theoretical underpinnings of the study are drawn from
cognitive and usage-based approaches to language learning, and the dissertation
incorporates ideas from both dynamic and systemic views of learner language
development. This study also provides suggestions for teaching and assessing
language.

The dissertation consists of three original publications and this summary. The
data comprise written and spoken productions by 31 Finnish-speaking university
students. A variety of syntactic complexity measures targeting different linguistic
levels (e.g. the clausal level) are used. The first two substudies explore differences
in absolute and relative complexity at two proficiency levels (lower vs higher) using
quantitative analyses. The first substudy examines the effect of learners' language
proficiency on the syntactic complexity and accuracy demonstrated in their language
use, as well as on the relationship between these aspects of learner language in
written production, whereas the second substudy investigates the differences in
syntactic complexity between speech and writing. The third substudy is an
individual-level enquiry that focuses on absolute complexity. In this study,
complexity profiles in learners' productions are explored using a mixed-methods
approach.

The results of the dissertation corroborate earlier findings demonstrating that
syntactic complexity is a complex feature of learner language. Somewhat
unexpectedly, the results of the first substudy do not indicate any statistically
significant correlations between measures of syntactic complexity and accuracy. In
line with several earlier studies, the results show that the learners at the higher
proficiency level produced more accurate structures. However, in contrast with some
earlier research, the learners at the lower proficiency level were able to produce

relatively complex structures in their writing. As the second substudy shows, the difference in syntactic complexity between the two proficiency levels was more noticeable in speech, indicating that the mode of production employed has a strong effect on the resulting syntactic complexity. At the lower proficiency level, the written productions seemed more syntactically complex than the spoken productions, whereas the syntactic complexity in the two modes of production did not differ to a great extent at the higher proficiency level. In line with many previous studies, production in the spoken mode seemed more cognitively demanding for the learners at the lower proficiency level. Based on the results of the first two substudies, individual style of producing language seems to be one of the key determinants of syntactic complexity. The results from the third substudy, conducted at the individual level, validate this assumption: while some learners have distinctively individual styles of producing structures, others show more variation in their preferred complexification strategies. The results of this dissertation highlight the importance of considering the methodological choices when interpreting the results of both group- and individual-level studies, as not every aspect of syntactic complexity can be identified using only traditional quantitative measures of complexity.

This dissertation critically discusses the operationalisation of syntactic complexity and the problem of choosing suitable complexity measures. Future studies would benefit from analyses that focus on several different linguistic levels and, ideally, combine quantitative and qualitative approaches. The new method of categorising productions into complexity profiles explored in the third substudy could be further developed in future research to shift the focus from statistically significant differences between groups of learners to the individual-level examination of syntactic complexity.

Individual-level variation is especially important in language teaching and assessment. By examining the different complexity profiles together with learners, teachers can guide the learners and their productions accordingly. Language learners should also be aware of the fundamental differences between speech and writing and recognise how these differences affect their use of syntactically complex structures in these two modes of production. Equally important for teachers and learners is to acknowledge the relationship between complexity and other features of learner language, such as accuracy. Due to the non-linearity of development and the constant interaction between complexity and accuracy, for instance, these features can show simultaneous progress and regress in learner language. All in all, it is crucial to regard syntactic complexity as a multifaceted construct in both teaching and research.

TIIVISTELMÄ

Väitöstutkimuksessani tarkastelen syntaktista kompleksisuutta oppijanruotsissa. Syntaktista kompleksisuutta lähestytään sekä absoluuttisesta että suhteellisesta näkökulmasta. Absoluuttinen kompleksisuus viittaa lingvistiseen, rakenteelliseen kompleksisuuteen, kun taas suhteellista kompleksisuutta tutkitaan suhteessa oppijaan. Näin ollen syntaktinen kompleksisuus määritellään monipuoliseksi pitkien ja alisteisten rakenteiden käytöksi, mutta myös vaativien rakenteiden tuottamiseksi. Syntaktinen kompleksisuus osana oppijankieltä on monitahoinen ilmiö, johon vaikuttavat monet tekijät. Tässä väitöstutkimuksessa huomion kohteena ovat kielitaidon tason, tuotostavan ja yksilövariaation vaikutus. Teoreettinen viitekehys pohjautuu kognitiivisiin ja käyttöpohjaisiin näkemyksiin kielenoppimisesta, ja tutkimus yhdistää ajatuksia dynaamisista ja systemaattisista kielenoppimiskäsityksistä. Tutkimustuloksia voidaan soveltaa niin kieltenopetuksessa kuin kielitaidon arvioinnissa.

Väitöskirja koostuu yhteenveto-osan lisäksi kolmesta osatutkimuksesta, joissa käytetty aineisto sisältää suullisia ja kirjallisia tuotoksia 31:ltä suomenkieliseltä yliopisto-opiskelijalta. Osatutkimuksissa hyödynnetään kompleksisuusmittareita, jotka mittaavat syntaktista kompleksisuutta eri tasoilla (esim. lausetaso). Kahdessa ensimmäisessä osatutkimuksessa tutkitaan eroja kahden kielitaitotason (alempi ja ylempi) välillä kvantitatiivisesti ja tarkastellaan sekä absoluuttista että suhteellista kompleksisuutta. Ensimmäisessä osatutkimuksessa kartoitetaan kielitaidon vaikutusta syntaktiseen kompleksisuuteen ja tarkkuuteen sekä näiden väliseen suhteeseen kirjallisessa tuotoksessa, ja toisessa osatutkimuksessa tarkastellaan eroja kompleksisuudessa suullisen ja kirjallisen tuotoksen välillä. Kolmas osatutkimus keskittyy absoluuttiseen kompleksisuuteen yksilötasolla ja kartoittaa kompleksisuusprofiileja oppijoiden tuotoksissa. Lähestymistapa viimeisessä osatutkimuksessa on monimenetelmällinen, sillä kvantitatiivisia analyysejä täydennetään tuotosten kvalitatiivisella tarkastelulla.

Väitöskirjan tulokset vahvistavat aiemman tutkimuksen näkemystä siitä, että syntaktinen kompleksisuus on monimutkainen osa oppijankieltä. On yllättävää, ettei syntaktisen kompleksisuuden ja tarkkuuden välille voitu osoittaa tilastollisesti merkitsevää korrelaatiota ensimmäisessä osatutkimuksessa. Yhteneväisesti aiemman tutkimuksen kanssa tulokset osoittavat, että kielitaidolla on selvä vaikutus tuotoksen tarkkuuteen. Toisaalta myös alemmalla taitotasolla oppijat kykenevät

tuottamaan suhteellisen kompleksisia rakenteita kirjallisesti. Ero syntaktisessa kompleksisuudessa taitotasojen välillä on kuitenkin selvempi suullisessa tuotoksessa, kuten toisen osatutkimuksen tulokset osoittavat. Alemmalla taitotasolla kirjallinen tuotos on selvästi suullista tuotosta kompleksisempaa, kun taas ylemmällä taitotasolla merkittävää eroa tuotostapojen välillä ei voida havaita. Tulokset tukevat aiempia havaintoja siitä, että suullinen tuotos on usein kognitiivisesti kuormittavampaa kielenoppijoille alemmilla kielitaitotasoilla. Ensimmäisten osatutkimusten tulokset viittaavat myös siihen, että syntaktinen kompleksisuus voi vaihdella huomattavasti yksilötasolla. Kolmannen osatutkimuksen tulokset vahvistavat tämän havainnon viitaten siihen, että osalla oppijoista on vahva yksilöllinen tyyli tuottaa kieltä, kun taas toiset osoittavat suurempaa vaihtelua rakenteissaan. Keskeistä tässä tuloksessa on myös se, ettei syntaktista kompleksisuutta voi aina tarkastella pelkästään perinteisten kompleksisuusmittareiden avulla. Metodologisten valintojen huomiointi tuloksia tulkittaessa on äärimmäisen tärkeää sekä ryhmä- että yksilötason tutkimuksissa.

Väitöskirjassa pohdin kriittisesti syntaktisen kompleksisuuden operationalisointia ja mittareiden valinnan vaikutusta tutkimustuloksiin. Tulevassa tutkimuksessa tulisi keskittyä syntaktisen kompleksisuuden ilmentymiseen useilla lingvistisillä tasoilla, ja kvantitatiivista ja kvalitatiivista analyysia tulisi yhdistää tutkimusasetelmissa. Tässä tutkimuksessa testattua uutta kategorisointimenetelmää, jonka mukaan yksilön profiili voidaan määrittää ryhmätason analyysien avulla, voidaan jatkossa kehittää edelleen. Tulevassa kompleksisuustutkimuksessa tulisi enenevässä määrin keskittyä myös yksilötasoon ryhmätason erojen sijaan.

Yksilötason variaatio on erityisen tärkeää huomioida kielenopetuksessa ja arvioinnissa. Kompleksisuusprofiilien tarkastelu yhdessä oppijan kanssa voi auttaa opettajia ohjaamaan oppijaa tarkoituksenmukaisesti. Lisäksi oppijan tulisi olla tietoinen suullisen ja kirjallisen viestinnän eroista sekä erojen vaikutuksesta tuotokseen. Tärkeää on myös huomioida syntaktisen kompleksisuuden ja tarkkuuden sekä muiden oppijankielen osa-alueiden välinen suhde. Oppijankielen osa-alueet voivat kehittyä epälineaarisesti, ja niiden välinen dynaaminen vuorovaikutus saattaa johtaa eri osa-alueiden yhdenaikaiseen kehitykseen ja taantumiseen. Yhteenvetona voidaan todeta, että syntaktinen kompleksisuus tulisi nähdä monitahoisena ilmiönä sekä opetuksessa että tutkimuksessa.

ABSTRACT

I denna doktorsavhandling undersöks syntaktisk komplexitet i inlärarsvenska. Syntaktisk komplexitet granskas genom både absolut och relativ komplexitet. Med absolut komplexitet avses lingvistisk, strukturell komplexitet, medan relativ komplexitet syftar till komplexitet i förhållande till inläraren. Därmed definieras syntaktisk komplexitet som en strukturell egenskap i inlärarspråket som framkommer via ett mångsidigt bruk av långa och inbäddade strukturer samt i bruket av utmanande strukturer. Syntaktisk komplexitet är mångdimensionell till sin natur och den påverkas av ett flertal faktorer. Denna undersökning granskar effekten av språkfärdighetsnivå, modalitet och individuell variation. Den teoretiska referensramen baserar sig på kognitiva och bruksbaserade perspektiv på språkinlärning och i avhandlingen kombineras dynamiska och systematiska uppfattningar om utvecklingen av språkkunskaper. Studien erbjuder möjligheter till att tillämpa resultaten i språkundervisning och i bedömning.

Avhandlingen innehåller tre delstudier utöver detta sammandrag. Materialet består av muntlig och skriftlig produktion av 31 finskspråkiga universitetsstuderande. I studierna används flera mått på syntaktisk komplexitet, som mäter komplexiteten på olika språkliga nivåer (t.ex. satsnivå). I de två första delstudierna undersöks skillnaderna i absolut och relativ komplexitet mellan två språkfärdighetsnivåer (lägre vs högre) med hjälp av kvantitativa analyser. Den första delstudien utforskar hur språkfärdigheten påverkar den syntaktiska komplexiteten och korrektheten samt relationen dessa emellan i skriftlig produktion, medan den andra delstudien granskar skillnaderna i syntaktisk komplexitet mellan tal och skrift. Den tredje delstudien koncentrerar sig på individnivån och på absolut komplexitet. I den delstudien används blandade metoder för att undersöka komplexitetsprofiler i inlärares produktion.

Undersökningen styrker tidigare iakttagelser om att syntaktisk komplexitet är en mångsidig komponent i inlärarspråk. Något överraskande visar resultaten i den första delstudien inga statistiskt signifikanta korrelationer mellan syntaktisk komplexitet och korrekthet. I enlighet med tidigare undersökningar producerar inlärarna på den högre färdighetsnivån fler målspråksenliga strukturer. Däremot tyder resultaten på att inlärarna också på den lägre färdighetsnivån kan producera relativt komplexa strukturer i skrift, vilket avviker från resultaten i flera tidigare studier. Den andra delstudien avslöjar att skillnaden i syntaktisk komplexitet mellan färdighetsnivåerna

är mer tydlig i muntlig produktion. På den lägre färdighetsnivån uppvisar inlärarna högre komplexitet i den skriftliga produktionen än i den muntliga. På den högre färdighetsnivån visar däremot jämförelsen mellan muntlig och skriftlig produktion inte på några stora skillnader i komplexitet. Detta resultat överensstämmer med flera tidigare forskningsresultat och tyder på att modaliteten spelar en stor roll för den syntaktiska komplexiteten. Den muntliga produktionen verkar vara kognitivt mer krävande för inlärare med begränsade språkkunskaper. Därtill tyder resultaten i både den första och andra delstudien på att skillnaderna kan vara betydande på individnivå. Resultaten från den tredje delstudien bekräftar detta antagande: somliga inlärare verkar ha ett relativt stabilt sätt att producera språk, medan andra visar mer variation i sina strukturer. Resultaten i denna avhandling visar hur viktigt det är att beakta de metodologiska valen när man tolkar resultaten från både grupp- och individnivåstudier. Det är inte möjligt att observera alla aspekter av syntaktisk komplexitet med hjälp av traditionella, kvantitativa komplexitetsmått.

I denna doktorsavhandling diskuterar jag kritiskt valet av komplexitetsmått och operationaliseringen av syntaktisk komplexitet. I framtida studier skulle det vara nyttigt att koncentrera sig på flera språkliga nivåer samt att kombinera kvantitativa och kvalitativa analyser. I den tredje delstudien används en ny metod för att utgående från analyser på gruppnivå kategorisera inlärares produktion med hjälp av profiler. Denna metod skulle kunna utvecklas vidare i framtida forskning och därmed bidra till att rikta fokus mot individnivån.

Det är speciellt viktigt att beakta variationen på individnivå i språkundervisningen och i bedömningen. Genom att granska komplexitetsprofilerna tillsammans med inlärarna kan läraren handleda dem på ett mer ändamålsenligt sätt. Språkinlärare borde också vara medvetna om de grundläggande skillnaderna mellan tal och skrift och känna till hur dessa skillnader påverkar deras produktion. Det är viktigt för både lärarna och inlärarna att vara medvetna om relationen mellan komplexitet och övriga aspekter av inlärarspråket, såsom korrekthet. Dessa aspekter kan utvecklas icke-linjärt och interaktionen mellan dem kan leda till att inlärarspråket visar tecken på samtidig progression och regression. Sammanfattningsvis kan man konstatera att det är ytterst viktigt att man både i undervisning och i forskning inser att syntaktisk komplexitet är en mångfacetterad del av ett inlärarspråk.

NYCKELORD: individuell variation, inlärarsvenska, modalitet, språkfärdighet, språkinlärning, syntaktisk komplexitet

Acknowledgements

I discovered my passion for languages during my school years. Early on, I knew that I wanted to become a teacher. Despite my growing interest in languages, I was always scared of making mistakes, especially when speaking in another language. I often found myself in situations in which I would have liked to express something in a certain way, and I would try to formulate an error-free utterance in my head, only to realise that the discussion had moved on without me and my fancy, flawless formulation of thought. 'How do they do it?' I wondered. I guess that was the initial spark for my later research on learner languages. This spark led me to fulfil my dream of becoming a teacher and ultimately – after some extra fuel was added to the flames in the form of external persuasion – to write this PhD dissertation.

It is said that writing a doctoral dissertation is a lonely endeavour. While this might have been true for many parts of the process, there are numerous people to whom I am greatly indebted for their contributions to my work. First, I want to express my deepest gratitude to the external examiners of this dissertation, Professor Emerita Gisela Håkansson and Professor Iwona Kowal. Their outstanding work in the field of second language acquisition has been a major inspiration during my years of research. I would also like to thank Professor Kowal for kindly agreeing to act as my opponent in the public examination. I want to extend my gratitude to the anonymous referees of the substudies included in this dissertation. All the comments undoubtedly took my research to the next level.

To my two supervisors, Professor Emerita Sinikka Lahtinen and Professor Pekka Lintunen, I owe gratitude beyond expression. Sinikka was the one who convinced me to turn my love for languages into a career in research, for which I am immensely thankful. In addition to providing scientific guidance, she has always been prepared to listen, to help, and to support me. Without her nurturing care and profound belief in my abilities, I would not have been able to complete this dissertation. Pekka has kept my feet on the ground and, at the same time, made me believe that I can achieve anything I set my mind to. He has provided me with valuable lessons about academia and life in general. I thank Pekka for his outspokenness and constructive criticism, and for the unwavering encouragement, especially during the hectic homestretch.

Table of Contents

Tables

Figures

List of Original Publications

This dissertation is based on three original publications.

I Mäkilä, M. (2021). Syntaktisk komplexitet och korrekthet i finska universitetsstudenters skriftliga inlärarsvenska: en jämförelse mellan två färdighetsnivåer. *Nordand, 16(2)*, 134–156.

II Mäkilä, M. (2022). Syntaktisk komplexitet i inlärarsvenska: en jämförelse mellan finska universitetsstudenters muntliga och skriftliga produktion på två olika färdighetsnivåer. *Folkmålsstudier, 60*, 73–108.

III Mäkilä, M. (manuscript). Individual Variation in Learner Language: a multiple extreme case analysis of syntactic complexity in L2 Swedish.

The original publications have been reproduced with the permission of the copyright holders.

1   Introduction

Based on evidence presented in earlier research, language learners start to produce more complex structures as they become more proficient in the target language (see, e.g. Wolfe-Quintero, Inagaki & Kim 1998; Housen, Kuiken & Vedder 2012a). However, several factors determine the structures that learners produce. For instance, learners' language proficiency can influence the interplay between complexity and other aspects of learner language, such as accuracy. Additionally, learners' ability to perform in written and spoken tasks can differ (see Skehan & Foster 2012; Kuiken & Vedder 2012a, 2012b). Furthermore, individual learners' styles of producing language may affect the complexity of their productions (see, e.g. Norrby & Håkansson 2007). Moreover, methodological choices and study designs influence the results gained from studies of complexity. These factors make syntactic complexity a complex and fascinating topic of research.

Ideas and thoughts can be expressed in multiple ways using structures of varying syntactic complexity, and these structures should suit the context in which the language is produced. For instance, the syntactic structures used when writing a scientific article differ from those employed when taking part in an informal conversation, as these contexts are very dissimilar. This implies that language use is highly context-bound. Therefore, choosing an appropriate level of formality, and of complexity, can be important when interacting with different people in various situations. Accordingly, second language[1] (L2) learners are expected to master the use of context-appropriate language by employing, for instance, a range of syntactic structures in their linguistic productions. These expectations can be seen as the basis for studies of syntactic complexity in learner language. As noted above, several factors can shape how structures of varying complexity are learned and applied in

---

[1]   In this study, the term 'second language', or L2, refers to all types of additional language acquisition, regardless of the learning context or the chronological order in which the languages are learned (cf. foreign language or third language; L1 = first language). Although Swedish is an official language in Finland, it is often learned after English and has a prominent status only in some parts of Finland, mainly the southern and western coasts (for more information about Swedish in Finland and in Finnish schools, see, e.g. Åberg 2020: 24–32; see also Section 2.3.2.5).

different contexts. These factors include **the language proficiency of the learner, the mode of production, and the learner's individual style of producing language**. This dissertation examines syntactic complexity in the productions of Finnish learners of Swedish, focusing on the effects of the abovementioned contributing factors.

1.1   Background of the study

In the field of Second Language Acquisition (hereafter SLA), researchers study the nature of learner language and describe how it develops. Three principal dimensions of learner language – complexity, accuracy, and fluency (the CAF triad) – have been widely used to investigate learner language (Wolfe-Quintero et al. 1998; Housen et al. 2012a; see also Section 2.3). These dimensions can be studied as distinct components of learner language, or the interplay between the dimensions can be examined by inspecting several components together (Housen & Kuiken 2009). In simple terms, complexity refers to the varied use of language suited to the context (see Section 2.3.1), accuracy refers to the need to produce language according to the normative rules of the given target language, and fluency refers to the ability to produce the target language smoothly, i.e. relatively rapidly without unnecessary pauses or corrections (Housen, Kuiken & Vedder 2012b). In this study, the focus lies on complexity – specifically, syntactic complexity. Accuracy is examined in the first substudy, but only in a minor role in relation to the dissertation as a whole, as the general aim of the first substudy is to explore the relationship between complexity and accuracy (see Section 1.2).

To date, there is no general agreement on the definition of syntactic complexity, or of complexity in general. Due to its multidimensional nature, syntactic complexity can be defined, and therefore operationalised, in various ways depending on the perspective a study adopts (Ortega 2003; Norris & Ortega 2009; see also Pallotti 2015; Bulté, Housen & Pallotti 2024). In the current study, syntactic complexity is investigated from different perspectives, and it is therefore studied according to different definitions (see *absolute complexity* and *relative complexity* in Section 2.3.1) and operationalised through various measures, such as length-based and specific metrics (see Section 3.3.2). Thus, in this dissertation, syntactic complexity is defined as **a characteristic of learner language that manifests itself in the use of varied language as evaluated by measures at different linguistic levels but also as the use of linguistically or cognitively demanding structures**.

All constructs in the CAF triad are multifaceted in that they contain several subsystems and can therefore be studied in various ways depending on the perspective of the study (see, e.g. Wolfe-Quintero et al. 1998). The operationalisation of accuracy and fluency can be relatively straightforward

compared to that of complexity, as these aspects of learner language have more fixed definitions and a desired outcome: L2 learners are expected to produce fewer errors and be more fluent in the target language as they become more proficient (see, e.g. Housen et al. 2012a). This does not mean that all L2 learners must learn to produce the target language flawlessly or fluently at the level of a native speaker, as each learner has their own objectives when learning a language (see, e.g. Ellis 2019: 53). Furthermore, it is not always perfectly clear what is meant by being more fluent in a given language, and the comparison with native speakers has been questioned, especially in the field of fluency studies (see, e.g. Larsen-Freeman 2006: 591; see also Lintunen, Mutta & Peltonen 2020). Nevertheless, it is possible to measure the constructs of accuracy and fluency, for instance, based on the presence of unwanted features, such as errors and lengthy pauses.

When it comes to complexity, operationalisation is more complicated, as this construct lacks a universal definition, and there is no similarly desired outcome in its development (see, e.g. Ferrari 2012: 282–283). L2 learners are expected to learn to produce a range of structures of varying syntactic complexity, but it is not expected that the end result will be highly complex language (e.g. Ferrari 2012; Vercellotti 2019). There are different expectations regarding syntactic complexity in different contexts. As shown in earlier research, syntactic complexity can be affected by, for instance, learners' language proficiency[2] (e.g. Bulté & Housen 2018; Kuiken, Vedder, Housen & De Clercq 2019; see also Section 2.3.2.1). However, as previous studies have yielded partly contradictory results, the relationship between language proficiency and syntactic complexity, as well as the link between proficiency and the other constructs in the CAF triad, remains partly unclear. On the one hand, there is ample evidence of a positive relationship between overall language proficiency and many individual complexity measures, and especially between proficiency and accuracy (Wolfe-Quintero et al. 1998). On the other hand, some earlier studies indicate that although language proficiency and the systems in the CAF triad are often related in learners' productions, there is no clear causal connection between, for instance, higher proficiency and greater complexity (e.g. Granfeldt 2008). Possible reasons for these contradictory findings include differing modes of production and learners' individual preferences in language production (e.g. Kuiken & Vedder 2012a; Vyatkina 2012; see also Sections 2.3.2.3 and 2.3.2.4).

The current study fills several gaps left open in past research on syntactic complexity. To begin with, in contrast to L2 English, there is little information about syntactic complexity in other languages, such as L2 Swedish (see, however, Norrby

---

[2]   Language proficiency can be defined in various ways in different contexts. In the current study, the language proficiency of the learners was assessed based on their performance in certain tasks (for further discussion, see Section 3.2).

& Håkansson 2007; Kowal 2016). As earlier studies have demonstrated that learner language can differ in syntactic complexity depending on the (target) language, it is important to investigate target languages other than English (Gyllstad, Granfeldt, Bernardini & Källkvist 2014; Bernardini & Granfeldt 2019). Second, studies that directly compare syntactic complexity across modes in written and spoken L2 production are rare, especially in languages other than English (see, e.g. Kormos & Trebits 2012; Lambert & Kormos 2014; see, however, e.g. Håkansson & Norrby 2007; for further discussion, see Section 2.3.2.3). Given the crucial role of the mode of production in language use, particularly in relation to syntactic complexity in L2 production, more studies focusing on this issue are needed. The current multidimensional study is among the first to provide a comprehensive picture of syntactic complexity in L2 Swedish in both modes of production. Finally, earlier studies on syntactic complexity in learner language mainly focused on the group level (van Dijk, Verspoor & Lowie 2011; see, however, e.g. Norrby & Håkansson 2007; Vyatkina 2013). Group-level studies are informative and useful when the focus is on comparing general-level tendencies at different proficiency levels or on developmental patterns in syntactic complexity (see, e.g. Biber, Gray & Staples 2016; Kuiken & Vedder 2019), but these studies should be combined with individual-level research if the aim is to provide a thorough picture of syntactic complexity, as earlier studies have revealed great individual variation even within the same proficiency level (e.g. Vyatkina 2012; Kowal 2016; Tagarelli, Ruiz, Vega & Rebuschat 2016; see also Section 2.3.2.4). In the current study, syntactic complexity is investigated both at the group level and the individual level. The written and spoken data come from the same learners, and the substudies are based on the same datasets (see Section 3.2). Furthermore, in this dissertation, cross-sectional comparisons are combined with a longitudinal study of syntactic complexity at the individual level (see Section 3.4).

The importance of individual variation in language use and development has been supported by recent studies that draw on the Dynamic Systems Theory framework (henceforth DST; see, e.g. Larsen-Freeman 2006; Lesonen, Steinkrauss, Suni & Verspoor 2020). As an approach to studying learner language, DST can be contrasted with more traditional approaches, which often depict a linear and stage-like picture of development in learner language. In this dissertation, both traditional approaches and an approach based on the main principles of DST are adopted (see Section 2.2). The aims of the dissertation and the substudies are presented in more detail in the following section.

1.2 Research aims, substudies, and outline of the dissertation

The purpose of this dissertation is threefold. First, it provides important information about syntactic complexity in L2 Swedish, focusing on the effects of learners' language proficiency, mode of production, and individual styles of producing language. Second, the dissertation contributes to the methodological debate on the operationalisation of syntactic complexity. Finally, by adopting a multilevel approach to syntactic complexity, it provides a conceptual framework for L2 teaching purposes.

The dissertation contains three substudies in addition to this summary section. This summary describes syntactic complexity as a multifaceted construct of learner language by inspecting and critically discussing the results gained from the three independent substudies.

The main research aims of the dissertation are as follows:

1. to examine how language proficiency affects

   a) syntactic complexity,
   b) syntactic accuracy, and
   c) the interplay of syntactic complexity and accuracy

2. to examine how the mode of production used affects syntactic complexity

3. to examine the role of individual variation in syntactic complexity

The main research aims and their connections to the substudies are presented in Table 1.

Table 1. The main research aims pursued in the substudies.

Research aim	Article I	Article II	Article III
to examine how language proficiency affects syntactic complexity syntactic accuracy the interplay of syntactic complexity and accuracy	X X X	X	
to examine how mode of production affects syntactic complexity		X	X
to examine the role of individual variability in syntactic complexity			X

The effect of learners' language proficiency on the syntactic complexity of the language they use is explored in the first two substudies. In the first substudy, the effect of proficiency on both syntactic complexity and accuracy is studied, and the overarching objective is to determine whether the relationship between syntactic

complexity and accuracy differs at two unequal proficiency levels. In the second substudy, the effect of language proficiency on syntactic complexity in spoken productions is studied, and syntactic complexity in written and spoken productions is compared at two proficiency levels. In the third substudy, the focus is on the effect of individual styles of producing language on both written and spoken productions. The specific foci and research designs of the substudies are illustrated in Figure 1.
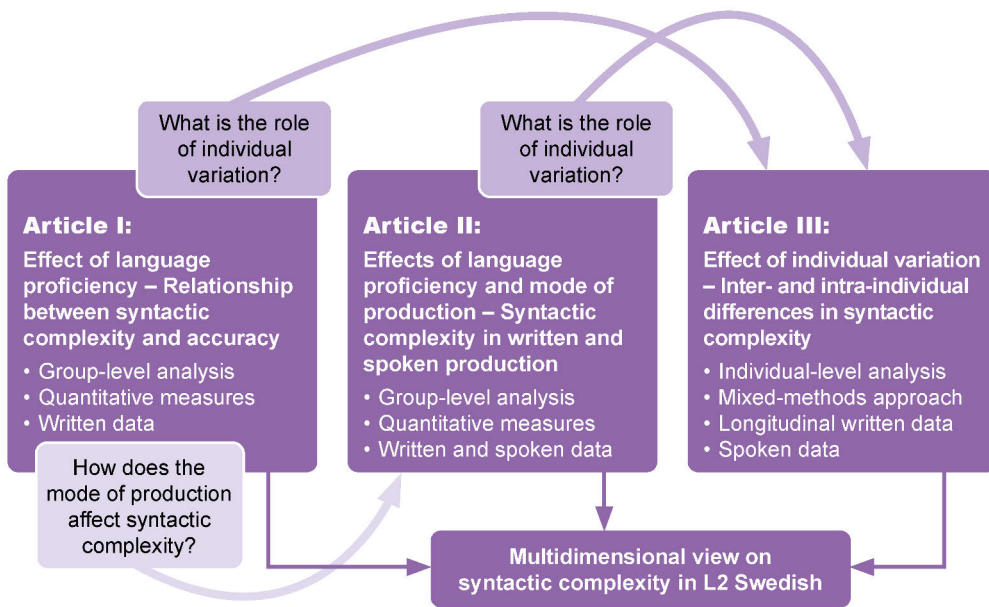
Figure 1. The main foci and research designs of the substudies.

The first two substudies focus on syntactic complexity at the group level, and the analyses are based on quantitative measures of syntactic complexity (see Section 3.3.2). Results from both substudies suggest that although some tendencies of the impact of language proficiency and modality can be detected at the group level, there is notable individual-level variation. Therefore, in the third substudy, the focus shifts from the group level to the individual level. Adopting a mixed-methods approach to syntactic complexity, the study aims to determine how syntactic complexity varies in learners' productions both synchronically and diachronically. The various methods used in these studies contribute to the methodological debate on operationalisations of syntactic complexity, and the multilevel analyses performed in the substudies provide a conceptual framework for analysing syntactic complexity. With its comprehensive investigation, the dissertation provides a multidimensional account of syntactic complexity in L2 Swedish. In the following paragraphs, the substudies are presented in more detail.

In Article I, the effect of proficiency on syntactic complexity, accuracy, and the relationship between syntactic complexity and accuracy is examined. Results from earlier research indicate that learners' language becomes increasingly complex and accurate as they become more proficient in the target language (e.g. Wolfe-Quintero et al. 1998; Kuiken & Vedder 2012a), but some studies have yielded mixed findings, indicating that the development of these constructs is not always linear (e.g. Bulté & Housen 2018). Furthermore, contradictory evidence has been reported regarding the nature of the relationship between the systems in the CAF triad (see, e.g. Housen et al. 2012b; Kormos & Trebits 2012; Kowal 2016), for instance, between complexity and accuracy (see Section 2.3.2.2). These earlier studies often lent support to one of the two main views on the aforementioned relationship. According to the *Limited Attentional Capacity* model (Skehan 1998), the relationship between complexity and accuracy is competitive because learners have limited cognitive resources, whereas the *Multiple Resources Attentional* model (Robinson 2001) argues for a possible supportive relationship between the systems on the basis that multiple attentional pools are available for learners. Researchers have demonstrated that language proficiency might have a significant impact on the interconnection between complexity and accuracy (e.g. Spoelman & Verspoor 2010; Gunnarsson 2012; Lesonen et al. 2020), as proficiency is closely related to learners' resources.

In the first substudy, 31 university students were divided into two groups according to their performance on a written task, and syntactic complexity and accuracy in these learners' written productions were studied at the group level. The general aim was to determine whether the relationship between syntactic complexity and accuracy represents a trade-off to a greater extent at the lower proficiency level than at the higher proficiency level. The following research questions were addressed in this study:

RQ1) To what extent do a) syntactic complexity and b) accuracy differ at two different proficiency levels?

RQ2) How does syntactic complexity stand in relation to accuracy at different proficiency levels?

This study adopted a traditional approach to studying syntactic complexity, which was considered in terms of both absolute (*linguistic*, *structural complexity*; see, e.g. Pallotti 2015) and relative (*cognitive complexity*; see, e.g. Michel 2017) complexity (for further discussion, see Section 2.3.1) to reveal a wider scope of differences between the proficiency levels (see, e.g. De Clercq & Housen 2017). Following suggestions made in earlier research (Ortega 2003; Norris & Ortega 2009), several independent measures were applied in the analyses. Syntactic

complexity was evaluated using both general measures of absolute complexity, that is, length-based measures, and a measure based on embedding, and specific measures of relative complexity, including occurrences of dependent clauses with a sentence adverbial and topicalised structures (see Section 3.3.2). The study of accuracy was limited to syntactic accuracy, that is, target-like word order in the structures produced, and accuracy was measured based on error-free clauses, error-free dependent clauses, error-free dependent clauses with a sentence adverbial, and the use of inversion after topicalised structures (see Section 3.3.3).

In Article II, the role of language proficiency in spoken productions by the participants discussed in Article I is studied. The aim was to determine whether language proficiency affects syntactic complexity differently in the written and spoken modes. The assumption was that producing syntactically complex structures in the spoken mode might be more challenging for learners with limited knowledge of the target language, whereas the ability to plan in the written mode might reduce the cognitive load on the learner and thus facilitate the production of complex structures (Skehan & Foster 2012; Tavakoli 2014; see also Section 2.3.2.3). Furthermore, it was assumed that automaticity at the higher proficiency level enables the production of complex structures, even with limited time for planning. The study aimed to answer the following research questions:

RQ1) What differences in syntactic complexity can be found in spoken productions by learners at two proficiency levels?

RQ2) What differences in syntactic complexity can be found between written and spoken productions a) at the lower proficiency level and b) at the higher proficiency level?

RQ3) How do the findings reflect the different demands that speech and writing impose on language learners?

As in the first substudy, a traditional approach to studying syntactic complexity was adopted, and the focus was on both absolute and relative complexity at the group level. Additionally, corresponding general and specific complexity measures were used in the analyses to study the differences in syntactic complexity between written and spoken productions at the two proficiency levels. In this study, the fundamental differences between speech and writing (e.g. Halliday 1989; Miller & Fernandez-Vest 2006; Lintunen & Mäkilä 2014; see also Section 2.3.2.3) were discussed theoretically and taken into consideration in the selection of the measures and in the implementation of the analyses.

Article III focuses on individual variation in syntactic complexity, but the research design also considers the effect of the mode of production. As already mentioned, most studies of syntactic complexity have focused on the group level, often ignoring learners' individual styles of producing structures (see, e.g. Vyatkina, Hirschmann & Golcher 2015). The aim of the third substudy was to show how syntactic complexity differs synchronically between the two modes of production and diachronically between two data gatherings separated by an interval of approximately six months through a multilevel analysis of productions by three individuals. The study provides answers to the following research questions:

RQ1)   To what extent is the categorisation of learner productions into complexity profiles according to general measures of syntactic complexity consistent a) synchronically, or between written and spoken productions, and b) diachronically, or between two data collection periods six months apart?

RQ2)   What differences and similarities in noun phrase complexity can be found in learners' productions?

RQ3)   What do the results pertaining to complexity profiles and noun phrase complexity reveal about variation in the use of complexification strategies?

RQ4)   How does the qualitative examination of learners' productions complement the study of their complexity profiles?

In the third substudy, a DST-based approach to studying syntactic complexity was adopted (see Section 2.2). Articles I and II provided the basis for the analyses carried out in Article III. Three learners were chosen to participate based on the group-level analyses conducted in the first substudy in accordance with the principles of extreme case sampling (Dörnyei 2007). A learner whose first written production embodied the use of relatively complex syntactic structures was selected from the higher extreme, and a learner who contributed relatively simple structures was selected from the lower end of the complexity continuum based on group-level descriptive statistics. Additionally, a third learner who produced structures of average syntactic complexity was included in the study. In the analyses, syntactic complexity in these productions was compared to syntactic complexity in the productions in the spoken mode and in the second written productions by the same individuals.

The analyses in most research on L2 complexity in general and syntactic complexity in particular are based on a limited number of quantitative measures (see, e.g. Kuiken et al. 2019), although some attempts have been made to use more fine-grained measures (see, e.g. Lambert & Kormos 2014; Kyle & Crossley 2018; Kuiken & Vedder 2019). In the third substudy, a multilevel analysis of syntactic complexity in the learners' productions was conducted, and general measures, as well as more fine-grained measures, were employed together with a qualitative examination of the structures employed in the learners' productions (see Section 3.4). By using an exploratory method of categorising individuals' productions based on group-level statistics and conducting a multilevel analysis of syntactic complexity, the study explores the differences between learners in a more straightforward manner and offers thorough insights into syntactic complexity in their productions.

This summary contains five chapters. In Chapter 2, the theoretical framework of the dissertation is presented. Cognitive and usage-based approaches to language learning are introduced in Section 2.1, and in Section 2.2, dynamic and systemic views on L2 development are discussed. The notion of syntactic complexity and its nature in learner language is discussed in Section 2.3, with a special emphasis on the different types of complexity (Section 2.3.1) and the contributing factors that are in focus in the dissertation, namely, the effect of language proficiency, modality, and individual variation (Section 2.3.2). Additionally, Section 2.3.2 provides a brief introduction to how syntactically complex structures emerge in L2 Swedish and how syntactic complexity is addressed in the Common European Framework of Reference for Languages (henceforth CEFR). Chapter 3 describes the data and methods used in the study. In Section 3.1, the participants are introduced, followed by a description of the data used in the current study (Section 3.2). Section 3.3 presents the measures chosen for the study. The section begins with a discussion of the segmentation units implemented (Section 3.3.1), followed by a presentation of the complexity measures (Section 3.3.2) and accuracy measures used (Section 3.3.3). In Section 3.4, the analyses conducted in the substudies are discussed. The results of the substudies are presented in Chapter 4. The independent substudies are discussed separately in Sections 4.1, 4.2, and 4.3, and Section 4.4 summarises the results. Chapter 5 concludes this summary with a discussion of the results and their methodological and theoretical contributions (Section 5.1). This chapter also identifies the implications of the research for L2 teaching and assessment (Section 5.2) and the limitations of the study, together with some suggestions for future directions in complexity research (Section 5.3). Concluding remarks are made in Section 5.4.

2   Theoretical framework

The theoretical standpoint of this dissertation is based on the assumption that language and language learning are complex and multidimensional phenomena that cannot be explained with simplistic and isolated theories. This assumption is supported by the myriad of theories and approaches applied to SLA in its relatively short history as an independent research field (De Bot & Larsen-Freeman 2011: 5; Verspoor & Behrens 2011: 25–26; see also, e.g. Segalowitz & Lightbown 1999; Eskildsen & Cadierno 2015; Ellis 2019). A variety of approaches can be combined to form a comprehensive picture of language and language learning (see Littlemore & Juchem-Grundmann 2010; Atkinson 2011; Ellis & Wulff 2014; MacWhinney 2015). Therefore, this dissertation draws on several related theoretical frameworks.

This chapter describes the theoretical framework of the dissertation. First, the core aspects of cognitive and usage-based theories of language learning are presented in Section 2.1. Second, two approaches based on these theories, namely, DST-based approaches and those based on Processability Theory (hereafter PT), are discussed in Section 2.2. Additionally, this section discusses how these theoretical views are combined in the dissertation. Thereafter, the CAF framework and syntactic complexity as a part of learner language are discussed in Section 2.3. Different types of complexity are introduced in Section 2.3.1. Section 2.3.2 focuses on the relationship between language proficiency and syntactic complexity (Section 2.3.2.1), the interplay between syntactic complexity and accuracy (Section 2.3.2.2), the differences in syntactic complexity between speech and writing (Section 2.3.2.3), individual variation in syntactic complexity (Section 2.3.2.4), syntactically complex structures in L2 Swedish (Section 2.3.2.5), and syntactic complexity in CEFR (Section 2.3.2.6).

2.1   Cognitive and usage-based approaches to language learning

While early research on language learning adopted behaviourist approaches (for an introduction to earlier theories in SLA, see, e.g. VanPatten & Williams 2014), most current SLA research has been based on cognitive and usage-based theories of language learning (e.g. Atkinson 2011; Eskildsen & Cadierno 2015; Ellis 2015,

2019). Cognitive linguistics contains several independent but related theories, views, and models (e.g. constructionist and connectionist views, exemplar- and statistic-based emergentism, and complexity theory; see, e.g. Ellis 2003; Ellis, Römer & O'Donnell 2016), and usage-based approaches combine ideas from various of these approaches (Wulff & Ellis 2018: 37; Ellis 2019: 50). These approaches also have connections to cognitive psychology in relation to topics such as learning, attention, and their relation to language and language use. The overarching aim of using these approaches is to examine the mechanisms that underlie language learning and production, determine how language is represented in learners' minds, and discover how these representations are visible in language use (Ellis 2019: 40). This section discusses central principles that the views have in common and how these views are implemented in the current study (for a comprehensive summary of current theories in SLA, see, e.g. Ortega 2013, 2014; see also the discussion in Littlemore & Juchem-Grundmann 2010).

In earlier approaches to language learning, language was considered modular, and a division was made between language competence (cognition) and actual performance (behaviour) (the Chomskian or generative view; van Dijk et al. 2011: 56; see also the discussions in, e.g. Bybee & Hopper 2001; Goldberg 2006; Bybee 2008; Atkinson 2011). More recent views on SLA hold that language is nonmodular, and that both language use and experience with language are cognitively organised in the learner's mind (e.g. Eskildsen & Cadierno 2015; Ellis et al. 2016; see also the discussion in Bulté & Housen 2020a: 213). In essence, according to cognitivist views, learning is accomplished in the mind of the learner through information processing (Atkinson 2011: 3). Therefore, **cognition** is central. Cognition is seen as a mental tool that processes *input* from the surrounding world and forms representations according to given goals (Behrens 2009: 388; Ellis 2019: 40). Cognition and language shape each other, as it is through language that people form experiences and representations and then organise and process this information (Ellis 2015: 49–50). Language, on the other hand, is interpreted as a tool for communication – for encoding and decoding units of thought (Atkinson 2011: 5; see also Eskildsen & Cadierno 2015). Language as a system operates at several levels (e.g. the brain and neural levels as well as different linguistic levels) on multiple time scales (e.g. interactional and diachronic), and it concerns many agents in different configurations (Ellis & Wulff 2014: 90; Wulff & Ellis 2018: 42; see also MacWhinney 2015).

Cognitive and usage-based approaches share a similar view of the language-learning process: learning is fundamentally seen as a cognitive process in which learners form representations and process the target language through versatile language use and meaningful interaction (see, e.g. Verspoor & Behrens 2011: 26; see also Ellis 2015). Furthermore, these approaches both argue that all cognitive

processing, including that of language, is similar in nature (Wulff & Ellis 2018: 37; see also Littlemore & Juchem-Grundmann 2010). They suggest that language representations, even new and complex ones, can emerge from simple representations through simple cognitive mechanisms guided by the specific need for communicating (Behrens 2009: 387; Verspoor & Behrens 2011: 27; Wulff & Ellis 2018: 41–42; Ellis 2019: 46; see also Bybee & Hopper 2001; Eskildsen 2009). Language use is considered a prerequisite for language learning (Eskildsen & Cadierno 2015). All linguistic knowledge – from simple and concrete to complex and abstract – emerges from actual language use and from past experiences with the target language (Behrens 2009: 385; Ellis 2015: 49, 2019: 49; Ellis et al. 2016: 24; see also Bybee & Hopper 2001; Littlemore & Juchem-Grundmann 2010). These experiences are central to the current study, as the participants were adult learners with extensive experience with both the target language and other languages in addition to their L1 (see Section 3.1).

According to Ellis (2019), language and language use are in constant interaction. This means that when a learner uses a target language, the status of their language changes, as the representations and constructions employed are affected by the language use (see also Ellis et al. 2016: 23; Lesonen et al. 2020: 2). Accordingly, when the learner is exposed to the target language during interactions, the linguistic input modifies the existing representations and constructions in the learner's language system, shaping their future language use. Furthermore, there are considerable individual differences in how learners' past experiences interact with new target forms (Ellis 2019: 153) and how these new forms further interact with one another in the learning process (De Bot & Larsen-Freeman 2011: 16; Verspoor & Behrens 2011: 29; Lambert & Kormos 2014: 611; Lesonen et al. 2020: 3–4). Therefore, the emergence of target language structures in learners' language use is highly individual and dynamic (see Section 2.2). This idea is central to the present study, as the enquiry combines cross-sectional comparisons with individual-level investigations of syntactic complexity.

To develop a thorough understanding of how language is learned, it is necessary to determine what the learners need to learn (Behrens 2009: 384). It has been suggested that language is learned through the acquisition of *constructions*, which are used to produce language (Ellis 2019: 48–49; see also Verspoor & Phuong Nguyen 2015). Constructions are linguistic mappings of form–function relationships in the target language (Ellis 2015: 66–68, 2019: 49). Grammar and lexis are seen as inseparable and as forming a continuum (Littlemore & Juchem-Grundmann 2010: 3; Wulff & Ellis 2018: 39). Learning constructions and the relations both within and between these constructions is essential to language learning (Ellis 2003: 68–69, 74–84; Littlemore & Juchem-Grundmann 2010: 3; see also Goldberg 2006), and this can only be done by using language (Ellis 2019: 48–49). These form–function pairings

contain both smaller parts of language, such as morphemes or words with a semantic or discursive purpose, and larger sets of linguistic material, such as syntactic frames that serve a functional purpose in the target language (Ellis 2015: 50, 2019: 49–50; Ellis et al. 2016: 26–28; Wulff & Ellis 2018: 38; see also Bybee & Hopper 2001: 14; Goldberg 2006: 6; Behrens 2009: 390). **Constructions in the sense of abstract syntactic frames** (e.g. X–V–S in Swedish, i.e. inverted word order when the structure does not start with a subject; see Section 2.3.2.5; see *slot-and-frame pattern* in, e.g. Ellis 2002; see also Pienemann & Lenzing 2014: 166–167, 169) are the most important type in the current study.

Especially in usage-based approaches, linguistic input is seen as the principal driving force of language learning (Wulff & Ellis 2018: 37; Lesonen et al. 2020: 2). Languages are learned by processing linguistic input to discover regularities and form abstract cognitive representations based on that input (see, e.g. Eskildsen 2009). Frequent language use can strengthen these representations, which can lead to language development (Ellis et al. 2016: 24; Saadat & Alavi 2017: 62).

In the constructionist view on language learning, constructions, such as the abstract syntactic frames, are learned by grasping *frequency principles* in the concrete input of the target language (for more information on frequency effects on syntactic development, see, e.g. Thompson & Newport 2007; Williams & Rebuschat 2012; see also the criticism presented in Ellis & Wulff 2014: 84). The most frequently recurring target forms have the most influence on the learning process, as learners recognise patterns in these repeated constructions (Eskildsen 2009: 335–336; Wulff & Ellis 2018: 40–41; see also *categorisation* and *schema formation* in, e.g. Behrens 2009: 386). This is the core idea behind approaches that draw on *statistical learning* (Behrens 2009: 392; for a more comprehensive discussion of the principles behind statistical learning, see Bybee 2006; Onnis 2012; Rebuschat & Williams 2012). A topicalised structure (i.e. a structure that does not begin with a subject; see Teleman, Hellberg & Andersson 2001: 69) is an example of a common construction in Swedish (see the discussion below and in Section 2.3.2.5). The cognitive mechanisms that allow us to make use of frequencies in the surrounding world are not specific to learning languages, and it has been argued that statistical learning does not require conscious attention to the learning process, as all statistical learning happens implicitly without such attention (Behrens 2009: 389; Onnis 2012: 206; Rebuschat & Williams 2012: 2–3; Wulff & Ellis 2018: 41; see, however, Ellis et al. 2016 and the discussion of L2 learning below).

A distinction is made between type and token frequency. A high frequency of a certain type of construction can facilitate the recognition of similarities between constructions (e.g. topicalised structures in Swedish), whereas a high token frequency forms strong traces in memory and can lead to multiword units or chunks being stored in memory as unanalysed wholes (e.g. certain words frequently used in

certain structures, e.g. *Idag ska vi = Today will we* in a topicalised structure) (Bybee & Hopper 2001: 9; Ellis 2003: 71–72, 2015: 50–52; Behrens 2009: 386, 399; Verspoor & Behrens 2011: 28–29; see also Eskildsen 2009). This indicates that learners can learn certain types of structures more quickly and effortlessly depending on the token frequencies and the variation of types in the target language (Behrens 2009: 395). Furthermore, it is beneficial for the learning process if the exemplars are used in various contexts, including in the learners' *output* (see the discussion in, e.g. Ortega 2018). Expertise requires a great amount of practice, and strong representations require significant input and output (Ellis 2019: 48; see also Bybee 2008). In addition to pure frequencies, the *availability*, *reliability*, and *validity* of the occurring pattern affect the learning process (Verspoor & Behrens 2011: 27; see also MacWhinney 2015: 20). It is also to note that linguistic input can vary both in quantity and quality – for instance, in natural vs educational settings – and that the input received by child and adult learners is often drastically different (Littlemore & Juchem-Grundmann 2010: 3; Wulff & Ellis 2018: 50). Furthermore, the learner does not fully process and internalise all the input available to them, and parts of the input, especially grammatical cues and closed-class words, are not necessarily part of the learner's *intake* or output (Wulff & Ellis 2018: 42; see also Corder 1967; Krashen 1985; Gilabert, Manchón & Vasylets 2016). In the current study, the effect of input frequency on the use of topicalised structures and dependent clauses with a sentence adverbial is discussed alongside the results of the first and second substudies (see Sections 4.1 and 4.2; see also Section 2.3.2.5).

The effect of input frequency is essential, for instance, in the emergentist, connectionist, and exemplar-based views on language learning (see, e.g. Behrens 2009: 390; Spoelman & Verspoor 2010: 536; Ellis & Wulff 2014: 78–79; Wulff & Ellis 2018: 39–40; Ellis 2019: 50; see also Behrens 2009: 387–388; Onnis 2012: 203). In cognitive and usage-based theories of SLA, the importance of input frequency is also recognised, but the effects of frequency are considered more limited in L2 learning (see, e.g. Van Patten & Williams 2014; see also Ortega 2014: 262–266 and the discussion in Ellis & Wulff 2014: 84–85). These theories hold that much learning can take place implicitly through frequent exposure to the target language, but they also highlight the essential role of explicit teaching (Ellis & Wulff 2014: 86–87; Van Patten & Williams 2014: 9–11).

Even though L1 and L2 learning share many cognitive elements and are in many ways qualitatively similar cognitive processes, the processes of learning first and additional languages differ in some respects (Ellis 2015: 72–74; Wulff & Ellis 2018: 50; possible differences in the development of L1 systems are discussed in Larsen-Freeman 2006: 601). In addition to the effects common to L1 and L2 learning, L2 is informed by a variety of interrelated factors, such as learner aptitude and motivation, the learning context (including quantity and quality of input), the learner's L1 and

possible additional languages, and several cognitive processes, including perception, attention, learning, and memory (see, e.g. De Bot, Lowie & Verspoor 2007: 14–15; Ellis 2015: 50; Lowie & Verspoor 2019: 186–188; see also Bybee 2008). When learning a new language, the language learner must learn new linguistic elements, and various factors can affect the learning process (see, e.g. DeKeyser 2005; Wulff & Ellis 2018). Three phenomena are often associated with the challenges of learning new linguistic structures: the *salience of the target form* (i.e. the most noticeable stimuli are processed and learned more easily; Wulff & Ellis 2018: 43–45), the *contingency of form–function associations* (i.e. the most reliable associations are processed and learned more easily; Wulff & Ellis 2018: 45–47), and *learned attention* (i.e. the learners' prior associations with language[s] affect the process of learning a new target language; Wulff & Ellis 2018: 47–50; see also De Bot et al. 2007: 14; Verspoor & Behrens 2011: 30).

The effect of L1 on the structures that learners produce is discussed later alongside the results of the first and second substudies (see Sections 4.1 and 4.2). Unlike in some earlier conceptions of the influence of L1 on L2 learning, this effect is no longer seen as the most important, but both positive and negative effects of both L1 and other L2s on a target L2 are acknowledged in SLA research (Verspoor & Behrens 2011: 30–31; Ellis & Wulff 2014: 88; see also the discussion in Spoelman & Verspoor 2010: 534). As L2 learners have prior experience with language(s), these learners operate with systems of constructions in several languages, which impacts the learning process (Ellis & Wulff 2014: 82; Wulff & Ellis 2018: 50; see also Ellis 2003; Onnis 2012; and the discussion in Ortega 2014: 254–256). L2 learners must relearn, for instance, cognitive habits, categorisation patterns, the use of attention-directing devices, and the relationships between function and form (Littlemore & Juchem-Grundmann 2010: 2; Ellis 2015: 59, 61; Ellis & Wulff 2014: 82–83; Wulff & Ellis 2018: 40; see also MacWhinney 2015). This relearning draws on multiple complex cognitive mechanisms (see, e.g. *blocking* and *transfer* in Ellis 2015; Wulff & Ellis 2018). Therefore, it has been argued that, in contrast to L1 acquisition, additional languages cannot be learned entirely implicitly and seem to require explicit teaching and conscious processing (see, e.g. Ellis & Wulff 2014: 83, 88; Ellis 2015: 59–60; Wulff & Ellis 2018: 38, 42, 50–51; see also the discussions in Segalowitz & Lightbown 1999: 48–49; Verspoor & Nguyen 2015: 312). Explicit instruction can facilitate learners' conscious processing of new target language forms, which can facilitate future implicit processing of these forms (Wulff & Ellis 2018: 51; for further information on, for instance, usage-based L2 teaching, see Ellis 2019: 53; see also Ortega, Tyler, Park & Uno 2016). L2 learners can balance their current representations of the target language with the information gained through explicit instruction (Ellis 2019: 46–47; see also Ortega 2014: 262–266). These points

will be important to keep in mind when the pedagogical implications of the current study are discussed (see Section 5.2).

Research on the effects of explicit instruction has shown that learners can benefit from conscious processing of a target language, but, for many reasons, some related to processability (see Pienemann 1998; Pienemann & Lenzing 2014; Gilabert et al. 2016; see also Ortega 2014: 262–266 and the discussion in Section 2.2), learners cannot always use the explicit knowledge they gain in future language production (Ellis & Wulff 2014: 88). Therefore, there is a difference between knowing a structure in the target language – for instance, an abstract syntactic frame – and being able to apply that structure in communication (Ellis 2019: 52). According to some views, declarative knowledge (i.e. explicit knowledge, know-what, facts about the target language) precedes procedural knowledge (i.e. implicit knowledge, know-how, automaticity) in language learning, and only declarative knowledge can be accessed through conscious awareness (for further discussion, see Krashen 1982; DeKeyser 2003; Hulstijn 2005; Ellis, R. 2009; Åberg 2020). These different types of knowledge are relevant to this study's discussion of the differences in syntactic complexity between learners' written and spoken productions.

The approaches discussed in this section acknowledge that there are a multitude of interacting variables at several levels of language learning, such as the target language itself, learners' processing of the target language, and meaning construction in interaction. Thus, these approaches can find common ground in DST (De Bot et al. 2007: 7; Verspoor & Behrens 2011: 28) but also share some views with the more traditional approaches, such as those based on PT. In the following section, dynamicity and systematicity in learner language and language learning are discussed from the perspectives of dynamic systems and processability.

2.2 Dynamic and systemic approaches to language learning

There is no general agreement in theories of language learning about whether the development of learner language is more dynamic or systematic, or about whether variability should be considered important for the learning process (van Dijk et al. 2011: 56–60; Lowie & Verspoor 2015: 64, 82). Two common views on this matter – those rooted in DST and PT – offer differing explanations of how learner language develops and, ultimately, can be described, but these views also agree on several characteristics of learner language (see the discussion in, e.g. Baten & Håkansson 2015; Dyson 2021). In this section, the approaches based on DST (including the *Dynamic Usage-based Approach*) and PT are presented, and their similarities and differences are discussed.

In recent years, many SLA studies have adopted the DST framework (see De Bot & Larsen-Freeman 2011; Kuiken et al. 2019; Bulté & Housen 2020a; Rosmawati 2020) and applied a combination of DST- and usage-based theories (i.e. the Dynamic Usage-based Approach, see, e.g. Lesonen et al. 2020; Lowie, Michel, Keijzer & Steinkrauss 2020) in studies of, for instance, the interaction between systems and subsystems (e.g. complexity and accuracy) in L2 learners' production and the development of L2 learners' proficiency (see, e.g. Vyatkina et al. 2015; Bulté & Housen 2020b). The origins of DST lie in mathematics, and its application to SLA can be traced back to Larsen-Freeman (1997) (see, e.g. De Bot & Larsen-Freeman 2011: 9). In DST, **systems are defined as 'groups of entities or parts that work together as a whole', while subsystems are nested within these systems at a lower level in the hierarchy** (De Bot & Larsen-Freeman 2011: 8). The nonlinearity of the learning process stems from the constant self-reorganisation of these interactive systems and subsystems (Lesonen et al. 2020: 3). In addition to the interaction between the systems and subsystems, all of these systems are in constant interaction with the learners' resources, which also interact with each other (De Bot et al. 2007; De Bot & Larsen-Freeman 2011). In this complex, dynamic system, a change in one system or subsystem can affect other systems and subsystems, possibly resulting in comprehensive changes in the system as a whole (e.g. Lesonen et al. 2020: 3). Furthermore, these complex interactions between the systems and subsystems are prone to change over time (Spoelman & Verspoor 2010; see also Section 2.3.2.2). Therefore, DST interprets L2 learning as a **complex, dynamic process** that takes place in various linguistic dimensions and affects a multitude of linguistic aspects of learners' productions (De Bot et al. 2007; De Bot & Larsen-Freeman 2011; Lowie 2013). The effect of dynamicity on the interconnectedness of systems is essential in the first substudy (see Section 4.1), but it should also be considered when interpreting the results of the other substudies (see Section 5.3).

Additionally, learners themselves are seen as dynamic, complex systems (see, e.g. De Bot & Larsen-Freeman 2011: 18; De Bot, Lowie, Thorne & Verspoor 2013: 200). Inter- and intraindividual variability add to the dynamicity of learner language development (see, e.g. Larsen-Freeman 2006: 596–597). As learner language and language learners impose innumerable dependent variables on the process of language learning, this process can be seen as both complex and dynamic, as well as highly **individual** (Lowie & Verspoor 2015: 67, 69, 2019: 191; see also De Bot et al. 2013). This is the core idea behind the third substudy, which focuses on individual variation in syntactic complexity. Individual variation is considered especially important in the views based on DST, as it is variation that is assumed to drive L2 learning forward (see, e.g. Verspoor, Lowie & van Dijk 2008: 217; Spoelman & Verspoor 2010: 535; Lowie & Verspoor 2015: 75–76, 2019: 202). Apparent regularities in L2 development are seen as descriptions, generalisations, or

'statistical abstractions' (Larsen-Freeman 2006: 598; De Bot & Larsen-Freeman 2011: 19; Ellis & Wulff 2014: 87; see also Ellis 2007; Lowie & Verspoor 2015).

Emphasising individual variation is nothing new in the field of SLA, but the emphasis has varied over time (see, e.g. Larsen-Freeman 2006: 591; Kuiken et al. 2019: 162; Lowie & Verspoor 2015: 74–75, 2019: 186; Bulté & Housen 2020a: 212–213; Dyson 2021: 2–3; see also the discussion in Pallotti 2015). As opposed to many other views, in DST, individual variation is not considered unwanted 'noise' but rather viewed as the sound of a changing system (De Bot et al. 2007: 14; Spoelman & Verspoor 2010: 533; see also Bulté & Housen 2020a). In other words, DST studies argue that not all variability can be explained by external factors, such as task-related ones, as variability is 'an intrinsic and central property of a self-organizing, dynamic system' (Verspoor et al. 2008: 229; see also Spoelman & Verspoor 2010: 550–551). The DST scholars agree that learner language is characterised by a high degree of both inter- and intraindividual variation (Bulté & Housen 2020b: 53; for a summary of a selection of studies, see Section 2.3.2.4). Therefore, according to DST, individual variation should not be excluded from the analysis but seen as important evidence of development in the learner language system.

Despite the growing interest in this theoretical framework, DST-based studies have been criticised on the grounds that the mathematical foundation of DST does not carry over well to SLA, and thus, the application of DST in SLA studies has predominantly been superficial (e.g. Bulté & Housen 2020a: 208–209, 217, 230). However, the main principles of DST (see De Bot et al. 2007: 8, 15; see also De Bot & Larsen-Freeman 2011: 9) have a clear connection to language and language learning, as it is widely agreed that these phenomena form complex, dynamic systems. DST does not consider L2 learning a product, or a set of target forms to be acquired; rather, it is an emergent cognitive process that takes shape in the constant interaction between social, cognitive, and environmental factors (see, e.g. Larsen-Freeman 2006; Verspoor et al. 2008; Lowie 2013), that is, through the interaction between input and output, and between the learners' resources, the social dimension, and the environment (De Bot et al. 2007: 13, 19; Bulté & Housen 2020b: 50–51; see also the discussion in Section 2.1). Therefore, according to DST, learner language development is more accurately characterised by constant fluctuation and nonlinearity than by linear and systematic development (Larsen-Freeman 2006: 590; Verspoor et al. 2008: 215; De Bot & Larsen-Freeman 2011: 10–12; Lesonen et al. 2020: 3; see also Lowie 2013). This change is maintained by a flow of energy and, at the same time, limited by the restricted resources of the learner (Larsen-Freeman 2006: 593; De Bot et al. 2007: 14; Verspoor et al. 2008: 214).

Learners' resources, whether internal (e.g. learning capacity) or external (e.g. input from the linguistic environment), are interconnected in the learning process (De Bot et al. 2007: 11–12; Spoelman & Verspoor 2010: 535; De Bot & Larsen-

Freeman 2011: 13–14; Verspoor & Nguyen 2015: 309; Lesonen et al. 2020: 3; on the interaction between learner systems, see De Bot et al. 2007: 14). These resources can be limited and are often distributed across several subsystems (Larsen-Freeman 2006: 593; De Bot et al. 2007: 11–12; Verspoor et al. 2008: 214; Spoelman & Verspoor 2010: 535). A lack in one domain may be compensated for by efforts in other domains, and changes in some resources can affect other resources. All in all, learners' resources are closely connected to the interactions between the linguistic subsystems in a given moment. This will be a key aspect of learner language to remember for the first substudy, which focuses on the effect of language proficiency on the interplay between syntactic complexity and accuracy (see Sections 2.3.2.2 and 4.1).

It is generally assumed that the interactions between linguistic systems and subsystems – for instance, between complexity and accuracy – can vary in nature (Larsen-Freeman 2006: 592–593, 597). For instance, growth in certain systems and subsystems can precede growth in other systems (De Bot et al. 2007: 19; Lesonen et al. 2020: 2). This is called a precursor or conditional interaction (Lesonen et al. 2020: 3). In a competitive relationship, the same resources are needed for the development of different systems, resulting in competition, whereas in a supportive relationship, different systems can support each other's growth (Spoelman & Verspoor 2010: 535–536; Lesonen et al. 2020: 3–4). The interaction between systems is dynamic, as is the process of L2 learning as a whole. Therefore, the interaction can be either weak or strong, and it is prone to changes over time (Spoelman & Verspoor 2010: 536; Lesonen et al. 2020: 3–4; see also Verspoor et al. 2008). These types of interactions are further discussed in Section 2.3.2.2.

This dissertation adopts the main principles of DST when it comes to the characteristics of learner language and the roles of inter- and intraindividual variation (Larsen-Freeman 2006: 591–594; De Bot & Larsen-Freeman 2011: 9; Bulté & Housen 2018: 149; Kuiken et al. 2019: 162; Rosmawati 2020: 114–115, 125). It is fully acknowledged that in SLA studies in general, and in the current study in particular, the DST framework is applicable only to a certain extent. As the literature shows (see, e.g. Bulté & Housen 2020a, 2020b), the field has only recently begun to explore the various methodologies that can be useful within this framework (see also the discussion of variability studies with established methodologies in Lowie & Verspoor 2015). In fact, it has been suggested that, in the field of SLA, DST could be interpreted as a view of change in learner language, not as a specific theory (Spoelman & Verspoor 2010: 534; Lowie & Verspoor 2015: 80–81).

As noted above, learner language development has been examined using diverse theories. In the context of L2 Swedish, PT has commonly been applied to study developmental paths in syntax (Pienemann 1998; see also, e.g. Pienemann & Håkansson 1999; Håkansson 2004; Håkansson & Norrby 2007; Norrby &

Håkansson 2007; Rahkonen & Håkansson 2008; Baten & Håkansson 2015). Originally, this theory was based on spoken learner language data, but later, it was applied to both modes of L2 production (Pienemann 1998; see also Håkansson & Norrby 2007: 82, 91). PT is designed to illuminate how target language is processed at various phases of L2 development and describes L2 development using *stages of processability*, which seem to be shared by many learners, regardless of their age and learning context (Pienemann & Håkansson 1999: 408–417; see also Ellis & Wulff 2014: 87; 89), in all L2s. It also accounts for the apparent variety, for instance, of grammar systems in different languages (Pienemann & Lenzing 2014: 159–160, 173; see also the discussion in Lowie & Verspoor 2015).

PT assumes a **processability hierarchy** according to which grammatical information can be transferred both within and between phrases (Pienemann & Lenzing 2014: 161; Dyson 2021: 4, 18; see the discussion of developmental sequences in, e.g. Baten & Håkansson 2015). To transfer grammatical information at a given level, the learner must have reached a certain level of processability and developed the procedures required to form the target language construction (Pienemann & Lenzing 2014: 161–162, 164; see also Pienemann & Håkansson 1999). These procedures are learned in a hierarchical order; to master a subsequent procedure, the learner must master the preceding one, and this procedure cannot be skipped, even with explicit teaching (Pienemann & Lenzing 2014: 163, 176–177; Baten & Håkansson 2015: 522; Dyson 2021: 4, 18; see the *developmental ladder*, Larsen-Freeman 2006: 590, and *conditional development*, Lesonen et al. 2020: 3). In other words, the theory holds that learners can only process the linguistic material that is processable at a given stage of development (Pienemann 1998; Pienemann & Håkansson 1999; Pienemann & Lenzing 2014), according to the following order: First, no procedures are applied to language production, and the learner produces only simple words. Next, words can be categorised accordingly, for instance, by means of conjugation, after which grammatical information can be transferred – first within noun phrases and then within verb phrases. Subsequently, it becomes possible for the learner to transfer grammatical information within a simple sentence and, finally, between a main clause and a subordinate clause. Despite its apparent focus on the development of target-like morpho-syntactic structures (e.g. Dyson 2021: 2, 4; see, however, Baten & Håkansson 2015: 541), PT is considered important for the current study in relation to the emergence of cognitively more demanding structures in learners' productions (see Section 2.3.2.5).

The focus of PT-inspired studies is often the differences between developmental stages, but many studies also account for individual variation in developmental trajectories (Pienemann & Lenzing 2014: 160; Dyson 2021: 1). Although the developmental stages are considered hierarchical and obligatory, and the grammatical rules are characterised by similar processing, the underlying principles

in PT allow for individual variation in the developmental paths (see *hypothesis space*, e.g. Pienemann & Lenzing 2014: 164; Baten & Håkansson 2015: 521–522, 529–530; Dyson 2021: 5). For instance, within a developmental stage, learners can produce different variants of the target construction. Additionally, when attempting to produce structures that require higher-level procedures, learners use varying strategies to formulate structures, resulting in individual variation (Pienemann & Lenzing 2014: 160, 164, 170–172; Dyson 2021: 1, 5). In the current study, this variation can be seen as an integral part of production by learners at the lower proficiency level.

The DST-based and PT-based approaches can be contrasted, for instance, in their views on systematicity in language development (see, e.g. Baten & Håkansson 2015; Lowie & Verspoor 2015; see also Ortega 2014: 251). While PT considers L2 development to be a stepwise, systematic, and linear process that can be described using predictable cause-and-effect models, DST-based views hold that the developmental process is not stepwise, systematic, or linear due to self-organisation and constant interaction between subsystems (see the discussion above). Contrasting interpretations of both the PT- and DST-based approaches can also be found in earlier discussions of various factors related to language development (see, e.g. Baten & Håkansson 2015; Lowie & Verspoor 2015; see also the *all-or-nothing position*, De Bot et al. 2007: 17), often generating conceptions of full systematicity or full randomness. However, there are some resemblances between the theoretical views of these approaches, enabling less-extreme interpretations. Even though the linguistic patterns in the developmental paths are considered highly systematic in PT, the interaction between the subsystems of the target language is acknowledged; therefore, language development can be seen as a dynamic process (Dyson 2021: 1). Thus, both DST-based theories and PT argue for dynamicity, at least to some extent, in language development. The difference is that DST holds that the variation is unsystematic, whereas PT considers it systematic. It is noteworthy that there is some systematicity in the unpredictable, dynamic, complex system of language learning, and that complexity and chaos in language development should not be confused with randomness (De Bot et al. 2007: 14). However, in DST, these regularities are explained by reference to, for instance, similar conditions in the early stages of development, attractor and repeller stages, and similar input and environmental factors, and these regularities in development are seen as mere generalisations (Larsen-Freeman 2006: 615; Lowie & Verspoor 2019: 192; Lesonen et al. 2020: 3; see also the discussion above). Additionally, DST holds that even during stable situations (e.g. attractor states), there is variability (Spoelman & Verspoor 2010: 535). It is also specified that just because a given learner is capable of doing something, they may not necessarily produce a certain structure in a predetermined manner (Larsen Freeman 2006: 591; De Bot et al. 2007: 16). Nevertheless, it is

possible that the learner language of this individual has developed according to the developmental stages (see the discussion in Baten & Håkansson 2015: 519). In other words, regularities can be masked by individual variation, and these approaches can be considered two sides of the same coin (Baten & Håkansson 2015: 539; see also Dyson 2021).

According to DST, stage-like tendencies in language use can describe how learner language develops 'in the big picture', and these tendencies can be studied, but they rarely predict an individual's developmental path or language production in different contexts (Ellis 2019: 47; see also De Bot et al. 2007). This statement aptly describes the perspective endorsed in the present study, which recognises the stepwise complexification of structures but considers variation to be a dynamic phenomenon in which the learner has an active role (see Larsen-Freeman 2019), not a mere choice between structures on a continuum (Dyson 2021: 2). As the focus of the current study is not on the developmental stages of morphological and syntactic development but rather on individual variation and developmentally meaningful changes in syntactic complexity at the group level (see the discussion below), the study does not adopt the research methodologies favoured in PT studies (see *emergence criterion* and *implicational scaling*, Dyson 2021: 115–116; see also Baten & Håkansson 2015 for a comparison between DST- and PT-inspired research methodologies). In this dissertation, it is assumed that **syntactic complexity in learner language progresses following a certain developmental path, but individual learners fluctuate in implementing different structures,** as well as target- and nontarget-like forms of these structures, in their productions for various reasons. Therefore, the big-picture developmental path may be the same for all learners, but these paths at the individual level are not entirely linear or consistent (Larsen-Freeman 2006: 601, 615; see also Lowie & Verspoor 2015).

In the discussion of the differences between DST and PT in earlier literature, it became evident that various misconceptions can be conveyed due to the terminology used in different approaches to SLA (see the discussion in, e.g. Housen et al. 2012a; Baten & Håkansson 2015; Dyson 2021). Therefore, it is important to specify how *development*, and the related terms *change* and *variation*, are defined in the current study. Based on previous research, there are clear differences between these notions. Nevertheless, they are often used interchangeably and, in many cases, not defined in enough detail (e.g. Dyson 2021: 3; see also Kowal 2016).

In many SLA studies, the aim is to trace language development. In these studies, particularly the PT-inspired ones, development has often been interpreted as a diachronic process of gradual progress in the (sub-)systems of learner language (see, e.g. Dyson 2021: 1, 4; see also Kowal 2016: 16, 20, 22). SLA researchers seek appropriate tools to measure this development in its traditional meaning (however, see Kowal 2016: 18, 23). For instance, it is firmly believed that CAF measures can

be used to evaluate language development (e.g. fewer errors imply improved accuracy), although the picture is complicated by the fact that the multidimensional and dynamic nature of language learning leads to nonlinearity (Housen et al. 2012b: 1–2, 5, 7–8; Polat & Kim 2014: 186; Kowal 2016: 18; see also Baten & Håkansson 2015). It seems that several researchers, especially those who hold to the DST and PT approaches, often describe development as linear and change as nonlinear (cf. Spoelman & Verspoor 2010: 536; Lesonen et al. 2020: 3). In De Bot et al. (2007: 19), learner language is described as a complex, dynamic system characterised by change rather than development (see also Lowie & Verspoor 2015: 74). Thus, the notions of change and development are considered separate or even opposed phenomena in language learning (cf. Kowal 2016: 20). When variation is discussed in the context of development, it is often seen as systematic, especially in PT-based views (see, e.g. Pienemann & Lenzing 2014), whereas in the context of change, variation is described as more dynamic, especially in DST-oriented views (see, e.g. Lowie & Verspoor 2015).

Sometimes, the term 'development' is replaced by 'gain' or 'growth' (Kowal 2016: 20; see also Baten & Håkansson 2015; Housen et al. 2012a; for the difference between development and acquisition, see De Bot & Larsen-Freeman 2011: 5–6). Particularly in complexity studies, it is not uncommon to find arguments against the assumption that complexity increases together with L2 proficiency (see, e.g. Pallotti 2015). Here, again, it would be necessary to define in more detail what exactly is meant by 'development'. In DST, it is often claimed that there is no direction to development, and there is only change (Verspoor et al. 2008: 217; De Bot & Larsen-Freeman 2011: 13). Changes in learner language can be both progressive and regressive, and both forms of change can be seen as variability in the learner language and considered development (Verspoor et al. 2008: 217; Lesonen et al. 2020: 2–3; see also De Bot et al. 2013). On the contrary, Baten and Håkansson (2015: 539) point out that certain measures frequently used to measure L2 development, especially measures of embeddedness (e.g. dependent clause ratio), can reveal variation in learner language but cannot properly be used to measure development. They state that the notions of variation and development seem to coincide in DST-based studies and argue further that the methods used in studies based on PT better suit accounts of development.

In this dissertation, development, change, and variation are considered separate but closely interrelated phenomena (see, e.g. Dyson 2021). The focus lies on change and variation, not on development in its traditional meaning. In other words, the aim is not to show how syntactic complexity develops per se but rather to investigate proficiency-related differences in learners' use of syntactically complex structures. A distinction is made between **developmentally meaningful change** (Articles I and II) and **change in the form of individual variation** (Article III). The former closely

resembles the traditional meaning of development in that it is linked to the differences between proficiency levels, but in line with DST, the change does not need to be progressive in order to be seen as developmentally meaningful. The latter is seen as variation caused by individuals' preferred styles of producing language. In addition, however, developmental and contextual factors, such as the mode of production, can have a major effect on the individual variation considered in the present study.

In sum, the theoretical underpinnings of this dissertation are drawn from multiple approaches to language learning. In line with the theoretical premises discussed in this section, it is assumed that when learning a new language, the learner receives **input** from the language community or other linguistic materials in different contexts, processes it through **earlier experiences**, identifies **patterns** in the frequently occurring structures, develops **hypotheses** about the target language system, and finally, broadens and deepens their knowledge of the target language by testing these hypotheses (e.g. abstract syntactic frames) in their linguistic **output** and interactions (Littlemore & Juchem-Grundmann 2010: 1; van Dijk et al. 2011: 84; see also Ellis 2015, 2019). In this dissertation, the effects of context, such as social and physical context, on language learning and language use are also acknowledged (for more information about the sociocultural theory as well as interactional and social factors, see, e.g. Cadierno & Ellis 2015; see also Ellis 2015: 61–63). Physical context is regarded as especially important, as the mode of production is one of the major sources of differences in the use of syntactically complex structures both between and within individuals (see, e.g. Verspoor, De Bot & Lowie 2011). Interactional context (see, e.g. Bybee 2010) is also recognised, but the effects of interaction fall outside the scope of the current study.

This study further assumes that learning processes are not independent of one another but rather in constant mutual interaction (Ellis 2019), and that these processes are also affected by **learners' limited resources**, which form a part of the interactional system (Larsen-Freeman 2006: 593; Verspoor et al. 2008: 214). In line with the main principles of DST, this study argues that learner language contains multiple **systems and subsystems** that are **interconnected**, that both **progress** and **attrition** can be seen as development, and that all **learners learn and use language differently** (De Bot et al. 2007; De Bot & Larsen-Freeman 2011; De Bot et al. 2013; Lowie 2013; Lowie & Verspoor 2015, 2019). The DST framework plays the most prominent role in the last substudy, where the focus lies on individual variation. More traditional approaches are considered suitable when focusing on, for instance, developmentally meaningful differences between proficiency levels. As the first two substudies examine group-level tendencies, a more traditional approach is adopted. Therefore, the study combines more traditional methods with DST-based approaches

to studying syntactic complexity. Syntactic complexity as a part of the CAF triad and as a multifaceted aspect of learner language is discussed in the following section.

2.3 Complexity, accuracy, and fluency (CAF) and syntactic complexity

This section provides an overview of the role of syntactic complexity in the CAF research tradition and as a feature of learner language. The section begins with a presentation of the CAF triad, followed by a critical discussion of the definitions of (syntactic) complexity in Section 2.3.1. In Section 2.3.2, various factors that affect syntactic complexity are discussed along with a selection of earlier research. Additionally, the emergence of syntactically complex structures in L2 Swedish is discussed. The section concludes with an overview of how syntactic complexity is referred to in CEFR (2001, 2007, 2020).

The CAF triad provides a framework for analysing proficient language use – one that suits the multidimensional nature of learner language (Polat & Kim 2014: 186). As all CAF dimensions, or systems, are multidimensional, the definitions of these systems vary, but the fundamental meanings can be summarised as follows: Complexity refers to the use of varied language with structures of differing length and composition suited to the context in which the language is produced, whereas accuracy refers to the need to produce language according to the normative rules of the given target language, and fluency refers to the ability to produce the target language relatively rapidly without unnecessary pausing or repair (see, e.g. Wolfe-Quintero et al. 1998; Housen et al. 2012a). Accuracy and fluency were already established areas of SLA research when complexity was added as a third component to systems of language proficiency (see, e.g. Housen et al. 2012b: 10; Bulté et al. 2024: 4). As noted earlier, complexity has been investigated from several points of view, and studies have focused on complexity both individually and together with other CAF systems (Bulté & Housen 2018: 148; Kuiken et al. 2019: 162–163; see also Housen & Kuiken 2009). In previous research, complexity has been seen as a dynamic property of learner language performance, and its development has been studied as part of increasing overall L2 proficiency. The interaction between complexity, accuracy, and fluency, as well as between the subsystems that comprise these systems, such as complexity at the clausal and phrasal levels, has been a special area of interest. The aim of many of these earlier studies was to examine how syntactic complexity is affected by various factors, of which L2 proficiency, modality (spoken vs written), and individual variation are highlighted in this dissertation (see also Kuiken et al. 2019: 163–164).

The CAF triad and its use in SLA research have also been criticised, especially when CAF measures are used to study learner language development (see, e.g.

Pallotti 2009; Baten & Håkansson 2015). Complexity is considered the most complicated dimension of the triad, and as will be shown in this section, its investigation as a part of L2 learning and use has proven rather complicated, mainly due to its multidimensionality and challenges in the operationalisation of this construct (Kuiken et al. 2019: 162; see further discussion in Section 3.3.2). Many researchers have concluded that there is a clear correlation between more complex language and overall L2 proficiency, but researchers disagree, for instance, on whether all subdimensions of complexity develop simultaneously, what the most significant influences on complexity are, and how complexity interacts with accuracy and fluency.

According to the CAF hypotheses, learner language becomes increasingly complex (to a certain extent; see also Section 2.3.2.1), accurate, and fluent as the learner becomes more proficient (see, e.g. Wolfe-Quintero et al. 1998; Housen et al. 2012a). Previous studies have shown that many learner-related factors, such as age and previous experience with languages, can influence how CAF dimensions develop in their language use. Learners can start with the production of simple structures that first become accurate, and then, by reorganising and combining the structures acquired in their productions, they can start to produce more complex language (CEFR 2007: 111; Kowal 2016: 39–40, 204–205, 216; see also Verspoor & Behrens 2011). According to Kowal (2016: 216), complexity is the dimension that requires the most time to develop, whereas accuracy and fluency seem to show rapid growth even at the beginning of the learning process. This order seems to apply especially to younger or beginner-level learners, as they tend to have a narrower repertoire of structures at their disposal in the initial stages of the learning process. On the other hand, learners may already possess a wide repertoire of structures at the beginning that eventually become more accurate (see, e.g. Housen et al. 2012b: 7). For instance, Lesonen et al. (2020: 2, 3) state that, based on earlier research, complexity in certain linguistic features precedes accuracy in the development of L2 English and L2 Finnish (see also Baten & Håkansson 2015: 4). This might be more typical of older and more experienced learners. These learners have more experience with language(s), and they can already express themselves in a complex way in one or several languages (Norrby & Håkansson 2007: 45; see *cognitive maturity*, Foster, Tonkyn & Wigglesworth 2000: 355; Vyatkina et al. 2015: 42). However, adult learners in particular might consciously prioritise accuracy at the expense of complexity to avoid making errors (e.g. Norrby & Håkansson 2007: 46, 55).

Furthermore, complexity is a context-bound construct, meaning that different types of structures are expected to be used in various contexts (see, e.g. Biber et al. 2016). Moreover, the expected outcome for complexity differs from those of the other CAF systems. For accuracy and fluency, the desired result is accurate and fluent language, whereas with complexity, the aim is not to produce highly complex

language at all linguistic levels. In fact, learners at higher proficiency levels learn to balance their use of structures according to the context, and it is possible that intermediate learners produce more complex language than advanced learners at the level of clausal-level complexity (see, e.g. Ferrari 2012). When it comes to complexity, there are neither unwanted characteristics unequivocally expected to decrease in number (cf. e.g. errors for accuracy and unnecessary pauses for fluency) nor wanted ones that are expected to increase in number (cf. e.g. error-free structures for accuracy and more words between pauses for fluency) as the learner becomes more proficient (see, however, the discussion of fluency features in Peltonen 2020).

2.3.1 Types of complexity

Various aspects of learner language have been described in terms of complexity, and as already mentioned, definitions of complexity vary depending on the research objectives and the point of view adopted in the study (see, e.g. Ortega 2003; Norris & Ortega 2009; Bulté et al. 2024). This variation is primarily a result of the inherently multidimensional nature of complexity in learner language. Due to this multidimensionality, the construct is approached from multiple perspectives and studied at different linguistic levels (De Clercq & Housen 2017: 316; Bulté & Housen 2018: 148; see also Housen et al. 2012a). The use of varying definitions and, for instance, disagreement on the role of relative complexity (see the discussion below) have resulted in major differences among earlier study designs and operationalisations of syntactic complexity. These differences make it difficult to compare the findings gained from the continuously growing body of research, and many aspects of complexity thus remain unclear (see, e.g. Kuiken et al. 2019: 163). For this reason, it is important to place special emphasis on definitions when studying complexity in learner language.

Since complexity is a multidimensional construct, it can be divided into several subsystems. To begin with, we can distinguish between *structural*, *cognitive*, and *developmental complexity* (Bulté & Housen 2012: 23–24; De Clercq & Housen 2017: 316; Wijers 2019: 23; see also Pallotti 2015). Structural complexity is a 'quantitative notion' that refers to 'the number and variety of parts or elements in an entity or system, as well as to the relationships and interactions between the constituent parts' (Kuiken 2023: 84). In simpler terms, structural complexity concerns systems that contain several interacting parts (Pallotti 2009: 593; Bulté & Housen 2018: 149). A related term, *linguistic complexity*, has also been used when discussing this type of complexity. According to Bulté and Housen (2012), linguistic complexity can be seen as a dynamic property of the L2 system, that is, *global* or *system complexity*, or as a more stable property of the relevant linguistic features, structures, and rules, that is, *local* or *structure* complexity. The former refers to the

size and richness of the learner's linguistic repertoire, that is, their range of different structures, whereas the latter refers to the individual linguistic features and their structural complexity (Bulté & Housen 2012: 25).

Structural or linguistic complexity can be divided into lexical and grammatical complexity, and grammatical complexity can further be divided into morphological and syntactic complexity (e.g. Bulté & Housen 2012: 27). In the current study, the focus lies solely on syntactic complexity. In some categorisations, structural and linguistic complexity have been combined under the term *absolute complexity*; this term is used in the current study to designate all types of system- and structure-related complexity (see Figure 2 below).

Complexity has also been associated with cognitively demanding structures and thus referred to as *cognitive complexity*. This type of complexity refers to the processing demands that a given structure imposes on the L2 learner (see, e.g. Di Domenico, 2017: 1–2; Michel 2017: 52; Bulté & Housen 2018: 148; Wijers 2019: 21). Since this demandingness is interpreted in relation to the learner, cognitive complexity is also referred to as *relative complexity* (or *comparative complexity*). Relative complexity can be determined by feature-, context-, and learner-related factors (Housen & Simoens 2016: 164). Feature-related complexity is associated with the absolute complexity of the linguistic feature but also with its input frequency (see Section 2.1), whereas context-related complexity is caused by factors related to, for instance, the learning environment, and learner-related complexity by factors such as attention and processing capacity (see Sections 2.1 and 2.2). Feature- and learner-related features are central to the present study (see Section 2.3.2 and Figure 2).

The term 'relative complexity' is used in this dissertation. This is complexity defined 'in relation to the language user and the cognitive cost or difficulty invoked in processing and acquiring certain linguistic structures' (De Clercq & Housen 2017: 316). As in this citation, relative complexity is often associated with difficulty (see also Bulté et al. 2024: 9, 19–25). Therefore, the reasoning behind the definition is circular: complex structures are difficult, and structures are difficult because they are complex (see, e.g. Pallotti 2015; see also Baten & Håkansson 2015: 521). Defining complexity in this manner is rather controversial due to the circular reasoning involved (e.g. Kowal 2016: 41; see, however, the discussion in De Clercq & Housen 2017: 316; Di Domenico 2017: 1–2). The notions of 'complex' and 'difficult' (and the related notion of 'complicated') should be kept separate (see the discussion in Dahl 2004: 39; Wijers 2019: 21), as not all linguistically complex structures are difficult to process (Pallotti 2015: 119; Michel 2017: 52).

In some studies, a linear positive correlation is assumed between measures of relative complexity and the learner's proficiency level, which can lead to tautological findings concerning this relationship. It is of note that *more* should not be considered

better or *more advanced* in the case of complexity. In other words, a mere high number of complex structures cannot be directly linked with a higher proficiency level (see, e.g. Bulté & Housen 2018: 149; Kuiken & Vedder 2019: 195). As shown in previous research and discussed later in this dissertation, when it comes to complexity, diversity in the structures produced can be a more important sign of mastery than the number of elements in the production (Iwashita 2006: 164; Ferrari 2012: 282–283, 293). However, based on evidence from earlier research, certain structures are more difficult for L2 learners (see Section 2.3.2.5 for a discussion of such structures in L2 Swedish), and this difficulty has been a focus of complexity studies (De Clercq & Housen 2017: 316; Wijers 2019: 22, 30; see also Bulté et al. 2024). In other words, the relationship between complexity in L2 performance and the learner's proficiency level can be studied but not assumed (e.g. Pallotti 2015: 118–119; Bulté & Housen 2020b: 52). In these cases, relative complexity is investigated, as in the current study. One of the major limitations of the investigation of relative complexity in this study is that the results cannot be generalised to other target languages or compared with results concerning other L2s, as the measures are (partly) language-specific (see Section 3.3.2). On the other hand, the examination of relative complexity with the help of these measures can provide more detailed information about syntactic complexity in a given target language.

Complexity has also been defined as structures that learners acquire late in the learning process (Baten & Håkansson 2015: 521; Di Domenico, 2017: 1–2; Wijers 2019: 21). The basic idea is that learners at higher proficiency levels use more complex structures, as these structures are acquired later. This has obvious connections to relative complexity, as the complexity of a structure is defined in relation to time points in the developmental path of the learner, and the same remarks on tautological findings apply to both relative complexity and developmental complexity. Pallotti (2015: 118) states that assumptions about the relationship between complexity and L2 development should be avoided altogether. It can be assumed, however, that complexity in a given performance can be positively correlated with the learner's current developmental state (e.g. Bulté Housen 2020b: 52). Here, one must remember the wide range of additional factors, such as modality and individual choices, that affect complexity in L2 production. Furthermore, as already discussed, complexity in language use is highly context-bound, and the development of such a multifaceted construct is not straightforward. It has been demonstrated in previous research (see, e.g. Vyatkina et al. 2015) that during L2 development, structural diversity increases, and cognitively more demanding structures become more frequent, whereas the use of less demanding structures can decline (see also the discussion in Section 2.3.2.1). These results depict a complicated picture of complexity in learner language. However, developmental timing can be at least partly connected with cognitive complexity, for instance,

through processability-related factors (see Section 2.2). Consequently, in this dissertation, developmental complexity is understood as part of relative complexity.

Another important type of complexity is *task complexity*. Task complexity refers to the cognitive demands that a given task imposes on the learner (see Skehan & Foster 2001: 194–195, 2012: 215–216, 218; see also Dahl 2004 and the discussion in Section 2.3.2.2). Complex tasks impose heavier cognitive requirements, whereas simple tasks are lighter in their cognitive load. Task complexity encompasses factors such as planning and use of additional information in the production (see, e.g. Skehan 1998; Robinson 2001, 2005; Yuan & Ellis 2003). Of course, whether the learner experiences a complex task as difficult depends on the individual, but this does not affect the internal complexity of the task (Dahl 2004: 39). In fact, some scholars argue that high task complexity can support learners' performance on a task (Robinson 2001; see also Section 2.3.2.2). In short, task complexity plays an important role in all studies on complexity, as it evidently affects the structures that L2 learners produce.

In line with the DST-based views on language learning, complexity is understood here as a system that contains several interconnected subsystems (Bulté & Housen 2018: 148, 2020b: 52; see also Section 2.2). Syntactic complexity is one of the subsystems implicated in complexity, and this subsystem contains many subsystems, such as syntactic complexity at the clausal and phrasal levels. Beyond looking at the number of interacting systems and subsystems, complexity can be understood through quantitative and qualitative measures, such as those based on length and diversity (Ferrari 2012: 282–283; Vyatkina et al. 2015: 42; see also Section 3.3.2). Based on the remarks made in earlier research, it is important that the different types of complexity be kept apart (see the discussions in, e.g. Dahl 2004; Pallotti 2015). However, as this dissertation demonstrates, these different aspects of complexity can and should be studied in parallel to gain more thorough insights into the use of complex structures in learner language. The types and levels of complexity considered in the current study are illustrated in Figure 2.
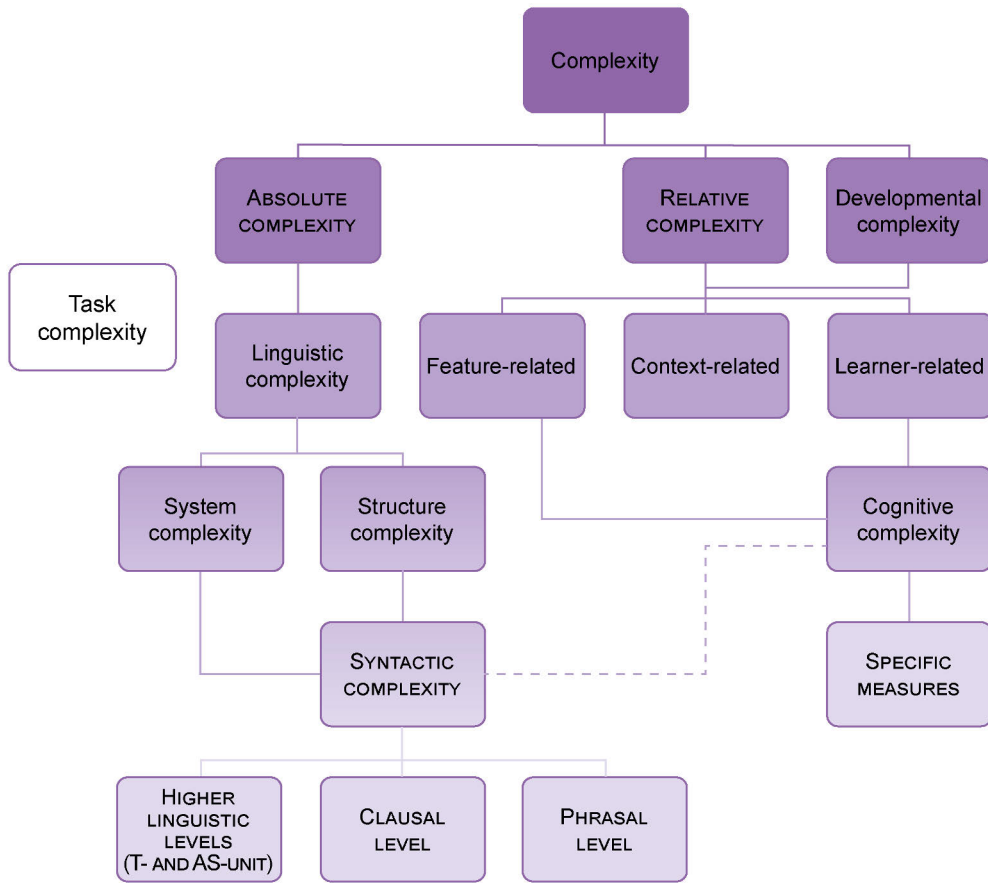
Figure 2. Types and levels of complexity in the current study.

In this dissertation, two types of complexity – absolute and relative (see, e.g. Miestamo, Sinnemäki & Karlsson 2008) – are investigated in learner language, and task complexity is seen as a contributing factor affecting both complexity types. Absolute complexity is seen as objectively indicative of linguistic complexity, and this notion encompasses both system- and structure-related complexity. Absolute syntactic complexity is measured at higher linguistic levels – the levels of *Minimal Terminable Units* (T-units, consisting of a main clause and all possible dependent clauses, as well as all nonclausal units or fragments attached to it, Hunt 1965: 20, 49) and *Analysis of Speech Units* (AS-units, consisting of a main clause, or a subclausal unit, and all possible dependent clauses associated with it, Foster et al. 2000: 365) – and at the clausal level, as well as the phrasal level (see Section 3.3.2). Developmental complexity is interpreted as part of relative complexity, but the focus is restricted to feature- and learner-related factors that affect cognitive complexity. It is seen as a type of syntactic complexity that is separate from absolute complexity,

and it is examined with the help of specific measures (see Section 3.3.2). In the first two substudies, both absolute and relative approaches to syntactic complexity are adopted. Even though complexity is discussed using a relative definition in these studies, and learners at different proficiency levels are compared in terms of their use of syntactically complex structures, the purpose is not to try to show how the learners' language develops, but to examine the use of different syntactic structures in their own right – that is, not as a sign of L2 development. In other words, although relative complexity is operationalised through linguistic features that, based on earlier studies, are demanding for L2 learners of Swedish and acquired late in the learning process, the relationship between relative complexity and proficiency is interrogated, not assumed. In the third substudy, syntactic complexity is approached through a multilevel analysis encompassing the three linguistic levels (see Figure 2) and a qualitative analysis of the learners' productions (see Section 3.4).

Based on the discussion in this section, complexity should be understood as a combination of intrinsic properties of L2 that can be defined in different ways. Due to the multidimensionality of complexity, it is understandably challenging to provide a single, clear-cut definition that covers this expansive notion. Earlier definitions of (syntactic) complexity referred to such phenomena as diversity, variability, and sophistication, as well as the development of the use of complex language (see, e.g. Wolfe-Quintero et al. 1998: 69; Skehan 2003: 8; Ellis & Barkhuizen 2005: 139). A common definition of complexity is 'the ability to use a wide and varied range of sophisticated structures and vocabulary in the L2' (Housen et al. 2012a: 2). While a variety of definitions have been suggested in earlier research, in this dissertation, syntactic complexity is defined as **a structural property of learner language that manifests itself in long and embedded units of language and in the diverse use of these structures** (absolute complexity). Furthermore, **structures that, according to earlier research, are perceived as difficult by L2 learners and acquired late in the learning process are operationalised as measures of syntactic complexity** (relative complexity). From these perspectives, syntactic complexity is examined in order to describe learners' performance and to study both the developmentally meaningful differences between proficiency levels and individual learners' differing styles of producing syntactically complex structures.

2.3.2   Syntactic complexity in learner language

In this section, syntactic complexity in learner language is discussed along with a selection of earlier research. The section begins with a discussion of the effect of language proficiency on syntactic complexity. Additionally, the relationship between language proficiency and syntactic accuracy is briefly discussed, as accuracy is examined in the first substudy. Thereafter, the interplay between

syntactic complexity and accuracy is examined, and various factors that shape syntactic complexity in learner language are presented. Furthermore, the use of syntactically complex structures by L2 learners of Swedish is discussed. Finally, an overview of how syntactic complexity is understood in CEFR (2001, 2007, 2020) is provided.

2.3.2.1 Effect of language proficiency on syntactic complexity and accuracy

Syntactic complexity is often interpreted as an indicator of L2 proficiency. Therefore, in many earlier studies, learners' productions at different proficiency levels were compared in terms of their syntactic complexity (Bulté & Housen 2018: 148; Kuiken et al. 2019: 163, 166; see also Wolfe-Quintero et al. 1998; Housen et al. 2012a). Most of these studies suggest that there is a clear link between syntactic complexity and overall language proficiency, but some studies yielded mixed or even opposed findings (Kuiken & Vedder 2019: 195, 207; see, e.g. Larsen-Freeman 2006; Verspoor et al. 2008; Spoelman & Verspoor 2010; Vyatkina 2012, 2013; Vyatkina et al. 2015; Kuiken et al. 2019). Complexity seems to correlate positively with proficiency level, for instance, in Berggreen and Sørland (2016), as well as in Gyllstad et al. (2014), and for most complexity measures in Kuiken and Vedder (2019). However, Granfeldt (2008) and Kuiken & Vedder (2012a) found partly mixed results. It is extensively agreed upon that, to exhibit development in L2 learning, the learner must acquire, for instance, new syntactic structures (Ferrari 2012: 282), but it remains unclear how this learning takes place and what factors contribute to both the learning and the use of these structures. The relationship between syntactic complexity and the learner's proficiency level is discussed in this section, and possible influences on this relationship are explored with the help of previous research.

The basic assumption behind theories of the developmental path of syntax is that, in the beginning, learners produce simple language, whereas later, they start to produce more complex structures. Verspoor and Behrens (2011: 38) state that learners begin with simple structures, be they certain forms of tense or simple ways to connect clauses, and these structures might be overused in the early phases of learning. Gradually, learners will start to use other structures more variedly, and at the advanced levels, learners can balance between simpler and more complex structures in their productions (see also Ferrari 2012; Khushik & Huhta 2022). Accordingly, Lesonen et al. (2020) state that learners must try out, or even overuse, certain structures to make progress. However, learners at lower proficiency levels sometimes exhibit relatively high complexity in their productions. This might be a

result of a chunk-like use of the target language (Vyatkina 2012: 594; see *formulaic sequences* in Myles 2012).

As they begin to complexify their production, learners can either add more words and phrases or elaborate at the clausal level, that is, produce more embedded clauses (Ferrari 2012: 282). Based on evidence from earlier research, it can be assumed that learners first complexify their productions at the clausal level; therefore, the use of dependent clauses increases at intermediate levels, whereas, at higher proficiency levels, learners tend to complexify their productions at the phrasal level by producing more complex phrases, resulting in longer clauses (Biber, Gray & Poonpon 2011: 26–27; Ferrari 2012: 282–283, 284, 293; Vyatkina 2013: 24; Lambert & Kormos 2014: 608; Kuiken & Vedder 2019: 193–195; Vercellotti 2019: 237; see also the discussions in Verspoor et al. 2008; Kuiken et al. 2019; Khushik & Huhta 2022; Sarte & Gnevsheva 2022). When learners start to elaborate more on the phrasal level, the use of dependent clauses usually decreases. This can be explained by the fact that clausal complexity in learners' productions reaches a plateau, after which complexity measures based on embedding can show a decrease in complexity (Lambert & Kormos 2014: 608; see also Ferrari 2012; Kuiken 2023). Therefore, as certain complexity features become less common while other features increase, the development of complexity can be observed in both the decline and growth of complexity measures, and 'less complex' can also be seen as an indicator of development (Ferrari 2012: 293).

Consistent with this, Lambert and Kormos (2014: 612) state that it may be necessary to reconsider the fundamental assumption that greater complexity in learners' productions implies higher language proficiency. They state that complexity, measured with certain features, can decline after the intermediate levels as the use of complex structures is optimised. Learners at intermediate levels might overuse complex structures, whereas more advanced learners can accomplish tasks with fewer linguistic resources and express complex ideas with efficiency rather than complexity. As previously noted, the use of complex structures is highly context-bound, as in some contexts, such as academic texts, complex structures are expected, whereas in other contexts, such as informal letters, simpler structures are more appropriate (see, e.g. Ferrari 2012: 283; see also Biber et al. 2011; Biber et al. 2016). At higher proficiency levels, learners may have mastered the use of context-appropriate language. Therefore, they may exhibit less complexity in their productions than their intermediate-level peers. Moreover, at these higher levels of proficiency, the importance of learners' stylistic choices becomes more evident (Ferrari 2012: 291; Vyatkina 2013: 25; see also Gyllstad et al. 2014: 22–23 and Section 2.3.2.4). Therefore, in addition to increases and decreases in the number of certain complex structures, the development of complexity can be identified based on increased variation in these structures.

Biber et al. (2011) created an index of the developmental stages of complexity based on large-scale corpora of academic writing and speech. With the help of this index, it is possible to measure differences in complexity not only between proficiency levels (Biber et al. 2016: 645) but also between different types of productions (see, e.g. Ansarifar, Shahriari & Pishghadam 2018; Kyle & Crossley 2018; Sarte & Gnevsheva 2022). In the index, the lower-level complexity features include complexification at the clausal level, which is understood as more common in spoken production, whereas the higher-level features include complex phrases – which are more typical of academic writing – with different stages of complexity depending on the modifier type (see the discussion below of phrasal-level complexity features in this index). According to the index, complexity proceeds from the use of finite dependent clauses to non-finite dependent clauses, and afterwards, to complex phrases featuring a diverse use of modifier types. These general developmental stages seem to hold for most learners, even though large interindividual differences are observed. As pointed out by Kuiken and Vedder (2019: 195), the stages partially overlap in the developmental process; therefore, in actual L2 performance, the stages are not discrete.

Many studies have demonstrated a positive correlation between L2 proficiency and syntactic complexity in **written production** (Bulté & Housen 2018: 148). For instance, Gyllstad et al. (2014) examined the differences in complexity as measured with T-unit length, clause length, and dependent clauses per T-unit between the CEFR levels ranging from A1 to B2 in written L2 English, L3 French, and L4 Italian by 120 subjects whose L1 was Swedish. The final dataset consisted of 211 texts. The results show, for instance, that although there are some differences between the measures and the target languages, there seems to be a statistically significant, linearly positive correlation between CEFR level and syntactic complexity in all languages.

In a similar vein, Kuiken and Vedder (2019) investigated the variation in syntactic complexity across proficiency levels, L1 and L2, and different target languages, using different types (both general and more fine-grained) of complexity measures. The data were argumentative essays written by subjects with Dutch, Italian, and Spanish as their L1 or L2s, and the learners' proficiency levels in the L2s ranged from CEFR levels A2 to B1 (see Kuiken & Vedder 2019: 199–200). The results indicate that, at the group level, the learner's proficiency level and the syntactic complexity of their productions are correlated, although not always significantly. The findings revealed variations across proficiency levels and languages. Based on these results, Kuiken and Vedder (2019) argue for the use of both general and fine-grained measures in studies of syntactic complexity (see the discussion of fine-grained measures below).

Previous studies have demonstrated that there can also be a positive correlation between L2 proficiency and syntactic complexity in the **spoken mode** (see, e.g. Iwashita 2006: 153–154; Vercellotti 2019: 243; see also De Clercq & Housen 2017; Lambert & Nakamura 2019). For instance, the findings of De Clercq and Housen (2017) show that complexity in spoken production correlates with increasing language proficiency. They studied the development of syntactic complexity to find differences between L2 English and L2 French among L1 Dutch speakers. In total, 200 subjects at different levels of L2 performance were chosen for the study based on a larger set of data (see De Clercq & Housen 2017: 320–322). They used both general measures of syntactic complexity (e.g. AS-unit length) and measures that indicate the variety of structures used by the learners (e.g. AS-unit length diversity, i.e. standard deviation of AS-unit length in a single text). The results show, for instance, that L2 English and L2 French learners begin to use more dependent clauses and more varied types of clauses as their language proficiency increases, indicating a positive relationship between these measures of complexity and the learner's proficiency level.

Along the same lines, Vercellotti (2019) reported a positive correlation between learners' proficiency levels and the complexity measures of AS-unit length, clause length, and subordination. They studied the development of and variation in syntactic complexity in the spoken productions of L2 learners of English with multiple L1s. All measures indicated increasing complexity. Furthermore, the findings suggest that these subsystems of syntactic complexity do not compete; that is, there is no trade-off between these complexity measures (see also Section 2.3.2.2).

As noted above, some studies suggest that the positive relationship between complexity and proficiency is not self-evident (Bulté & Housen 2018: 148; see also, e.g. Golden, Kulbrandstad & Tenfjord 2017). Possible reasons for the contradictory findings include the modality (e.g. Kuiken & Vedder 2011; Biber et al. 2016; Vercellotti 2019) and measures chosen (e.g. Iwashita 2006; Norris & Ortega 2009; Kormos 2014; Bulté & Housen 2018), but the task type, the target language, and some learner-related factors may also affect the results (see, e.g. Kuiken & Vedder 2019: 206).

Few studies have directly compared syntactic complexity in the written and spoken modes (see, e.g. Håkansson & Norrby 2007; Kormos & Trebits 2012; Lambert & Kormos 2014; see also Section 2.3.2.3). Kuiken et al. (2019: 168) point out that, thus far, the effect of modality has mainly been studied at the group level, although some studies have pointed to notable individual-level differences between written and spoken L2 performance (see also Section 2.3.2.4). Research on the differences in syntactic complexity between speech and writing has often concluded that syntactic structures tend to be more complex in the written mode (Kuiken et al. 2019: 164). Findings on the effect of the learner's language proficiency are more

diverse. On the one hand, Kuiken and Vedder (2012a: 160; see also Kuiken & Vedder 2011) argue that the difference in syntactic complexity between learners at higher and lower proficiency levels is evident in written production, while there are no differences in syntactic complexity between proficiency levels in speech. On the other hand, results in, for instance, Vercellotti (2019) demonstrate that learners' proficiency correlates positively with the syntactic complexity of their productions in the spoken mode.

The results of previous research have been affected by the measures used to analyse complexity. This seems to have been especially true when measures traditionally used for the analysis of complexity in written production were adopted to analyse complexity in spoken production (Iwashita 2006: 154–155, 160; Kormos 2014: 198; Lintunen & Mäkilä 2014: 383, 385–387; Kuiken et al. 2019: 164; Vercellotti 2019: 237; see also Norris & Ortega 2009). In this context, it is also important to pay attention to the segmentation units used in the analyses (see also Section 3.3.1). Lintunen and Mäkilä (2014) investigated how the segmentation units used affect the difference in syntactic complexity between written and spoken L2 English (N = 18). According to Lintunen and Mäkilä (2014: 392–393, 395), it is crucial to precisely define all segmentation units and consider their effects on the findings, as their results show that, especially when comparing production in the two modes, these units can have a significant impact on the result. In particular, clauses are often defined differently, and definitions are sometimes missing altogether (Lintunen & Mäkilä 2014: 382–383; Kuiken & Vedder 2019: 198; see, e.g. Foster et al. 2000; Kormos 2014). In the current study, when comparing the two modes of production, different but comparable units are used (see Section 3.3.2).

One contributing factor to the choice of measures is the differences in complexity between proficiency levels. As discussed above, learners at different proficiency levels complexify their productions at different linguistic levels: Clausal-level complexity is higher at intermediate levels, whereas phrasal-level complexity measures indicate higher complexity at more advanced proficiency levels. To date, phrasal-level complexity is a relatively neglected area of research, especially compared to the number of studies that focus on overall (length-based) and clausal-level complexity (see, e.g. Biber et al. 2011; Ansarifar et al. 2018; Sarte & Gnevsheva 2022). Thus far, many studies have concentrated on differences between text types, and in studies of proficiency-related differences, the effect of text type is often integrated into the study design. Phrasal-level complexity can be studied using the complexity index introduced by Biber et al. (2011). In this index, different modifier types of varying complexity and their developmental order are presented. The lower-level modifiers include, for instance, attributive adjectives, whereas modifiers such as complement clauses are considered more complex (for further information, see Section 3.3.2).

Although they are relatively scarce, most studies on this topic point to a positive relationship between phrasal-level complexity and learners' proficiency (see, e.g. Biber et al. 2011; Lahuerta Martínez 2018; Lan, Lucas & Sun 2019; Sarte & Gnevsheva 2022), but opposite findings on the relationship between phrasal-level complexity and proficiency have also been obtained (see, e.g. Lambert & Nakamura 2019; Khushik & Huhta 2020). According to both longitudinal and cross-sectional studies, high phrasal-level complexity is typical of advanced academic writing (see Sarte & Gnevsheva 2022: 2). Lan, Lucas and Sun (2019) reported on studies that suggest a positive correlation between learners' (assumed, often based on differences in educational levels) proficiency and noun phrase complexity, for instance, in the number of different modifier types used. Ansarifar et al. (2018) show that lower-level students use significantly fewer attributive adjectives, pre-modifying nouns, -ed clauses, and prepositional phrases as modifiers than expert writers. The results of Lahuerta Martínez (2018) indicate that the noun phrase length correlates with higher scores on writing tasks. Sarte and Gnevsheva (2022: 1) point out that complex (i.e. longer and more varied) noun phrases are more common at higher proficiency levels, whereas complexification at the clausal level is less common. In contrast, Lambert and Nakamura (2019: 258–259) argue for a negative correlation between several phrasal-level complexity measures and learners' proficiency, with a decrease in noun phrase complexity associated with increased proficiency. It is of note that, in accordance with DST-based principles, the proficiency-related differences may seem linear at the group level, but the variation at the individual level can be notable. For instance, Vyatkina et al. (2015: 41, 43) demonstrated that there can be significant interindividual differences and variations in the use of modifier types.

Researchers have explored the use of traditional, general measures as well as more specific measures in the analysis of syntactic complexity. Kyle and Crossley (2018) compare general (e.g. T-unit length) and more fine-grained measures (both at the clausal and phrasal levels) of syntactic complexity in predictions of proficiency in L2 English based on a large corpus of argumentative texts written as part of the Test of English as a Foreign Language. They argue for the superiority of fine-grained phrasal-level complexity measures over both the more traditional measures and fine-grained clausal-level measures. Kuiken and Vedder (2012a: 165), on the other hand, state that general and specific measures can be mutually complementary. While the analyses conducted in the first two substudies of this dissertation mainly focus on general measures of complexity (and two specific measures of relative complexity), in the third substudy, both general and fine-grained measures are employed (see Section 3.3.2).

As shown in this section, the task type (used here to cover all task-related variations, such as modality, genre, and text type) influences the structures used in a given production. Large-scale analyses of comprehensive corpora have shown that

there are great differences, for instance, between academic writing and daily conversation (see, e.g. Biber et al. 2011; Biber et al. 2016). These differences are partly attributable to the fundamental differences between speech and writing (discussed in Section 2.3.2.3), but other task and text type characteristics also influence syntactic complexity. Kuiken et al. (2019: 164–166) argue that differences in task types (e.g. descriptive and argumentative) and genres (e.g. newspaper articles and narratives) result in differences in syntactic complexity, but studies on the effect of task type and genre are still rare (see, however, e.g. Bernardini & Granfeldt 2019). Research has demonstrated, for instance, that both L1 and L2 users demonstrate higher syntactic complexity in argumentative texts than in narrative styles (Kuiken & Vedder 2019: 197), and that the frequent use of dependent clauses is often connected with narrative style, whereas complexification at the phrasal level is seen as characteristic of academic styles (see, e.g. Biber et al. 2011). In this dissertation, the productions examined are responses to tasks that contained a mixture of descriptive, narrative, and argumentative task types (see Section 3.2), and the influence of task type is critically examined (see Section 5.3), although this influence is not a focus of the current study.

Some studies have adopted a cross-linguistic perspective in studying syntactic complexity, as research has shown that there are differences in such complexity between (target) languages (e.g. Kuiken & Vedder 2012a, 2019; Gyllstad et al. 2014; Bernardini & Granfeldt 2019; Kuiken et al. 2019). It is of note that length-based measures are highly affected by typological differences between languages (e.g. Swedish and Finnish). However, other measures of syntactic complexity – for instance, measures based on embedding, but also phrasal-level measures – can vary between languages due to the differences in how different dependent clauses and modifiers are used (Kuiken & Vedder 2019: 199; Kuiken et al. 2019: 165). Kuiken and Vedder (2012a; see also Kuiken & Vedder 2019) report results from three studies of L2 Italian and French that examine, for instance, the relationship between proficiency level and syntactic complexity (number of clauses per T-unit and number of dependent clauses per clause). In the case of Italian, the results indicate a significant correlation between proficiency and complexity, whereas in French, there was no significant correlation, indicating differences in syntactic complexity between these languages (Kuiken & Vedder 2012a: 155, 164; see, however, the discussion in Kuiken & Vedder 2019: 206).

Furthermore, several learner-related factors can affect syntactic complexity and the relationship between complexity and proficiency. For instance, the complexity of a learner's production is affected by the type of knowledge (explicit or implicit) they have of the target language (see Housen et al. 2012b: 7). The learner's age and degree of exposure to languages in general can also affect the structures they produce. In the early phases of the learning process, learners learn new words,

leading to longer production units, and new structures, leading to more frequent embedding, whereas learners at later time points use more variation in syntactic complexity; therefore, increasing complexity can possibly be better detected in studies of the early phases of learning (see, e.g. Ferrari 2012; Khushik & Huhta 2022). Adult L2 learners are more cognitively advanced and have prior experience with structures in their L1 and possibly other L2s. Therefore, they often feel the need to express themselves in a more complex way, as they do in their L1, even though their knowledge of the target language might be limited (Vyatkina 2012: 578; Vyatkina et al. 2015: 42). This type of conflict between learners' relative cognitive maturity (Foster et al. 2000: 355; Vyatkina et al. 2015: 42) and limited L2 resources might lead to a competitive relationship between complexity and accuracy in their productions (see Section 3.2.2.2). In this study, the participants were exposed to education in Swedish for a relatively long time, and all subjects were familiar with linguistic structures in several languages (see Section 3.1).

It has been pointed out, particularly in studies drawing on the DST principles, that even if group-level results indicate a positive correlation between proficiency and complexity measures, the progression is not always linear at the individual level (see, e.g. Kuiken 2023: 90). For instance, Larsen-Freeman (2006) longitudinally studied the CAF dimensions embodied in written and spoken productions of five L1 Chinese intermediate learners of English using a mixed-methods approach. Written and spoken tasks were administered four times over a six-month period, and some general syntactic complexity measures (mean length of T-unit and number of clauses per T-unit) as well as lexical complexity and overall accuracy measures were employed in the analyses. The central finding of the study is that every CAF domain showed improvement at the group level, but there was considerable individual variation (Larsen-Freeman 2006: 598–560). Vyatkina (2012) studied the development of complexity in written L2 German both cross-sectionally and longitudinally. In line with the results of Larsen-Freeman (2006), the group-level analysis indicates a relatively linear increase in syntactic complexity, whereas the individual-level findings point to great inter- and intraindividual variation (Vyatkina 2012: 583–589, 594). This variation seems to have been caused by the varying strategies used to complexify the production (Vyatkina 2012: 590; see also Vyatkina 2013 and Section 2.3.2.4).

Verspoor et al. (2008) examined the written productions of one advanced L1 Dutch learner of English. Their analysis considered 18 written samples produced over three years and focuses on the development in vocabulary use and syntactic complexity. Both progress and regress were discovered in this development, and there was notable variability even at advanced proficiency levels. Furthermore, according to the findings, the interaction between subsystems of lexical and syntactic complexity is dynamic (for further information on the interaction between

complexity and accuracy, see Section 2.3.2.2), meaning that both supportive and competitive interactions between the subsystems are possible. Therefore, the results indicate that learners' resources are limited, and this can lead to a competitive relationship between, in this case, lexical and syntactic complexity – but also that aspects of complexity can grow together, resulting in a supportive relationship, between, here, noun phrase length and ratio of number of words to finite verbs.

Following Verspoor et al. (2008), Spoelman and Verspoor (2010) longitudinally investigated dynamicity in the relationship between subdimensions of complexity. They analysed the written beginner-level L2 Finnish productions of an L1 Dutch student. The data were gathered over three years and comprise 54 texts. In the study, both complexity and accuracy, as well as the interaction between complexity and accuracy, are examined using a variety of methods. In accordance with the results from Verspoor et al. (2008), the findings demonstrate that both complexity and accuracy can exhibit rapid alternating progression and momentary regression, and that the interaction between these subsystems is complex and dynamic. In accordance with the principles of DST, the study highlights that the development of learner language is closely connected with the degree of variability therein, and that a competition for resources appears both within and across subsystems (see also Lesonen et al. 2020).

Bulté and Housen (2020b) examined the interaction between the subsystems of complexity from a developmental point of view in L2 English productions by ten L1 Dutch learners. Their data came from a larger longitudinal corpus, and the study reviews various methods employed within the Dynamic Usage-based framework (Bulté & Housen 2020b: 61–63). The study demonstrates that there was notable inter- and intraindividual variation in the sample, indicating both variable and non-parallel developmental paths of complexity and dynamicity in the interconnections between different subsystems of complexity. The findings show no clear evidence of either a supportive or competitive relationship between the subsystems, with the exception of a slight tendency towards competition between certain measures, that is, complexification at the clausal and phrasal levels. According to Bulté and Housen (2020b), these findings accentuate the need to include multiple independent measures in research on complexity (see also, e.g. Norris & Ortega 2009; Lowie & Verspoor 2019).

Kowal (2016) investigated the development of CAF constructs and their interagency in written L2 Swedish texts by 15 L1 Polish learners. In accordance with the studies discussed above, the results show that syntactic complexity, measured by dependent clause ratio and diversity of structures used, in individual productions does not increase linearly together with the group mean (Kowal 2016: 209). These results point to high interindividual variation, indicating that some learners clearly prefer complex structures, while others settle for simpler constructions. These

differences seem to remain the same or even become greater as progress is made at the group level. It seems that learners who exhibit persistently low levels of syntactic complexity tend to achieve an attractor state and stay at this level, whereas learners who progress rapidly at the beginning of the learning process can either stay at the same level for a longer time or demonstrate regress or further progress in syntactic complexity. The low intraindividual variability in the results indicates that learners might represent different profiles in their productions (see the discussion of learner profiles in Section 2.3.2.4).

This discussion suggests that **differences in learners' performance can only partly be explained by their language proficiency**, and learners at the same proficiency level can differ vastly in their use of syntactically complex structures (e.g. Bulté & Housen 2018: 148; see also Dörnyei 2005: 125; Tagarelli et al. 2016: 294). One's individual style of producing language might, in fact, be more important than one's proficiency when it comes to explaining complexity.

Although this dissertation is mostly concerned with complexity, accuracy plays a minor role in the first substudy, where the effect of language proficiency on syntactic accuracy and the relationship between syntactic complexity and accuracy are investigated. Despite the fact that accuracy is studied in the first substudy, the aim is not to highlight errors but rather to investigate how accuracy interacts with proficiency and syntactic complexity in the learners' productions (see the discussion in Ellis 2019: 53). As mentioned above, it is easy to give a simplified definition of accuracy: A learner is able to use the target language according to that language's norms (Housen et al. 2012b: 3). However, it is not easy to operationalise accuracy since, especially in learner language produced by learners at lower proficiency levels, different types of errors can be extremely frequent (see, e.g. Kuiken & Vedder 2012a: 146; Kowal 2016: 46–47). Since the focus of the present study is syntactic features, only syntactic accuracy is considered, and accuracy is operationalised with the help of errors in word order (see Sections 2.3.2.5 and 3.3.3; see also Kowal 2016: 211).

There is ample evidence of a positive relationship between proficiency and accuracy, but some studies have indicated the opposite (see, e.g. Wolfe-Quintero et al. 1998: 37–38; Kuiken & Vedder 2012a: 160). Thus, since accuracy is multidimensional and interconnected with other systems in learner language, the relationship between proficiency and accuracy is not always linear or progressive. In the following section, the interplay between syntactic complexity and accuracy is discussed.

2.3.2.2   Interplay between syntactic complexity and accuracy

The dimensions in the CAF triad are considered (sub)systems of learner language, and their interaction has been studied extensively (Lesonen et al. 2020: 2–3; see,

e.g. Verspoor et al. 2008; Spoelman & Verspoor 2010; Kowal 2016). As all the systems and subsystems are interconnected, it can be assumed that there is some form of interaction between the dimensions in learners' productions (Kowal 2016: 139). Research on the nature of the interaction between the dimensions has yielded partly mixed results (Kuiken et al. 2019: 163; see also, e.g. Kormos & Trebits 2012). According to these studies, the correlation between the CAF dimensions can be either negative or positive (see also Section 2.2 and *neutral*, *supportive*, *conditional*, or *competitive* relationships in Lesonen et al. 2020: 3–4) depending on several factors, such as the learner's proficiency level. In other words, the dimensions can either compete or develop simultaneously (Bulté & Housen 2018: 148; see also Skehan 1998; Robinson 2001). In this section, the interaction between syntactic complexity and accuracy is discussed in light of two opposing views – the Limited Attentional Capacity model (Skehan 1998; see also the *Trade-Off Hypothesis*) and the Multiple Resources Attentional model (Robinson 2001; see also the *Cognition Hypothesis* and the *Triadic Componential Framework*), – along with an overview of relevant research (for a more comprehensive discussion of these models, see, e.g. Salimi & Dadashpour 2012). In this dissertation, the assumptions made in these models are investigated in the first substudy, where the interrelatedness between syntactic complexity and syntactic accuracy is examined at two proficiency levels (see Section 4.1). The aim of this substudy is not to test the hypotheses behind these opposed models but to investigate whether different complexity measures exhibit competitive or supportive relationships with overall syntactic accuracy.

According to the Limited Attentional Capacity model, L2 learners must **distribute their attentional resources between the systems and subsystems** of a target language when producing the language due to their limited cognitive resources (see Skehan 1998, 2009; Skehan & Foster 1999; see also, e.g. Tavakoli 2014: 219; Kowal 2016: 213; Ellis 2019: 52). Consequently, learners might have to prioritise one or two CAF dimensions at the expense of others. Due to this prioritisation, **trade-off effects** between the dimensions are possible – for instance, between complexity and accuracy (for more information on the interaction between the CAF systems, see, e.g. Housen et al. 2012b: 5–8; Kowal 2016: 175–176, 182, 184, 213). In other words, learners might pay more attention to complexity, which can, because of the trade-off effect, lead to more frequent non-target-like structures in their productions, or they might prioritise accuracy and, as a result, produce simpler structures. The trade-off effects are linked to the learners' reliance on different types of linguistic knowledge. This knowledge consists of automated structures – that is, structures that learners produce as wholes (chunks) – and abstract rules that can be used to produce an unlimited number of structures. In principle, learners can either use their automated knowledge and produce relatively simple and error-free

language, or they can experiment with their structures using abstract rules and risk producing non-target-like structures.

Some studies offer support for trade-off effects in learner language production, partly supporting the assumption that learners' resources can be limited. For instance, in Ferrari (2012: 284, 288, 293), some learners' accuracy decreased when they started to experiment with complex and varied structures, resulting in a U-shaped development (i.e. non-target-like structures first become target-like but are again produced in the non-target-like form due to the trade-off effects). Kowal (2016: 213) reports on two versions of trade-offs, that is, a softer and a stronger version of competition between the CAF dimensions. In the softer version, learners prioritise one or two dimensions at the expense of the other (two), but all dimensions show development, whereas in the strong version, the trade-off leads to a decline in one or two dimension(s). Furthermore, learners' prioritisation of the different dimensions and the relationship between the dimensions can vary over time (Kowal 2016: 213–214; see also Spoelman & Verspoor 2010). Therefore, the trade-off effects are dynamic and individual in nature, making it impossible to predict individual developmental paths (Kowal 2016: 214, 215). Despite this, some general trends in learners' productions can emerge. These similarities in learners' productions enable their categorisation into learner profiles (see Section 2.3.2.4).

Skehan and Foster (2012) note that the interaction between systems and subsystems can be affected by several factors, and that trade-off effects are not always present. In other words, under certain circumstances, complexity and accuracy can exhibit a supportive relationship (Skehan & Foster 2012: 215). For instance, different external factors (e.g. the type of planning; see Yuan & Ellis 2003) can affect the relationship between complexity and accuracy. In Foster and Skehan's (1999) study, the focus of planning (content vs language) had no impact on the interaction between complexity and accuracy in learners' productions, but various sources of planning (teacher, group, or individual) yielded different kinds of results (see also Skehan & Foster 2012: 216). These findings suggest that learners are not always able to direct their attention effectively, but, for instance, teacher-assisted planning can lead to the use of more complex structures, as well as target-like language use.

According to Robinson's (2001) Multiple Resources Attentional model, **learners can pay attention to different CAF dimensions at the same time** (see also Robinson 2007; Robinson & Gilabert 2007). In this model, it is assumed that learners have multiple pools of attentional systems; therefore, paying attention to one aspect of a linguistic feature does not affect the resources in other attentional pools (see also Tavakoli 2014: 219). Cognitive factors in task performance can be either resource-directing or resource-dispersing, and through the cognitive demands that these factors impose on learners, the task can direct learners to pay attention to

several systems and subsystems of language simultaneously (see also Kormos & Trebits 2012: 444–445).

Evidence from earlier research suggests that learners can take several dimensions into consideration simultaneously in accordance with Robinson's model. Results in Michel, Kuiken, and Vedder (2007), for instance, partially support the hypothesis of multiple pools of attention (see also Kormos 2011). Earlier findings showed that, particularly at higher proficiency levels, learners are able to pay attention to both complexity and accuracy (see, e.g. Spoelman & Verspoor 2010). Studies that have tested the predictions of the Multiple Resources Attentional model have also yielded contradictory evidence on the effects of task complexity (see the discussion below) on the different CAF systems. For instance, in Kuiken and Vedder (2012a: 162–164), support for increased accuracy was found, but no influence on syntactic complexity was detected.

Thus, research indicates that both trade-offs and supportive interaction between the systems in the CAF triad are possible, as well as between the subsystems within these systems – for instance, between clausal and phrasal complexity (Kormos & Trebits 2012: 445; Vercellotti 2019: 244; Kuiken et al. 2019: 163; Bulté & Housen 2020a: 226; see, e.g. Verspoor et al. 2008; Spoelman & Verspoor 2010; Lesonen et al. 2020). Research has further indicated that the interaction is dynamic (Bulté & Housen 2018: 148; Lesonen et al. 2020: 3–4). At the beginning of the learning process, the CAF dimensions are distinct in the learners' minds, which can result in frequent trade-off effects between the dimensions (Kowal 2016: 214). With increasing language proficiency, these dimensions can become more integrated. Accordingly, Spoelman and Verspoor (2010) highlight that a learner's proficiency plays an important role in the interplay between the CAF dimensions (see also Gunnarsson 2012). In the early phases of language learning, it can be challenging for learners to focus on all dimensions simultaneously. This helps explain why trade-off effects between the dimensions can be notable and frequent in productions by learners with lower proficiency. Contrarily, at higher proficiency levels, when learners' language use becomes increasingly automatised and, as a result, the cognitive load on working memory diminishes (e.g. Tagarelli et al. 2016: 295–296), it can become easier for learners to divide their attentional resources between linguistic features (Spoelman & Verspoor 2010: 550). Myles (2012) further points out that the use of chunks in the early phases of learning can lead to relatively high complexity in beginner-level productions, indicating a possible supportive relationship between complexity and accuracy, even at lower proficiency levels.

In both models of the relationship between the CAF systems, the cognitive demands that tasks impose on learners play a central role (Kuiken & Vedder 2012a: 165; Skehan & Foster 2012: 215–216, 218). Skehan (1998; see also Skehan & Foster 2001: 194–195) examined three aspects of task complexity: code complexity,

cognitive complexity, and communicative stress. Code complexity pertains to linguistic features, such as linguistic complexity and vocabulary variety; cognitive complexity refers to both the familiarity of the task (e.g. the topic) and the cognitive processing (e.g. additional information) carried out during the task; and communicative stress includes such factors as time pressure and mode of production. Robinson (2001, see also Robinson 2005), on the other hand, distinguishes between the interrelated factors that comprise the dimensions of a task, including task complexity (i.e. the actual cognitive demands of the task, e.g. chance to plan the output), difficulty (i.e. learner-related factors, e.g. language proficiency), and conditions (i.e. interactive factors, e.g. one- or two-way participation). In previous research, L2 performance was compared under different task conditions in accordance with the factors that affect task complexity. For instance, task characteristics such as planning time (a resource-dispersing factor) and the number of elements to be taken into account (a resource-directing factor) have been manipulated in earlier studies, and their effect on L2 performance has been studied (see Robinson 2005: 5; Kuiken & Vedder 2012a: 148; Tavakoli 2014: 219–220). Task-related factors have been widely investigated, especially in spoken L2 data (see, e.g. Kuiken & Vedder 2012a; Levkina & Gilabert 2012).

Studies of the effects of task complexity on syntactic complexity, on the one hand, and on the relationship between complexity and accuracy, on the other hand, have yielded mixed findings (e.g. Kuiken & Vedder 2012a: 144, 160; Kuiken et al. 2019: 163). According to these studies, the relationship between the systems and subsystems of learner language seems to be different in spoken and written production (see, e.g. Skehan & Foster 1997; Kormos & Trebits 2012; Kormos 2014; see, however, the discussion in Kuiken & Vedder 2012a: 162–164). Trade-off effects can be expected to be more frequent in spoken productions due to the limited time available for planning and the lack of opportunities to revise the utterance (see Section 2.3.2.3). Results in Tavakoli (2014: 230) indicate that task complexity appears to affect syntactic complexity (T-unit length and subordination) differently in written and spoken L2 productions. Learners seem to produce complex language even in simple written tasks, and the difference in complexity between simple and complex spoken tasks is greater than in the written mode (Tavakoli 2014: 228–229). In Kuiken and Vedder (2012a: 162, 164; see also Kuiken & Vedder 2011), spoken productions based on the more complex task contained somewhat fewer dependent clauses, indicating that learners resort to simpler language when their cognitive resources become overloaded. Kormos and Trebits (2012) found that, in tasks with light cognitive demands, learners can produce more target-like structures in writing than in speech, and that syntactic complexity increases only in written tasks with higher cognitive demands. For the written tasks, the results can be seen as evidence of a supportive relationship between complexity and accuracy, whereas the results

concerning spoken data can be interpreted as support for a competitive relationship. In the following section, the differences between written and spoken productions are discussed in more detail.

2.3.2.3 Effect of mode of production on syntactic complexity

Syntactic complexity is a well-researched dimension of CAF in L2 writing. In recent decades, researchers have also begun to study complexity in the spoken mode (Iwashita 2006: 152; see also Vercellotti 2019). However, the differences in complexity between speech and writing at both the group and individual levels are still under-researched in the field of SLA (Kuiken & Vedder 2012b: 364–365; Kormos 2014: 194; Bulté & Housen 2018: 148). The few studies available have offered mixed results regarding both complexity in spoken L2 production and the differences between the two modes of production (see, e.g. Kuiken & Vedder 2012b: 365, 369; Kormos 2014: 198). These mixed results are partly due to the task type (see, e.g. Kormos 2014; Tavakoli 2014) and some external factors (Kuiken & Vedder 2012b: 375; see also Granfeldt 2008), but the measures chosen and the definition of production units adopted could also have had a major effect on the results (see, e.g. Lintunen & Mäkilä 2014).

As already noted, there are **fundamental differences between speech and writing** that affect production in general and syntactic complexity in particular (Halliday 1989; Miller & Fernandez-Vest 2006). The situations in which written and spoken communication take place are usually dissimilar in nature – for instance, in terms of the presence of recipients (see, e.g. Tanskanen 2006: 74–75). In spoken communication, the interlocutors are often in direct contact, whereas in writing, there is typically no such contact with the recipient of the message (see, e.g. Gilabert et al. 2016: 119). In some cases, it is not even known if the text will be read or who will read it. Additionally, when speaking, one must remember what has been said, while it is possible to go back in a written text and read what has already been written (Kuiken & Vedder 2011: 92, 2012b: 365–366). The most important difference between these processes is that more extensive planning is possible in traditional written production, whereas spoken communication takes place in real time, as both planning and production take place simultaneously (Kormos & Trebits 2012: 446; Kormos 2014: 196; Tavakoli 2014: 223; Vercellotti 2019: 234; see also Yuan & Ellis 2003). Therefore, spoken production offers limited opportunities for planning. The time constraint in spoken production can lead to the use of simpler structures due to the L2 learner's limited resources for processing the target language (e.g. Håkansson & Norrby 2007: 92; Gilabert et al. 2016: 127; Vercellotti 2019: 234). In addition, some syntactically simpler structures – for instance, at the phrasal level – are generally more common in

spoken communication than in writing (e.g. Miller & Fernandez-Vest 2006: 13; see also Biber et al. 2011; Biber et al. 2016).

However, some complex structures are more typical of spoken discourse. For instance, certain types of dependent clauses and complex phrases (e.g. complex prepositional phrases) are more common in speech (Miller & Fernandez-Vest 2006: 15; Biber et al. 2016: 645). Moreover, these structures are used differently in different languages (see, e.g. Gyllstad et al. 2014; Bernardini & Granfeldt 2019; Kuiken et al. 2019). Furthermore, when considered in terms of length-based measures of complexity, spoken productions can be more complex than written ones. This is due to differences in how the message is produced. In spoken production, the message is produced simultaneously with the planning process; therefore, thoughts can be formulated in a complex manner, and production units can be long (Beaman 1984: 50–51; cf. Jörgensen 1976: 43). On the other hand, unfinished utterances are more frequent in speech. In writing, thoughts are usually expressed in a more concise manner. Planning and revision during the writing process facilitate the formulation of concise thoughts (see Kormos 2014: 197; Tavakoli 2014: 231). Because of the fundamental differences between the processes of speech and writing, learners may direct their attention differently in the two modes of production (Gilabert et al. 2016: 119–120). This can significantly affect complexity in L2 production.

When comparing complexity in written and spoken productions, it is important to separate the fundamental differences from the differences caused by formality (Beaman 1984: 51; Tanskanen 2006: 75; Miller & Fernandez-Vest 2006: 15–16). As discussed earlier, different types and genres of texts have their own distinguishing characteristics. Therefore, the use of syntactically complex structures can differ greatly, for instance, in academic articles vs text messages, or in prepared presentations vs informal conversation (see, e.g. Jörgensen 1976; Biber et al. 2016). Based on the level of complexity, these types of productions form a continuum in which, for example, informal conversations lie at the oral end of the continuum, marked by simpler syntax, whereas academic articles lie at the literary end and involve more complex structures (Miller & Fernandez-Vest 2006: 9–10; Gilabert et al. 2016: 120–121; see also Biber et al. 2011; Larsson & Kaatari 2020).

Writing and speaking are demanding in different ways. For instance, writing can be seen as more demanding than speech in the sense that more attention needs to be paid to accuracy (García Mayo & Azkarai 2016: 244). In speech, errors are usually more readily tolerated (see also Weinert, Basterrechea & Garcia Mayo 2013: 161). In both written and spoken communication, language users need to access their linguistic resources efficiently, but in spoken production, quick access to these resources plays an especially important role (e.g. Håkansson & Norrby 2007: 91; Leonard & Shea 2017: 179–183, 190; see also Skehan & Foster 2012). However, the difficulty of processing and producing language in these two modes of production

depends on the language user. Producing complex written structures can be more demanding than doing so in the spoken mode in the case of L1 production by children, as the writing process is cognitively more demanding for children than it is for adults (Silva, Sánchez Abchi & Borzone 2010: 50, 54; see also Bourdin & Fayol 1994). Silva et al. (2010) studied differences in syntactic maturation between written and spoken L1 productions by young children (approximately 6–8 years old). Syntactic maturity was defined as the ability to produce complex syntactic units, and subordination is seen as one of its most important features (Silva et al. 2010: 48). They found that the spoken data were more complex than the written responses, but the difference was not statistically significant in terms of the number of clauses per T-unit, and the difference in complexity between spoken and written production was smaller at the higher proficiency level. The results indicate that as lower-level processes in writing become more automatic, more complex structures can be produced (see also Chafe & Tannen 1987: 384).

In contrast, adults may find it easier to produce syntactically complex writing, as the cognitive load on one's working memory is lower in writing than in speech (see, e.g. Gilabert et al. 2016: 127; Vercellotti 2019: 234). Especially in L2, spoken production can be more challenging than writing due to the limited resources of L2 learners (see the discussion above). In written production, learners have more control over their productions (Kuiken & Vedder 2012b: 365–366; Gilabert et al. 2016: 119–120; see also Granfeldt 2008), and according to earlier studies, the chance to plan out one's productions in L2 seems to result in more complex language (Håkansson & Norrby 2007: 92; see also Skehan & Foster 2012). One's proficiency level has a major effect on the differences in complexity between one's written and spoken productions. As a learner's processing of the target language becomes more automatic with increasing language proficiency, they can produce more complex structures in the spoken mode as well (see, e.g. Leonard & Shea 2017). Therefore, the difference in complexity between speech and writing can be assumed to be more notable in productions by learners at lower proficiency levels than by those with greater proficiency.

Due to the lighter cognitive load and the fact that it is possible to plan one's output and utilise explicit knowledge more effectively in writing, it can be assumed that written L2 production is more complex than spoken production (Håkansson & Norrby 2007: 91; Granfeldt 2008: 87; Kuiken & Vedder 2012b: 365–366; Tavakoli 2014: 228–229). Indeed, Norrby and Håkansson (2007: 56) note that styles of producing structures do differ in speech and writing; the participants in their study seemed to produce more complex structures in the written mode than in the spoken mode, but the differences were relatively small (see also Håkansson & Norrby 2007). However, other studies have indicated that, by some measures, written and spoken L2 productions do not differ significantly, and spoken productions can even be more

complex than written ones (see, e.g. Granfeldt 2008; Kormos & Trebits 2012; Kormos 2014; Lintunen & Mäkilä 2014). For instance, in contrast to the findings of Norrby and Håkansson (2007), Granfeldt's (2008) results show that L2 French learners use more dependent clauses in speech than in writing. According to Granfeldt (2008: 92), this somewhat unexpected result can probably be explained by qualitative differences in the dependent clauses, as more diverse types of dependent clauses can be used in writing. Additionally, the number of learners (N = 6) in the study was limited, and the results may have been influenced by the learners' individual preferences (see also Weissberg 2000: 37, 44–45). Some learners favour spoken production as the primary platform for experimenting with new syntactic structures, while others are more comfortable with experimenting in writing (Granfeldt 2008: 88, 97; see also García Mayo & Azkarai 2016: 244).

The learning context (formal vs informal) as well as the quantity and quality of input are crucial for understanding the differences between written and spoken L2 production (Leonard & Shea 2017: 179; see also Håkansson & Norrby 2007). In L2 teaching, writing has traditionally dominated curricula and teaching materials used around the world (see, e.g. Zhang 2013: 836), whereas, more recently, an increasing focus has been placed on spoken communication (see, e.g. Peltonen 2020: 1). The learning contexts and emphases of teaching greatly affect the skills that learners acquire in different domains. L2 learning outside the target language community and heavy reliance on written materials might result in good explicit knowledge of rules, whereas meaningful communication in the target language can lead to the rapid development of both receptive and productive oral skills (Ortega 2008: 80).

In summary, the cognitive load on working memory is lighter in writing than in spoken communication, as written production rarely requires the target language to be produced and processed in real time. Thus, the demands that speech and writing impose on the learner are fundamentally different, and this is why learners' spoken productions can differ greatly from their productions in the written mode (Kuiken & Vedder 2012a: 150; Kormos 2014: 195–197; Leonard & Shea 2017: 181). Despite the partly contradictory findings in earlier studies, it can be concluded that, **when learners have the chance to plan their productions, and to concentrate on the form in addition to the content, the production tends to exhibit higher complexity** (Håkansson & Norrby 2007: 81–82, 92–93; Tavakoli 2014: 224, 230–231; see also Kuiken & Vedder 2011, 2012b). As noted above, few studies have directly compared syntactic complexity in written and spoken L2 productions (Kuiken & Vedder 2012b: 364–365; Kormos 2014: 194). Vercellotti (2019: 234) points out that the process of producing language is reflected more directly in spoken productions, and more research is therefore needed on syntactic complexity in spoken L2 productions – particularly on individual variation in this mode of production. This dissertation aims to fill these gaps. In addition to the factors

discussed above, individual variation is important when studying syntactic complexity. In the following section, these individual-level differences are discussed in more detail, and the focus turns to learner profiles – that is, individual learners' styles of producing complex structures.

2.3.2.4   Role of individual variation in syntactic complexity

As discussed in Section 2.2, recent SLA studies, especially those that have drawn on usage-based and DST-based theories, have highlighted **the role of individual variation** in L2 learning in general and in learner language complexity in particular (e.g. Bulté & Housen 2018: 148; see also Verspoor et al. 2008: 217; Spoelman & Verspoor 2010: 535; Lowie & Verspoor 2015: 75–76, 2019: 202). Group-level studies have indicated that complexity increases together with L2 proficiency according to certain stages in a linear manner. The disadvantage of these group-level studies is that they have ignored participants' differing styles of processing and producing language (Verspoor & Behrens 2011: 36; Vyatkina 2012: 577–578). The individual-level results of earlier studies demonstrate that the correlation between complexity and proficiency is, in fact, nonlinear, and that the interaction between both higher-level systems and subsystems in learner language is dynamic (Kuiken & Vedder 2019: 195; see also, e.g. Verspoor et al. 2008). Variation is seen as a central component of L2 development (Verspoor et al. 2008: 217; Tagarelli et al. 2016: 294; Kuiken et al. 2019: 162). Verspoor et al. (2008: 217) argue that 'only when learners have access to a variety of forms are they able to select those that help them develop, so the more different forms from which they can select, the more likely development is to take place'.

Variation is considered especially important for understanding complexity, as learners experiment with varying structures, both at the intermediate levels at which they learn new structures, and later, when they begin to optimise the level of complexity employed in their productions (Ferrari 2012: 282–283, 293; Kuiken & Vedder 2019: 195, 197). Along the same lines, Vyatkina (2013: 25) states that individual variation in complexity can grow over time (see also Kowal 2016). A linear, stepwise development of complexity features is possible at the beginning of the learning process, but later, individual learners may exhibit diverging developmental paths. Learners may also start experimenting with their language relatively early in the learning process, and this can lead to major differences at the individual level due to limited resources – but these differences also play an important role at higher proficiency levels (Vyatkina 2013: 25).

Individual variation can be detected both between learners, as the same meaning can be expressed with the help of different structures by different language users (i.e. interindividual variation), and in the same learner's performance in different contexts

and at different times, as language users produce structures of varying complexity depending on the relevant conditions, resulting in synchronic and diachronic (i.e. intraindividual) variation (Vyatkina 2013: 15; Kuiken et al. 2019: 162, 195). Stage-like descriptions of learner language complexity argue that complexity emerges at different linguistic levels in accordance with proficiency (see Sections 2.2 and 2.3.2.1), but many scholars have pointed out that individuals at the same proficiency level may differ in their syntactic elaboration and complexification strategies (e.g. Verspoor et al. 2008; Kuiken et al. 2019). For instance, the results in Bulté and Housen (2018) show that learners at the same proficiency level display major differences in terms of how they produce syntactically complex structures, and that the development of syntactic complexity is different for individual learners. This indicates that not all variability in syntactic complexity is linked to proficiency, and there are other contributing factors, such as learner profiles (Bulté & Housen 2018: 160).

Kuiken and Vedder (2019: 195) discuss the importance of individual variation and point out that the use of structures of varying complexity can at least partly be explained by learners' individual or stylistic choices (i.e. *idiosyncratic choices*), which again may be a sign of greater mastery of the target language. Some learners favour simpler structures, while others seem to consistently produce more complex ones, and preferences for clausal- and phrasal-level elaboration vary. Therefore, the complexity of syntactic structures varies by proficiency level and the needs a certain context creates – but also depending on the language user. L2 learners form a special group of language users, and especially within this group, there is notable variation in how syntactically complex structures are used, not only because of varying proficiency and task-related factors, but also because L2 users have individual preferences for expressing themselves and can draw on strategic resources from several languages (Forbes 2018: 3–4).

L2 learners differ in how they process and produce language. Learners can be either analysis-oriented, following rule-based systems, or memory-oriented, following memory-based systems (Dörnyei 2005: 152). The former type has organised, rule-based representations of language, and they regularly restructure and complexify their underlying interlanguage systems, whereas the latter possess a wide range of lexicalised exemplars that can be modified to suit communication in real time (see the discussion in Section 2.1). These types of learners differ, for instance, in their reliance on grammar and their control over the language produced (i.e. *monitoring*). These differences can be partly attributed to language proficiency, but monitoring can also depend on individual styles of processing a language (Dörnyei 2005: 152–153). These traits are central determinants of learner profiles. Additionally, many personality traits, such as creativity, affect learner profiles and the complexity of their productions, especially when it comes to productive complexity and length measures (Vyatkina 2012: 594; see also Albert 2011).

In this enquiry, learner profiles are understood with reference to *cognitive style* (Dörnyei 2005; see also *learning styles* and *learning strategies*, Forbes 2018: 2; Vyatkina 2013: 14–15). Cognitive style is defined as a learner's individual manner of processing and producing language (Dörnyei 2005: 124). The cognitive style of a learner encompasses a set of characteristics pertaining to both personality and intelligence. Therefore, according to Dörnyei (2005: 125), it is important to distinguish a learner's ability, or level of performance, from their cognitive style, or manner of performance (cf. competence and performance in van Dijk et al. 2011: 56). In this dissertation, cognitive style receives greater attention. The third substudy focuses on differences in syntactic complexity between learners with comparable language proficiency (see Section 4.3). Cognitive style is understood as an influence on syntactic complexity that creates a continuum between productions exhibiting low and high complexity, and no assumptions are made about relative superiority (see also Dörnyei 2005). In other words, depending on the context, both ends of this continuum can be linked with successful performance.

A learner's profile is influenced by a variety of factors, some more stable than others. Therefore, learner profiles may be relatively constant, but they can also change rapidly, even during a single production, indicating high dynamicity (Kuiken et al. 2019; see, however, *wanderers* in Kowal 2016; see also De Bot & Larsen-Freeman 2011). Learners' differing styles of processing and producing a target language are connected to their writing and writer profiles, which can describe, for instance, how learners plan and revise their productions (e.g. Van Waes & Schellens 2003). In this study, writing and writer profiles are not the primary focus, but they are interpreted as contributing factors when discussing individual variation. Learner profiles can differ – for instance, in the two modes of production, as learners prefer different types of structures in the written and spoken modes (Weissberg 2000: 37, 44–45; Granfeldt 2008: 88; García Mayo & Azkarai 2016: 244) – whereas others exhibit a more consistent style in both modes. While some research has examined the differences between written and spoken complexity at the group level, only a few studies have investigated individual variation in these two modes of production (see, e.g. Håkansson & Norrby 2007). Furthermore, profiles can differ between target languages, for instance, due to differences in the context of acquisition (such as informal vs formal learning; see Ortega 2003: 498, 500, 2008: 80).

The use of different *complexification strategies* – that is, elaboration at different linguistic levels – is a central part of the learner profile. According to earlier research, learners make use of different complexification strategies at different proficiency levels as their repertoire of structures grows; however, other factors, such as personal preferences, also affect the use of these strategies (Vyatkina 2013: 15; Forbes 2018: 3). In this dissertation, in line with Vyatkina (2013), complexification strategies are defined by the use of certain syntactic structures, where it is understood that the same

meanings can be expressed in a variety of forms. For instance, elaboration at the clausal or phrasal level, as well as the use of prepositional or infinitive phrases as modifiers in noun phrases, can be seen as complexification strategies. The choice between linguistic forms in a learner's repertoire can be made on the basis of both linguistic and nonlinguistic factors (Kuiken et al. 2019: 162). The current study examines nonlinguistic factors, namely, mode of production and individual preferences.

Categorising learners into profiles according to the syntactic aspects of their productions is not new in the field of SLA (see, e.g. Bardovi-Harlig & Bofman 1989). Similarities and differences in L2 productions at the group level have been investigated, and on this basis, learners have been categorised into learner profiles (e.g. Norrby & Håkansson 2007; Kowal 2016; see also Forbes 2018 and *developmental* [*performance*] *profiles* in Vyatkina 2013). Norrby and Håkansson (2007) studied the relationship between complexity (e.g. sentence length and subordination) and learners' PT levels (morphosyntactic development; see Section 2.2). The results reported by Norrby and Håkansson (2007) show that, on the one hand, individual variation influences the relationship between complexity and accuracy, and on the other hand, similarities between learners can be detected, and profiles can be formed, based on these similarities. The profiles articulated in Kowal (2016) are differentiated based on the relationship between the CAF dimensions, and similar categories are used to describe learners' different styles of producing a target language.

In Norrby and Håkansson (2007: 62), the *risk-taker* profile represents learners whose PT level is low but who formulate highly complex productions. A text produced by a risk-taker consists of complex structures, such as long sentences with frequent subordination. Risk-takers try to produce ideas in their L2 at the same level of complexity as they would in their L1 (see Vyatkina 2012: 578; Vyatkina et al. 2015: 42; Wijers 2019: 111). This strategy can result in inaccurate productions, as the learner might not have the linguistic resources needed to produce such complex language (Kowal 2016: 215; cf. the trade-off between complexity and accuracy in Skehan 1998 and Skehan & Foster 2012). In Kowal (2016: 157), learners with high complexity scores are categorised as either risk-takers in cases of low accuracy or as *smart learners*, whose accuracy is not negatively affected by high complexity. Another profile, described in Norrby and Håkansson (2007: 64), is the *recycler*. This profile represents learners with a high PT level but a low complexity score. Recyclers tend to use only familiar structures in their productions; consequently, their language often consists of simple structures. This means, for instance, that the sentences in the productions are short, and the productions contain few dependent clauses. In some cases, this can result in highly accurate language (see *careful learners* in Kowal 2016: 139, 157), while some learners produce language exhibiting both low

complexity and low accuracy (see wanderers in Kowal 2016: 157). In the categorisation by Norrby and Håkansson (2007: 63), *careful and thorough* learners produce language that matches their current PT level in terms of complexity. The structures that these learners produce are more varied than those of recyclers but are not at the level of risk-takers.

Kowal (2016: 139) points out that, in the case of syntactic complexity, and in the relationship between syntactic complexity and accuracy, the distinction between the profiles is not as clear as it is in lexical–semantic systems. Research has demonstrated that there is considerable overlap between profiles and variation in the use of complexification strategies, even within productions, making their categorisation highly ambiguous in most cases (Wijers 2019: 111; see also Kowal 2016). Therefore, in this dissertation, a different perspective on profiles is adopted. The focus lies solely on syntactic complexity, and complexity profiles are examined based on productions instead of learners so that both synchronic and diachronic variations in learners' style of producing syntactically complex structures are investigated (see Section 4.3). In the following section, the emergence of syntactically complex structures in L2 Swedish is discussed.

2.3.2.5   Syntactically complex structures in L2 Swedish

Swedish is the second national language of Finland (alongside Finnish), and although only a minority of the population has Swedish as their L1, all Swedish-speaking citizens have the right to use their L1 in official contexts (see, e.g. Åberg 2020: 24). This has natural consequences for the role of Swedish in the Finnish educational system, as all state personnel are required to be sufficiently competent in both official languages. In the Act on the Knowledge of Languages Required of Personnel in Public Bodies (i.e. the Language Skills Act, **§**424/2003), it is stated,

> State personnel who are statutorily required to have an academic degree are required, in bilingual authorities, to have an excellent ability to speak and write the language of the majority in the authority's district and a satisfactory ability to speak and write the other language. The requirement in a unilingual authority is an excellent ability to speak and write the language of the authority and a satisfactory ability to understand the other language.

In accordance with these requirements, Swedish is a mandatory subject in the Finnish school system (see Juurakko-Paavola & Palviainen 2011: 13; see also Ministry of Education and Culture 2018). In Finnish schools, pupils normally start studying Swedish in the sixth grade (in the case of the participants in the current study, Swedish was introduced in the seventh grade; see Section 3.1) and learn Swedish as

a B1[3]-syllabus language (OPH 2014). However, some pupils may start studying Swedish somewhat earlier as an A[4]-syllabus language (e.g. Åberg 2020: 28; see also Section 3.1; for information on the relevant learning objectives and achievements, see Åberg 2020: 28–32).

In line with the usage-based theoretical premises of the present study, it is assumed that learners begin to produce syntactically complex structures by processing linguistic input and recognising patterns in this input. Based on the regularities they notice, learners form hypotheses and slot-and-frame patterns (e.g. XVS patterns; see Section 2.1 and the discussion below), which they subsequently test in their output (see the discussion in Section 2.1). The frequency of these input patterns plays an important role in the learning process, as frequent exemplars are likelier to be used in forming constructions. Additionally, L2 learners seem to require explicit teaching to master, for instance, target-like word order in different syntactic structures. This discussion examines the acquisition of syntactic structures in L2 Swedish (for a comprehensive discussion of structures in both written and spoken L1 Swedish, see, e.g. Jörgensen 1976, 1978; Henrichsen & Allwood 2005).

Researchers have argued that syntactic structures of varying complexity emerge in L2 Swedish in a given order, and that the development of word order in these structures follows certain stages (Rahkonen & Håkansson 2008: 142; Baten & Håkansson 2015: 523; see also, e.g. Hyltenstam 1977; Håkansson & Nettelbladt 1993; Pienemann 1998; Paavilainen 2015). This developmental order begins with **simple structures that follow the basic word order** in which the subject is the first constituent (SVX, that is, subject–verb–other constituent [e.g. object]). Thereafter, learners start to produce **topicalised structures**, first with topicalised adverbs, initially **without inversion** (XSV, that is, other constituent–subject–verb), resulting in non-target-like word order (e.g. *Igår han spelade fotboll* = Yesterday he played football). Gradually, learners begin to position the **verb in the second place**, and the topicalised structures get their target-like form (XVS, that is, other constituent–verb–subject, e.g. *Igår spelade han fotboll* = Yesterday played he football). Finally, the **use of dependent clauses and the appropriate word order** are acquired in varying ways depending on whether main, auxiliary, or modal verbs are used (Baten & Håkansson 2015: 525–526, 536; see also Hyltenstam 1977; Paavilainen 2015).

Word order in Swedish is relatively fixed (SAG 4: 5) when compared to, for instance, the order in Finnish (see, e.g. Åberg 2020: 33, 34, 238). In principle, clauses in Swedish can be divided into two types according to the word order: *af* and *fa*

---

[3]   B1 stands for a compulsory, intermediate syllabus that, prior to autumn 2016, started in Grade 7 (see Åberg 2020: 28). This abbreviation is not to be confused with CEFR level B1.

[4]   A-syllabus languages start in the lower grades of primary school (see Åberg 2020: 28).

clauses (SAG 4: 7; see also SAG 4: 6; Wijers 2019: 14). In af clauses, the sentence adverbial is placed before the finite verb. This is typical of dependent clauses (SAG 4: 7, 674; see the discussion below). In fa clauses, the finite verb precedes the sentence adverbial. This order is typical of declarative main clauses (SAG 4: 7, 674; for exceptions, see SAG 4: 676; see also Åberg 2020: 34–35). Declarative main clauses can begin with either a subject or another constituent, such as an object or adverb. If the main clause begins with a constituent other than a subject, an inverted word order is used so that the finite verb stands in the second position (XVS; see, e.g. Teleman et al. 2001: 69). Therefore, in these clauses, the finite verb also precedes the subject if the latter does not stand in the initial position as the fundament of the clause. In other words, when it comes to declarative main clauses, Swedish is a verb-second language, which means that in these clauses, the finite verb stands in the second position (the *V2 rule*; see, e.g. Baten & Håkansson 2015: 519; Åberg 2020: 34).

Topicalised structures are common in both spoken and written L1 Swedish, and in both official and informal contexts (Håkansson & Nettelbladt 1993: 133, 136; see also Jörgensen 1976: 101, Lindgren 2020: 188–189; for discourse-pragmatic and semantic aspects, see Bohnacker 2010). According to Bohnacker and Lindgren (2014), 30%–40% of structures in written Swedish are topicalised. Therefore, although topicalised structures are relatively frequent, the direct word order, in which the subject is the first constituent of the sentence, is more common. It is of note that the text type affects the frequency of topicalised structures. For instance, in narratives, topicalised temporal adverbs are especially frequent (Lindgren 2020: 198). According to the assumptions of this dissertation (see Section 2.1), topicalised structures, such as structures with topicalised adverbs of time (e.g. *Idag ska vi städa* = Today will we clean), should be relatively easy to learn, as these syntactic patterns are frequent in learners' input (see the discussion in Section 2.1).

However, despite their relatively high frequency, previous studies have shown that L2 learners start to produce topicalised structures at higher proficiency levels (Rahkonen & Håkansson 2008; see also Bohnacker 2010: 133). Furthermore, earlier studies of the development of syntax in L2 Swedish have demonstrated that, at the beginning of the learning process, learners use only the direct word order, even when they begin to produce topicalised structures. Therefore, L2 learners of Swedish seem to have long-lasting difficulties with learning to use inversion after topicalised structures (e.g. Rahkonen & Håkansson 2008; Åberg 2020). The learner's L1 can impact this learning process. In the case of the current study, the learners' L1, Finnish, has a relatively free word order (see, e.g. Åberg 2020: 33, 34, 238), which might affect how these learners pay attention to word order (see also Section 2.2). Learners of Swedish whose L1 is Finnish may not automatically notice these differences in word order patterns. Thus, explicit instruction plays an important role

in redirecting learners' focus and teaching target-like word order (see, e.g. Åberg 2020). Bohnacker (2010) suggests that native German speakers learning Swedish seem to first acquire the use of topicalised structures in spoken production. This is somewhat surprising, as learners have more control over their productions in writing (see Section 2.3.2.3).

In Swedish, word orders differ in main clauses and dependent clauses when it comes to the placement of the sentence adverbial (e.g. negation). Most dependent clauses are af clauses; that is, in dependent clauses, the sentence adverbial comes before the finite verb (e.g. SAG 4: 465–468; see also Wijers 2019: 13; Åberg 2020: 35–36). Therefore, the V2 rule does not apply to dependent clauses, and instead, the V3 rule is applied (Baten & Håkansson 2015: 519, 524). In practice, learners can differentiate between main and dependent clauses by relying on lexical clues – for instance, based on a subordinating conjunction at the beginning of dependent clauses – and apply the different word orders accordingly (Baten & Håkansson 2015: 523).

Dependent clauses can function as object of verbs in the main clause (nominal function) or as attributes relative to noun phrases (relative function); they can also be clauses with an adverbial conjunction and express, for instance, time or causality (adverbial function) (SAG 4: 468, 469–472; Lambert & Nakamura 2019: 249–250; Wijers 2019: 12–13). Therefore, dependent clauses cannot usually convey meaning independently, but this does not mean that the message in the dependent clause is somehow of lesser importance (SAG 4: 475). In line with this, Jörgensen (1978: 183) points out that in complex sentences (i.e. sentences containing a main clause and a dependent clause), dependent clauses often carry the most important information (e.g. the information after matrix clauses, such as *Jag tycker att* = I think that).

Dependent clauses are part of the natural and varied use of Swedish. For instance, relative dependent clauses are frequent in spoken L1 Swedish (Håkansson & Norrby 2007: 92). According to previous research, there is a clear connection between input and output in the frequency of dependent clauses (e.g. Wijers 2019: 115–124). In certain learning contexts, the frequency of dependent clauses in learners' input can be relatively restricted, as traditionally, these clauses do not frequently occur in teaching materials (see Wijers 2019; Åberg 2020). Other types of input, such as input based mainly on spoken communication, can depict a rather one-sided picture of dependent clauses, as certain types of dependent clauses are typical of spoken discourse (see Section 2.3.2.3). Again, as with topicalised structures, the effect of the text type on the frequency of dependent clauses therein must be remembered (e.g. Wijers 2019: 50–51; see also Kormos 2014). It can be expected, for instance, that argumentative texts contain many dependent clauses, as arguments are often formulated with the help of dependent clauses. Furthermore, there are differences

between the frequencies of dependent clauses in the two modes of production (see, e.g. Biber et al. 2011). According to Håkansson and Norrby (2007), learners of Swedish seem to produce dependent clauses somewhat more frequently in writing than in speech.

Wijers (2019) studied the use of dependent clauses in written L2 Swedish by 21 L1 Dutch learners. Both the proportion of dependent clauses and the frequency of different clause types were examined. According to the results, learners achieve native-like proportions of dependent clauses relatively early in the learning process (see also Baten & Håkansson 2015), while the differences between L1 and L2 users in the use of different clause types are more persistent (see also Granfeldt 2008). In contrast, the results reported by Kowal (2016: 86) demonstrate that learners start to complexify their productions by adopting embedding later in the learning process. As research has also shown that the amount of embedding used can decrease at higher proficiency levels (see Section 2.3.2.1), the study of dependent clauses as part of developing L2 Swedish syntax seems rather complicated. Based on earlier findings, it can be concluded that the use of embedded structures is not challenging per se for L2 learners of Swedish. In fact, due to the stricter word order and more limited variation, dependent clauses can be considered less complex than main clauses (Wijers 2019: 33–34; cf. absolute complexity in Section 2.3.1).

However, the difference between word orders in main and dependent clauses (V2 and V3 rules) can cause long-lasting difficulties, especially when it comes to dependent clauses containing modal verbs and sentence adverbials (Norrby & Håkansson 2007: 50; Baten & Håkansson 2015: 525, 531–532, 536, 538; see also Paavilainen 2015; Åberg 2020). Target-like word order in dependent clauses with sentence adverbials emerges relatively late in L2 Swedish, regardless of the learning context (Rahkonen & Håkansson 2008: 149, 151; see also Åberg 2020). These structures are also relatively rare in L1 Swedish (Jörgensen 1978: 188; Rahkonen & Håkansson 2008: 153, 154–156; Wijers 2019: 119–120), which presumably contributes to the late emergence of these structures in L2 production (see Section 2.1). The difficulty of mastering the different word orders can lead to the use of avoidance strategies, which, in turn, can generate lower complexity if measures such as dependent clause ratio are used to examine syntactic complexity (Baten & Håkansson 2015: 538).

Based on these remarks, it can be concluded that both **topicalised structures and dependent clauses with a sentence adverbial can be challenging structures for learners of Swedish**. Thus, these structures are examined in the current study in relation to relative complexity (see Section 3.3.2). In the following section, the role of syntactic complexity in CEFR (2001, 2007, 2020) is discussed.

2.3.2.6   Syntactic complexity in the Common European Framework of Reference for Languages

CEFR was originally published in 2001 by the Council of Europe to provide a common basis for the development of the learning and teaching of L2s (CEFR 2007: 1, 2020: 13, 27–29), and it has been widely used ever since for creating syllabuses and learning materials, as well as for testing language proficiency both in and beyond Europe (Gyllstad et al. 2014: 1–2). Later, further materials were created, and in 2018, an extended version with more illustrative descriptors was introduced in response to criticism of the empirical validity of CEFR and growing demands among language educators and researchers (Gyllstad et al. 2014: 2, 4–5; CEFR 2020: 13–15, 23–24). CEFR contains, among other tools, a conceptual model of language proficiency and language-independent descriptive scales of language proficiency and use in the different domains of language (CEFR 2020: 31–38). The scales consist of levels A1, A2, B1, B2, C1, and C2 (and, in the latest addition, pre-A1), which can be grouped into the categories of Basic User (levels A1 and A2), Independent User (levels B1 and B2), and Advanced User (levels C1 and C2) and further divided into sublevels, which are used in the Finnish education system (CEFR 2020: 36; OPH 2014). The descriptions in the CEFR levels are 'can do' statements that describe, for instance, what communicative activities a learner can carry out and how different linguistic elements are used to accomplish different tasks (CEFR 2020: 35, 43; see also Gyllstad et al. 2014: 3 and Appendix 2). The correlation between these descriptions and actual L2 performance has been widely studied in the field of SLA (see, e.g. Bartning, Martin & Vedder 2010).

In CEFR, syntactic complexity is referred to as part of the learner's developing linguistic knowledge of an L2 (see, however, the discussion in Gyllstad et al. 2014: 5–6). It is argued that learners' productions proceed from isolated words, simple phrases, and short sentences to the context-appropriate level of complexity by integrating these elements into increasingly complex structures (CEFR 2020: 62, 67). Complexity is referred to as the range of language at the learner's disposal, and this repertoire of structures is said to expand from 'memorised phrases to a very wide range of language [used] to formulate thoughts precisely' (CEFR 2020: 130). In Table 2, some excerpts pertaining to syntactic complexity in the CEFR descriptions from the General Linguistic Range scale are presented (see Appendix 2 for the descriptions used in this dissertation).

Table 2. Syntactic complexity based on CEFR descriptions (CEFR 2020: 130–131, emphasis added).

CEFR level	Description
C2	Can exploit a comprehensive and reliable mastery of **a very wide range of language** to formulate thoughts precisely, give emphasis, differentiate and eliminate ambiguity.
C1	Can select an appropriate formulation from **a broad range of language** to express themselves clearly, without having to restrict what they want to say. Can use **a broad range of complex grammatical structures appropriately** and with considerable flexibility.
B2	Has **a sufficient range of language** to be able to give clear descriptions, express viewpoints and develop arguments without much conspicuous searching for words/signs, using **some complex sentence forms** to do so.
B1	Has **enough language to get by**, with sufficient vocabulary to express themselves with some hesitation and circumlocutions on topics such as family, hobbies and interests, work, travel and current events, but lexical limitations cause repetition and even difficulty with formulation at times.
A2	Has **a limited repertoire of short, memorised phrases** covering predictable survival situations; frequent breakdowns and misunderstandings occur in non-routine situations. Can use **basic sentence patterns** and communicate with **memorised phrases, groups of a few words/signs and formulae** about themselves and other people, what they do, places, possessions, etc.
A1	Can use **some basic structures in one-clause sentences** with some omission or reduction of elements. Has **a very basic range of simple expressions** about personal details and needs of a concrete type.
Pre-A1	Can use **isolated words/signs and basic expressions** in order to give simple information about themselves

According to these descriptions, learners at the lower proficiency levels – that is, pre-A1, A1, and A2 – produce isolated words and have a (very) limited repertoire of memorised expressions and basic sentence patterns. This indicates that language produced at this level is chunk-like and mainly consists of simple structures. At the intermediate levels, that is, B1 and B2, a learner's language repertoire is large enough to complete (simple) tasks, and some complex sentence patterns emerge, whereas at the highest levels, C1 and C2, learners possess a (very) broad range of complex structures and can apply these in their productions appropriately according to a given context.

The interplay between syntactic complexity and accuracy is also discussed in CEFR. The emergence of new forms (i.e. range) in learner language is seen as more important than their mastery (i.e. control), and it is acknowledged that attempts to produce more complex structures can diminish accuracy in learners' productions (CEFR 2020: 130). Formulating complex thoughts or completing demanding tasks requires that learners use their cognitive resources effectively. It is stated that when

learners devote their attentional capacity to fulfilling a complex task, accuracy can be negatively affected, and that earlier research on L2 English, French, and German indicates that this is common at around CEFR level B1, as it is at this level that learners begin to experiment with structures (CEFR 2020: 132). The relationship between syntactic complexity and accuracy at the different CEFR levels is described as follows (CEFR 2020: 132): At level A1, learners are expected to show 'only limited control of a few simple grammatical structures and sentence patterns'. Learners at level A2 can produce some simple structures according to the target language rules, but basic mistakes are common, whereas learners at level B1 can produce relatively accurate language in familiar contexts. At level B2, learners have good control of simple structures and of some more complex forms. However, at this level, errors can still be expected if complex language is produced. Beginning at level C1, learners' productions are described as highly accurate, and at level C2, learners maintain 'consistent grammatical control of complex language, even while attention is otherwise engaged'.

It has been argued that although syntactic complexity is addressed relatively frequently in CEFR, the references are unsystematic and lack empirical evidence (Gyllstad et al. 2014: 5–6; Kowal 2016: 34–35). With this in mind, some SLA studies have investigated differences in syntactic complexity in L2 productions at different CEFR levels. Gyllstad et al. (2014) found that the CEFR level of a given production seems to correlate with its syntactic complexity in written L2 English, L3 French, and L4 Italian. However, there appear to be some differences between target languages when it comes to syntactic complexity. Khushik and Huhta (2022) study syntactic complexity in written L2 English by L1 Finnish students. In line with earlier studies of the development of syntactic structures in L2, they state that the lower CEFR levels (A1–A2) are clearly separated by length-based measures, whereas at the higher levels (B1–B2), the clearest separators between the levels are, for instance, clausal- and phrasal-level measures (for similar results, see, e.g. Lahuerta Martínez 2018).

Yamaguchi (2018) studied both written and spoken L2 English samples by 60 L1 Japanese learners and investigated, for instance, the development of syntactic structures according to PT levels, as well as their relation to CEFR levels. According to the results, PT levels seem to be linearly connected with CEFR levels in both modes of production, but learners at the highest PT levels are not necessarily at the highest CEFR levels. Maeda (2021) examined the differences in complexity between CEFR levels in spoken L2 English by 153 L1 Japanese students. The syntactic complexity measures used in the study – namely, the ratio of subordinate clauses per AS-unit and the mean length of AS-unit – seem to distinguish the CEFR levels relatively efficiently (see Maeda 2021: 39–40). In this dissertation, the focus is not on the validity or feasibility of using CEFR levels as reference points for L2

proficiency, but as **the productions used as data in the current study were assessed according to CEFR levels**, and a modified version of the descriptions was used in the evaluation (see Section 3.2), the remarks of previous researchers must be kept in mind.

To conclude this chapter, we can summarise the main theoretical views of this dissertation. In line with the discussion in this section, and following the theoretical framework presented earlier, it is assumed in the current study that

a) language learning happens through language use

b) by processing linguistic input (following the frequency principles), learners form patterns that they test in their productions

c) learners' cognitive resources are limited, and this affects the structures produced, especially in the spoken mode at lower proficiency levels

d) the automatisation of processing and earlier experiences with language(s) can increase syntactic complexity in learners' productions

e) learners' individual styles of processing and producing language have a major effect on the structures produced

In the following section, the data and methods used in this study are presented.

3　　Data and methods

In this chapter, the methods of the current study are presented along with the data analysed. Section 3.1 introduces the study participants, and in Section 3.2, the data collection procedures, as well as the written and spoken data collected for the study, are described. In Section 3.3, the measures used to analyse the data are discussed. Section 3.3.1 describes the segmentation units adopted. In Section 3.3.2, the complexity measures chosen and their relation to earlier research are discussed, followed by a brief presentation of the accuracy measures used in the first substudy in Section 3.3.3. Thereafter, in Section 3.4, the analyses conducted and the principles followed in these analyses are presented. The chapter concludes with a summary of the data used and the analyses carried out in the separate substudies.

3.1　　Participants

The ethical guidelines of the Finnish National Board on Research Integrity (see TENK 2023) and the European data protection rules (under the Data Protection Directive 1995; see also the General Data Protection Regulation, GDPR 2018) were followed in the implementation of the current study. The data were gathered with permission from both the heads of the educational units and the participants. Participation in the research was voluntary, and all participants gave their informed consent to contribute data. They were also informed that they could terminate their participation in the study at any time. As all participants were adults, no parental permission was needed. The participants' anonymity was ensured during all phases of the project. All data were pseudonymised, and the participants were assigned individual codes[5]. The assistants who helped transcribe the spoken data signed a

---

[5]　The codes refer to the learner's proficiency level – FN1 (FN = *färdighetsnivå*, 'proficiency level' in English) for the lower proficiency level and FN2 for the higher level (see the discussion in Section 3.2) – and code (01–48). For instance, FN1.01 stands for Participant 1, which is at the lower proficiency level. In the third substudy, the learners are referred to using participant codes without proficiency levels, as all participants in the study belong to the higher proficiency group in the written data, e.g. for P04, P = the participant, 04 = their individual code).

contract that protects the participants' anonymity. Furthermore, secure storage of the data was ensured.

The data were collected at a university in Southwest Finland. Some of the participants were **majoring in Scandinavian languages**, while the others were **participating in a Swedish course** at the Centre for Language and Communication Studies. The latter had varying majors, such as law and computer science. In the first meeting with the participants, the research project was briefly introduced. Only essential information about the study was shared with the participants so that the aims of the study would not affect their contributions. In these meetings prior to the data collection, with the help of background information questionnaires, basic information about the participants, such as their gender and age, as well as some central information about their linguistic background, such as possible multilingualism and the use of L2s outside the educational context, was recorded.

A total of 48 students consented to participate in the first data gathering phase, and 35 students contributed both written and spoken productions. From this dataset, four bilingual Finnish–Swedish students were excluded, as productions in Swedish by these students were not considered L2 samples. One participant was bilingual in Finnish and Estonian, and one identified as bilingual in Finnish and English because they frequently used English in everyday life. The productions by these students were included in the dataset. Thus, **31 students** were included in the first two substudies. These participants were, on average, 22.1 years old (min 19, max 34). The gender distribution between females and males was relatively equal, with 16 females and 14 males, as well as one non-binary participant. The subjects for the third substudy, which focused on the individual level, were chosen from the larger dataset based on more specific criteria for syntactic complexity in their productions (see Section 3.4). These participants included one 22-year-old male and two 20-year-old females.

Only students majoring in Scandinavian languages participated in the second round of data collection. This dataset consists of written and spoken production by 13 students who also participated in the first data collection phase. An additional two students who were not part of the first dataset participated in the second data collection stage. In this dissertation, only parts of the written longitudinal data are used (see Section 3.4).

According to the background questionnaires, the learners were rather homogenous in their linguistic backgrounds, except concerning the role of Swedish in their studies and everyday lives. Those majoring in Scandinavian languages reported using Swedish outside the educational context more often than the other participants. The older students had, of course, been using Swedish for a longer time than the younger students. Many participants reported using languages other than Finnish in both formal and informal contexts. For instance, several participants

reported using English with their foreign friends. Additional languages besides Swedish and English were relatively frequent, and the most commonly studied languages were German (14/31 learners), French (9/31), and Spanish (7/31); other languages, such as Russian and Japanese, were also studied. These additional languages were somewhat more common among those majoring in languages. Most participants in both groups of learners had studied Swedish for approximately six years as a B1-language (i.e. from Grade 7 onwards), but some students had studied Swedish somewhat longer as an A-language (from Grade 3 or 5 onwards). Again, A-syllabus studies in Swedish were more common in the group majoring in Scandinavian languages. Specific numbers are not reported here, as not all participants specified the syllabus used in their Swedish studies in the background information questionnaire.

3.2   Written and spoken data

The current study is based on samples of **written and spoken L2 Swedish** contributed by L1 Finnish university students. At the beginning of the data collection phase, teachers were contacted at both the Department of Scandinavian Languages and the Centre for Language and Communication Studies. Meetings were scheduled with these teachers, and at these meetings, the possibility of collecting data in their courses and the practical arrangements for doing so were discussed. After the teachers' approval to collect data during lessons in different courses was received, the heads of the educational units were contacted to secure their permission.

Before gathering the data, both the tasks and practical arrangements were carefully planned with the help of research literature and earlier study designs (see, e.g. Chaudron 2003; Kormos 2014). **Informal essays** and **monologues** were considered ideal for the purposes of the current study, as these types of productions are similar, and the use of similar structures can be expected (Kormos & Trebits 2012: 462). In the spoken mode, tasks based on monologues were used instead of dialogues, as including interaction in a task can have undesirable effects (e.g. interrupted sequences of speech) and therefore negatively affect the complexity of the structures used (Leonard & Shea 2017: 181; Vercellotti 2019: 237). In dialogue, successful interaction may even require the use of simple and short structures (Clercq & Housen 2017: 320), and as Ferrari (2012: 278) noted, in monologues, learners are likelier to produce syntactically complex structures. The spoken productions in the current study can be categorised as spontaneous speech with some features of planned language.

Cartoon strips (see Appendix 1) were used to prompt both the written and spoken responses, as visual material is often used as the basis for gathering this type of data (see, e.g. Kormos & Trebits 2012; for studies of fluency, see Segalowitz

2010: 43; see also Lennon 1990). In tasks based on visual material, such as a cartoon strip with a predetermined storyline, learners do not need to use their resources to create a story. In other words, the planning effort is decreased in such tasks, and this reduces the cognitive load in the planning phase (Kormos 2014: 198–199). Furthermore, using visual prompts can make productions more comparable, as the same storyline is followed by every participant (see, e.g. Segalowitz 2010: 43).

Based on previous experiences with similar study designs, it was deemed necessary to include **additional questions** in the tasks so that the structures produced would be more varied, and the productions would be inspired by both visual material and spontaneous discourse (see the discussion in Segalowitz 2010: 44). The questions were written in the participants' L1, Finnish. With the help of the questions, the participants were asked to describe what happened in the cartoon strip, but also to talk about their own experiences with the topic and discuss it more generally. Therefore, the tasks consisted of a **descriptive** section, a **narrative** section, and an **argumentative** section; the different parts of the productions mainly follow this division. A mixture of text types enables the production of diverse structures of varying complexity. The disadvantage of this study design is that the effect of the text type is not controlled for, and it is difficult to compare the results with earlier findings, as task and text types affect the structures used in a given production (Jörgensen 1976: 42–44, 125; De Clercq & Housen 2017: 320; Kuiken & Vedder 2019: 197; Kuiken et al. 2019: 164–165; Wijers 2019: 50–51; see also Kormos 2014 and the discussion in Section 5.3). Due to this hybrid nature, no specific communicative goal was set in the tasks. The same task design was used to collect both written and spoken productions. This means that differences in complexity between these types of productions could be reduced, as similar tasks elicit structures of the same type (Kormos & Trebits 2012: 462). Therefore, the productions in the two modalities should correspond with one another relatively well and elicit comparable types of structures. Different cartoon strips with varying topics were used in all tasks to prevent the participants from, on the one hand, trying to repeat the same production in the written and spoken tasks, and on the other hand, trying to remember their first productions in the second stage of data collection.

The next step in planning the data collection was to decide on the topics to be used in the cartoon strips and on additional questions. Topics that could be expected to be familiar to the students were chosen (see, e.g. Baaijen & Galbraith 2018: 201). In the first written task, the topic of the assignment was the student's dream job; in the first spoken task, the cartoon strip and the questions dealt with a normal day as a student. In the second stage of data collection, the participants discussed the use of money in the written task, and the use of smart devices and social media in the spoken task. The effect of the participant's (at this point assumed) proficiency level was taken into consideration in choosing the topics, as learners with both lower and

higher language proficiency took part in the first data collection phase. An appropriate level of difficulty was ensured by consulting a teacher of the less proficient group. Based on further discussions with more experienced researchers, it was deemed necessary to add **a word list** with some key words for the students' aid. The rationale here was that, in tasks based on a specific topic and using a detailed cartoon strip, the learners are obliged to exploit their (possibly) limited vocabulary (see the discussion in Jung 2017: 193–194). If the learner is unfamiliar with the vocabulary needed to complete the task, the language used in the production might be less complex. Although the processing load in the tasks is reduced by the fact that the productions are based on predetermined content (as noted above), the need to express this content using limited resources can increase the cognitive load (Kormos 2014: 199). Without a word list, less proficient learners could experience difficulties with expressing themselves.

After the chosen topics were discussed with the supervisors of this dissertation, the cartoon strips were planned together with an illustrator. The topics were discussed, and based on the discussions, the illustrator made the first drafts. The supervisors commented on the drafts, after which the final versions were drawn. The illustrator consented to the use of the cartoon strips for research purposes. Finally, comments on the cartoon strips, questions, and word lists were requested from other experienced researchers. Based on these comments, the word lists were slightly modified (Appendix 1 contains the cartoon strips, questions, and word lists used in the present study).

The first data collection phase was carried out in autumn 2016, at the beginning of the academic year, and the second in spring 2017, at the end of that academic year. The teachers of the courses, the supervisors of this dissertation, and an assistant provided help during the data collection sessions. Two groups of learners – those majoring in Scandinavian languages and the students taking the Swedish course at the Centre for Language and Communication Studies – participated in the first round of data collection. The majors completed the tasks during one 90-minute lesson – first the written task and then the spoken task – whereas the other students first took part in the written data collection and, about a week later, the spoken tasks were performed. The same data collection procedure was followed for both groups. Only the Scandinavian language majors participated in the second data collection phase. Again, the tasks were completed during one lesson, with the written part preceding the spoken part.

Those who taught these students' courses assisted with the practical arrangements by booking the rooms for data collection. The spoken data were audio-recorded in language laboratories, and the written data were gathered in IT classrooms. The keystroke logging programme ScriptLog was used in the collection of the written data. The programme was downloaded on the computers used by the

IT service staff. ScriptLog records what users type and information about process-related features of the writing activity, such as pauses and deletions (see, e.g. Strömqvist, Holmqvist, Johansson, Karlsson & Wengelin 2006; for process-oriented studies in L2, see, e.g. Gunnarsson 2012; Jung 2017), and enables the subjects' writing to be analysed both as a process and an end-product (Kowal 2016: 56). The writing process is not studied in this dissertation, but the use of a keystroke logging programme in the data collection enables further research on, for instance, how participants revise syntactically complex structures in their productions (see Section 5.3).

Each participant had approximately 40 minutes to plan and write the essay. For the spoken task, no specific time limit was set, but, when necessary, they were reminded that the lesson was about to end. The participants were given the opportunity to plan their production before the tasks by familiarising themselves with the cartoon strip, questions, and word list for about five minutes per task (see *pre-task planning*, Yuan & Ellis 2003). However, no tools were provided for note-taking. In the written task, it was possible to plan the text in the writing programme, but in the spoken task, planning could only take place mentally. The participants had the papers with the cartoon strip, questions, and word list at their disposal, but no additional tools (e.g. dictionaries) were allowed during the tasks. For the written part, a goal of about 100–150 words was set with no upper limit, as this was considered feasible for both groups of learners. In the spoken task, the participants were encouraged to produce as much speech as possible.

At the beginning of the sessions, the participants received general instructions together with the tasks. In the written task, these instructions contained information, for instance, on how to start and end the writing process and how to fill in a questionnaire in ScriptLog. For the spoken task, the participants were told, for instance, to turn off their cell phones and say their individual code out loud when the recording began. Additionally, they were told to keep their headphones on when they were ready with their productions, as all participants completed the spoken task in the same room at the same time. After completing the written and spoken tasks, the Scandinavian language majors participated in a group interview about the tasks and the data collection in general. This interview was arranged to ensure that there were no major problems with completing the tasks.

The participants' productions were **evaluated** by two independent raters according to **modified CEFR descriptions** (see Appendix 2). Lexical repertoire and the ability to convey information were emphasised in the criteria, whereas complexity- and accuracy-related descriptions were excluded to minimise the risk of circular reasoning between the evaluations and research results. Although complexity and accuracy were not part of the assessment criteria, these features in the learners' productions may have influenced the evaluations, especially since the

raters were experienced in assessing learners' productions and, as a result, familiar with several types of criteria used in the evaluation of L2 output. In the case of disagreement on proficiency levels, the raters discussed their evaluations to come to a mutual understanding, or a third rater was consulted. The CEFR levels of the productions studied ranged from A1 to C1. Based on the evaluations, the learners were divided into two groups: **a lower proficiency level** consisting of productions at levels A1 and A2 (Basic User, CEFR 2020: 36), and **a higher proficiency level** consisting of productions at levels B1 and B2 (Independent User, CEFR 2020: 36) and one written production at level C1[6]. In the first and second substudies, these proficiency levels were compared (for a similar approach, see, e.g. Gyllstad et al. 2014: 17). From the second data collection phase, only the productions included in the longitudinal analysis of complexity in the written mode (see Section 4.3) were evaluated.

The written and spoken contributions differ by group, and the groups were formed based on the proficiency levels exhibited in these productions. The written data from the less proficient learners comprise productions by 17 participants, and the higher level includes productions by 14 participants. The group with lower proficiency (n = 17) included data from 5 females, 11 males, and 1 nonbinary participant, and the average age in this group was 21.8 years (min 19, max 27), whereas the group with higher proficiency (n = 14) included 11 females and 3 males with an average age of 22.5 (min 19, max 34). The number of participants at the lower level is somewhat greater in the spoken data than in the written data, as the productions by four learners (FN1/2.01, FN1/2.08, FN1/2.11, and FN1/2.13) were evaluated at a lower proficiency level in the spoken mode. The spoken data include 21 less proficient learners (9 females, 11 males, 1 nonbinary; average age 22.0, min 19, max 34), whereas the group with higher proficiency contains 10 students (7 females and 3 males; average age 22.2, min 19, max 30). Table 3 illustrates the descriptive data, which include the mean[7] number of words (M), median[8] value, minimum (min) and maximum (max) values[9], and standard deviation[10] (std dev) of the written and spoken productions at the two proficiency levels.

---

[6]   This production does not differ in complexity from the other productions assigned to the higher proficiency level. Therefore, it is included in the analysis of the higher-level responses.

[7]   The mean is the average value of a group of values (see, e.g. Larson-Hall 2016: 478)

[8]   The median is the centre of the values, where 50% of the values are higher and 50% are lower (see, e.g. Larson-Hall 2016: 80, 478).

[9]   The minimum and maximum values are the smallest and largest values in a group of data.

[10] The standard deviation indicates the variation around the mean (see, e.g. Larson-Hall 2016: 483).

Table 3.   Descriptive statistics of the written and spoken data.

	Lower proficiency level	**Higher proficiency level**
Written task	**(n = 17)**	**(n = 14)**
mean number of words	160.1	303.7
median	158	280
min	67	185
max	279	492
std dev	61.2	93.8
Spoken task	**(n = 21)**	**(n = 10)**
mean number of words	153.6	529.6
median	108	116
min	56	268
max	479	1108
std dev	111.5	300.2

As shown in Table 3, there is a notable difference in length of production (as measured by number of words) both between and within the participants at the two proficiency levels. For instance, there is a considerable difference in the number of words in spoken productions between the lower (M = 153.6) and the higher proficiency levels (M = 529.6), and the shortest spoken production at the higher proficiency level is 268 words, whereas the longest production is 1,108 words long (std dev 300.2). Additionally, the written productions (M = 160.1) at the lower proficiency level are, on average, somewhat longer than the spoken productions (M = 153.6), whereas the written productions (M = 303.7) at the higher proficiency level are clearly shorter on average than the spoken productions (M = 529.6). This variation in the number of words should be considered when interpreting the results, as most of the complexity measures used in this study are affected by the number of occurrences in a given production (see also Section 5.3). The longer the production is, the more possibilities there are for a certain structure to be used. Despite this effect, length was not considered, and the productions are analysed as wholes (not delimited to, for instance, 100 words), as the focus of this study is learners' concrete ability to produce structures in L2 Swedish (see *ecological validity*, Skrzypczak 2021).

3.3   Measures

As discussed earlier, the study of syntactic complexity can be problematic due to the multifaceted nature of complexity itself. The main challenge faced by researchers lies in operationalising complexity. Following the recommendations in previous research (see, e.g. Norris & Ortega 2009; Bulté & Housen 2012, 2020b), this study

adopted a multidimensional approach to syntactic complexity: Both absolute and relative complexity (see Section 2.3.1) are investigated, and both general and fine-grained measures of syntactic complexity are used in the analyses. In this section, the segmentation units (Section 3.3.1) and all measures used – general, phrasal-level, and specific (relative complexity) – are presented along with a critical review of the operationalisation of L2 syntactic complexity in earlier studies (Section 3.3.2). Additionally, the syntactic accuracy measures used in the first substudy are introduced (Section 3.3.3).

3.3.1 Segmentation units

Segmentation units function as the basis for the complexity and accuracy measures used in this study. As discussed earlier, the use of segmentation units and their definitions vary greatly in existing research, which complicates the interpretation of results and makes it difficult to compare findings from earlier studies. Therefore, it is necessary to define these units explicitly. In this section, the segmentation units used in the current study – words, clauses, dependent clauses, T- and AS-units, and noun phrases – are introduced. Additionally, the specific structures studied – that is, topicalised structures and dependent clauses with a sentence adverbial – are defined. These units are summarised in Table 4 (for the symbols used in the transcriptions, see Appendix 3).

Table 4. Segmentation units.

Unit	Written (Wr) /Spoken (Sp)	Definition	Abbre-viation	Example
Word	Wr+Sp	a segment of production separated by two spaces or a space and a punctuation mark or a unit transcribed as a whole and separated by two spaces or other characters (see, e.g. Biber, Johansson, Leech, Conrad & Finegan 2021: 53, 56–57)	W	*betalningar / att / ta / hand / om* = bills to take care of (FN2.15) *å (= och) / jag / skriver* = and I write (FN1/2.08)
Clause	Wr+Sp	a unit minimally consisting of a subject and a finite verb (SAG 4: 3–4)	C	*Nu tror jag* **C ---** = Now, I think --- (FN2.15) *å jag (0.7) skriver (0.5) essäer (2.5) å läser (1.1) skolböcker (5.4)* **C** = and I write essays and read schoolbooks (FN1/2.08)

Unit	Written (Wr) /Spoken (Sp)	Definition	Abbre- viation	Example
Dependent clause	Wr+Sp	an embedded clause generally beginning with a dependent marker word followed by a subject (sentence adverbial) and a finite verb as well as other possible constituents (SAG 4: 5, 462, 468)	DC	*Han har ingen flickan **C** eftersom han är inte trevlig. **DC C** =* He doesn't have a girl because he isn't nice. (FN1.12) *den här är en (.) pojke **C** som vaknar (.) klockan sju (.) ööö (9.8) **DC C** =* this is a boy who wakes up at seven o'clock (FN1.27)
T-unit	Wr	a main clause and all possible dependent clauses, as well as all nonclausal units or fragments attached to it (Hunt 1965: 20, 49)	T	*Han är intresserad av anatomin **T** och han vill hjälpa andra människor. **T =** He is interested in anatomy, and he wants to help other people. (FN1/2.13) Det kan hända att jag måste jobba på något annat yrke än lärare eller översättare. **T** =* It is possible that I have to work in another profession than teaching or translating. (FN2.04)
AS-unit	Sp	a main clause, or a subclausal unit, and all possible dependent clauses associated with it (Foster et al. 2000: 365)	AS	*och öö sen ska hon (.) till föreläsningen (.)*pt* *hh* **AS** och ää hon studerar hela dan↑ (1.3) **AS** =* Then she goes to the lecture and she studies all day. (FN2.04) *när jag kom från skolan jag rastar min hund (.) e:n- (.) och (4.9) ser min (.) kompisa:r (11.0) **AS** =* when I come from school I walk my dog and meet my friends (FN1.27)
Noun phrase	Wr+Sp	a phrase consisting of a possible article or another determiner, a head noun or a pronoun, and/or a modifier (see, e.g. SAG 1: 202, SAG 3: 11–183; see also Biber et al. 2011)	NP	*så många stora betalningar att ta hand om =* so many big bills to take care of (FN2.15) *min pojkvän =* my boyfriend (FN2.11)
Dependent clause with a sentence adverbial	Wr+Sp	a dependent clause containing a sentence adverbial (see, e.g. Rahkonen & Håkansson 2008)	DC+	*Ibland även när man redan är vuxen **DC C** =* Sometimes even when you are already an adult (FN2.12) *om jag inte ha (1.0) jobb↑ (0.8) på den dag **DC C** =* if I don't have work that day (FN2.06)
Topicalised structure	Wr+Sp	a declarative T- or AS-unit with another constituent than a subject in the initial position (see, e.g. Teleman et al. 2001)	TOP	*I gymnasiet drömde han att han vill bli en läkare. =* In upper secondary school, he dreamed that he wants to become a doctor (FN1.09) *i dag (0.4) jag (.) vaknade (2.7) klockan nio↑ (2.0) och (1.8) ät:tade (5.9) smörgås↑ (2.2) och te↑ (1.2) *pt* (0.9) =* Today I woke up at nine and ate a sandwich and tea (FN1.16)

Words are commonly used as the basis for length-based measures of complexity. 'Words' can be generally defined as the occurrences that make up a spoken or written production, and the concept rarely leads to misunderstandings (see the discussion in Biber et al. 2021: 53, 56–57). However, it is often necessary to specify which linguistic materials are counted as words. In the present study, a word is defined as a written whole separated by two spaces or a space and a punctuation mark in written data, or as a unit transcribed as a whole and separated by two spaces or other characters in spoken data. All occurrences of words were included in the analyses, including numbers written as digits, and possible compounds were counted as they occurred in the productions, whether they were target-like or not. In the spoken data, the words occurring in unfinished utterances were excluded from the word count (see Section 3.4).

Clauses can be divided into main and **dependent clauses** (also called independent and subordinate clauses). Main clauses can stand independently from other clauses, whereas those embedded in other clauses are dependent clauses (SAG 4: 5, 468; on different functions, see SAG 4: 469–472; Lambert & Nakamura 2019). In Swedish, these different types of clauses are often distinguished by differing word orders (see Section 2.3.2.5). Typically, a dependent clause begins with a subordinating conjunction, a relative pronoun, or an interrogative pronoun or adverb, and dependent clauses in Swedish usually follow the af word order (i.e. both the subject and possible sentence adverbials precede the finite verb) (SAG 4: 462). Simple structures can be made up of one main clause, whereas complex structures contain either several coordinated main clauses or a main clause and one or more dependent clause(s).

The definition of 'clause' varies, and in some studies, no definition is offered (Bulté & Housen 2012: 39; Lintunen & Mäkilä 2014: 382–383; Kuiken & Vedder 2019: 198). This is especially alarming when the number of studies that have measured the use of dependent clauses as an indicator of syntactic complexity is considered. According to some definitions, a clause must contain a finite predicate verb, whereas in other definitions, this demand is not made (see, e.g. Foster et al. 2000: 366; Kuiken & Vedder 2019: 198). In the current study, only clauses with finite verbs are counted. Non-target-like versions of a finite verb (e.g. an infinite verb in contexts where a finite verb is required) are interpreted as finite verbs. Additionally, following Hunt (1965: 15), a clause must contain a subject. This demand is based on the fact that structures with listed activities (e.g. *Och sen kom jag hem, tittade på tv, lyssnade på musik och bara slappnade av resten av dagen* = And then I came home, watched tv, listened to music and just relaxed the rest of the day) form one syntactically complex structure with the same subject, and they are not considered separate clauses. Thus, in the present study, a clause is understood to contain both a subject and a finite verb (see, e.g. SAG 4: 3–4). On one hand, this

definition excludes the study of non-finite clauses, which are often considered complex structures, but on the other hand, the definition eliminates distortion in the results gained with both length-based measures, such as mean clause length, and with the dependent clause ratio.

In earlier studies of syntactic complexity, macrosyntagms were used as segmentation units for segmenting written productions (Norrby & Håkansson 2007: 48). A macrosyntagm refers to a main clause with all possible dependent clauses attached to it. Thus, this unit corresponds to a syntactic sentence (and, to some extent, also to T- and AS-units, as discussed below). In addition, orthographic sentences, that is, sections beginning with a capital letter and ending with sentence punctuation (Hunt 1965: 49; SAG 1: 177), have functioned as comparative units (i.e. a unit functioning as the basis for a measure of the occurrence of a given linguistic feature, e.g. clauses per *sentence,* Foster et al. 2000: 354). Due to inconsistencies in the use of capital letters and punctuation, especially in productions by younger learners, units based on syntax are often more appropriate. Furthermore, the segmentation units used should correspond to the psycholinguistic planning process of the learner as closely as possible (Crookes 1990), and according to earlier research, units based on syntax suitably represent this planning process (Foster et al. 2000: 362, 365).

A syntax-based unit, the **T-unit** (*minimal terminable unit*), has been widely used in complexity studies. The T-unit is defined as a section of production that contains a main clause and all possible dependent clauses as well as all nonclausal units (i.e. units lacking the properties of a clause), or fragments, attached to it (Hunt 1965: 20, 49). In the present study, coordinated verb phrases, in which the subject is shared but omitted after the first verb phrase, belong to the preceding T-unit (see Foster et al. 2000: 363–364). The T-unit has received some criticism, for instance, because it ignores the complexity of coordinated structures (see, e.g. Gaies 1980: 54; Biber et al. 2011: 12; Lintunen & Mäkilä 2014: 383; see also Bardovi-Harlig 1992). For instance, two coordinated main clauses can be segmented as one orthographic sentence or two T-units (Norrby & Håkansson 2007: 48–49). For the purposes of this study, the T-unit was considered most appropriate for segmenting the written data.

Although originally meant for use in analysing written data, the T-unit has also been used in studies of syntactic complexity in spoken productions (Foster et al. 2000: 360; Lintunen & Mäkilä 2014: 383). This can be problematic, as spoken L2 productions contain many features that units based on written data ignore (see also Foster et al. 2000: 354; Ellis & Barkhuizen 2005: 147). While a variety of alternative segmentation units have been proposed, such as the tone unit (i.e. a stretch of speech between unit boundaries defined using pauses and other phonetic elements) and the C-unit (i.e. a stretch of speech containing one meaning), the **AS-unit** (Analysis of Speech Unit) was adopted as the segmentation unit for spoken productions in the

current study (see the discussion in, e.g. Foster et al. 2000). The AS-unit is defined with the help of syntactic features as an utterance containing a main clause, or a subclausal unit, and all possible dependent clauses associated with it (Foster et al. 2000: 365). In spoken communication, especially in dialogue, subclausal units and other incomplete utterances with ellipted elements are common (Foster et al. 2000: 366). However, in the current study, these did not affect the analyses, as monologues were used as data, and all unfinished utterances were excluded from the analysis (see Section 3.4). In this study, coordinated verb phrases without a repeated subject belong to the preceding AS-unit (see the discussion above; cf. Foster et al. 2000: 367).

In addition to the syntactic definition, intonation is crucial in AS-units (Foster et al. 2000: 365–368), and this unit has received criticism due to the relatively difficult segmentation process that it entails (e.g. Iwashita 2006: 153, 157). In the current study, more emphasis was placed on the syntactic features (Foster et al. 2000: 358), as pausing and intonation can be considered more important for studies of L2 fluency. Pausing was not emphasised, as the fact that L2 learners need more time, for instance, to plan their productions should not affect the latter's syntactic complexity (see the discussion in Lintunen & Mäkilä 2014). In ambiguous cases, intonation was used to segment spoken productions.

Phrasal-level complexity is studied by way of **noun phrase** complexity in the third substudy. In this study, a noun phrase is defined as a phrase that consists of a possible article or another determiner, such as a possessive pronoun or an indefinite pronoun, a head noun or a pronoun, and/or a modifier (see, e.g. SAG 1: 202, SAG 3: 11–183; Vyatkina et al. 2015: 29; Ansarifar et al. 2018: 61; see also Biber et al. 2011; Lambert & Nakamura 2019; Biber et al. 2021). The modifiers can be either pre- or postmodifiers, such as premodifying nouns or prepositional phrases. Nouns or pronouns in apposition and noun phrases that contain several noun phrases are analysed as single noun phrases, and noun phrases that are part of prepositional phrases, as well as coordinated nouns without modifiers, are excluded from the analysis.

Along with the segmentation units discussed above, **topicalised structures** and **dependent clauses with a sentence adverbial** play central roles in the first and second substudies. Topicalised structures are defined as declarative units of production – that is, T- or AS-units – with a constituent other than a subject in the initial position (Teleman et al. 2001: 69; see also Section 2.3.2.5). In the current study, sequences that include direct speech followed by a reporting clause are considered topicalised, as these structures require inversion to be used in the reporting clause (cf. Jörgensen 1976: 22). 'Dependent clause with a sentence adverbial' as a term is rather self-explanatory. These are dependent clauses that contain a sentence adverbial, such as an adverb of time (e.g. *redan* = already) or a

negation (e.g. *inte* = no) – for instance, *det beror på alltså om jag <u>inte</u> har jobb på den dag* = it depends on like if I don't have work that day (FN2.06).

To summarise, the choice of segmentation units is of great importance to this study, as these function as **the basis for the measures used in the analyses**. As is evident from the discussion in this section, certain segmentation units are more problematic than others. As syntactic features are highlighted in the analysis of both written and spoken productions in the current study, the segmentation units and therefore also the measures in which these units are used (e.g. the mean length of T- and AS-units) are more compatible. In the following section, the measures used in the current study are discussed.

3.3.2   Complexity measures

Researchers have long debated the most valid measures of syntactic complexity (Iwashita 2006: 154; Tonkyn 2012: 222; Bulté & Housen 2018: 151; Kuiken & Vedder 2019: 194; Kuiken et al. 2019: 164; see also, e.g. Norris & Ortega 2009; Pallotti 2015). Differing definitions and emphases in research objectives as well as varying study conditions have led to different operationalisations of complexity, with a variety of measures introduced at several linguistic levels. Complexity is generally examined using objective quantitative measures, but subjective ratings have also been used (Bulté & Housen 2020b: 52).

It is widely agreed that a learner's repertoire of syntactic structures involves aspects that can be measured with **length-based measures**, **measures based on subordination**, and the **range of structures** employed (Ferrari 2012: 282–283; Bulté & Housen 2012: 35; see also Wolfe-Quintero et al. 1998; Housen et al. 2012a). Norris and Ortega (2009) recommend using several independent measures that target these areas of complexity when studying syntactic complexity in L2. Still, in most studies, the analyses are based on a limited number of general-level measures, mainly those targeting production unit length and use of subordination (Bulté & Housen 2018: 149; Kuiken & Vedder 2019: 193; Kuiken et al. 2019: 163; see also, e.g. Wolfe-Quintero et al. 1998; Norris & Ortega 2009). For instance, the dependent clause ratio has been used as a hybrid measure to target many aspects and levels of complexity, such as depth and relative complexity (Vyatkina et al. 2015: 42).

Recently, the use of general measures, particularly those based on subordination, has been criticised by SLA researchers (Kyle & Crossley 2018: 333–334, 345; Kuiken et al. 2019: 164; see also Norrby & Ortega 2009; Biber et al. 2011; Baten & Håkansson 2015). In their criticisms, many researchers highlight that while the dependent clause ratio, a frequently used measure of embedding, might indicate a decrease in complexity from intermediate- to advanced-level learners, more specific measures, such as those based on the type of dependent clause (i.e. nominal,

adverbial, and relative), can reveal a different picture of the differences between proficiency levels (see, e.g. Kowal 2016: 86). Previous studies have shown that learners at different proficiency levels produce different types of dependent clauses and that dependent clause types tend to emerge in learner language in a given order, with nominal and adverbial clauses preceding relative dependent clauses (see, e.g. Lambert & Nakamura 2019: 250; Vercellotti 2019: 239; see also Wijers 2019). However, in the current study, the types of dependent clauses are not examined, as the first two substudies focus on surface-level complexity, and in the third substudy, the participants were advanced learners whose use of dependent clause types revealed no clear differences, whether interindividual or intraindividual. Instead, phrasal-level complexity was studied in this project to compensate for the shortcomings of a measure based on embedding.

In many studies, researchers suggest using phrasal-level complexity measures in addition to clausal-level metrics, especially when studying syntactic complexity at higher proficiency levels. As a phrasal-level measure of complexity, mean clause length has often been adopted due to the assumption that clause length can reflect increasing phrase length (Kuiken & Vedder 2019: 198). However, this assumption has been criticised, as increasing clause length reflects not only lengthening at the phrasal level but also the use of various clausal-level features, such as expressions of manner, place, and time, that can increase the mean clause length. Thus, 'purer' phrasal-level measures of complexity – for instance, noun phrase complexity measures – have been employed in studies of phrasal-level complexity (see, e.g. Lahuerta Martínez 2018; Lan, Lucas & Sun 2019; Khushik & Huhta 2022; Sarte & Gnevsheva 2022). In the current study, both clause length and noun phrase complexity measures are used as phrasal-level measures.

The noun phrase complexity measures used in previous studies include mean length and the use of different modifier types. The use of modifiers is clearly linked with the length of a noun phrase, as an increase in the use of modifiers inevitably increases the length of the phrase (see, e.g. Lahuerta Martínez 2018: 9). Therefore, these measures can be seen as partly overlapping. According to Vyatkina et al. (2015: 29), modifiers can be seen as optional elements of noun phrases that indicate structural complexity, especially through their varied use in L2 production. Therefore, a qualitative examination of the modifiers used by learners can provide deeper insight into the complexity of the noun phrases the learners produce. The use of different modifiers is examined with the help of the index proposed by Biber et al. (2011). This index includes developmental stages of complexification at the phrasal level (see also, e.g. Ansarifar et al. 2018). As the aim of the present study was not to analyse the relationship between proficiency and noun phrase complexity as such, but rather to describe the variation in the use of these phrases, the focus was

on the range of modifier types instead of complexity stages (see also Vyatkina et al. 2015: 29, 32, 41).

Due to the criticism of the unbalanced use of different measures, combining general, global-level, and more specific, local-level measures of complexity has become more common in more recent studies of syntactic complexity (Kuiken & Vedder 2019: 193; Kuiken et al. 2019: 164). Norris and Ortega (2009) recommend measuring distinct but interrelated dimensions of syntactic complexity, for instance, at the levels of a) general complexity; b) complexity at the clausal level, i.e. via subordination; and c) complexity at the phrasal level. Lahuerta Martínez (2018), for instance, combines sentence-level measures with both clausal- and phrasal-level measures, investigating both mean sentence length and complexification from the perspectives of coordination and subordination as well as mean noun phrase length. Additionally, Norris and Ortega (2009: 567) state that the developmental timing of the different structures should be considered (cf. relative complexity in the current study; see Section 2.3.1).

In line with the recommendations made in previous research, measures for gauging complexity at higher linguistic levels (i.e. T- and AS-units), at the clausal as well as phrasal levels were adopted in the current study. Complexification through coordination was not included in the complexity measures due to the relatively high proficiency level of the participants (see Hunt 1965: 21–22; Norris & Ortega 2009: 562). Even though the aim of the current study was not to trace development in syntactic complexity (cf., e.g. Kuiken & Vedder 2019: 193; Vercellotti 2019: 237; Bulté & Housen 2020b: 77), using a broad repertoire of distinct measures was expected to provide a more thorough account of syntactic complexity in the analysed data. These measures are presented in Table 5.

As noted above, absolute complexity refers to the number of linguistic elements and the relationship between these elements, whereas relative complexity is determined in relation to the learner (see Section 2.3.1). The absolute measures used in the current study are traditional CAF measures that have been widely used by researchers (see, e.g. Bulté Housen 2020b: 52). Baten and Håkansson (2015: 548) argue that CAF measures, such as dependent clause ratio, should not be used in measuring learner language development at all. As the current study focuses on developmentally meaningful differences and individual variation instead of developmental paths in syntax, these measures are considered suitable. However, following Baten and Håkansson (2015: 519, 521), it is considered crucial to investigate how these measures have previously been interpreted, and how the use of different measures influences the findings of the present study.

Table 5.   Syntactic complexity measures.

Type of measure		Written measure	Abbreviation	Spoken measure	Abbreviation
Absolute complexity	**General measures**				
	length-based	words per T-unit	W/T	words per AS-unit	W/AS
	clausal-level – embedding	dependent clauses per clause	WrDC/C	dependent clauses per clause	SpDC/C
	Phrasal-level				
	length-based	words per clause words per noun phrase	WrW/C WrW/NP	words per clause words per noun phrase	SpW/C SpW/NP
	type of noun phrase	NP with: determiner modifier determiner+modifier	DET MOD DET+MOD	NP with: determiner modifier determiner+modifier	DET MOD DET+MOD
	type of noun phrase modifier	NP with: attributive adjective premodifying noun genitive attribute relative clause prepositional phrase infinitive phrase apposition comparative phrase	ADJ PRE.noun GEN REL PP INF APP COMP	NP with: attributive adjective premodifying noun genitive attribute relative clause prepositional phrase infinitive phrase apposition comparative phrase	ADJ PRE.noun GEN REL PP INF APP COMP
Relative complexity	**Specific measures**	topicalised structures per declarative T-unit	TOP/T(P)	topicalised structures per declarative AS-unit	TOP/AS(P)
		dependent clauses with a sentence adverbial per T-unit	DC+/T	dependent clauses with a sentence adverbial per AS-unit	DC+/AS

The general measures pertain to **production unit length** (W/T, W/AS) as well as **complexification at the clausal level**[11] (WrDC/C, SpDC/C). The advantage of these measures is that they are relatively simple to apply, even to larger sets of data. The use of length-based measures has been criticised because they do not reveal the level of complexification, or the internal structure – that is, what exactly is considered complex – in the unit (see, e.g. Biber et al., 2016: 640, 649; see also

---

[11]   In some studies, the dependent clause ratio is not seen as a general-level complexity measure but as an index of subordination (Kuiken & Vedder 2019: 200).

Jörgensen 1976: 39). Regardless, general-level, length-based measures are among the most commonly used complexity measures, as they efficiently distinguish proficiency levels. Regarding complexification at the clausal level, it has been argued that complexity studies have relied too much on measures grounded in embedding (see the discussion above). For instance, Iwashita (2006) notes that using measures based on embedding can be problematic in studies of spoken productions due to the difference in the features that are typical of a given production in the two modes (see the discussion in Section 2.3.2.3). Despite this, the dependent clause ratio (DC/C) is one of the most widely used measures of complexity in spoken productions. The general measures, including both length-based metrics and the dependent clause ratio, are used in all the substudies, but they are complemented by other measures, and the problems with these measures and their effects on the results are acknowledged.

The more fine-grained measures used in this study target **syntactic complexity at the phrasal level**. As discussed above, mean clause length (WrW/C, SpW/C) has been considered a phrasal-level (or subclausal) measure. In addition to this somewhat controversial measure, measures that target both the length and **range** (i.e. diversity and variety) of noun phrases – that is, measures that show the extent to which learners produce different types of noun phrases and how different modifier types are used in these phrases – are included in the analyses of the third substudy. These measures are used due to the relatively high proficiency level of the participants in that study. As discussed previously, it can be expected that learners at higher proficiency levels complexify their productions more through phrasal-level elaborations than at the clausal level (Ferrari 2012: 282–283; Lambert & Kormos 2014: 608; Vyatkina 2013: 24). Additionally, according to previous research, there is considerable individual variation in the use of complexification strategies at the clausal and phrasal levels (Vyatkina 2012: 590, 2013: 25). The phrasal-level analysis in the third substudy enables a comparison of the use of complexification strategies at these linguistic levels.

Although they are criticised for producing tautological findings (see Section 2.3.1), **relative measures** were used to analyse the differences between the proficiency levels in the first two substudies. The relative measures used – topicalised structures per T-unit and per AS-unit (TOP/T[P] and TOP/AS[P]) and dependent clauses with a sentence adverbial per T-unit and per AS-unit (DC+/T and DC+/AS) – were chosen based on the findings of earlier research (e.g. Rahkonen & Håkansson 2008; Åberg 2020; see also Norris & Ortega 2009: 562 and the discussions in Sections 2.3.1 and 2.3.2). As discussed earlier, the applicability and impact of relative complexity measures are limited in the current study by the fact that these measures can be language specific (see the discussion of the operationalisation of difficulty in Bulté et al. 2024: 25–29). However, the use of

these measures was considered important for the study, as they enabled a more detailed investigation of the differences in the structures produced at the two proficiency levels (see, e.g. De Clercq & Housen 2017: 316–317).

The same segmentation units and measures have often been used in studies of complexity in both speech and writing (e.g. Iwashita 2006: 151, 153; Kuiken & Vedder 2012a: 146–147), although the structures in these two modes of production can differ significantly (see Section 2.3.2.3). In the current study, the measures used are the same, but they are informed by different segmentation units (see Tables 4 and 5), making them as comparable as possible.

3.3.3   Accuracy measures

Accuracy was examined in the first substudy alongside complexity. As with syntactic complexity, several measures have been used in earlier studies of L2 accuracy (see, e.g. Wolfe-Quintero et al. 1998). As discussed earlier, operationalising accuracy can be challenging due to frequent and versatile errors in L2 production. In the current study, the examination of this aspect is limited to syntactic accuracy, that is, **target-like word order** in the structures produced by the learners. The target-like use of word order was evaluated by a native speaker of Swedish. The accuracy measures used in the study are presented in Table 6.

Table 6.   Syntactic accuracy measures.

Type of measure	Measure	Abbreviation
overall measure	error-free clauses per total number of clauses	EFC/C
specific measures	error-free dependent clauses per total number of dependent clauses	EFDC/DC
	error-free dependent clauses with a sentence adverbial per total number of dependent clauses with a sentence adverbial	EFDC+/DC+
	use of inversion in topicalised structures per total number of topicalised structures	EFTOP/TOP

To analyse syntactic accuracy in the first substudy, four measures were used: one overall accuracy measure (EFC/C) and three specific ones targeting different types of syntactically complex structures: dependent clauses (EFDC/DC), dependent clauses with a sentence adverbial (EFDC+/DC+, focusing on the placement of sentence adverbials in dependent clauses), and topicalised structures (EFTOP/TOP, focusing on the use of inversion after topicalised structures). In some studies (see, e.g. Håkansson & Norrby 2007), the differentiation between different word orders has been seen as an indication of grammatical processability (see Section 2.2). In the

current study, the placement of sentence adverbials in dependent clauses and the use of inversion after topicalised structures were included as specific accuracy measures to enable comparisons between the corresponding relative complexity measures (i.e. the use of sentence adverbials in dependent clauses and the use of topicalised structures). The overall accuracy measure is used to examine the interplay between syntactic complexity and accuracy (see Section 4.1), as this measure provides an overview of accuracy in the learners' productions.

3.4 Analyses

Quantitative studies are ideal for investigating differences between groups of learners at different proficiency levels. In this dissertation, the analyses in the first two substudies were quantitative and conducted at the group level. In these substudies, the level of analysis is referred to as 'surface level'. The surface-level, quantitative analyses were complemented with qualitative elements in the third substudy, and this individual-level enquiry combines cross-sectional comparisons with a longitudinal study of syntactic complexity in the learners' productions. A mixed-methods approach is useful in studies focusing on individual-level differences in syntactic complexity. In this section, the analyses carried out in the present study are presented.

The spoken data were transcribed with the help of assistants who were advanced students of Scandinavian languages. These assistants were given detailed instructions about the level of precision (orthographic transcription with some phonetic details) expected and the symbols used in the transcriptions (see Appendix 3). In the transcriptions, all spoken words, as well as other noises made by the participant (e.g. laughter), were transcribed. Additionally, for instance, pause lengths and micropauses (defined here as pauses under 0.4 seconds[12]) as well as intonation (rising/falling), increased tempo, and emphatic stress within a word were marked in the transcriptions. Prior to the transcription process, the assistants familiarised themselves with the programme Transcriber, which was used to facilitate the process. The productions transcribed by the assistants were double-checked.

Some remarks concerning the manipulation of the spoken data in **the transcription phase** are necessary. As mentioned earlier, spoken productions can contain many fragmentary and unfinished structures. These structures influence some complexity measures, especially those based on length. Prior to the

---

[12]   The limit for micropauses varies in studies of spoken L2 production. Often, this limit is set at 0.25 seconds (see, e.g. Peltonen 2020: 45). As the focus in the current study was on syntactic complexity instead of fluency, a somewhat higher limit for these pauses was considered appropriate.

segmentation of the spoken data, false starts and repetitions were excluded from the transcriptions (see Foster et al. 2000: 368, 374). In other words, the transcripts contain only the final, repaired versions of the utterances (see Lambert & Nakamura 2019: 254). However, some reformulations were included in the analysis if they formed a whole with the final version of the utterance. These data were manipulated to facilitate a focus on the structures that the participants were able to produce as complete utterances and to eliminate unfinished utterances that could distort the results concerning complexity (Foster et al. 2000: 370–371), thus making the written and spoken data more homogenous and enabling a more reliable comparison between the two types of data (see the discussion in Lintunen & Mäkilä 2014).

Filler words (i.e. productions during filled pauses; see, e.g. Peltonen 2020: 26) were excluded from the transcripts. As filler, they do not carry meaning but have other functions, such as showing that the production is continuing. A common filler word in the data is *och* (= and), and especially with this filler, the context and pauses determined whether it was included in the word count (see, e.g. Beaman 1984: 59). When *och* occurs as filler, the pronunciation is often elongated. When *och* was used as a coordinating conjunction, showing that two structures are at the same syntactic level (see, e.g. Biber et al. 2021: 81), the occurrences were naturally included in the transcriptions. This was also the case when *och* was used to relate new information to the preceding production. However, when occurrences of *och* were clearly separated from other output, that is, surrounded by pauses, they were left out of the transcriptions. Furthermore, the participants' comments on the language forms (e.g. *ett tal (3.0) eller en tal jag är jag är jag är inte säker >men spelar ingen roll<* = a [neutrum] speech or a [utrum] speech I am not sure but it doesn't matter [FN2.12]) and comments on the task (e.g. when the participant translated the questions in the task and said them out loud) were excluded from the transcripts, as the latter are not part of the task itself, and the former are not interpreted as the participant's own production.

In the first substudy, **quantitative analyses** were carried out to examine the relationships between the learners' proficiency levels and the syntactic complexity of their productions, the relationship between proficiency level and syntactic accuracy, and between the differences in the interplay between syntactic complexity and accuracy at the two proficiency levels. The analysis phase began with the segmentation of the data. The written data from the first data collection round were manually segmented according to T-units, clauses, and dependent clauses. Additionally, the total number of words in every production was calculated, and all occurrences of topicalised structures and dependent clauses with sentence adverbials were excerpted from the productions. Thereafter, the accuracy of the clauses, dependent clauses (with and without a sentence adverbial), and topicalised structures was analysed, and the ratios of target-like structures were calculated. All these data

were entered in Excel, and the values for the measures used in the study (see Sections 3.3.2 and 3.3.3) were computed. Additionally, descriptive statistics, including averages and median values, minimum and maximum values, standard deviations, and interquartile ranges[13], were calculated for every measure.

The statistical analyses were performed with the help of SPSS Statistics. The first step in these analyses was to determine whether the data were normally distributed (see, e.g. Larson-Hall 2016: 479). Shapiro–Wilk's test of normality (see, e.g. Larson-Hall 2016: 109–110) revealed some instances of non-normal distribution in the variables, and thereafter, non-parametric tests were used. These tests were used in all analyses to ensure easier comparisons, and because these tests better suit the relatively small samples included in the study (see, e.g. Tähtinen & Isoaho 2001: 16, 88; Larson-Hall 2016: 74). The Mann–Whitney U-test, the non-parametric equivalent of the independent sample t-test (see, e.g. Larson-Hall 2016: 286, 478), was used to examine the relationships between proficiency and syntactic complexity and between proficiency and syntactic accuracy. For every measure, effect sizes[14] (r effect sizes; see, e.g. Kormos 2014: 204; see also the discussion in Bulté & Housen 2020b: 76) and confidence intervals[15] (CI 95%; see Larson-Hall 2016: 149) are reported. Spearman's rho, that is, the rank-order correlation coefficient and the non-parametric equivalent of Pearson's correlation (Larson-Hall 2016: 208, 483), was used to study the relationship between the measure of syntactic complexity and overall accuracy – that is, the proportion of error-free clauses to total clauses in the production (see Sections 3.3.2 and 3.3.3) – at the two proficiency levels. When multiple comparisons are performed, the p-values (see, e.g. Larson-Hall 2016: 481) must be corrected to control for false positives (familywise error rate and type I error in statistical analyses; see, e.g. Benjamini & Hochberg 1995: 291, 298; Larson-Hall 2016: 287, 319). The approach used in the current study, namely the false discovery rate (FDR) method, differs from other traditional methods, such as Bonferroni correction, and allows a higher risk of false positives (see, e.g. Larson-Hall 2016: 287, 476; for a comprehensive description of this method, see Benjamini & Hochberg 1995). Some examples of the differences in complexity and accuracy between the proficiency levels are illustrated with the help of box plots, and scatter plots are used to illustrate differences in the interplay between complexity and

---

[13]   Interquartile range (IQR) is the non-parametric equivalent to standard deviation. It indicates the variation around the median and the distance between the lower and upper quartiles, that is, 50% of cases (Larson-Hall 2016: 477).

[14]   Effect sizes measure the importance of the result (see Larson-Hall 2016: 208, 475).

[15]   'The range of values around a statistic such as the mean that defines the range where the true population value of the statistic will be found on repeated testing of the research question, with 95% confidence' (Larson-Hall 2016: 474).

accuracy measures. Furthermore, excerpts from the learners' productions are used to illustrate the findings.

The second substudy examined the effect of proficiency on syntactic complexity in spoken productions and the differences in syntactic complexity between written and spoken productions. The analyses of the written data completed in the first substudy were used in the second substudy. Again, the analysis phase of the second substudy began with the manual segmentation of the data. The spoken data were segmented into clauses, dependent clauses. and AS-units; the total number of words in the productions was calculated; and the occurrences of topicalised structures and dependent clauses with sentence adverbials were excerpted from the productions. The data were entered into Excel, the complexity measures were computed, and descriptive statistics were calculated for every measure.

Once again, statistical analyses were carried out using SPSS. The tests for normality showed that, as with the written data, the spoken data were not normally distributed, and all variables were therefore examined non-parametrically. These tests were considered more suitable than their parametric alternatives because the sample sizes were also relatively small in the second substudy. The Mann–Whitney U-test was used to examine whether the productions at the different proficiency levels differed in syntactic complexity. The differences in complexity between the proficiency levels were examined using the Wilcoxon signed-rank test, which is the non-parametric equivalent to the dependent samples (paired) t-test (see, e.g. Larson-Hall 2016: 478, 485). These tests are appropriate for comparing the results of two tasks completed by the same participant (Tähtinen & Isoaho 2001: 86). In these tests, the measures of written productions were compared with those of spoken productions.[16] This analysis was conducted separately for the groups of learners with different proficiency levels to examine the differences in syntactic complexity between the written and spoken productions from each proficiency group. The participants who exhibited different proficiency levels in written and spoken productions (FN1/2.01, FN1/2.08, FN1/2.11, and FN1/2.13) were excluded from this analysis. Here again, to account for the multiple comparisons and control for the risk of false positives, the FDR method was used in these quantitative analyses (see the discussion above). The results are illustrated using box plots and excerpts from the learners' productions.

The third substudy differs substantially from the first two. In this study, the focus shifts from the group level to the individual level, and elements of **quantitative and**

---

[16] The comparisons included length of T-units and of AS-units, clause length in written and spoken productions, dependent clause ratio in written and spoken productions, ratio of dependent clauses with a sentence adverbial in written and spoken productions, and ratio of topicalised structures per T-unit and per AS-unit.

qualitative analyses are combined. In addition, the cross-sectional comparisons conducted in the earlier studies are complemented by a longitudinal study of syntactic complexity in the written samples. The focus of this study is complexity profiles, and variation in the use of syntactic structures is studied both synchronically and diachronically.

As DST was mainly used as a theoretical framework in the third substudy (see, e.g. Bulté & Housen 2020a: 216), and only two data points were available for the longitudinal study of syntactic complexity in the learners' productions, the approach taken to studying inter- and intraindividual variation in this substudy differed from those of earlier DST-inspired research (see also Vyatkina et al. 2015). While earlier studies with dense data points have used various methods that, for instance, can visualise the developmental paths of complexity (e.g. Verspoor et al. 2008), the third substudy aimed to categorise individuals' productions according to their complexity and study both synchronic and diachronic variation. This variation was examined according to complexity profiles, differences in the use of complexification strategies at the clausal and phrasal levels, and the stylistic choices revealed by the qualitative analysis.

The sampling procedure was based on the analysis carried out in the first substudy. As the aim was to study productions with contrasting complexity profiles, extreme case sampling (Dörnyei 2007: 128) was employed to choose the initial data. Based on the descriptive statistics from the first substudy, three learners were chosen for this multiple case study. From the lower extreme, a production with relatively simple syntactic structures was selected, and from the higher end, a production with more complex structures was chosen. This type of sampling enabled the examination of individual variation across learners with productions that represented the extremes of the syntactic complexity continuum. As these productions represent extreme cases, the results should not be generalised to average learners (see the discussion in Section 5.3). In addition to the extreme samples, a third production with an average complexity score was included in the analysis. The average profile was integrated to compensate for the extremities and to represent a production at the centre of the syntactic complexity continuum.

These productions were categorised into **complexity profiles** in an explorative manner **based on group-level descriptive statistics** (cf., e.g. Vyatkina 2013). The categories were determined based on general measures of syntactic complexity – mean length of T- and AS-units, mean clause length, and dependent clause ratio – as they are thought to represent the different linguistic levels of syntactic complexity (see the discussion in Section 3.3.2). The data were categorised on the basis of threshold values determined with the help of group statistics. The upper threshold was set at the third quartile (i.e. the 75th percentile, the median value of the upper half of the data) of each measure. In other words, if a measure had a value higher

than this threshold, the measure supported categorisation as a complex production. The lower threshold value was set at the first quartile (i.e. the 25th percentile, the median value of the lower half of the data), and complexity measures below this threshold value supported categorisation as a simple production. Values between these thresholds supported categorisation as an average production.

Due to the multidimensionality and dynamicity of syntactic complexity, a production can rarely be categorised unambiguously into a certain profile, based on the measures used in this study (see also Kowal 2016: 142). Therefore, it was deemed sufficient if two out of the three general complexity measures were consistent with a given profile, and one measure was permitted to deviate from the others. For instance, if two measures suggested complexity below the lower threshold, the production was categorised as simple, even though one measure indicated higher complexity. The productions for which the measures indicated contradictory values – one measure above the upper threshold, one between the threshold values, and one below the lower threshold value – were categorised as average. After categorising the first productions in the sampling procedure, the written sample gathered in the second data collection phase, and the spoken sample by the same learners, were categorised into these profiles following the same principles. The profiles were examined for diachronic and synchronic variation. The categorisation of the spoken productions was based on the group statistics from the second substudy.

Furthermore, the productions were analysed for phrasal-level complexity in the third substudy. Measures of noun phrase complexity were used. As noted above, increasingly long noun phrases reflect a rise in the use of determiners and modifiers (see, e.g. Lahuerta Martínez 2018; Sarte & Gnevsheva 2022; see also Lambert & Nakamura 2019). Therefore, in addition to the mean length of noun phrases, the types of noun phrases (i.e. whether it contains only determiners, only modifiers, or both) and modifiers used therein were investigated in all productions (see Section 3.3.3). An index of noun phrase modifiers proposed by Biber et al. (2011) was used as a starting point to analyse the types of modifiers used. The contents of this index were slightly modified to suit the data used in the current study. Modifier types that were included in the analysis but are not part of the original index are discussed separately. This analysis of phrasal-level complexity in the learners' productions enabled comparisons in the use of different complexification strategies, both between learners (interindividual variation) and between productions by the same learner (intraindividual variation).

The third substudy adopted **a mixed-methods approach** to illuminate individual variation in syntactic complexity (see, e.g. Gilmore & Gánem-Gutiérrez 2020: 186–189). It was grounded on a triangulation of analyses of syntactic complexity in the learners' productions (for more information on triangulation and a discussion of the link between the mixed-methods approach and triangulation, see, e.g. Angouri 2018;

Flick 2018; Morgan 2018). The main methodology used in the third substudy was quantitative, as the categorisation of the productions into complexity profiles was premised on quantifying complexity; the phrasal-level analysis also included numerical measures. As the sample size in the third substudy did not permit statistical analyses, the quantitative analyses were solely based on descriptive statistics. In this substudy, the quantitative observations were complemented with a qualitative examination of the types of noun phrases and modifiers employed, which were then also analysed quantitatively. Furthermore, the productions were qualitatively examined to find patterns in the learners' use of structures (on QUAN+qual designs, see Hashemi & Babaii 2013: 832–833; Angouri 2018: 41; on macro- and micro-level perspectives, see Larsen-Freeman 2006: 597). In this qualitative analysis, special emphasis was placed on the frequent or repetitive use of certain structures (see also Wijers 2019: 110). This complementary method facilitated the investigation of the stylistic choices made by the learners and the examination of possible similarities between productions by the same learner. These characteristics and similarities can impart important information about a learner's individual style of producing language that cannot be detected using quantitative analyses alone (see, e.g. Bulté & Housen 2020b: 77).

The data used and analyses carried out in the three substudies are summarised in Table 7.

Table 7.   Research designs employed in the substudies.

	Article I	**Article II**	**Article III**
data	written	written and spoken	written and spoken
level of analysis	group	group	individual
analyses	QUAN	QUAN	QUAN+qual
	absolute, general-level complexity measures	absolute, general-level complexity measures	absolute, general-level complexity measures
	absolute, phrasal-level complexity measure	absolute, phrasal-level complexity measure	absolute, phrasal-level complexity measures
	relative, specific measures of complexity	relative, specific measures of complexity	qualitative analysis

While the first two substudies were based on group-level quantitative analyses of syntactic complexity, the third adopted a mixed-methods approach to understanding individual-level variation in syntactic complexity. The first substudy was solely concerned with written production, whereas both written and spoken productions were analysed in the second and third substudies. In the next chapter, the main results of these three independent substudies are presented.

4   Results

In this chapter, the findings of the three substudies included in this dissertation are presented. The results are discussed one study at a time in Sections 4.1, 4.2, and 4.3. In Section 4.4, the main results of the three enquiries are summarised.

The primary aim of the current study was to examine syntactic complexity in L2 Swedish from different perspectives to develop a more thorough account of this multidimensional construct. The specific aims were to investigate the following factors:

1. the effect of language proficiency on

    a) syntactic complexity
    b) syntactic accuracy
    c) the interplay of syntactic complexity and accuracy

2. the effect of mode of production on syntactic complexity

3. the role of individual variation in syntactic complexity

The first substudy focused on the effect of language proficiency on syntactic complexity and syntactic accuracy, as well as on the interplay between these aspects (Aim 1). In this study, written productions by two proficiency groups – a group with lower proficiency and a more proficient group – were compared in terms of their complexity and accuracy, and differences in the relationship between complexity and accuracy within their contributions were examined. The effect of proficiency (Aim 1) was also part of the second substudy, in which spoken productions by the two proficiency groups were compared, and the differences in syntactic complexity between the written and spoken productions were investigated (Aim 2). In the third substudy, the focus shifted from the group level to the individual level, and the role of individual variation in syntactic complexity was explored (Aim 3). The variation in the individuals' use of syntactically complex structures was examined synchronically in speech and writing (see Aim 2) and diachronically in writing between two data collection periods about six months apart.

4.1 Results of Article I

In the first substudy, the effects of language proficiency on syntactic complexity, accuracy, and the interplay of complexity and accuracy in written data were investigated (Aim 1). The analyses carried out in this substudy were based on absolute general-level measures and relative specific measures of complexity and focused on group-level comparisons between unequal proficiency levels (see Sections 3.3.2 and 3.3.3).

The results based on descriptive statistics representing the differences in syntactic complexity between the proficiency levels indicated that the productions by the more proficient learners were slightly more complex, especially according to the length-based measures, i.e. T-unit length and clause length (see Table 4 in Article I, p. 144). However, the difference was slight regarding the use of dependent clauses and topicalised structures. Variation in the measures was relatively high, especially concerning the T-unit lengths in both groups (min 6.49 W/T and max 12.36 W/T at the lower proficiency level, and min 8.30 W/T and max 13.77 W/T at the higher proficiency level) but also in the dependent clause ratio (min 0.06 DC/C and max 0.55 DC/C) and the use of topicalised structures (min 0.11 TOP/T[P] and max 0.86 TOP/T[P]) at the lower proficiency level. The learner who produced the highest proportion of both dependent clauses and topicalised structures belonged to the less proficient group. The difference in syntactic complexity between the proficiency levels was statistically significant in only one measure, which was determined using the Mann–Whitney U-test: mean T-unit length (see Table 5 in Article I, p. 144). The learners in the higher proficiency group produced significantly longer T-units ($p = 0.005$, $r = -0.60$) than the less proficient students, whereas no significant differences after FDR correction were found in clause length, dependent clause ratio, use of dependent clauses with a sentence adverbial, or use of topicalised structures. The result concerning the difference in mean T-unit length is generalisable to some extent.

These results confirm previous findings that length-based measures can efficiently distinguish between proficiency levels, as there was a significant difference in mean T-unit length between the proficiency groups, and clause length also showed a difference between the groups. The relatively small difference between the proficiency levels in the use of dependent clauses can be expected, considering the differences in clausal- and phrasal-level elaboration between proficiency levels discussed earlier in this dissertation (see Section 2.3.2.1). This difference in elaboration between the proficiency levels is bolstered by the relatively large difference in mean clause length. The fact that learners in both proficiency groups produced topicalised structures relatively frequently indicates that they were all familiar with this structure in Swedish and willing to experiment with it in their output. It can be assumed that the relatively high frequency of topicalised structures

in L1 Swedish contributed to the high frequency of this structure in the data (see Section 2.1). This result clashes with earlier findings indicating a late emergence of this structure in the developmental path. The frequent use of this form and other complex structures at the lower proficiency level could be an indication of the cognitive maturity of the participants. The insignificant difference in the use of dependent clauses with a sentence adverbial is somewhat surprising. However, this structure was rarely used in both groups (21 occurrences at the lower proficiency level and 38 at the higher level), which is consistent with the results of earlier studies. This was expected to some degree, as dependent clauses with a sentence adverbial are relatively infrequent in the input that learners receive (see Sections 2.1 and 2.3.2.5). Overall, however, the lack of differences between the proficiency groups was surprising.

In the first substudy, the analysis of accuracy was limited to syntax. The descriptive statistics demonstrated clear differences between the two proficiency levels, with the learners at the higher proficiency level producing more accurate structures (see Table 6 in Article I, p. 146). The measures indicated greater variation in the productions by the learners at the lower proficiency level than at the higher proficiency level, except for the measure of dependent clauses with a sentence adverbial. When it comes to these structures, it is important to remember that both groups contained learners who did not produce any dependent clauses with a sentence adverbial (6 learners at the lower proficiency level and 2 at the higher level), and it is difficult to draw any reliable conclusions based on this limited number of occurrences. According to the Mann–Whitney U-test, the ratios of error-free clauses ($p = 0.000$, $r = -0.80$) and error-free dependent clauses with a sentence adverbial ($p = 0.001$, $r = -0.59$) as well as the use of inversion after topicalised structures ($p = 0.000$, $r = -0.77$) differed significantly between the proficiency levels (see Table 7 in Article I, p. 146). These results are generalisable when it comes to the production of error-free clauses and the use of inversion after topicalised structures, and to some degree, to the production of error-free dependent clauses with a sentence adverbial. Altogether, these results indicate that learners at the higher proficiency level produce more error-free structures compared to less proficient students, reinforcing the findings of several earlier studies.

As the difference in syntactic complexity was, for the most part, small between the groups of learners, and the accuracy of syntactically complex structures clearly differed between these groups, some degree of interplay between complexity and accuracy could be expected. However, as the analyses based on Spearman rank-order correlation coefficients showed, there were no statistically significant correlations between the overall measure of syntactic accuracy (i.e. error-free clauses per total number of clauses) and the complexity measures used in the study (i.e. mean length of T-unit, clause length, dependent clause ratio, topicalised structures, and dependent

clauses with a sentence adverbial per total number of dependent clauses) after the FDR correction at either proficiency level (see Table 8 in Article I, p. 147). Therefore, the results indicated no trade-off effects between complexity and accuracy in the data. The fact that these measures did not point to significant correlations between complexity and accuracy at the group level does not exclude the possibility of a trade-off, that is, a negative correlation, or a supportive relationship, that is, a positive correlation, at the individual level or between other measures of complexity and accuracy.

The results of the first substudy did reveal a tendency towards lower accuracy in the context of certain complex structures, but the correlations were not statistically significant. At the lower proficiency level, the use of topicalised structures seemed to result in lower overall accuracy in the productions (see Figure 4 in Article I, p. 148), whereas a higher number of dependent clauses with a sentence adverbial was linked to lower overall accuracy in the productions by learners at the higher proficiency level (see Figure 3 in Article I, p. 148). Thus, according to these results, only the relative complexity measures seemed to have a connection with the overall accuracy measure.

As noted above, one of the most intriguing findings in the first substudy was the small differences in syntactic complexity between the proficiency levels. As this result was obtained with written data, the next step was to examine the effect of language proficiency on syntactic complexity in spoken data. Furthermore, although the first substudy was intended solely to measure group-level differences, the role of individual variation became evident when interpreting the results (see Section 4.3).

4.2   Results of Article II

The second substudy focused on the effect of mode of production (writing vs speech) on syntactic complexity (Aim 2). In this study, syntactic complexity was compared in spoken productions by learners at the lower and higher proficiency level (see Aim 1), and these results were compared with the results from the first substudy pertaining to syntactic complexity in written productions (see Section 4.1). Again, the analyses were based on absolute general measures and relative specific measures of complexity to ensure comparability between the two modalities. It is of note that the groups of learners with different proficiency levels in the written and spoken data were not identical, as the spoken contributions by some of the learners were deemed less proficient than those in writing (see Section 3.2 and the discussion below).

The descriptive statistics indicated that, in the spoken samples, the learners at the higher proficiency level produced more complex structures than those at the lower level (see Table 3 in Article II, p. 89). Again, the variation within the measures was notable, especially when it comes to AS-unit length at the higher proficiency level.

The differences in syntactic complexity between the lower and the higher proficiency levels were tested using the Mann–Whitney U-test (see Table 4 in Article II, p. 91), indicating that the differences in mean length of AS-units (p = 0.000, r = -0.75), dependent clause ratio (p = 0.001, r = -0.59), and use of dependent clauses with a sentence adverbial (p = 0.000, r = -0.65) were statistically significant.

Thus, the results indicate that the groups differed especially in mean AS-unit length, whereas clause lengths did not differ greatly between the groups. Additionally, the more proficient learners seemed to produce more dependent clauses, both with and without sentence adverbials. Again, as with the written productions, dependent clauses with a sentence adverbial were extremely rare at both the lower and higher proficiency levels. Topicalised structures were produced frequently in both proficiency groups.

The difference in syntactic complexity between written and spoken productions was investigated separately for the lower and higher proficiency levels to study the effect of language proficiency on this difference (see Aim 1). Four learners from the lower proficiency level were excluded from these analyses, as their proficiency differed in the written and spoken modes. The descriptive statistics for the difference in syntactic complexity between written and spoken productions at the lower proficiency level revealed that three measures, namely, mean T- and AS-unit length, dependent clause ratio, and use of dependent clauses with a sentence adverbial, indicated higher complexity in the written mode, and two measures, that is, mean clause length and use of topicalised structures, indicated somewhat higher complexity in the spoken mode (see Table 5 in Article II, p. 92). Additionally, these learners demonstrated more variation according to the written measures than based on the spoken ones, except for the dependent clause ratio. In spoken productions, some learners used no dependent clauses, whereas all learners used dependent clauses in the written task. Similarly, all learners produced topicalised structures in the written task, but not all spoken productions contained these structures. As noted above, the frequency of sentence adverbials in dependent clauses was extremely low in both types of data.

According to the descriptive statistics representing the difference in syntactic complexity between written and spoken productions at the higher proficiency level, the learners showed somewhat higher syntactic complexity in the spoken mode in terms of mean AS-unit length, clause length, use of dependent clauses with a sentence adverbial, and use of topicalised structures. The only measure that indicated higher complexity in the written mode was dependent clause ratio. All these differences were small, especially for dependent clause ratio and use of dependent clauses with a sentence adverbial. As opposed to the learners at the lower proficiency level, the more proficient students employed more variation in the spoken mode than in the written mode, except according to the relative complexity measures – use of

dependent clauses with a sentence adverbial and use of topicalised structures – in which the written variation was greater.

The fact that mean clause length at the lower proficiency level and both mean clause length and AS-unit length at the higher proficiency level indicated higher complexity in spoken productions is consistent with the assumption that production units tend to be longer in speech than in writing due to differences in how the message is formulated (see Section 2.3.2.3). There was a small difference in the use of dependent clauses in writing between the learners at different proficiency levels, whereas in the spoken mode, the difference was clearly larger. In the first substudy, it was assumed that the more proficient learners would elaborate more at the phrasal level, and the difference in dependent clause ratio between the groups of learners was accordingly rather small (Section 4.1). However, as the difference was clearer in the spoken mode, it can be concluded that the more proficient learners still frequently use dependent clauses in spoken tasks. The relatively frequent use of dependent clauses at the higher proficiency level might also be influenced by the task type chosen and the fundamental differences between speech and writing (see Section 2.3.2.3). The frequent use of topicalised structures in both modes at both proficiency levels is most likely linked to the task type (see Sections 3.2 and 5.3).

Syntactic complexity in the written and spoken productions was compared using the Wilcoxon signed-rank test one proficiency level at a time (see Tables 7 and 8 in Article II, pp. 93, 97). At the lower proficiency level, the difference in syntactic complexity between the modes of production was statistically significant for T-unit vs AS-unit length ($p = 0.005$), clause length ($p = 0.008$), dependent clause ratio ($p = 0.000$), and the use of dependent clauses with a sentence adverbial ($p = 0.003$). It is important to remember the relative scarcity of dependent clauses with a sentence adverbial when interpreting these results. Other measures indicated greater complexity in the written mode, but clauses were significantly longer in the spoken productions than in writing. For instance, at this proficiency level ($n = 17$), dependent clauses were produced more frequently in writing by 16 learners, and 1 learner used equally many dependent clauses in speech and writing, whereas no one used dependent clauses more frequently in the spoken productions. Regarding clause length, 13 learners produced longer clauses in speech than in writing, whereas 3 used longer clauses in writing, and 1 employed equally long clauses in the two modes of production. Overall, based on these results, the written productions by the less proficient learners were more syntactically complex than the spoken productions for the most part.

As discussed in Article II, there could be many explanations for the differences in length-based measures between the written and spoken productions. For instance, the use of dependent clauses increases the length of both T- and AS-units, and as coordinated verb phrases are considered to belong to the same unit, the frequent use

of coordination within clauses increases the length in both clauses and in T- and AS-units. However, it could also be assumed that these learners can produce more complex structures in writing because they have the chance to plan and edit their productions, whereas producing complex structures in speech might be more challenging due to the demands and cognitive load imposed by real-time processing (see Section 2.3.2.3).

In contrast to the results from the lower proficiency level, the Wilcoxon signed-ranks tests indicated no statistically significant differences in syntactic complexity between written and spoken productions at the higher proficiency level. The greatest difference was found in the use of topicalised structures, which were somewhat more frequent in the spoken task, but the difference was not statistically significant. This lack of differences can be interpreted as an indication that more proficient learners can produce syntactically complex structures even with limited time for planning, and that a heavier cognitive load does not affect their production in the same manner as it does among the less proficient.

Based on the results of this substudy, it can be concluded that proficiency has a greater effect on syntactic complexity in speech than in writing (see Section 4.1) and thus plays an important role in demarcating the differences between written and spoken production. As with the first substudy, the interpretation of these group-level results inspired an interest in the role of individual variation, which is examined in the third substudy.

4.3   Results of Article III

The main purpose of the third substudy was to describe synchronic and diachronic individual variation in syntactic complexity (Aims 2 and 3) by analysing the structures produced, complexification strategies preferred, and stylistic choices made by the learners. An exploratory method was used to categorise productions by three learners into complexity profiles. The analyses carried out in Article I formed the basis for this substudy, as the subset of learners was selected based on group-level descriptive statistics from the first substudy using extreme case sampling. The group-level statistics from both Articles I and II were used in the categorisation procedure. The categorisation of the productions by different learners described the interindividual variation in the data, and the comparison of profiles by the same learner described the intraindividual variation. Furthermore, in this substudy, the analyses of general-level complexity in the previous substudies were complemented with analyses of phrasal-level complexity and with qualitative analyses of the structures in the data. Thus, the third substudy adopted a mixed-methods approach to analysing syntactic complexity. The aim was to deepen our understanding of individual variation in the use of complex language in L2 speech and writing.

The results showed that there may be considerable inter- and intraindividual differences, even within proficiency levels, in both general-level (i.e. length-based measures and dependent clause ratio) and phrasal-level (noun phrases) syntactic complexity. To illustrate the differences in complexity profiles according to general measures of syntactic complexity, Table 8 summarises the results.

Table 8.   Complexity profiles based on general syntactic complexity measures.

Participant Initial complexity profile (1st written)	Synchronic change in complexity profile (1st written vs spoken)	Diachronic change in complexity profile (1st written vs 2nd written)
P04 simple production	no change simple production	change average production
P11 average production	change simple production	no change average production
P15 complex production	change simple production	no change complex production

The results of the analyses of synchronic changes showed that the profile remained the same in the production of one learner, P04. On the other hand, the productions by Participant P15 demonstrated great differences in syntactic complexity in the written and spoken modes, both at the general and phrasal levels (see below for the phrasal-level differences). The general-level measures indicated high complexity in the writing but lower values in the spoken sample (see Table 4 in the manuscript, p. 9). Thus, whether due to a (conscious or subconscious) personal preference or another contributing factor (e.g. task effect), there was a notable difference in complexity between the written and spoken productions by P15. It is noteworthy that the spoken productions were categorised as simple in the case of every learner. This could be partly explained by the measures adopted and the fundamental differences between speech and writing; however, as the differences between the modes of production were not equally prominent across all learners, the productions might have been affected by other factors, such as individual styles of producing language.

Overall, the differences in syntactic complexity between the production modes (synchronic differences) were larger at both the general and phrasal levels than those between the two data collection stages (diachronic differences). The analysis of diachronic intraindividual variation indicated that the complexity profiles in the two written productions seemed similar. In the productions by Participants P11 and P15, there was no change in complexity profiles, and the general measures in the second written production by P04 indicated contradictory values (see Table 6 in the manuscript, p. 11). The lack of diachronic differences could be due to the learners'

relatively high proficiency levels. Furthermore, the threshold values used in this study were relatively strict, and the changes in these values had to be substantial enough to meet the criteria for the extreme profiles (see Section 5.3). However, the lack of differences could also be attributed to the stability of the complexity profile.

In the case of Participant P15, the analysis of noun phrase complexity was consistent with the categorisation according to the general measures of syntactic complexity, whereas noun phrase complexity only partly coincided with the profiles in the productions by P04 and P11 (see Table 5, p. 10; Table 7, p. 12; and Table 9, p. 13 in the manuscript). These results indicate that there could be some differences in the use of complexification strategies at the clausal and phrasal levels, both between learners (interindividual) and between productions by the same learner (intraindividual). As discussed above, the written and spoken productions by P15 demonstrated major differences at the phrasal level in addition to the general-level differences, but the differences in phrasal-level complexity between the written productions were small. The noun phrases in the productions by P15 were clearly longer and more varied in the written than in the spoken mode, but the use of noun phrases differed modestly between the written productions. This indicates that this learner made use of similar complexification strategies in the written mode; that is, the written productions were complexified at both the clausal and phrasal levels, whereas in the spoken mode, P15 did not use complex structures at either level. Dependent clauses were scarcely used, and the phrasal-level measures pointed to low complexity in the spoken task. As the phrasal-level analyses indicated greater complexity than the general measures in the cases of P04 and P11, it can be assumed that these learners preferred phrasal-level elaboration to some extent. However, even compared to the complex productions by P15, P11 produced a relatively high number of dependent clauses in the written tasks (see Table 8 in the manuscript, p. 13), indicating further elaboration at the clausal level. The complexification strategies of P04 varied throughout their productions, which could indicate that the learner had mastered the context-appropriate use of syntactically complex structures (see Section 5.3 for further discussion).

The stylistic choices made by the learners were qualitatively examined. These qualitative analyses pointed to major differences in how the complexification strategies were employed and how certain structures were used both within and across productions. Perhaps the most striking finding of the third substudy was the extent to which the learners differed in the consistency of their use of syntactically complex structures. According to the results, some learners have strong personal styles of producing language that can be consistent diachronically and/or in the different modes of production, whereas others show more variability in their use of syntactically complex structures. For instance, the repetitive use of certain structures that were otherwise rare in the data was indicative of a strong personal style. The

qualitative analysis indicated that Participant P11 in particular had a distinctive personal style of producing language, and this could also be seen in the relatively consistent complexity profiles.

Together, the general- and phrasal-level quantitative results and the qualitative analysis of the intra-individual differences confirmed the hypothesis that a complexity profile can be consistent, a stable characteristic of a learner, but can also be affected by the mode of production employed, among other factors. Therefore, variations in complexity can be both learner- and production-specific. Furthermore, the results demonstrate that participants who consistently used similar complexification strategies or a distinctive style, whether it be consistent or inconsistent in other productions, were easier to categorise into complexity profiles (P15), while others produced language in a more varied manner within single productions (P04), complicating the categorisation. On the whole, it can be concluded that general- and phrasal-level analyses can reveal a great amount of information about the complexification strategies used in a production. However, qualitative analyses enable a more in-depth investigation of the stylistic choices made and the structures used. Therefore, in the categorisation of learners' production into complexity profiles, a mixed-methods approach is extremely useful.

4.4   Summary of the results

The first substudy examined the effect of language proficiency on syntactic complexity, syntactic accuracy, and the relationship between these constructs, contributing to the first research aim. According to the results in **Article I**, **proficiency affects accuracy** in particular, while **complexity can also be influenced by other factors**, such as the learner's experience with languages and structures in a given language. The results indicate that learners at lower proficiency levels can also produce structures of relatively high complexity, which could be partly explained by the **relative cognitive maturity** of the participants in the current study.

No statistically significant correlation, neither positive nor negative, between syntactic complexity and accuracy was demonstrated at either proficiency level. However, there were **some tendencies in the results that point to trade-off effects** between the use of dependent clauses with a sentence adverbial and overall accuracy at the higher proficiency level, and between topicalised structures and accuracy at the lower proficiency level. The lack of statistically significant differences in both syntactic complexity and in the interplay of complexity and accuracy can be partly explained by large individual variation. Therefore, the results do not exclude, for instance, the possibility of a trade-off or a supportive relationship between complexity and accuracy at the individual level.

In the first substudy, syntactic complexity was examined with reference to both absolute complexity (i.e. structural, linguistic complexity) and relative complexity (i.e. complexity in relation to the learner, cf. difficulty). The study of relative complexity is particularly useful for examining differences in productions by learners at different levels of proficiency. Somewhat surprisingly, the relative complexity measures did not point to significant differences between the proficiency groups. As discussed above, this lack of differences could be due to the relative maturity of the learners, but also due to the task effects. However, these measures proved fruitful in studying the relationship between syntactic complexity and accuracy. In fact, the only measures that indicated tendencies in the relationship between these systems were **the measures of relative complexity**. As noted earlier, relative complexity measures often lead to ungeneralisable results, but when they are used alongside absolute complexity measures, they can provide important information about structures in a given target language.

As Article I revealed that proficiency was not a central influence on syntactic complexity in written productions, it was considered important to study the effects in spoken samples by the same groups of learners. Therefore, in **Article II**, the first and second research aims of the dissertation, which concern the effects of language proficiency and mode of production, are addressed. The second substudy examined the effect of proficiency on syntactic complexity in speech and investigated the differences in syntactic complexity between the learners' written and spoken productions.

The second substudy demonstrated that **proficiency has a greater effect on syntactic complexity in speech than in writing**. According to the results, the learners at the higher proficiency level used structures in their spoken productions that were notably more complex than those used by the less-proficient students. As in the first substudy's analysis of the written data, the relative complexity measure based on topicalised structures failed to reveal differences between the proficiency groups' spoken data. This might be due to the relative cognitive maturity of the learners, but the result is also attributable to task effects. The lack of differences in clause length was also somewhat unexpected. However, this result can at least partly be explained by the fundamental differences between speech and writing, and by the definition of 'clause' used in the study.

The results concerning differences in syntactic complexity between the written and spoken productions at the lower proficiency level showed that the written samples tended to be more complex than the spoken productions, whereas among the more proficient learners, for the majority of the measures, the spoken productions seemed somewhat more complex than the writing samples, but none of the measures revealed statistically significant differences between the production types at the higher proficiency level. Therefore, the difference in

syntactic complexity between written and spoken productions may be greater at the lower proficiency level than at the higher proficiency level. The results indicate that **the cognitively more demanding process of producing language in the spoken mode** has more influence on productions by learners at the lower proficiency level, whereas the **increasingly automated ability to produce language** facilitates the production of syntactically complex structures in speech by the more proficient learners.

In addition to language proficiency, individual preferences can affect the structures used in speech and writing. Hence, in response to the insights from Articles I and II regarding the role of individual variation, it was deemed necessary to study syntactic complexity at the individual level in Article III. The third substudy addressed the second and third research aims by examining synchronic and diachronic variation in syntactic complexity at the individual level.

The results of **Article III** demonstrate that even within a given proficiency level, **interindividual variation in syntactic complexity can be notable**. Additionally, the same learner can demonstrate considerable variation both across and within productions, indicating **great intraindividual differences**. Both inter- and intraindividual variation can be reflected in different complexity profiles (complex, average, and simple) and in the varying use of clausal- and phrasal-level complexification strategies. Overall, certain complexification strategies seem to be preferred by certain learners, but there are major disparities in whether the use of these strategies is consistent across and within productions. Thus, individual variation can be either **learner- or production-specific**. The qualitative analyses of the stylistic choices made by these learners contribute to a more nuanced understanding of how learners differ in their personal styles of producing syntactically complex structures. Therefore, a mixed-methods approach proves beneficial when studying individual variation in syntactic complexity. The exploratory method of using group-level statistics in individual-level analyses provided encouraging findings, as this method seemed to work relatively well for categorising learners' productions into complexity profiles.

In conclusion, all the substudies in this dissertation deepen our understanding of **syntactic complexity as a multidimensional property of learner language** – in this case, L2 Swedish. By examining developmentally meaningful differences in group-level tendencies and individual variation, the three substudies provide a holistic and nuanced overview of syntactic complexity. Furthermore, by including more fine-grained quantitative measures and qualitative analysis of syntactic structures in the third substudy, a more in-depth picture of individual variation in syntactic complexity was depicted. Together, these substudies underline that when it comes to syntactic complexity, both group-level and individual-level studies are needed. However, it is important to acknowledge that results from individual-level

studies cannot be generalised to a larger group, and it is equally dangerous to make generalisations about individual learners based on group-level studies. This problem is further discussed in the following chapter, which concludes this summary.

5   Discussion

In this chapter, the main findings of the current study are discussed in more detail, with an emphasis on their methodological and theoretical contributions (Section 5.1). Additionally, the implications of the current study for L2 teaching and assessment are considered (Section 5.2). Furthermore, the limitations that should be considered when interpreting the results are discussed, and some suggestions are made for future research (Section 5.3). The chapter concludes with concluding remarks about the dissertation as a whole (Section 5.4).

5.1   Methodological and theoretical contributions

The aim of this dissertation was to examine syntactic complexity from different perspectives and provide a thorough account of this multidimensional construct in L2 Swedish. The substudies focused on the effects of language proficiency, modality, and individual variation on syntactic complexity in Finnish university students' productions. These substudies contribute to the methodological debate on operationalising syntactic complexity by adopting **different definitions of complexity** and by employing **general, specific, and fine-grained measures** in the analyses, as well as by complementing the quantitative analyses with a **qualitative examination** of syntactic structures in productions by a small subset of learners. The first two substudies focused on group-level tendencies and developmentally meaningful differences between two proficiency levels, whereas the third substudy examined syntactic complexity at the individual level. Furthermore, in the third substudy, an **exploratory approach** to categorising the learners' productions into complexity profiles was piloted.

With the help of a thorough investigation of syntactic complexity, the three substudies were designed to fill various gaps in previous research. To begin with, studies on syntactic complexity have predominantly focused on English as the target language (see, however, e.g. Kuiken & Vedder 2019), and **studies of L2 Swedish are rare**. Additionally, to date, most studies of syntactic complexity in L2 Swedish have focused on written productions (e.g. Kowal 2016), and research focusing on **spoken L2 Swedish** from this point of view is still lacking. This paucity of research on the spoken mode holds true for other target languages as well (see, e.g. Lambert

& Kormos 2014), and up to now, only a few studies have directly compared the use of syntactically complex structures in written and spoken productions by the same learners. Furthermore, despite the demonstrated importance of **individual variation in syntactic complexity** (see, e.g. Vyatkina 2012, 2013; Bulté & Housen 2018), most previous research has primarily focused on group-level differences, whether between proficiency levels or task types (e.g. van Dijk et al. 2011; see also Biber et al. 2016). As the results of the current study demonstrate, both group-level and individual-level studies are needed in order to develop a comprehensive account of syntactic complexity in learner language (see also, e.g. Vyatkina et al. 2015). Moreover, studies of syntactic complexity in general have yielded partly contradictory findings, mostly due to **variations in the operationalisation of complexity** and **differences in study designs** (e.g. Kuiken & Vedder 2012a; Kormos 2014; Bulté & Housen 2018). These contributing factors are thoroughly discussed in this dissertation.

The current study is situated in the CAF research tradition (see Housen et al. 2012a). This means that complexity, accuracy, and fluency are seen as the main dimensions, or systems, of learner language, and these systems can be investigated to describe learners' performance on a given task. Earlier research has shown that these systems form a dynamic system, in which all systems and subsystems are in constant interaction (see, e.g. Lesonen et al. 2020). However, the systems in CAF have been studied both together and separately (Housen et al. 2012b), just as the current study does. While the first substudy included an investigation of the interplay between syntactic complexity and accuracy, the second and third substudies focused exclusively on syntactic complexity. It is crucial to acknowledge the interconnectedness of these systems when interpreting the results of the current study (see Section 5.3), as the first substudy only targeted the relationship between syntactic complexity and accuracy, excluding the analysis of the interplay between these systems and fluency, and between the various subsystems, that is, complexity and accuracy at different linguistic levels (see, e.g. Kowal 2016: 216). However, it is argued here that the CAF systems and subsystems can be investigated separately as long as one remains aware of the possible effects that the **interaction both between and within the systems** has on the results.

The theoretical framework for the present study was based on cognitive usage-based approaches and adopted features and ideas from a variety of theoretical views on language learning (see Sections 2.1 and 2.2). More traditional approaches and more recent DST-based approaches were combined in that developmentally meaningful changes at the group level were examined alongside individual variation. The investigation of change at the group level was grounded on the assumption that L2 productions can show stepwise tendencies in the use of syntactically complex structures (see, e.g. Pienemann 1998; see also Section 2.2), whereas individual

variation was explored because DST-inspired studies underline that no predictions regarding the emergence or use of complex structures can be made, as development is guided by individual variation (see, e.g. De Bot & Larsen-Freeman 2011; see also Section 2.2). According to DST-based views, learner language performance consists of multiple systems, including complexity, and all these systems contain several subsystems, such as syntactic complexity at different linguistic levels (Housen et al. 2012b: 5). As these systems and subsystems are in constant interaction with one another, learner language tends to develop dynamically and nonlinearly (Kowal 2016: 216; see also, e.g. Lesonen et al. 2020). The more traditional approaches are often strictly contrasted with DST-based approaches (see, e.g. Baten & Håkansson 2015: 539), as they differ especially in their view on language development. The current study argues that **these views can be complementary**, particularly when the focus extends beyond the developmental aspect of syntactic complexity.

The first substudy explored the extent to which proficiency affects syntactic complexity, accuracy, and the relationship between these constructs (Aim 1). The findings are consistent with earlier studies in that the learners at the higher proficiency level seemed to produce more accurate structures (see, e.g. Wolfe-Quintero et al. 1998). Furthermore, the low frequency of dependent clauses with a sentence adverbial was in line with many previous studies (see, e.g. Åberg 2020), and this result was expected, given the low frequency of this structure in the learners' input (e.g. Rahkonen & Håkansson 2008: 153; see also Section 2.1). In this study, it was assumed that the more prominent types and features of input have a stronger effect on learners' output (e.g. Ellis 2019: 47). For instance, topicalised structures are frequent in the data from both proficiency levels. This might be because this syntactic frame (XVS) also occurs frequently in these learners' input (see, e.g. Bohnacker & Lindgren 2014; see also Section 2.3.2.5). However, qualitative differences in the use of topicalised structures between the proficiency levels are probable. Less proficient learners tend to topicalise adverbs of time, whereas more proficient learners exhibit more variation in the use of topicalised structures.

On the other hand, the results differed from earlier findings in that both the more and less proficient learners demonstrated relatively high complexity in their productions (see, e.g. Gyllstad et al. 2014). In fact, the learner with the most complex productions in terms of both dependent clause ratio and use of topicalised structures belonged to the lower proficiency level. This indicates that **the relative cognitive maturity** of the learners at the lower proficiency level and their extensive **experience with the target language structures** (as well as with structures in other languages) might have affected their production of syntactically complex structures (see Ellis 2019: 47; Kuiken & Vedder 2019: 207). According to the results of this study, one could argue that, in the case of adult learners, earlier experiences with language might be more important than proficiency level when it comes to syntactic

complexity. It seems that some learners were capable of producing complex structures in Swedish despite apparent limitations in their knowledge of the target language's grammar and vocabulary. It is important to note that this finding does not contravene the assumption of a stage-like development of syntax but rather shows that different factors play more important roles among different learners. The frequency of certain structures in the learners' productions could also be partly explained by task-related factors (see Section 5.3) and by the learners' L1, Finnish, which has a relatively free word order when compared to Swedish (see, e.g. Åberg, 2020: 33, 34, 238; see also Ferrari 2012: 282, 293).

The analyses of the interplay between syntactic complexity and accuracy revealed no statistically significant correlations within the groups at different proficiency levels. This was somewhat surprising, as the analyses of the effects of language proficiency on syntactic complexity and accuracy showed that less proficient learners produce relatively complex structures, but that these structures are more frequently inaccurate than those at the higher proficiency level. The results did demonstrate that there might be a tendency towards a trade-off between relative complexity and overall accuracy in learners' productions. This indicates that the use of cognitively demanding structures might bring with it higher inaccuracy (see, e.g. Skehan 1998). Furthermore, the potential trade-off effects between proficiency levels differed. The frequent use of topicalised structures by the less proficient learners seemed to result in somewhat lower overall accuracy, and the use of dependent clauses with a sentence adverbial affected overall accuracy negatively at the higher proficiency level. Additionally, there might be aspects of the interconnection between syntactic complexity and accuracy at the individual level that could not be detected with the group-level analyses. Some learners might be more willing to take risks, while others may avoid using complex structures, for instance, to avoid making mistakes (see, e.g. Norrby & Håkansson 2007: 46, 55). Whether this is **a conscious strategy of the learner or a result of limited resources** (see Skehan 1998), it may have affected the results regarding the relationship between syntactic complexity and accuracy in the current study. As Ellis (2019: 53) suggests, all learners have their own goals when it comes to language production. Some learners might strive for grammatical accuracy, whereas others focus more on conveying meaning.

All in all, it can be concluded that syntactic complexity is a complex system in which several factors interact with one another. Based on the frequency of such occurrences in the data, one could argue that the structures representing relative complexity (that is, dependent clauses with a sentence adverbial and topicalised structures) correspond to their relative frequency in target language input. Despite the differences in proficiency, all participants in this study were familiar with the target language due to their relatively extensive education on the subject. This might

influence how the participants produced structures with their (partly) limited target language knowledge. Furthermore, the data analysed in the first substudy indicated large individual variation within the proficiency levels.

The second substudy examined the effect of both proficiency and modality on syntactic complexity (Aims 1 and 2). In this study, the two proficiency levels were compared in terms of syntactic complexity in the spoken mode, and these results were then compared with the findings gained from the first substudy. The results of this study lend support to the assumption that language proficiency has a major effect on syntactic complexity in speech (see, e.g. Vercellotti 2019). However, the difference in clause length between the proficiency levels was relatively small. The result concerning clause length was somewhat surprising, as research has demonstrated clear differences in clause length between lower and higher proficiency levels (see Kuiken & Vedder 2012a: 145; see also Kormos 2014). The fact that the proficiency levels did not differ to a great extent regarding clause lengths raises questions about the operationalisation of complexity through this length-based measure in the spoken mode (cf. Vercellotti 2019: 235). On the other hand, relatively long clauses can be expected in spoken discourse, as the planning and production of an utterance occur simultaneously (Beaman 1984: 50–51). As discussed earlier, several factors, such as listing coordinated verb phrases, could have affected clause lengths.

The difference between the proficiency levels was small also in use of topicalised structures. Again, as in the written data, topicalised structures were used frequently at both proficiency levels. This can be partly explained by the task type chosen for the current study (see the discussion in Section 5.3), but it could also be seen as an indication that the learners had noticed the relatively common structure in their input, formed a syntactic frame, and used this frame frequently in their productions (see the discussion above). It is important to remember that the qualitative differences in these structures between the proficiency levels may be notable. In the spoken productions, the number of dependent clauses with a sentence adverbial was low, as it was in the written data, aligning with earlier studies (see, e.g. Åberg 2020) and with the idea that the low frequency of this structure in the learners' input, and in Swedish in general, likely results in low frequency in the learners' output.

The syntactic complexity of the written and spoken productions differed significantly at the lower proficiency level, according to most of the measures used. Written productions were more syntactically complex than spoken productions when T- and AS-unit length, dependent clause ratios and use of dependent clauses with a sentence adverbial in the written and spoken tasks were compared. These results are consistent with some earlier studies (see, e.g. Kuiken & Vedder 2011), but not when it comes to the difference in the use of dependent clauses (see, e.g. Kormos & Trebits 2012; see also the discussion of the fundamental differences between speech and

writing in Section 2.3.2.3). Overall, these findings corroborate the assumption that L2 productions can be affected by the mode of production employed. The **cognitive demands** and the **load on working memory** are greater in the spoken mode, whereas the **opportunities to plan the production** and **utilise explicit knowledge** more efficiently in the written mode facilitate the production of syntactically complex structures (see Skehan & Foster 2012; Tavakoli 2014).

At the higher proficiency level, no statistically significant differences in syntactic complexity were found between the written and spoken productions. The lack of statistically significant differences indicates that these learners can express themselves by using equally complex structures in both the written and spoken modes, despite speech being more cognitively demanding. It appears that the production of structures at this proficiency level is more **automatised**, and learners are therefore able to produce complex structures, even with limited time allowed for planning (see, e.g. Gilabert et al. 2016).

As with the results of the first substudy, the results of the second substudy point to individual variation in the use of syntactically complex structures. Therefore, in the third substudy, the aim was to describe this variation (Aim 3). In this substudy, a three-phase mixed-methods approach (i.e. general-level quantitative analysis, phrasal-level quantitative and qualitative analysis, and qualitative analysis of the structures used in the productions) to studying syntactic complexity in the learners' productions was adopted, and an exploratory method (i.e. an analysis based on group statistics) for analysing individual variation was used. To the best of the author's knowledge, this method has not been used in previous studies of individual variation in learner language. The theoretical approach taken in this substudy was based on DST (see, e.g. Verspoor & Behrens 2011). However, contrary to the general focus in DST studies on examining development with the help of dense data points, the learners in this study were only partially tested longitudinally. The purpose of this study was to examine whether syntactic complexity in productions by the same individuals differs synchronically in the written and spoken modes and diachronically between two data collection periods.

The findings indicate clear inter- and intraindividual differences in the data, revealing that **learners' idiosyncratic choices play a major part in determining the syntactic complexity of their productions** (see also Ferrari 2012: 295). In accordance with Kowal (2016), the learners seemed to exhibit consistent complexity profiles diachronically, as there were no great differences in syntactic complexity between the written productions gathered across an interval of six months. However, the differences in profiles were greater when it came to the productions in the written and spoken modes, and every spoken production was categorised as simple. This may be partly due to the fundamental differences between speech and writing, and the use of complex structures might have been suited to the context (see, e.g. Ferrari

2012; see also Weinert et al. 2013: 172), but as there were some differences between the spoken productions of the different learners, complexity profiles seem to have affected the productions as well. Furthermore, the fact that all productions were categorised as simple indicates that different learners' productions in the larger dataset represent the profile of complex production in the spoken mode (cf. P15 in the written data).

A comparison of the general-level and phrasal-level analyses revealed interesting **differences in the use of complexification strategies**. Some learners seemed to elaborate at both the clausal and phrasal levels, while others preferred one type over the other (see, e.g. Bulté & Housen 2020b: 75; see also Vyatkina 2012, 2013). Furthermore, there are differences in how these strategies were employed throughout the productions and in whether the use of strategies varied between the modes of production. These differences could be studied in more detail with the help of qualitative analysis that reveals, for instance, a repetitive use of certain structures in some learners' productions, which could indicate strong personal styles of producing language. Based on the results, it can also be concluded that, due to the dynamic nature of learner language, syntactic complexity in learners' productions can be influenced by several factors and vary greatly even within a production (cf. Kowal 2016; see also Lesonen et al. 2020). Complexity profiles can help us better understand this variation and direct learners' attention appropriately while performing in different situations (see Section 5.2).

On the whole, this dissertation describes group-level tendencies but also highlights the role of individual variation in syntactic complexity. The results contribute to the methodological debate on operationalising syntactic complexity. Following Bulté and Housen (2012: 27; see also Vyatkina et al. 2015), three levels of the examination of syntactic complexity in the current study can be distinguished: **theoretical**, **observational**, and **operational**. At the theoretical level, this study focuses solely on syntactic complexity and defines it with reference to both absolute complexity, that is, the number of elements used and their relations (Pallotti 2009: 593; Bulté & Housen 2018: 149; Kuiken 2023: 84), and relative complexity, that is, complexity in relation to the learner (Di Domenico, 2017: 1–2; Michel 2017: 52; Bulté & Housen 2018: 148; Wijers 2019: 21; see also Section 2.3.1 and Bulté et al. 2024). The manifestations of these two types of complexity were evaluated according to general-level and more fine-grained measures (absolute complexity) as well as specific measures (relative complexity). These complexity measures target complexification at the higher linguistic levels (i.e. T- and AS-units), the clausal level, and the phrasal level. These comprise the observational level of syntactic complexity in the current study. The general (absolute) measures gauged complexification using length-based measures and a measure based on embedding (ratio), and the phrasal-level measures included measures based on both phrase

length and frequency of qualitatively different phrase types. With the help of the specific (relative) measures, demanding structures in L2 Swedish were examined. Together, these measures form the operational level of the study.

The results of the study demonstrate that, although the investigation of **absolute complexity** can provide an account of syntactic complexity that is nonrelative and, therefore, applicable across learners and languages, the study of **relative complexity** can provide more in-depth, language-specific knowledge of the use of syntactically complex structures and unveil interesting differences between learners at different proficiency levels (see, e.g. De Clercq & Housen 2017: 316–317). According to the results of the first substudy, this is especially true when it comes to the interplay between complexity and accuracy. As noted earlier, criticisms of circular reasoning in relative complexity are common (see, e.g. Baten & Håkansson 2015: 521), because the different aspects of complexity are mixed up. When kept separate, **the two types of complexity** (absolute and relative) **can complement one another**.

The development of complexity can be observed in both the increased use of more demanding structures and the decreased use of less advanced features (e.g. Vyatkina et al. 2015: 42; Kuiken et al. 2019: 164), which poses great challenges for the operationalisation of complexity. In accordance with Norris and Ortega (2009), the results of the current study underline the importance of **using several complexity measures that target different linguistic levels** in production samples (see also Bulté & Housen 2020b: 77). If possible, the results for syntactic complexity at various linguistic levels should be interpreted together. For instance, the interpretation of T-unit length together with that of dependent clause ratio is far more informative than T-unit length alone. In this context, it is important to consider the overlap between these measures, as frequent use of dependent clauses naturally affects mean T-unit length. Similarly, when measuring noun phrase complexity, the effect of the use of modifiers – especially longer modifiers, such as relative clauses – on mean noun phrase length must be acknowledged. To gain a proper understanding of noun phrase complexity in a production, both length-based measures and modifier types should be examined. Although it has been considered less effective to use **overlapping measures**, especially those that are to some extent mutually exclusive (e.g. measures targeting complexity at the clausal and phrasal levels), these measures can also be seen as complementary (see Norris & Ortega 2009: see also Bulté & Housen 2020a: 215). Linguistic levels can also be studied separately, but due to the interconnectedness of subsystems, the interpretations should be solely based on the linguistic level studied, and the results should not be generalised to other linguistic levels, due to possible trade-off effects, differences in the complexification strategies preferred, or other factors.

When it comes to the choice of measures, the current study demonstrates that the relative cognitive maturity of the learners can affect the results. Some measures

might better suit the investigation of syntactic development in productions by younger, beginner-level learners, whereas cross-sectional comparisons of cognitively mature learners might require different measures, such as complexity indexes or scales based on complexity scores of different structures (see, e.g. Vercellotti 2019: 239). Furthermore, as noted above, high complexity can be a product of several factors. For instance, a learner may repeat the same syntactically complex structures throughout a production. This factor was inspected with the help of **qualitative analysis** in the third substudy. This method proved especially useful in revealing this type of repetitive use of structures, but also in studying other stylistic choices made by the learners that could not be detected with the help of general- and phrasal-level quantitative measures. An individual's stylistic choices are closely related to their style of producing structures of varying degrees of complexity. Therefore, it is recommended that, in individual-level studies, quantitative observations be complemented by qualitative analyses.

In agreement with earlier studies (see, e.g. Vyatkina et al. 2015: 43; Bulté & Housen 2020b: 74–75), the **combination of group-level and individual-level analysis** is considered extremely beneficial when it comes to complexity research. The group-level analyses in the substudies provided important information about tendencies in the data, whereas the individual-level analyses in the third substudy showed why the tendencies cannot be generalised to individual learners. As discussed above, syntactic complexity in learner productions can be studied with reference to a number of features, such as general measures, phrasal-level complexity, and range of structures. These features can be relatively constant in an individual's production or prone to change depending on various factors, such as modality (Weissberg 2000; Norrby & Håkansson 2007; De Bot & Larsen-Freeman 2011; Kowal 2016; Bulté & Housen 2018; Kuiken et al. 2019). The third substudy demonstrated how inter- and intraindividual variation can be studied by categorising learners' productions into complexity profiles with the help of group-level statistics. This exploratory approach was not intended to generalise the findings per se but rather to show how more general conclusions, here regarding individual variation in syntactic complexity, can be drawn based on individual cases (cf. Bulté & Housen 2020a: 233, 235). This new method seemed to work relatively well in categorising the productions according to their complexity profiles (see Section 5.3 for further discussion), although the dynamicity of syntactic complexity in learner language complicates the categorisation to some extent. This highlights once more the importance of qualitative analysis in individual-level studies of syntactic complexity.

Another facet to be considered in the difference between group- and individual-level studies is the use of different types of measures. Analyses in studies conducted at the group level are often based on general measures of syntactic complexity, as such measures are relatively simple to apply to large sets of data. These measures

can provide important information about tendencies within groups, but general-level measures alone, especially those based on length, fail to indicate levels of complexification, that is, the strategies used to make a structure more complex (see, e.g. Vercellotti 2019: 235; see also Kuiken et al. 2019). More fine-grained syntactic complexity measures can be more rigorous and therefore more suitable for individual-level analysis, especially when computed manually (see also the discussion in Section 5.3).

In sum, this dissertation is methodologically significant because it demonstrates the benefits of studying syntactic complexity by distinguishing different types of complexity, using various measures at different linguistic levels, and, when possible, complementing these efforts with qualitative analysis. Furthermore, the current study takes the first step towards exploring the use of group-level statistics to straightforwardly profile syntactic complexity in individual learners' productions.

This study drew on different theoretical views by combining more traditional approaches to studying syntactic complexity in learner language with a DST-inspired approach. The traditional approaches include the study of complexity within the CAF research tradition as well as those that approach language learning as a stepwise process, such as PT-inspired approaches (see, e.g. Pienemann 1998). Although the PT framework was not implemented in this study, it is considered important when it comes to understanding the stages of syntactic complexification. The DST approach was primarily used as a theoretical standpoint and not as a methodological framework (see the criticism presented in Bulté & Housen 2020a: 216). With this combination of approaches, the study aimed to show that these approaches can be complementary for understanding different perspectives on syntactic complexity— that is, individual variation and developmentally meaningful differences between proficiency levels. Even though the developmental path of syntax was not examined in the present study, it can be concluded that **individual variation and nonlinearity in learners' productions do not exclude the possibility of linearity in their development**. The underlying development of syntactic structures (mostly related to accuracy; see Lowie & Verspoor 2015: 66; see, however, the discussion in Baten & Håkansson 2015: 541) can take place in a linear, stepwise manner in line with the principles presented in PT, whereas variation can be part of the (developmentally meaningful) change that a learner's language use undergoes in the later phases (mostly relating to complexity), in accordance with DST. Additionally, learners' varying styles of producing syntactically complex structures can pass through stages in which certain strategies are favoured over others. As noted by several researchers, group-level tendencies can be detected but not generalised to the individual level.

Finally, from an empirical point of view, the current study provides further information on how **language proficiency, modality, and individual variation affect syntactic complexity**, and how these contributing factors can interact in

learners' productions. Although the relationship between language proficiency and various syntactic complexity features has been well researched, many aspects of syntactic complexity remain unclear. Previous studies have often assumed a positive correlation between proficiency and syntactic complexity (see, e.g. the discussion in Kuiken & Vedder 2019). However, as syntactic complexity is inherently multifaceted and affected by various factors, this correlation can vary depending on the measures adopted and the study design. The findings in previous research have been mixed, and it is still unclear how proficiency affects, for instance, the syntactically complex structures used in speech and writing. The current study offers further evidence regarding the influence of modality and individual variation. Proficiency seems to play an important part in producing syntactically complex structures in speech (see, e.g. Spoelman & Verspoor 2010; see also Skehan 1998; Skehan & Foster 2012), but even here, complexity profiles can vary across learners. A complexity profile can also be a more stable trait of a given learner's production, and the similarities can be used to form learner profiles (see, e.g. Norrby & Håkansson 2007; Kowal 2016).

5.2 Implications for L2 teaching and assessment

The results of this study expand our understanding of the differences in syntactic complexity at different proficiency levels and in the two modes of production, as well as of learners' differing styles of producing varyingly complex language. These results can help teachers develop teaching and evaluation strategies and can guide learners in their language use (see the discussion in, e.g. Dörnyei 2005: 154). The overall results of the dissertation have many potential implications for L2 teaching and assessment. The main implications of the first substudy pertain to the assessment of learner productions, whereas the findings of the second and third substudies provide important insights into how syntactic complexity could be implemented in L2 teaching. In this section, the main implications of the independent substudies are discussed.

The starting point of the first substudy, which examines the relationship between syntactic complexity and accuracy, is that the interaction between systems and subsystems in language is guided by task-related factors and affected by learners' cognitive resources (see, e.g. Robinson 2001; Skehan & Foster 2012). Therefore, the interaction has a dynamic nature (see, e.g. Larsen-Freeman 2006; Spoelman & Verspoor 2010). Under certain circumstances, especially in productions by learners at lower proficiency levels, the interplay between complexity and accuracy can result in trade-off effects. In other words, complexity and accuracy can enter into a competitive relationship with one another, as learners are able to pay attention to either complexity or accuracy but cannot consider both aspects simultaneously. Due

to their limited attentional resources, **learners might have to prioritise either complexity or accuracy in their productions** (see the discussion in, e.g. Tavakoli 2014: 219; Kowal 2016: 213, 216). In practice, this could mean that learners either take a risk with complex structures and produce more non-target-like structures, or they concentrate on the accuracy of the structures and produce simple language. Adult learners in particular might use the latter as a conscious strategy in L2 production (Norrby & Håkansson 2007: 46, 55). On the other hand, learners who can express themselves using complex language in their L1 and possibly in other additional languages (e.g. in English) might wish to be able to use more complex language, even with limited resources and knowledge of the target language (see the discussion in Norrby & Håkansson 2007: 45). The relative cognitive maturity of learners and their use of varying communicative strategies enables, among other things, the production of long structures (Foster et al. 2000: 355). The findings of the first substudy accentuate the role of cognitive maturity in learner language produced by adult learners.

Even though the results of the first substudy do not indicate any statistically significant correlations between syntactic complexity and accuracy, teachers and learners alike should be aware of the possible trade-off effects between these constructs. In general, the use of more demanding structures, such as topicalised structures and dependent clauses with a sentence adverbial (the relative complexity measures used in this study), can diminish accuracy when learners start to experiment with these structures in their productions (see, e.g. Ferrari 2012: 284, 288, 293; see also Lowie & Verspoor 2015). The tendencies discovered in the written productions by learners at different proficiency levels support this assumption. Learners should not be punished for increased inaccuracy when their performance is assessed by teachers, as this can be a natural stage of language development, and teachers should not try to discourage this central part of the learning process (Lowie & Verspoor 2015: 83). Similarly, both increases and decreases in syntactic complexity can indicate progress in a learner language (Ferrari 2012: 293), and this dynamicity of complexity should also be recognised in the assessment of L2 production.

Traditionally, errors have had a major influence on the overall assessment of linguistic productions in educational contexts, and accuracy has been emphasised in L2 teaching. This has also influenced the way learners understand the different features of L2 production, such as complexity and accuracy. In recent years, more emphasis has been placed on communication, and accuracy has played a lesser role in L2 teaching and assessment. This can also be observed in the shift in attitudes towards linguistic prescriptivism, and in the lesser use of productions by natives as a reference point for L2 learners' proficiency assessments (see, e.g. Ellis 2019: 53; see also Lintunen et al. 2020). The multidimensionality of learner language, and the

interaction between the systems and subsystems therein, should continue to receive increasing attention in L2 teaching. This change in perspective must be rooted in education and assessment so that the emphasis on errors in general can be reduced. As discussed in Section 2.3.2.6, the interplay of systems and subsystems is acknowledged in CEFR, but further studies on this matter are needed to provide a comprehensive picture of the structures that might be influenced by trade-off effects. Target-language-specific descriptions of such structures based on empirical evidence, such as those presented in the current study, could be highly useful for L2 teachers.

The second substudy examines syntactic complexity in spoken productions. The results demonstrate that the mode of production is an important influence on the syntactic complexity produced (see also, e.g. Håkansson & Norrby 2007). Many reasons could be cited for the notable differences between the two modes. In addition to the fundamental differences between speech and writing, for instance, learners' cognitive resources affect the syntactic complexity of their spoken productions. Furthermore, based on earlier research, it can be assumed that advanced learners have learned to balance their use of syntactically complex structures and use context-appropriate structures in their productions. This does not necessarily mean that spoken output is automatically less complex. As discussed in Section 2.3.2.3, certain complex structures are common in speech; in particular, clauses tend to be long, as the processes of planning and production take place simultaneously (see also Rosmawati 2020: 126).

First and foremost, teachers should be aware of learners' **different styles of producing both spoken and written language** (see, e.g. Dörnyei 2005: 154). According to Vyatkina (2013: 25), the use of **context-appropriate structures** should be explicitly taught to learners (see also Vyatkina et al. 2015: 44; on the effects of explicit teaching, see, e.g. Ortega 2014). For instance, learners can be provided with level-appropriate specific syntactic features prior to a language production task to increase their awareness of what is expected. In this manner, learners can also practice using more demanding structures instead of relying too heavily on simple ones, or they can use more register-appropriate structures, for example, complex noun phrases in academic writing (Rosmawati 2020: 126; see also Ferrari 2012; Lan, Liu & Staples 2019; Larsson & Kaatari 2020). As Larsson and Kaatari (2020) argue, it is not logical to explicitly instruct learners to, for instance, produce longer sentences, but it is important to make them aware of the differences between genres and registers. Learners often underuse certain structures, or they may rely on structures that are atypical of a given text type. Increased knowledge of context-appropriate structures can expand the active role of the learner in the learning process (Rosmawati 2020: 126). To fully benefit from explicit instruction

regarding differences across registers, learners should also be made aware of their own styles of producing structures in different contexts.

As discussed earlier, different emphases in L2 teaching can influence the type of knowledge that learners acquire (see Section 2.3.2.3). In the recent past, spoken communication has received more attention in teaching, balancing the previously dominated emphasis on writing (see, e.g. Zhang 2013: 836). It is important that in L2 teaching, the fundamental differences between speech and writing and their effects on the structures commonly used in written and spoken communication be explicitly addressed. In this way, learners can become aware of their use of syntactic structures of varying degrees of complexity in written and spoken production. Increasing learners' awareness of variation in syntactic structures might help them notice new target language forms (see Dörnyei 2005: 154; Rosmawati 2020: 126; see also the discussion in Littlemore & Juchem-Grundmann 2010). In addition to promoting such awareness, learners should receive feedback on their use of syntactically complex structures in different contexts (Vyatkina et al. 2015: 44).

Although complexity is not considered a core component of spoken communication, and communicative skills are far more important than variation in syntactic structures (see, e.g. De Clercq & Housen 2017: 320), learners should be encouraged to vary their use of structures even in spoken discourse (see Yuan & Ellis 2003: 24, García Mayo & Azkarai 2016: 243). In general, learners pay more attention to the content in speech and to the form in writing (García Mayo & Azkarai 2016: 241, 244, 257; Kormos 2014: 208). Specific tasks that integrate spoken and written discourse can be used in L2 teaching to direct learners' attention to both content and form (García Mayo & Azkarai 2016: 257, 260). More research is needed on the extent to which learners can benefit from explicit teaching of the differences between modes and contexts (see also Gilabert et al. 2016). In conclusion, it is important that the context-appropriate use of structures continue to be a part of L2 teaching, and that learners are encouraged to use varyingly complex structures in speech.

Examining syntactic complexity in learners' productions through **multilevel analyses** provides a **conceptual framework** that could be useful in L2 teaching (see also Rosmawati 2020). The triphasic analysis used in the current study – which combines general-level, phrasal-level, and qualitative approaches – can be harnessed to analyse complexity in learners' productions in educational contexts. Based on the results of the third substudy, even though length-based measures alone do not fully reveal the strategy used in a given production, together with dependent clause ratios, these measures can relatively efficiently describe the syntactic complexity therein. For instance, long T- or AS-units together with a low dependent clause ratio indicate that the clauses in the production are long, and complexification is therefore probable at the phrasal level. However, this is not always the case, as can be seen in the spoken production by Participant P15, wherein long clauses are mainly a result of listing

subjectless phrases and noun phrases without determiners or modifiers. Thus, phrasal-level analysis can develop a more thorough account of the syntactic complexity of a production. However, the definitive interpretation can only be based on a qualitative analysis, as was the case with Participant P04. By conducting these multilevel analyses, teachers and learners can investigate different styles of producing syntactically complex structures in learner language.

As the results of the third substudy show, there are broad inter- and intraindividual differences in the use of syntactically complex structures, and learners can also differ in how consistent their style of producing language is both synchronically (between the modes of production) and diachronically (within the same mode). These differences can be understood with the help of complexity profiles. **Building profiles of learners' productions** can be an effective method of both recognising individual learners' differing styles of producing syntactically complex structures and **instructing learners to modify their use of these structures**. In other words, based on such profiles, teachers can guide learners accordingly: The learners who constantly use simple structures can be guided, in accordance with the given context, to use more syntactically complex language, whereas those whose productions contain an excessive use of complex structures can be instructed to pay attention to complexity and vary their use of these structures to a greater extent. By enriching their teaching materials, teachers can provide students with examples of more context-appropriate structures (Rosmawati 2020: 126).

Other aspects of learner language, such as accuracy, can be analysed in learners' productions in a similar manner with the help of profiles. Additionally, it is important to investigate the relations between these aspects and, if possible, build learner profiles based on these observations (see, e.g. Norrby & Håkansson 2007; Kowal 2016; see also Olkkonen, Mutta & Lintunen 2024). It would be beneficial for both teachers and learners to analyse patterns spanning productions and to determine whether this output points to a specific profile (e.g. +complexity and -accuracy; -complexity and +accuracy; +complexity and +accuracy; -complexity and -accuracy). This method could also be used as part of the continuous assessment of learners' progress.

5.3   Limitations and future directions

In this section, the limitations of the current study are addressed, and suggestions for future research are proposed. To begin with, the results have limited generalisability. The scope of the first and second substudies was limited in terms of the number of subjects, and with a **small sample size**, it is difficult to make generalisations about a larger population. The third substudy focused on the individual level and did not aim to generalise apart from the fact that some individuals can represent extremes

on the syntactic complexity continuum, and that these individuals differ from the group averages. The data sampling procedure used in the third substudy, extreme case sampling, albeit appropriate for the purpose of this study, yielded **accentuated differences** between learners. Therefore, the results should not be overgeneralised. This limitation was mitigated by including an average production in the analyses. Individual-level studies, in general, suffer from **limited generalisability** (e.g. Lowie & Verspoor 2015: 80; see also De Bot et al. 2007; Bulté & Housen 2020a; and the discussion below). In many studies, the authors warn their readers not to make generalisations about a larger population based on individual-level results. Conversely, in the current study, it is highlighted that making **generalisations about individuals based on large learner-language corpora is just as dangerous** (see also Bulté & Housen 2018: 160). Despite the small number of participants in the current study, the results offer important information on both group-level tendencies and individual variation in syntactic complexity in L2 Swedish. However, larger datasets with more individuals are needed to support or refute the results presented in the first and second substudies, and more research on individual-level differences is needed (see, e.g. Bulté & Housen 2018: 148).

Additionally, this study was unable to survey the entire scope of syntactic complexity, and the analyses were limited to **certain aspects at certain linguistic levels**. As the outcome of an analysis of complexity is affected by the measures applied (Bulté & Housen 2020b: 75–76), the results should be interpreted with caution. For instance, particularly in short productions, the number of linguistic structures produced influences the results that can be obtained with certain complexity measures. The shorter the production, the fewer the possibilities for a given structure – such as a topicalised structure – to occur. Furthermore, short productions contain fewer clauses than longer ones, and the opportunities to use dependent clauses are therefore limited compared to longer productions. In this study, the length of productions was not standardised, as the focus was on the learners' concrete capability to produce syntactically complex structures in Swedish (see Skrzypczak 2021). In future studies, however, such as those focused on a specific complexity feature, it would be beneficial to standardise the production length. Moreover, it is important to acknowledge the limitations of using relative complexity measures. In the current study, the syntactic features examined were highly **language-specific**. However, when combined with absolute complexity measures, these measures revealed intriguing details about the syntactic structures used by learners at different proficiency levels (see also De Clercq & Housen 2017: 316–317).

As the current study examined syntactic complexity and the interplay of syntactic complexity and accuracy from a limited number of perspectives, the interpretation of the learners' productions was also limited. It is important to remember that the

results reported only concern the linguistic levels examined in the substudies. Apart from the tentative findings gained in the first substudy, this enquiry does not address **the interconnectedness of systems and subsystems**, but it does acknowledge the existence of this interplay and the possible effects it had on the results. For instance, when interpreting the results regarding the use of noun phrases, one must keep in mind that this distinct aspect of syntactic complexity is interconnected with other facets and can even be considered mutually exclusive with respect to elaboration at the clausal level (see, e.g. Norris & Ortega 2009). The reader should also bear in mind that partly overlapping measures (e.g. T-/AS-unit length and clause length) were used in the analyses.

All in all, by delimiting the investigation of complexity to syntax and using a limited number of measures, the study risks ignoring other core features of complexity in learner language. This problem arises with all research on this topic, as the wide scope of complexity cannot be included in a single study, or even in a compilation thereof. However, the delimitation of the research objectives and the choice of measures in the current study were carefully considered, and the decisions were based on the findings presented in earlier research.

A crucial point to be made when interpreting the statements made in this dissertation is that **more is not necessarily better** when it comes to syntactic complexity. As the current study focuses on describing differences between proficiency levels and individuals instead of measuring development, the use of measures based on complexity indexes or weighted scales that circumvent the assumption that 'more is better' was not considered necessary (see the discussion in Vercellotti 2019). As already discussed, syntactic complexity is a highly context-bound construct, and proficient language users can often express themselves in a simpler but more qualitatively varied manner than those at lower proficiency levels (see, e.g. Granfeldt 2008; Lambert & Kormos 2014; Ferrari 2012). For this reason, in an ideal scenario, a multilevel analysis with a mixed-methods approach should be conducted to validly study the multifaceted construct of complexity (see the third substudy; see also Rosmawati 2020). Especially in future individual-level studies, qualitative analysis should play a more prominent role. Furthermore, based on the arguments here, it can be concluded that DST can serve as an especially fruitful framework for studies of learner language complexity.

Several other factors should be considered when interpreting the results. For instance, the fundamental differences between speech and writing played a crucial role in the comparison between the modes of productions. However, as discussed previously in this summary, the type of data (i.e. monologues and informal essays) used in the study and the analysis procedures adopted (see Foster et al. 2000: 370–371) increase the comparability between the two modes of production. In addition to the factors studied in this dissertation, influences on learners' production include, for

instance, the learner's attitude towards the task and various task-specific factors. These contributing factors can be considered **noise**, that is, unwanted effects from factors not encompassed within the study objectives (see, e.g. Bulté & Housen 2020a). This noise should be kept distinct from meaningful noise caused by individual variation (see, e.g. Verspoor & Behrens 2011: 37; van Dijk et al. 2011: 60–62; Bulté & Housen 2020a: 212). It is beyond the scope of this study to examine all the unwanted effects of these factors, and thus, they must be kept in mind.

The productions analysed in this study were based on tasks that consisted of three parts. The participants described actions in a cartoon strip, narrated how the topic related to their own lives, and discussed the topic on a more general level. This design was based on the idea that a **hybrid task type** enables the use of versatile syntactic structures. The disadvantage of the design is that the analyses were based on different task types, which elicit the use of different types of structures (see, e.g. Ferrari 2012: 288, 291; Kormos & Trebits 2012: 462; De Clercq & Housen 2017: 320; Kuiken et al. 2019: 164–165), so the study could not control for certain task effects. For instance, topicalised structures are frequent in the descriptive parts of tasks (see Lindgren 2020: 198), but all parts seem to evoke the use of dependent clauses (see, e.g. Kormos 2014: 209; Baten & Håkansson 2015: 538–539). As the different parts of the learners' productions varied in length, the frequency of these structures may have been affected accordingly. It appears that the argumentative sections of the tasks were especially demanding for the less proficient learners, as these responses were short in many productions and, in some cases, missing altogether. Additionally, as discussed in Section 3.2, a learner's vocabulary can shape the syntactic complexity of their productions. A broader vocabulary enables the use of more complex structures (see Tomasello 2003: 92), whereas a lack of vocabulary relevant to a given topic may restrict learners' ability to perform a task.

The **application of DST** in the current study is tentative. Some of the central principles of DST regarding, for instance, dense longitudinal data and the intricate methods used in studying learner language development, are not aligned with the study of developmentally meaningful changes in the first two substudies, or with the case study approach taken in the third substudy. Housen and Bulté (2020b: 76) point out that several time points are needed to reliably and validly study the dynamicity of systems in learner language. The current study demonstrates that dense data points are not always needed when the focus is not on development, as even cross-sectional and individual-level studies with limited data points can indicate considerable variation both within and across learners, although this change may be developmentally insignificant (see the discussion in De Bot et al. 2007: 15; Lowie & Verspoor 2015: 81–82; see also Murakami 2020). As the aim was not to study the development of complexity per se, naturally, nothing can be concluded about the

developmental aspect based on the substudies in this dissertation (see also Murakami 2020).

It is crucial to consider all the limitations discussed in this section when interpreting the results of the present study to avoid drawing faulty conclusions. Despite these limitations, the study offers important information about the effects of language proficiency, modality, and individual variation on syntactic complexity in L2 Swedish. Numerous topics require further examination to provide additional insights into both group-level tendencies and the role of individual variation in syntactic complexity. In order to study developmentally meaningful changes more reliably and develop a fuller picture of complexity in L2 Swedish at different proficiency levels, larger datasets are needed. However, as the use of more fine-grained measures is necessary in complexity studies, and these are extremely laborious to calculate manually, the next step should be to further develop programmes that enable the **automatic analysis of L2 Swedish** (see, e.g. Kyle 2016; see also Bulté et al. 2024: 5, 31). Such programmes could also provide a useful means for learners to receive feedback on their use of syntactically complex structures (see Section 5.2).

In future research, a **cross-linguistic approach to studying complexity profiles** and their relation to learner profiles would be beneficial. As discussed previously, languages can differ in terms of syntactic complexity, and a learner's L1 can affect the structures they produce (see, e.g. Gyllstad et al. 2014; De Clercq & Housen 2017; Bernardini & Granfeldt 2019; Kuiken et al. 2019; see also Åberg 2020). Therefore, it would be helpful to investigate complexity profiles in L1 and several additional languages by the same learners (see also Gyllstad et al. 2014: 23). Additionally, some of the individual variation in syntactic complexity might be due to how learners plan and edit their productions. Thus, in future, it would be interesting to study the **process** underlying written composition, as well as the cognitive processing of the individuals as they produce both written and spoken learner language. With the help of retrospective interviews, it would be possible to investigate the conscious choices learners make in their productions.

The new method piloted in the third substudy provides a first step towards exploring the **use of group-level statistics in building complexity profiles**. The findings concerning this method, while preliminary, provide a useful starting point for future research on complexity profiles. In the current study, a production was categorised into a certain profile if two out of the three general syntactic complexity measures were aligned with the given profile. This seemed to work well in distinguishing productions of differing complexity. However, there is considerable variation within these profiles, as each individual represents their profile in a different way and combines characteristics from different mixtures of profiles (see also Kowal 2016). Therefore, a larger number of measures in future studies could be

beneficial in developing this method of profiling. In even deeper analyses, for instance, of the relationship between different syntactic complexity measures (see, e.g. Bulté & Housen 2020a), perhaps in combination with the accuracy of the complex structures produced (see, e.g. Kowal 2016), more differences within and between the complexity profiles could be detected. In future studies, the aim should be to further develop these types of methods in order to establish a basis for the dynamic modelling of learners' differing styles of producing language based on complexity profiles. The use of such methods could help researchers and teachers alike **to study individual variation in syntactic complexity in a more straightforward manner**.

5.4   Concluding remarks

The aim of the present study was to describe syntactic complexity in L2 Swedish produced by L1 Finnish university students. In the first substudy, the effect of language proficiency on syntactic complexity and accuracy, as well as on the relationship between these systems of learner language, was examined, addressing the first main research aim. Somewhat surprisingly, the results showed that the two proficiency levels investigated in the study did not differ greatly in terms of syntactic complexity, and there was no statistically significant relationship between complexity and accuracy at either of the proficiency levels. In the second substudy, the effect of language proficiency on syntactic complexity was studied in spoken productions, contributing to the second main research aim of the dissertation. The results of the study confirm that a learner's proficiency plays an important role in the syntactically complex structures they produce in the spoken mode. The third substudy focused on individual variation in syntactic complexity, addressing the third and final main research aim of this dissertation. The results of this study demonstrate that individual variation can be considerable, even within a proficiency level, when it comes to syntactic complexity at both the general and phrasal levels. This is an important contribution to the growing body of individual-level research. Building profiles of learners' productions and use of qualitative analyses proved to be an effective method of investigating the role of individual variation in syntactic complexity.

By focusing on group-level tendencies and individual variation in both written and spoken data, the study provides a comprehensive account of syntactic complexity in learners' productions. As such, the study is among the first to target syntactic complexity in L2 Swedish from a variety of perspectives and therefore fills a gap in the field of complexity research. The methodological contributions of this study include, for instance, **the versatility of the approaches used to analyse syntactic complexity**. Even though the study of syntactic complexity can be

demanding at the group level due to individual variation and the multidimensionality of this construct, these studies offer valuable information that can be used as a starting point for individual-level research, as exemplified in this dissertation. In accordance with earlier research, it can be concluded based on the results of the substudies that complexity only partially correlates with proficiency. It is argued in the present study that an individual's style of producing a language might, in fact, have a more prominent influence on syntactic complexity. This highlights the need for more studies of individual variation in learner language complexity.

From a theoretical perspective, in line with De Bot et al. (2007), this study argues that we should adopt more open-minded approaches to studying syntactic complexity. **Seemingly contradictory views and approaches can, in fact, be complementary**, and group-level and individual-level studies should be conducted in parallel to explore both 'the wood and the trees' (Ellis 2007: 25, 2019: 53; see also Kowal 2016: 19; Kuiken 2023: 90). Traditional approaches can describe stepwise group-level tendencies, but these tendencies can be masked by individual variation. Even individual variation can reveal constancies, which can be studied by profiling learners' productions (see also Ellis 2019: 53).

In SLA research, the aim of CAF studies has been to assess and describe language proficiency (e.g. Kuiken 2023). Therefore, complexity, among other CAF components, has often been studied in productions by learners at a certain proficiency level (see, e.g. Biber et al. 2016; Lan, Lucas & Sun 2019). Generally, the aim of these studies is to discover statistical differences between learners' productions and correlations, for instance, between different measures of complexity and proficiency levels. To identify these differences and correlations, large datasets are needed to counterbalance the influence of individual variation (see Ortega 2003). Group-level studies based on large sets of data are important, but **the aim should not be to eliminate individual variation and disregard this important aspect of learner language as unwanted noise** (Bulté & Housen 2020b: 77; see also the discussion in Larsen-Freeman 2019). In other words, group-level case studies have their place in complexity research (see Bulté & Housen 2020a: 228–229), but these studies should not ignore the importance of individual variation that can be concealed behind group averages (Larsen-Freeman 2006: 598; Verspoor & Behrens 2011: 37; see also the *ergodic principle*, Lowie & Verspoor 2019: 185; see also Ferrari 2012: 295; Lowie & Verspoor 2015: 81; Bulté & Housen 2018: 160, 2020b: 56, 74, 77–78). Ultimately, this is a question of the research objectives and methods that a researcher adopts (Lowie & Verspoor 2015: 84; see also the discussion in Bulté & Housen 2020b: 77–78). Lowie and Verspoor (2015: 63) argued that group-level studies should be conducted and group trends investigated if the aim is to examine the big picture, but if the goal is to examine variations in multiple variables, individual-level studies with appropriate analyses are needed. In sum, these

approaches should be considered complementary, and group trends should not be generalised to individual learners (Larsen-Freeman 2006: 598; Lowie & Verspoor 2015: 78; see also the discussion in Bulté & Housen 2020b: 77–78).

In this study, it was assumed that syntactic structures are learned via processing of the target language in such a way that learners develop abstract frames of frequently occurring structures and use constructions based on these frames to produce the target language (see, e.g. Eskildsen 2009). Learners are able to restructure linguistic input in their own production with the help of explicit instruction (see, e.g. Ellis & Wulff 2014: 83, 88). The changes in a language system occur through self-organisation and are proceeded by variation (e.g. Lesonen et al. 2020: 3; Verspoor, Lowie & van Dijk 2008: 217). Furthermore, it was assumed that learners have limited cognitive resources that can easily become overloaded when performing a cognitively demanding task (see Skehan 1998). Due to the limited resources available at lower proficiency levels, learners can be obliged to limit their attention to certain linguistic aspects while neglecting other areas. This can be seen as a trade-off between the systems of learner language, such as complexity and accuracy. Through language use, learners can develop a more automatic ability to produce syntactically complex structures in the target language, even with the limited time for planning in speech (Gilabert et al. 2016). Moreover, learning and language use are highly individual and dynamic in nature (see, e.g. Lowie & Verspoor 2015: 67, 69). Therefore, they cannot be fully described by linear stepwise models, which, nevertheless, have their place in studies concentrating on the big picture.

Syntactic complexity is inherently multifaceted and context-bound. Complexity as a part of learner language is not an aspect that, like accuracy and fluency, develops towards a given goal that can be described in a relatively straightforward manner and measured based on unwanted and desirable features regardless of the context. The very nature of complexity changes as a learner becomes increasingly proficient, and many context-related factors determine which facets of complexity features will emerge (see, e.g. Ferrari 2012). In addition, as demonstrated in the current study, an individual's style of producing language has a strong influence on the structures they use (see also Norrby & Håkansson 2007; Kowal 2016).

In conclusion, this discussion and the results of the current study do not imply that syntactic complexity should not be approached using traditional methods and described as stepwise development. Instead, it demonstrates that insignificant results in group-level studies should not be automatically disregarded as meaningless due to the noise in the measurement, but rather seen as an indication of variation in the dynamic system. Group- and individual-level studies should complement one another. In addition to the methodological and theoretical contributions of this study,

the enquiry opens up many paths for applying the findings to L2 teaching. From a purely research-related perspective, this study contributes to **the shift in focus from developmental patterns to a more multidimensional and open-minded approach to syntactic complexity**, particularly in studies of syntactic complexity in L2 Swedish.

Abbreviations

AS-unit	Analysis of Speech Unit
C	Clause
CAF	Complexity, accuracy, and fluency
CEFR	The Common European Framework of Reference for Languages
DC	Dependent clause
DC+	Dependent clause with a sentence adverbial
DST	Dynamic Systems Theory
L1	First language
L2	Second language
NP	Noun phrase
PT	Processability Theory
SLA	Second Language Acquisition
TOP	Topicalised structure
T-unit	(Minimal) Terminable Unit
W	Word

List of References

Albert, A. (2011). When individual differences come into play. In P. Robinson (Ed.) (pp. 239–266). https://doi.org/10.1075/tblt.2.16ch9.

Angouri, J. (2018). Quantitative, qualitative, mixed or holistic research? Combining methods in linguistic research. In L. Litosseliti (Ed.), *Research Methods in Linguistics* (second edition) (pp. 35–51). London: Bloomsbury Publishing. https://doi.org/10.5040/9781350043466.ch-003.

Ansarifar, A., Shahriari, H., & Pishghadam, R. (2018). Phrasal complexity in academic writing: A comparison of abstracts written by graduate students and expert writers in applied linguistics. *Journal of English for Academic Purposes, 31*, 58–71. https://doi.org/10.1016/j.jeap.2017.12.008.

Atkinson, D. (2011). Introduction. Cognitivism and Second Language Acquisition. In D. Atkinson (Ed.), *Alternative Approaches to Second Language Acquisition*. (pp. 1–23). London: Routledge.

Baaijen, V. M., & Galbraith, D. (2018). Discovery through writing: Relationships with writing processes and text quality. *Cognition and Instruction*, *36(3)*, 199–223. https://doi.org/10.1080/07370008.2018.1456431.

Bardovi-Harlig, K. (1992). A second look at T-unit analysis: Reconsidering the sentence. *TESOL Quarterly, 26(2)*, 390–395. https://doi.org/10.2307/3587016.

Bardovi-Harlig, K., & Bofman, T. (1989). Attainment of syntactic and morphological accuracy by advanced language learners. *Studies in Second Language Acquisition*, *11(1)*, 17–34. https://doi.org/10.1017/s0272263100007816.

Bartning, I., Martin, M., & Vedder, I. (Eds.) (2010). *Communicative Proficiency and Linguistic Development: Intersections between SLA and Language Testing Research*. Eurosla Monographs Series 1. Amsterdam: European Second Language Association.

Baten, K., & Håkansson, G. (2015). The development of subordinate clauses in German and Swedish as L2s. *Studies in Second Language Acquisition, 37(3)*, 517–547. https://doi.org/10.1017/s0272263114000552.

Beaman, K. (1984). Coordination and subordination revisited: Syntactic complexity in spoken and written narrative discourse. In D. Tannen (Ed.), *Spoken and Written Language: Exploring Orality and Literacy* (pp. 45–80). Norwood, N.J: Ablex.

Behrens, H. (2009). Usage-based and emergentist approaches to language acquisition. *Linguistics*, *47(2)*, 383–411. https://doi.org/10.1515/ling.2009.014.

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(2)*, 211–232. https://doi.org/10.1111/ijal.12257.

Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly, 45(1)*, 5–35. https://doi.org/10.5054/tq.2011.244483.

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.

Biber, D., Johansson, S., Leech, G. N., Conrad, S., & Finegan, E. (2021). *Grammar of Spoken and Written English*. Amsterdam: John Benjamins. https://doi.org/10.1075/z.232.

Bohnacker, U. (2010). The clause-initial position in L2 Swedish declaratives: Word order variation and discourse pragmatics. *Nordic Journal of Linguistics, 33(2)*, 105–143. https://doi.org/10.1017/S033258651000017X.

Bohnacker, U., & Lindgren, J. (2014). Fundament, formellt subjekt och frekvens. Ordföljdsmönster i svenska, nederländska och hos vuxna inlärare av svenska [Fundament, formal subject and frequency: Word order patterns in Swedish, Dutch and in Swedish by adult learners]. *Språk och stil, 24,* 33–71.

Bourdin, B., & Fayol, M. (1994). Is written language production more difficult than oral language production? A working memory approach. *International Journal of Psychology*, *29(5)*, 591–620. https://doi.org/10.1080/00207599408248175.

Bulté, B., & Housen, A. (2012). Defining and operationalising L2 complexity. In A. Housen, F. Kuiken & I. Vedder (Eds.) (pp. 21–46). https://doi.org/10.1075/lllt.32.02bul.

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.

Bulté, B., & Housen, A. (2020a). A critical appraisal of the CDST approach to investigating linguistic complexity in L2 writing development. In G. Fogal & M. Verspoor (Eds.) (pp. 207–238). https://doi.org/10.1075/lllt.54.09bul.

Bulté, B., & Housen, A. (2020b). A DUB-inspired case study of multidimensional L2 complexity development: Competing or connecting growers? In W. Lowie, M. Michel, M. Keijzer & R. Steinkrauss (Eds.) (pp. 50–86). https://doi.org/10.21832/9781788925259-006.

Bulté, B., Housen, A., & Pallotti, G. (2024). Complexity and Difficulty in Second Language Acquisition: A Theoretical and Methodological Overview. *Language Learning*, early view, 1–42. https://doi.org/10.1111/lang.12669.

Bybee, J. (2006). From usage to grammar: The mind's response to repetition. *Language, 82(4)*, 711–733. https://doi.org/10.1353/lan.2006.0186.

Bybee, J. (2008). Usage-based grammar and Second Language Acquisition. In P. Robinson & N. C. Ellis (Eds.), *Handbook of Cognitive Linguistics and Second Language Acquisition* (pp. 226–246). New York: Routledge. https://doi.org/10.4324/9780203938560-18.

Bybee, J. (2010). *Language, Usage and Cognition*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511750526.

Bybee, J., & Hopper, P. (Eds.) (2001). *Frequency and the Emergence of Linguistic Structure*. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.45.

Byrnes, H., & Manchón, R. M. (Eds.) (2014). *Task-Based Language Learning – Insights from and for L2 Writing* (Task-Based Language Teaching 7). Amsterdam: John Benjamins. https://doi.org/10.1075/tblt.7.

Cadierno, T., & Eskildsen, S. W. (Eds.) (2015). *Usage-based Perspectives on Second Language Learning*. Berlin: De Gruyter Mouton. https://doi.org/10.1515/9783110378528.

CEFR (2001, 2007, 2020) = Council of Europe: *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Retrieved from https://www.coe.int/en/web/common-european-framework-reference-languages/level-descriptions.

Chafe, W., & Tannen, D. (1987). The relation between written and spoken language. *Annual Review of Anthropology, 16*, 383–407. https://doi.org/10.1146/annurev.an.16.100187.002123.

Chaudron, C. (2003). Data collection in SLA research. In C. J. Doughty & M. H. Long (Eds.) (pp. 762–828). https://doi.org/10.1002/9780470756492.ch22.

Corder, P. (1967). The significance of learner's errors. *International Review of Applied Linguistics in Language Teaching, 5*, 161–170. https://doi.org/10.1515/iral.1967.5.1-4.161.

Crookes, G. (1990). The utterance, and other basic units for second language discourse analysis. *Applied Linguistics, 11(2)*, 183–199. https://doi.org/10.1093/applin/11.2.183.

Dahl, Ö. (2004). *The Growth and Maintenance of Linguistic Complexity*. Amsterdam: John Benjamins.

Data Protection Directive. (1995). Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data. Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A31995L0046. Accessed 10 August 2016.

De Bot, K., & Larsen-Freeman, D. (2011). Researching second language development from a Dynamic Systems Theory perspective. In M. Verspoor, K. de Bot & W. Lowie (Eds.) (pp. 5–24). https://doi.org/10.1075/lllt.29.01deb.

De Bot, K., Lowie, W., & Verspoor, M. (2007). A Dynamic Systems Theory approach to Second Language Acquisition. *Bilingualism: Language and Cognition, 10(1)*, 7–21. https://doi.org/10.1017/S1366728906002732.

De Bot, K., Lowie, W., Thorne, S. L., & Verspoor, M. (2013). Dynamic Systems Theory as a comprehensive theory of second language development. In M. P. García Mayo, M. J. Gutiérrez Mangado & M. Martínez-Adrián (Eds.) (pp. 199–220). https://doi.org/10.1075/aals.9.13ch10.

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.

DeKeyser, R. (2003). Implicit and explicit learning. In C. J. Doughty & M. H. Long (Eds.) (pp. 313–348). https://doi.org/10.1002/9780470756492.ch11.

DeKeyser, R. (2005). What makes learning second-language grammar difficult? A review of issues. *Language Learning*, *55(S1)*, 1–25. https://doi.org/10.1111/j.0023-8333.2005.00294.x.

Di Domenico, E. (2017). Introduction. In E. Di Domenico (Ed.), *Syntactic Complexity from a Language Acquisition Perspective*. (pp. 1–27). Newcastle upon Tyne, England: Cambridge Scholars Publishing.

Doughty, C. J., & Long, M. H. (Eds.) (2003). *The Handbook of Second Language Acquisition*. Malden, MA: Blackwell Publishing. https://doi.org/10.1111/b.9781405132817.2005.00001.x.

Dyson, B. P. (2021). *Dynamic Variation in Second Language Acquisition: A Language Processing Perspective* (Processability Approaches to Language Acquisition Research & Teaching). Amsterdam: John Benjamins. https://doi.org/10.1075/palart.8.

Dörnyei, Z. (2005). *The Psychology of the Language Learner: Individual Differences in Second Language Acquisition*. New York: Routledge. https://doi.org/10.4324/9781410613349.

Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*. Oxford: Oxford University Press.

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*, 143–188. https://doi.org/10.1017/s0272263102002024.

Ellis, N. C. (2003). Constructions, chunking, and connectionism: The emergence of second language structure. In C. J. Doughty & M. H. Long (Eds.) (pp. 33–68). https://doi.org/10.1002/9780470756492.ch4.

Ellis, N. C. (2007). Dynamic systems and SLA: The wood and the trees. *Bilingualism: Language & Cognition, 10(01)*, 23–25. https://doi.org/10.1017/s1366728906002744.

Ellis, N. C. (2015). Cognitive and social aspects of learning from usage. In T. Cadierno & S. W. Eskildsen (Eds.) (pp. 49–74). https://doi.org/10.1515/9783110378528-005.

Ellis, N. C. (2019). Essentials of a theory of language cognition. *The Modern Language Journal, 103(S1)*, 39–60. https://doi.org/10.1111/modl.12532.

Ellis, N. C., Römer, U., & O'Donnell, M. B. (2016). Constructions and usage-based approaches to language acquisition. *Language Learning*, *66(S1)*, 23–44. https://doi.org/10.1111/lang.1_12177.

Ellis, N. C., & Wulff, S. (2014). Usage-based approaches to SLA. In B. VanPatten & J. N. Williams (Eds.) (pp. 75–93).

Ellis, R. (2009). Measuring implicit and explicit knowledge of second language. In R. Ellis, S. Loewen, C. Elder, R. Ertlam, J. Phil & H. Reinders (Eds.), *Implicit and Explicit Knowledge in Second Language Learning, Testing and Teaching* (pp. 31–64). Bristol: Multilingual Matters. https://doi.org/10.21832/9781847691767-004.

Ellis, R., & Barkhuizen, G. (2005). *Analysing Learner Language*. Oxford: Oxford University Press.

Eskildsen, S. W. (2009). Constructing another language – usage-based linguistics in Second Language Acquisition. *Applied Linguistics, 30(3)*, 335–357. https://doi.org/10.1093/applin/amn037.

Eskildsen, S. W., & Cadierno, T. (2015). Advancing usage-based approaches to L2 studies. In T. Cadierno & S. W. Eskildsen (Eds.) (pp. 1–16). https://doi.org/10.1515/9783110378528-003.

Ferrari, S. (2012). A longitudinal study of complexity, accuracy and fluency variation in second language development. In A. Housen, F. Kuiken & I. Vedder (Eds.) (pp. 277–298). https://doi.org/10.1075/lllt.32.12fer.

Flick, U. (2018). *Doing Triangulation and Mixed Methods*. London: SAGE Publications Ltd, https://doi.org/10.4135/9781529716634.

Fogal, G., & Verspoor, M. (Eds.) (2020). *Complex Dynamic Systems Theory and L2 Writing Development.* (Language Learning & Language Teaching 54). Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.54.

Forbes, K. (2018). The role of individual differences in the development and transfer of writing strategies between foreign and first language classrooms. *Research Papers in Education, 34(4)*, 445–464. https://doi.org/10.1080/02671522.2018.1452963.

Foster, P., & Skehan, P. (1999). The influence of source of planning and focus of planning on task-based performance. *Language Teaching Research*, *3(3)*, 215–247. https://doi.org/10.1191/136216899672186140.

Foster, P., Tonkyn A., & Wigglesworth G. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics, 21(3)*, 354–375. https://doi.org/10.1093/applin/21.3.354.

Gaies, S. J. (1980). T-unit analysis in second language research: Applications, problems and limitations. *TESOL Quarterly, 14(1)*, 53–60. https://doi.org/10.2307/3586808.

García Mayo, M. P., & Azkarai, A. (2016). EFL task-based interaction: Does task modality impact on language-related episodes. In M. Sato & S. Ballinger (Eds.), *Peer Interaction and Second Language Learning: Research Agenda and Pedagogical Potential* (pp. 241–266). Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.45.10gar.

García Mayo, M. P., Gutiérrez Mangado M. J., & M. Martínez-Adrián (Eds.) (2013). *Contemporary Approaches to Second Language Acquisition* (AILA Applied Linguistics 9). Amsterdam: John Benjamins.

GDPR = General Data Protection Regulation. (2018). Retrieved from https://gdpr-info.eu/. Accessed 9 September 2018.

Gilabert, R., Manchón, R. M., & Vasylets, O. (2016). Mode in theoretical and empirical TBLT research: Advancing research agendas. *Annual Review of Applied Linguistics, 36*, 117–135. https://doi.org/10.1017/S0267190515000112.

Gilmore, A., & Ganem Gutiérrez, G. (2020). Investigating complexity in L2 writing with mixed methods approaches. In G. Fogal & M. Verspoor (Eds.) (pp. 183–206). http://doi.org/10.1075/lllt.54.08gil.

Goldberg, A. E. (2006). *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199268511.001.0001.

Golden, A., Kulbrandstad, L. A., & Tenfjord, K. (2017). Evaluation of texts in tests, or: Where is the dog buried? In A. Golden, S. Jarvis & K. Tenfjord (Eds.), *Crosslinguistic Influence and Distinctive Patterns of Language Learning: Findings and Insights from a Learner Corpus* (pp. 231–271). Bristol: Multilingual Matters. https://doi.org/10.21832/9781783098774-011.

Granfeldt, J. (2008). Speaking and writing in L2 French: Exploring effects on fluency, complexity and accuracy. In S. Van Daele, A. Housen, F. Kuiken, M. Pierrard & I. Vedder (Eds.), *Complexity, Accuracy and Fluency in Second Language Use, Learning & Teaching* (pp. 87–98). University of Brussels. Retrieved from https://www.researchgate.net/publication/263125186_Speaking_and_Writing_in_French_L2_Exploring_Effects_on_Fluency_Complexity_and_Accuracy.

Gunnarsson, C. (2012). The development of complexity, accuracy and fluency in the written production of L2 French. In A. Housen, F. Kuiken & I. Vedder (Eds.) (pp. 247–276). https://doi.org/10.1075/lllt.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)*, 1–30. https://doi.org/10.1075/eurosla.14.01gyl.

Halliday, M. A. K. (1989). *Spoken and Written Language*. Oxford: Oxford University Press.

Hashemi, M. R., & Babaii, E. (2013). Mixed methods research: Toward new research designs in applied linguistics. *The Modern Language Journal, 97(4)*, 828–852. https://doi.org/10.1111/j.1540-4781.2013.12049.x.

Henrichsen, P. J., & Allwood, J. (2005). Swedish and Danish, spoken and written language: A statistical comparison. *International Journal of Corpus Linguistics, 10(3)*, 367–399. https://doi.org/10.1075/ijcl.10.3.05hen.

Housen, A., & Kuiken, F. (2009). Complexity, accuracy and fluency in Second Language Acquisition. *Applied Linguistics, 30(4)*, 461–473. https://doi.org/10.1093/applin/amp048.

Housen, A., Kuiken, F., & Vedder, I. (Eds.) (2012a). *Dimensions of L2 Performance and Proficiency. Complexity, Accuracy and Fluency in SLA* (Language Learning & Language Teaching 32). Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.32.

Housen, A., Kuiken, F., & Vedder, I. (2012b). Complexity, accuracy and fluency: Definitions, measurement and research. In A. Housen, F. Kuiken & I. Vedder (Eds.) (pp. 1–20). https://doi.org/10.1075/lllt.32.01hou.

Housen, A., & Simoens, H. (2016). Introduction: Cognitive perspectives on difficulty and complexity in L2 acquisition. *Studies in Second Language Acquisition, 38(2)*, 163–175. https://doi.org/10.1017/s0272263116000176.

Hulstijn, J. H. (2005). Theoretical and empirical issues in the study of implicit and explicit second language learning: Introduction. *Studies in Second Language Acquisition, 27(2)*. 129–140. https://doi.org/10.1017/s0272263105050084.

Hunt, K. W. (1965). *Grammatical Structures Written at Three Grade Levels*. Champaign, IL: National Council of Teachers of English.

Hyltenstam, K. (1977). Implicational patterns in interlanguage syntax variation. *Language Learning, 27(2)*, 383–410. https://doi.org/10.1111/j.1467-1770.1977.tb00129.x.

Håkansson, G. (2004). Utveckling och variation i svenska som andraspråk enligt processbarhetsteorin [Development and variation in Swedish as a second language according to the processability theory]. In K. Hyltenstam & I. Lindberg (Eds.), *Svenska som andraspråk – i forskning, undervisning och samhälle* [Swedish as a second language – in research, teaching, and society] (pp. 153–169). Lund: Studentlitteratur.

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.

Håkansson, G., & Norrby, C. (2007). Processability theory applied to written and oral L2 Swedish. In F. Mansouri (Ed.), *Second Language Acquisition Research: Theory-Construction and Testing* (pp. 81–94). Cambridge: Cambridge Scholars Press.

Iwashita, N. (2006). Syntactic complexity measures and their relation to oral proficiency in Japanese as a foreign language. *Language Assessment Quarterly, 2*, 151–169. https://doi.org/10.1207/s15434311laq0302_4.

Jung, J. (2017). Effects of task complexity on L2 writing processes and linguistic complexity: A keystroke logging study. *English Teaching, 72(4)*, 179–200. https://doi.org/10.15858/engtea.72.4.201712.179.

Juurakko-Paavola, T. & Palviainen, Å. (2011). Svenskans plats i den finska utbildningen. Lagstiftning och utbildningssystem [The place of Swedish in Finnish education. Legislation and education system]. In T. Juurakko-Paavola & Å. Palviainen (Eds.), *Svenskan i den finska skolan och högre utbildningen. Om kunskaper och motivation genom olika utbildningsstadier* [Swedish in Finnish schools and in higher education. On skills and motivation at different stages of education]. (pp. 9–22). Hämeenlinna: Hämeen ammattikorkeakoulu.

Jörgensen, Nils (1976). *Meningsbyggnaden i talad svenska* [The sentence structure in spoken Swedish]. Lund: Studentlitteratur.

Jörgensen, Nils (1978). *Underordnade satser och fraser i talad svenska* [Dependent clauses and phrases in spoken Swedish]. Lund: Studentlitteratur.

Khushik, G. A., & Huhta, A. (2020). Investigating syntactic complexity in EFL learners' writing across Common European Framework of Reference levels A1, A2, and B1. *Applied Linguistics*, *41(4)*, 506–532. https://doi.org/10.1093/applin/amy064.

Khushik, G. A., & Huhta, A. (2022). Syntactic complexity in Finnish-background EFL learners' writing at CEFR levels A1–B2. *European Journal of Applied Linguistics, 10(1)*, 142–184. https://doi.org/10.1515/eujal-2021-0011.

Kormos, J. (2011). Speech production and the Cognition Hypothesis. In P. Robinson (Ed.) (pp. 39–60). https://doi.org/10.1075/tblt.2.06ch2.

Kormos, J. (2014). Differences across modalities of performance: An investigation of linguistic and discourse complexity in narrative tasks. In H. Byrnes & R. M. Manchón (Eds.) (pp. 193–216). https://doi.org/10.1075/tblt.7.08kor.

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*. Krakow: Jagiellonian University Press. https://doi.org/10.4467/k9474.72/16.16.6986.

Krashen, S. D. (1982). *Principles and Practice in Second Language Acquisition*. Oxford Pergamon Press.

Krashen, S. D. (1985). *The Input Hypothesis: Issues and Implications*. New York: Longman.

Kuiken, F. (2023). Linguistic complexity in Second Language Acquisition. *Linguistics Vanguard*, *9(S1)*, 83–93. https://doi.org/10.1515/lingvan-2021-0112.

Kuiken, F., & Vedder, I. (2011). Task performance in L2 writing and speaking: The effect of mode. In P. Robinson (Ed.) (pp. 91–104). https://doi.org/10.1075/tblt.2.09ch4.

Kuiken, F., & Vedder, I. (2012a). Syntactic complexity, lexical variation and accuracy as a function of task complexity and proficiency level in L2 writing and speaking. In A. Housen, F. Kuiken & I. Vedder (Eds.) (pp. 143–170). https://doi.org/10.1075/lllt.32.07kui.

Kuiken, F., & Vedder, I. (2012b). Speaking and writing tasks and their effects on second language performance. In S. Gass & A. Mackey (Eds.), *The Routledge Handbook of Second Language Acquisition* (pp. 364–377). London: Routledge. https://doi.org/10.4324/9780203808184.ch22.

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(2)*, 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(2)*, 161–170. https://doi.org/10.1111/ijal.12255.

Kyle, K. (2016). *Measuring Syntactic Development in L2 writing: Fine Grained Indices of Syntactic Complexity and Usage-Based Indices of Syntactic Sophistication.* Doctoral dissertation. Georgia State University. Retrieved from https://scholarworks.gsu.edu/alesl_diss/35.

Kyle, K., & Crossley, S. A. (2018). Measuring syntactic complexity in L2 writing using fine-grained clausal and phrasal indices. *The Modern Language Journal, 102(2)*, 333–349. https://doi.org/10.1111/modl.12468.

Lahuerta Martínez, A. C. (2018). Analysis of syntactic complexity in secondary education EFL writers at different proficiency levels. *Assessing Writing, 35*, 1–11. https://doi.org/10.1016/j.asw.2017.11.002.

Lambert, C., & Kormos, J. (2014). Complexity, accuracy, and fluency in task-based L2 research: Toward more developmentally based measures of Second Language Acquisition. *Applied Linguistics, 35*, 607–614. https://doi.org/10.1093/applin/amu047.

Lambert, C., & Nakamura, S. (2019). Proficiency-related variation in syntactic complexity: A study of English L1 and L2 oral descriptive discourse. *International Journal of Applied Linguistics, 29(2)*, 248–264. https://doi.org/10.1111/ijal.12224.

Lan, G., Liu, Q., & Staples, S. (2019). Grammatical complexity: 'What does it mean' and 'So what' for L2 writing classrooms? *Journal of Second Language Writing, 46*, 100673. https://doi.org/10.1016/j.jslw.2019.100673.

Lan, G., Lucas, K., & Sun, Y. (2019). Does L2 writing proficiency influence noun phrase complexity? A case analysis of argumentative essays written by Chinese students in a first-year composition course. *System, 85*, 102116. https://doi.org/10.1016/j.system.2019.102116.

Larsen-Freeman, D. (1997) Chaos/complexity science and Second Language Acquisition. *Applied Linguistics, 18(2)*, 141–165. http://dx.doi.org/10.1093/applin/18.2.141.

Larsen-Freeman, D. (2006). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied Linguistics, 27(4)*, 590–619, https://doi.org/10.1093/applin/aml029.

Larsen-Freeman, D. (2019). On language learner agency: A Complex Dynamic Systems Theory perspective. *The Modern Language Journal, 103(S)*, 61–79. https://doi.org/10.1111/modl.12536.

Larson-Hall, J. (2016). *A Guide to Doing Statistics in Second Language Research Using SPSS and R* (second edition). New York: Routledge. https://doi.org/10.4324/9781315775661.

Larsson, T., & Kaatari, H. (2020). Syntactic complexity across registers: Investigating (in)formality in second-language writing. *Journal of English for Academic Purposes, 45*, 100850. https://doi.org/10.1016/j.jeap.2020.100850.

Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning, 40*, 387–417. http://dx.doi.org/10.1111/j.1467-1770.1990.tb00669.x.

Leonard, K. R., & Shea, C. E. (2017). L2 speaking development during study abroad: Fluency, accuracy, complexity, and underlying cognitive factors. *The Modern Language Journal, 101*, 179–193. https://doi.org/10.1111/modl.12382.

Lesonen, S., Steinkrauss, R., Suni, M., & Verspoor, M. (2020). Dynamic usage-based principles in the development of L2 Finnish evaluative constructions. *Applied Linguistics, 42(3)*, 442–472. https://doi.org/10.1093/applin/amaa030.

Levkina, M., & Gilabert, R. (2012). The effects of cognitive task complexity on L2 oral production. In A. Housen, F. Kuiken & I. Vedder (Eds.) (pp. 171–198). https://doi.org/10.1075/lllt.32.08lev.

Lindgren, J. (2020). Fundamentet i muntliga berättelser hos vuxna samt utvecklingen hos barn från 4 till 7 år [Fundament in spoken narratives by adults and development in production by children from the age of 4 to 7]. In S. Haapamäki, L. Forsman & L. Huldén (Eds.), *Svenskans beskrivning 37*. Turku: Skrifter från Svenska institutionen vid Åbo Akademi 8. Available online.

Lintunen, P., Mutta, M., & Peltonen, P. (2020). Synthesising approaches to second language fluency: Implications and future directions. In P. Lintunen, M. Mutta & P. Peltonen (Eds.), *Fluency in L2 Learning and Use.* (pp. 186–201). Bristol: Multilingual Matters. https://doi.org/10.21832/9781788926317-014.

Lintunen, P., & Mäkilä, M. (2014). Measuring syntactic complexity in spoken and written learner language: Comparing the incomparable? *Research in Language, 12(4)*, 377–399. https://doi.org/10.1515/rela-2015-0005.

Littlemore, J., & Juchem-Grundmann, C. (2010). Introduction to the interplay between cognitive linguistics and second language learning and teaching. *AILA Review*, *23(1)*, 1–6. https://doi.org/10.1075/aila.23.01lit.

Lowie, W. (2013). Dynamic Systems Theory approaches to Second Language Acquisition. In C. A. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics*. Chichester, England: Wiley Blackwell. https://doi.org/10.1002/9781405198431.wbeal0346.

Lowie, W., Michel, M., Keijzer, M., & Steinkrauss, R. (Eds.) (2020). *Usage-Based Dynamics in Second Language Development*. Bristol: Multilingual Matters. https://doi.org/10.21832/9781788925259.

Lowie, W., & Verspoor M. (2015). Variability and variation in Second Language Acquisition orders: A dynamic re-evaluation. *Language Learning, 65(1)*, 63–88. https://doi.org/10.1111/lang.12093.

Lowie, W., & Verspoor M. (2019). Individual differences and the ergodicity problem. *Language Learning*, *69(S1)*, 184–206. https://doi.org/10.1111/lang.12324.

MacWhinney, B. (2015). Multidimensional SLA. In T. Cadierno & S. W. Eskildsen (Eds.) (pp. 19–48). https://doi.org/10.1515/9783110378528-004.

Maeda, H. (2021). Speaking performance among high school students at different CEFR levels: A comparative study on the TEAP speaking section. *Annual Review of English Language Education in Japan, 32*, 38–48.

Michel, M. (2017). Complexity, accuracy and fluency (CAF). In S. Loewen & M. Sato (Eds.), *The Routledge Handbook of Instructed Second Language Acquisition*. London: Routledge. https://doi.org/10.4324/9781315676968-4.

Michel, M. C., Kuiken, F., & Vedder, I. (2007). The influence of complexity in monologic versus dialogic tasks in Dutch L2. *International Review of Applied Linguistics in Language Teaching*, *45(3)*, 241–259. https://doi.org/10.1515/IRAL.2007.011.

Miestamo, M., Sinnemäki, K., & Karlsson, F. (Eds.) (2008). *Language Complexity: Typology, Contact, Change.* Amsterdam: John Benjamins.

Miller, J., & Fernandez-Vest, M. M. J. (2006). Spoken and written language. In G. Bernini & M. L. Schwartz (Eds.), *Pragmatic Organization of Discourse in the Languages of Europe* (pp. 9–66). Berlin: De Gruyter Mouton. https://doi.org/10.1515/9783110892222.9.

Ministry of Education and Culture (2018). *Finnish Education in a Nutshell*. Retrieved from https://www.oph.fi/en/statistics-and-publications/publications/finnish-education-nutshell.

Morgan, D. L. (2018). After triangulation, what next? *Journal of Mixed Methods Research, 13(1)*, 6–11. https://doi.org/10.1177/1558689818780596.

Murakami, A. (2020). On the sample size required to identify the longitudinal L2 development of complexity and accuracy indices. In W. Lowie, M. Michel, M. Keijzer & R. Steinkrauss (Eds.) (pp. 29–49). https://doi.org/10.21832/9781788925259-005.

Myles, F. (2012). Complexity, accuracy and fluency: The role played by formulaic sequences in early interlanguage development. In A. Housen, F. Kuiken & I. Vedder (Eds.) (pp. 71–94). https://doi.org/10.1075/lllt.32.04myl.

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*, 555–578. https://doi.org/10.1093/applin/amp044.

Olkkonen, S., Mutta, M., & Lintunen, P. (2024). Fluency profiles in L2 monologue production: The role of underlying cognitive factors. *AFinLA-Teema, 17*, 167–190. https://doi.org/10.30660/afinla.136716.

Onnis, L. (2012). The potential contribution of statistical learning to Second Language Acquisition. In P. Rebuschat & J. N. Williams (Eds.) (pp. 203–236). https://doi.org/10.1515/9781934078242.203.

OPH (2014) = Opetushallitus [Ministry of Education]. Opetussuunnitelman perusteet [National Core Curricula for Education]. Retrieved from https://www.oph.fi/fi/koulutus-ja-tutkinnot/perusopetuksen-opetussuunnitelman-perusteet. Accessed 13 January 2024.

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.

Ortega, L. (2008). *Understanding Second Language Acquisition*. Abingdon, Oxon / New York: Routledge. https://doi.org/10.4324/9780203777282.

Ortega, L. (2013). SLA for the 21st century: Disciplinary progress, transdisciplinary relevance, and the bi/multilingual turn. *Language Learning, 63(S1)*, 1–24. https://doi.org/10.1111/j.1467-9922.2012.00735.x.

Ortega, L. (2014). Second language learning explained? SLA across 10 contemporary theories. In B. VanPatten & J. N. Williams (Eds.) (pp. 245–272).

Ortega, L. (2018). Second Language Acquisition as a road to bilingualism. In A. De Houwer & L. Ortega (Eds.), *The Cambridge Handbook of Bilingualism* (pp. 408–434). Cambridge: Cambridge University Press. https://doi.org/10.1017/9781316831922.022.

Ortega, L., Tyler, A. E., Park, H. I., & Uno, M. (Eds.) (2016). *The Usage-based Study of Language Learning and Multilingualism*. Washington, DC: Georgetown University Press. https://doi.org/10.1353/book45841.

Paavilainen, M. (2015). *Inlärning och behärskning av svenskans verb- och adjektivböjning samt negationens placering hos finska grundskoleelever* [Finnish primary school pupils' learning and mastery of verb and adjective conjugation, and placement of negation in Swedish] Doctoral dissertation. University of Jyväskylä. Retrieved from https://jyx.jyu.fi/handle/123456789/46800.

Pallotti, G. (2009). CAF: Defining, refining and differentiating constructs. *Applied Linguistics*, *30(4)*, 590–601. https://doi.org/10.1093/applin/amp045.

Pallotti, G. (2015). A simple view of linguistic complexity. *Second Language Research, 31*, 117–134. https://doi.org/10.1177/0267658314536435.

Peltonen, P. (2020). *Individual and Interactional Speech Fluency in L2 English from a Problem-solving Perspective: A Mixed-methods Approach*. Doctoral dissertation. University of Turku. Retrieved from https://www.utupub.fi/handle/10024/150313.

Pienemann, M. (1998). *Language Processing and Second Language Development: Processability theory* (Studies in Bilingualism 15). Amsterdam: John Benjamins.

Pienemann, M., & Håkansson, G. (1999). A unified approach towards the development of Swedish as L2: A processability account. *Studies in Second Language Acquisition, 21(3)*, 383−420.

Pienemann, M., & Lenzing, A. (2014). Processability Theory. In B. VanPatten & J. N. Williams (Eds.) (pp. 75–93).

Polat, B., & Kim, Y. (2014). Dynamics of complexity and accuracy: A longitudinal case study of advanced untutored development. *Applied Linguistics*, *35(2)*, 184–207. https://doi.org/10.1093/applin/amt013.

Rahkonen, M., & Håkansson, G. (2008). Production of written L2 Swedish – processability or input frequencies. In H-J. Kessler (Ed.), *Processability Approaches to Second Language Development and Second Language Learning* (pp. 135–161). Cambridge: Cambridge Scholars Publishing.

Rebuschat, P., & Williams, J. N. (Eds.) (2012). *Statistical Learning and Language Acquisition* (Studies in Second and Foreign Language Education 1). Berlin: De Gruyter Mouton. https://doi.org/10.1515/9781934078242.

Rebuschat, P., & Williams, J. N. (2012). Introduction: Statistical learning and language acquisition. In P. Rebuschat & J. N. Williams (Eds.) (pp. 1–12). https://doi.org/10.1515/9781934078242.1.

Robinson, P. (Ed.) (2001). *Cognition and Second Language Instruction*. Cambridge: Cambridge University Press. https://doi.org/10.1017/cbo9781139524780.

Robinson, P. (2001). Task complexity, cognitive resources, and syllabus design: A triadic framework for examining task influences on SLA. In P. Robinson (Ed.) (pp. 287–318). https://doi.org/10.1017/CBO9781139524780.012.

Robinson, P. (2005). Cognitive complexity and task sequencing: Studies in a componential framework for second language task design. *International Review of Applied Linguistics in Language Teaching*, *43(1)*, 1–32. https://doi.org/10.1515/iral.2005.43.1.1.

Robinson, P. (2007). Task complexity, theory of mind, and intentional reasoning: Effects on L2 speech production, interaction, uptake and perceptions of task difficulty. *International Review of Applied Linguistics in Language Teaching, 45(3)*, 193–213. https://doi.org/10.1515/iral.2007.009.

Robinson, P. (Ed.) (2011). Second Language Task Complexity: Researching the Cognition Hypothesis of Language Learning and Performance (Task-Based Language Teaching 2). Amsterdam: John Benjamins. https://doi.org/10.1075/tblt.2.

Robinson, P., & Gilabert, R. (2007). Task complexity, the Cognition Hypothesis and second language learning and performance. *International Review of Applied Linguistics in Language Teaching*, *45(3)*, 161–177. https://doi.org/10.1515/iral.2007.007.

Rosmawati, R. (2020). Profiling the dynamic changes of syntactic complexity in L2 academic writing. In G. Fogal & M. Verspoor (Eds.) (pp. 109–131). https://doi.org/10.1075/lllt.54.05ros.

Saadat, M., & Alavi, S. Z. (2017). Inter-individual variability in CAF: A case study of two individuals and two pairs' written productions. *International Journal of English Language & Translation Studies*, *5(2)*, 61–74.

SAG 1 = Teleman, U., Hellberg, S., Andersson, E. & Christensen, L. (1999). *Svenska Akademiens grammatik 1, Inledning. Register.* [Introduction. Register]. Stockholm: Svenska Akademien.

SAG 3 = Teleman, U., Hellberg, S., Andersson, E. & Christensen, L. (1999). *Svenska Akademiens grammatik 3, Fraser.* [Phrases]. Stockholm: Svenska Akademien.

SAG 4 = Teleman, U., Hellberg, S., Andersson, E. & Christensen, L. (1999). *Svenska Akademiens grammatik 4, Satser och meningar.* [Clauses and sentences]. Stockholm: Svenska Akademien.

Salimi, A., & Dadashpour, S. (2012). Task complexity and language production dilemmas (Robinson's Cognition Hypothesis vs. Skehan's Trade-off Model). *Procedia - Social and Behavioral Sciences, 46*, 643–652. https://doi.org/10.1016/j.sbspro.2012.05.177.

Sarte, K. M., & Gnevsheva, K. (2022). Noun phrasal complexity in ESL written essays under a constructed-response task: Examining proficiency and topic effects. *Assessing Writing*, *51*, 100595. https://doi.org/10.1016/j.asw.2021.100595.

Segalowitz, N. (2010). *Cognitive Bases of Second Language Fluency*. New York: Routledge. https://doi.org/10.4324/9780203851357.

Segalowitz, N., & Lightbown, P. M. (1999). Psycholinguistic approaches to SLA. *Annual Review of Applied Linguistics*, *19*, 43–63. https://doi.org/10.1017/s0267190599190032.

Silva, M. L., Sánchez Abchi, V., & Borzone, A. (2010). Subordinated clauses usage and assessment of syntactic maturity: A comparison of oral and written retellings in beginning writers. *Journal of Writing Research*, *2(1)*, 47–64. https://doi.org/10.17239/jowr-2010.02.01.2.

Skehan, P. (1998). *A Cognitive Approach to Language Learning*. Oxford: Oxford University Press.

Skehan, P. (2003). Task-based instruction. *Language Teaching, 36(1)*, 1–14. https://doi.org/10.1017/s026144480200188x.

Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics*, *30(4)*, 510–532. https://doi.org/10.1093/applin/amp047.

Skehan, P., & Foster, P. (1997). Task type and task processing conditions as influences on foreign language performance. *Language Teaching Research*, *1(3)*, 185–211. https://doi.org/10.1177/136216889700100302.

Skehan, P., & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language Learning*, *49(01)*, 93–120. http://dx.doi.org/10.1111/1467-9922.00071.

Skehan, P., & Foster, P. (2001). Cognition and tasks. In P. Robinson (Ed.) (pp. 183–205). https://doi.org/10.1017/cbo9781139524780.009.

Skehan, P., & Foster, P. (2012). Complexity, accuracy, fluency and lexis in task-based performance: A synthesis of the Ealing research. In A. Housen, F. Kuiken & I. Vedder (Eds.) (pp. 199–220). https://doi.org/10.1075/lllt.32.09fos.

Skrzypczak, W. (2021). Ecological validity in the study of language. *Theoria et Historia Scientiarum, 17*, 95–127. https://doi.org/10.12775/ths.2020.004.

Spoelman, M., & Verspoor, M. (2010). Dynamic patterns in development of 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.

Strömqvist, S., Holmqvist, K., Johansson, V., Karlsson, H., & Wengelin, Å. (2006). What keystroke logging can reveal about writing. In K. Sullivan & E. Lindgren (Eds.), *Computer Key-Stroke Logging and Writing* (Studies in Writing 18) (pp. 45–71). Leiden, The Netherlands: Brill. https://doi.org/10.1163/9780080460932_005.

Tagarelli, K. M., Ruiz, S., Vega, J. L. M., & Rebuschat, P. (2016). Variability in second language learning: The roles of individual differences, learning conditions, and linguistic complexity. *Studies in Second Language Acquisition, 38(2)*, 293–316. https://doi.org/10.1017/s0272263116000036.

Tanskanen, S.-K. (2006). *Collaborating Towards Coherence: Lexical Cohesion in English Discourse*. Amsterdam: John Benjamins. https://doi.org/10.1075/pbns.146.

Tavakoli, P. (2014). Storyline complexity and syntactic complexity in writing and speaking tasks. In H. Byrnes & R. M. Manchón (Eds.) (pp. 217–236). https://doi.org/10.1075/tblt.7.09tav.

Teleman U., Hellberg S., & Andersson E. (2001). *Inledning till grammatiken* [Introduction to grammar]. Stockholm: Svenska Akademien.

TENK (2023) = Tutkimuseettinen neuvottelukunta [Finnish National Board on Research Integrity]. The Finnish Code of Conduct for Research Integrity (PDF). Retrieved from https://tenk.fi/en/research-integrity-ri. Accessed 13 January 2023.

Thompson, S. P., & Newport, E. L. (2007). Statistical learning of syntax: The role of transitional probability. *Language Learning and Development*, *3*, 1–42. https://doi.org/10.1080/15475440709336999.

Tomasello, M. (2003). *Constructing a Language: A Usage-based Theory of Language Acquisition*. Cambridge: Harvard University Press. https://doi.org/10.2307/j.ctv26070v8.

Tonkyn, A. P. (2012). Measuring and perceiving changes in oral complexity, accuracy and fluency. In A. Housen, F. Kuiken & I. Vedder (Eds.) (pp. 221–244). https://doi.org/10.1075/lllt.32.10ton.

Tähtinen J., & Isoaho H. (2001). *Tilastollisen analyysin lähtökohtia. Ensiaskeleet kvantiaineiston käsittelyyn, analyysiin ja tulkintaan SPSS-ympäristössä* [Starting points for statistical analysis. First steps in processing, analysing, and interpreting quantitative data in SPSS]. Turku: University of Turku.

Van Dijk, M., Verspoor, M., & Lowie, W. (2011). Variability and DST. In M. Verspoor, K. de Bot & W. Lowie (Eds.) (pp. 55–84). https://doi.org/10.1075/lllt.29.04van.

VanPatten, B., & Williams, J. N. (Eds.) (2014). *Theories in Second Language Acquisition: An Introduction* (second edition). New York: Routledge. https://doi.org/10.4324/9780203628942.

Van Waes, L., & Schellens, P. J. (2003). Writing profiles: The effect of the writing mode on pausing and revision patterns of experienced writers. *Journal of Pragmatics*, *35(6)*, 829–853. https://doi.org/10.1016/s0378-2166(02)00121-2.

Vercellotti, M. L. (2019). Finding variation: Assessing the development of syntactic complexity in ESL Speech. *International Journal of Applied Linguistics, 29(2)*, 233–247. https://doi.org/10.1111/ijal.12225.

Verspoor, M., & Behrens, H. (2011). Dynamic Systems Theory and a usage-based approach to second language development. In M. Verspoor, K. de Bot & W. Lowie (Eds.) (pp. 25–38). https://doi.org/10.1075/lllt.29.02ver.

Verspoor, M., de Bot, K. & Lowie, W. (Eds.) (2011). *A Dynamic Approach to Second Language Development: Methods and Techniques* (Language Learning and Teaching 29). Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.29.

Verspoor, M., Lowie, W. & Van Dijk, M. (2008). Variability in second language development from a dynamic systems perspective. *The Modern Language Journal*, *92(2)*, 214–231. https://doi.org/10.1111/j.1540-4781.2008.00715.x.

Verspoor, M., & Phuong Nguyen, H. T. (2015). A dynamic usage-based approach to second language teaching. In T. Cadierno & S. W. Eskildsen (Eds.) (pp. 305–328). https://doi.org/10.1515/9783110378528-014.

Vyatkina, N. (2012). The development of second language writing complexity in groups and individuals: A longitudinal learner corpus study. *The Modern Language Journal*, *96(4)*, 576–598. https://doi.org/10.1111/j.1540-4781.2012.01401.x.

Vyatkina, N. (2013). Specific syntactic complexity: Developmental profiling of individuals based on an annotated learner corpus. *The Modern Language Journal, 97(S1)*, 11–30. https://doi.org/10.1111/j.1540-4781.2012.01421.x.

Vyatkina, N., Hirschmann, H., & Golcher, F. (2015). Syntactic modification at early stages of L2 German writing development: A longitudinal learner corpus study. *Journal of Second Language Writing, 29*, 28–50. https://doi.org/10.1016/j.jslw.2015.06.006.

Weinert, R., Basterrechea, M., & García Mayo, M. P. (2013). Investigating L2 spoken syntax. In M. P. García Mayo, M. J. Gutiérrez Mangado & M. Martínez-Adrián (Eds.) (pp. 153–176). https://doi.org/10.1075/aals.9.11ch8.

Weissberg, B. (2000). Developmental relationships in the acquisition of English syntax: Writing vs. speech. *Learning and Instruction, 10(1)*, 37–53. https://doi.org/10.1016/s0959-4752(99)00017-1.

Wijers, M. (2019). *Subordinate Clauses in Swedish as a Foreign Language: A Longitudinal Study*. Doctoral dissertation. University of Ghent. Retrieved from https://biblio.ugent.be/publication/8612221.

Williams, J. N., & Rebuschat, Patrick (2012). Statistical learning and syntax: What can be learned, and what difference does meaning make? In P. Rebuschat & J. N. Williams (Eds.) (pp. 237–264). https://doi.org/10.1515/9781934078242.237.

Wolfe-Quintero, K., Inagaki, S., & Kim, H.-Y. (1998). *Second Language Development in Writing: Measures of Fluency, Accuracy, and Complexity*. Honolulu: University of Hawai'i, Second Language Teaching & Curriculum Center.

Wulff, S., & Ellis, N. C. (2018). Usage-based approaches to SLA. In D. Miller, F. Bayram, J. Rothman, & L. Serratrice (Eds.), *Bilingual Cognition and Language: The State of the Science across Its Subfields* (Studies in Bilingualism 54) (pp. 37–56). Amsterdam: John Benjamins. https://doi.org/10.1075/sibil.54.03wul.

Yamaguchi, Y. (2018). Developmental stages and the CEFR levels in foreign language learners' speaking and writing. *Studies in English Language Teaching*, *7(1)*. https://doi.org/10.22158/selt.v7n1p1.

Yuan, F., & Ellis, R. (2003). The effects of pre-task planning and on-line planning on fluency, complexity and accuracy in L2 monologic oral production. *Applied Linguistics, 24(1)*, 1–27. https://doi.org/10.1093/applin/24.1.1.

Zhang B. (2013). An analysis of spoken language and written language and how they affect English language learning and teaching. *Journal of Language Teaching and Research, 4(4)*, 834–838. https://doi.org/10.4304/jltr.4.4.834-838.

Åberg, A.-M. (2020). *Effekten av explicit undervisning på inlärning av explicit och implicit kunskap om ordföljden i svenska* [The effect of explicit instruction on learning explicit and implicit knowledge of word order in Swedish]. Doctoral dissertation. University of Turku. Retrieved from https://www.utupub.fi/handle/10024/148971.

§ 424/2003. The Language Skills Act. Retrieved from https://www.finlex.fi/en/laki/kaannokset/2003/en20030424. Accessed 13 January 2024.

Appendices

Appendix 1. Tasks with cartoon strips and instructions.

Appendix 1A. Task in the first written production.
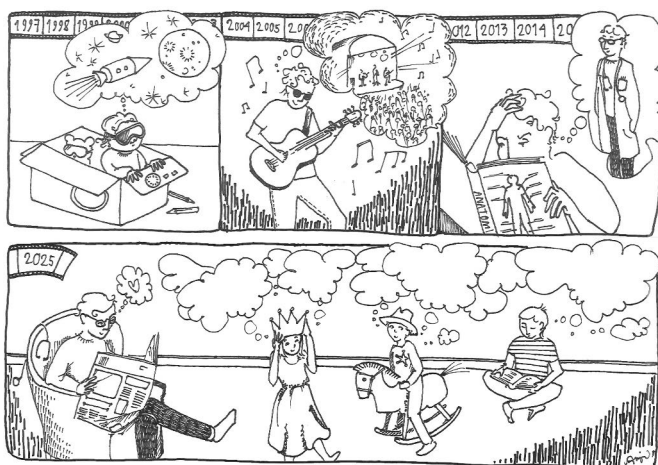
Ammatilliset haaveet (= Professional dreams)
1. Kerro, mitä kuvasarjassa tapahtuu. (= Describe the actions in the cartoon strip.)
2. Mikä on sinun unelmatyösi? Mistä haaveilit, kun olit lapsi? Missä näet itsesi kymmenen vuoden kuluttua? (= What is your dream job? What did you dream about as a child? Where do you see yourself in ten years?)
3. Pohdi, miten ja miksi ammatilliset haaveet muuttuvat. (= Discuss how and why the professional dreams change.)
Pyri tuottamaan tekstiä mahdollisimman paljon, vähintään 100–150 sanaa. Ylärajaa ei ole. (= Try to produce as much text as possible, at least 100–150 words. There is no upper limit.)

Tukisanoja: (= word list)
aikuinen = vuxen (adult)
ammatti = ett yrke (-t, -n, -na) (profession)
astronautti = en astronaut (-en, -er, -erna) (astronaut)
avaruus = en rymd (-en, -er, -erna) (space)
koulutus = en utbildning (-en, -ar, -arna) (education)
kuuluisa = berömd (-t, -da) (famous)
lapsi = ett barn (-et, - , -en) (child)

lääkäri = en läkare (-n, - , -arna) (doctor)
muusikko = en musiker (-n, - , -na) (musician)
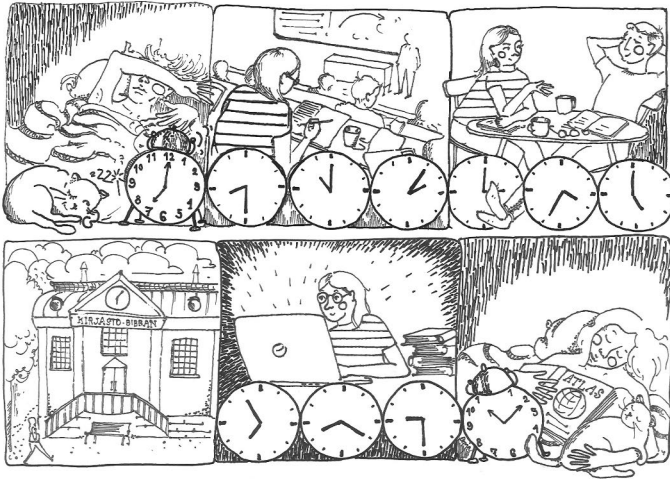opiskella = studera (-ar, -ade, -at) (to study)
palkka = en lön (-en, -er, -erna) (salary)
suunnitella = planera (-ar, -ade, -at) (to plan)
tulevaisuus = framtid (-en) (future)
unelmoida = drömma (-er, -de, -t) (to dream)

Appendix 1B. Task in the spoken production.

Opiskelijan päivä (= Day of a student)
1. Kerro, mitä kuvasarjassa tapahtuu. (= Describe the actions in the cartoon strip.)
2. Kerro, millainen oma arkipäiväsi on. Mistä päiväsi koostuu? (= Explain what your everyday life is like. What does your day consist of?)
3. Pohdi opiskelijaelämää ja opintojen kuormittavuutta. (= Reflect on life as a student and the burden of studying.)
Pyri tuottamaan puhetta mahdollisimman paljon. (= Try to produce as much speech as possible.)

Tukisanoja: (= word list)

harrastus = en hoddy (-n, -er, -erna) (hobby)

kirjasto = ett bibliotek (-et, - , -en) (library)

kirjoittaa esseetä = skriva (skriver, skrev, skrivit) essä (to write an essay)

lainata kirjoja = låna (-ar, -ade, -at) böcker (to loan books)

luento = en föreläsning (-en, -ar, -arna) (lecture)

lukupiiri = en läsecirkel (-n, -ar, -arna) (reading circle)opintopiste = en studiepoäng (-en / -et, - , -en) (study point)

opiskelija = en studerande (-n, - , -na) / en student, (-en, -er, -erna) (student)

pienryhmä = en smågrupp (-en, -er, -erna) (small group)

päntätä = plugga (-ar, -ade, -at) (to study)

tehdä kotitehtäviä = göra (gör, gjorde, gjort) hemuppgifter (to do homework)

tentti = en tentamen (- , tentamina, tentamina) (exam)

vapaa-aika = fritid (-en) (free time)

Appendix 1C. Task in the second written production.
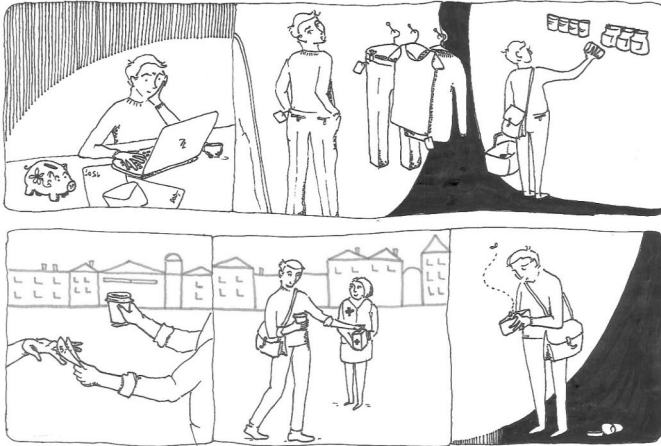
Rahan käyttöä ja säästöä (= To spend and save money)

1. Kerro, mitä kuvasarjassa tapahtuu. (= Describe the actions in the cartoon strip.)

2. Kerro omista tottumuksistasi rahankäytön ja säästämisen suhteen. Mihin käytät rahaa? Säästätkö? (= Explain your own habits when it comes to spending and saving money. What do you spend money on? Do you save money?)

3. Pohdi opiskelijoiden taloudellista tilannetta. Riittääkö opintotuki? Pitäisikö opintojen ohella voida tehdä töitä? (= Reflect on the financial situation of students. Is the financial aid for students enough? Should students be able to work while studying?)

Pyri tuottamaan tekstiä mahdollisimman paljon, vähintään 150 sanaa. Ylärajaa ei ole. (= Try to produce as much text as possible, at least 150 words. There is no upper limit.)

Tukisanoja: (= word list)

ansaita rahaa = tjäna (-ar, -ade, -at) pengar (to earn money)

antaa rahaa keräykseen = skänka (-er, -te, -t) pengar till insamlingen (to donate money to a fundraiser)

käydä ostoksilla = handla (-ar, -ade, -at) (to shop)

lompakko = en plånbok (-en, -böcker, -erna) (wallet)

luottokortti = ett kreditkort (-et, - , -en) (credit card)

maksaa laskuja = betala (-ar, -ade, -at) räkningar (to pay bills)

menot = utgifter (expenses)

opintolaina = ett studielån (-et, - , -en) (student loan)

opintotuki = ett studiestöd (-et, - , -en) (financial aid for students)

osa-aikatyö = ett deltidsarbete (-t, -n, -na) (part-time job)

ostaa luotolla = köpa (-er, -te, -t) på kredit (to buy on credit)

säästää rahaa = spara (-ar, -ade, -at) pengar (to save money)

toimeentulo = försörjning (-en) / uppehälle (-t) (livelihood)

tulot = inkomster (income)

vuokra = en hyra (-an, -or, -orna) (rent)

Appendix 2. Assessment criteria.

Appendix 2A. Criteria for written production (based on CEFR 2007: 63–64, 107, 109).

Nivå (= Level)	Deskriptorer (= Descriptions)
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 (= *Can write clear and fluent texts with an effective style, using an appropriate structure that helps the reader to understand the main points, has a wide linguistic repertoire [including colloquial and idiomatic expressions] and uses it to formulate ideas precisely, can emphasise certain points and bring out nuances*)
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 (= *Can write clear, well-structured texts, developing and supporting points of view in detail with supporting arguments and concluding appropriately, has a wide linguistic repertoire, can rephrase (if necessary) and does not need to limit what he/she wants to say*)
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 (= *Can write clear, detailed texts and descriptions, presenting ideas clearly in a coherent text, can argue thoughtfully, emphasising important points and providing relevant supporting facts, can explain the advantages and disadvantages of different positions, has a wide vocabulary on most general topics and can vary the wording to avoid repetition*)
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) (= *Can write coherent texts by putting together a series of elements in a logical sequence, can describe experiences and feelings, can write a detailed description of the cartoon strip; has a sufficiently large vocabulary to write about most everyday topics [e.g., family, hobbies, interests, work and current events]*)
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 (= *Can write a description in the form of separate sequences, can write about things in the everyday environment [e.g., people, places and study experiences], has a sufficiently large vocabulary (also outside the word list) to describe simple situations*)
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 (= *Can write separate sentences about him/herself and what he/she is doing, has a very basic vocabulary related to specific concrete situations [e.g., in the cartoon strip], describes the cartoon strip [mostly] using the words from the word list*)
under A1	

Appendix 2B. Criteria for spoken production (based on CEFR 2007: 60–61, 74, 109).

Nivå *(= Level)*	Deskriptorer *(= Descriptions)*
C2	Kan producera tydligt och flytande tal, kan ge detaljerade och utförliga beskrivningar med en logisk struktur samt argumentera genomtänkt, har ett mycket stort ordförråd, kan formulera tankar exakt, betona vissa saker och få fram nyanser (= *Can produce clear and fluent speech, give detailed and comprehensive descriptions with a logical structure and argue thoughtfully, has a very large vocabulary, can formulate thoughts precisely, emphasise certain points and bring out nuances*)
C1	Kan ge tydliga och detaljerade beskrivningar, utveckla särskilda punkter och avrunda med en lämplig slutsats, har ett stort ordförråd, kan omformulera för att täcka luckor i ordförråd, kan effektivt välja idiomatiska uttryck (= *Can give clear and detailed descriptions, develop specific points and finish with an appropriate conclusion, has a large vocabulary, can paraphrase to fill gaps in vocabulary, can effectively choose idiomatic expressions*)
B2	Kan ge tydliga, systematiskt utformade beskrivningar där centrala punkter betonas på lämpligt sätt, kan utveckla tankar med hjälp av understödjande detaljer och argumentera genomtänkt samt förklara fördelar och nackdelar med olika ställningstaganden, har ett stort ordförråd när det gäller de flesta allmänna ämnen, kan variera formuleringarna för att undvika upprepningar (= *Can give clear, systematic descriptions, emphasising key points appropriately, can develop ideas with supporting details and argue thoughtfully and explain the advantages and disadvantages of different positions, has a wide vocabulary on most general topics, can vary the wording to avoid repetition*)
B1	Kan relativt flytande ge beskrivningar och presentera dem i logisk följd som en linjär sekvens av punkter, kan detaljerat återge handlingen i bildserien och beskriva sina känslor och reaktioner, kan motivera och förklara åsikter och handlingar, har ett tillräckligt stort ordförråd för att uttrycka sig om de flesta vardagliga ämnen (såsom familj, hobbyer, intressen, arbete och aktuella händelser) (= *Can give descriptions relatively fluently and present them in a logical order as a linear sequence, can describe in detail the action in the cartoon strip and describe his/her feelings and reactions, can justify and explain opinions and actions, has a sufficiently large vocabulary to express him/herself on most everyday topics [such as family, hobbies, interests, work and current events]*)
A2	Kan beskriva handlingen i bildserien genom separata punkter, kan förklara vad han/hon (inte) tycker om en sak, har ett tillräckligt stort ordförråd (*också utanför stödordlistan*) för att beskriva enkla situationer (= *Can describe the action in the cartoon strip with the help of separate elements, can explain what he/she [doesn't] think[s] about a thing, has a sufficiently large vocabulary [also outside the word list] to describe simple situations*)
A1	Kan producera isolerade sekvenser av tal, kan berätta om konkreta angelägenheter och beskriva 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 (= *Can produce isolated sequences of speech, can talk about concrete topics and describe what he/she is doing, has a very basic vocabulary related to specific concrete situations [e.g., in the cartoon strip], describes the cartoon strip [mostly] using the word list*)
under A1	

Appendix 3. Symbols used in the transcriptions.

:	(colon)	a stretched sound, syllable or word, e.g., å:, me:d
_	(underline)	an accentuated word or part of a word, e.g., biblio<u>te</u>ket
(.)	(micro pause)	a short pause (under 0.4 seconds)
(x.x)	(pause)	the length of a pause, e.g., (1.2)
öö / mm / ää	(filled pause)	sounds produced during pauses, transcribed accordingly
↑↓	(intonation)	rising and falling intonation
-	(hyphen)	repair or unfinished word, e.g., på- till skolan
> <	(greater than, smaller than)	increase in tempo, e.g., >och efter det<
hh	(exhale)	audible exhalation
hihi	(laughter)	participant laughs or speaks with a smile
krhm	(cough / throat clearing)	participant clears his/her throat or coughs
snif	(sniff)	"sniffing" breaths through nose
pt	(lip smack)	participant smacks his/her lips
?	(inaudible)	hard to hear what is said / uncertainty

**TURUN
YLIOPISTO
UNIVERSITY
OF TURKU**