The complexity of linguistic complexity:

a corpus-based study on development of phrasal complexity in written L2-Swedish

Karin Rasmusson

Supervisors: Anna Flyman Mattsson, Joost van de Weijer
Centre for Language and Literature, Lund University
MA in Language and Linguistics, General Linguistics
SPVR01 Language and Linguistics: Degree Project – Master's (Two Years) Thesis, 30 credits

January 2023

Abstract

The aim of this study was to establish whether or not a development of the complexity of the noun phrase, i.e. an increase in noun phrase elaboration could be observed in written L2 Swedish over time. To provide an insight to said development, three on L2 English previously tested and strongly correlating phrasal complexity measures, MLNP, MNPDep and WNPCx (Bulté 2013), were employed. A fourth measure that measure the mean length of morphemes per noun phrase (MLNPm) was added to the study to investigate if such a measure could present a different picture than mean length of words per noun phrase (MLNP). The data consisted of the written production of ten L2 learners collected within the ASU-corpus. Forty essays, i.e. four essays per learner, were analysed. The written production were all descriptive texts since these were considered to potentially contain a high degree of noun phrases. The author and an experienced rater carried out the analyses and calculations manually in Excel spreadsheets. All four measures correlated strongly, suggesting that they do measure noun phrase complexity also for L2 Swedish. No evidence that MLNPm would provide a different picture than MLNP was found. The two measures followed similar developmental pathways. *T*-tests performed on the first and last measurement point for all four measures provided evidence suggesting that there was a development of NP complexity over time, and the results of the *t*-tests were significant. For half of the learners, the development was linear, and for the other half it was mostly linear from the first to the third measurement point, with the exception of one learner who displayed more non-linear and irregular developmental pathways. A decline from the third to the fourth measurement point was detected. This decline could be due to different reasons. One being that the learners no longer had any L2 instruction at the fourth measurement point. However, since Bulté (2013) also found a decline in his study, where the learners still had instructions, this explanation does not seem to hold. That the learners resorted to other constructions, i.e. what Verspoor et al. (2008) and Spoelman and Verspoor (2010) refer to as a 'competitive relationship' between noun phrase complexity and sentence complexity (e.g. subordination), seems more likely, and would be interesting to further examine in a multidimensional complexity study in order to account for all aspects of the development of complexity.

Keywords: Second language development, Complexity, Accuracy, and Fluency, Linguistic complexity, Syntactic complexity, Phrasal complexity, Noun phrase complexity, Morphological complexity

Acknowledgement

First and foremost I would like to express my absolute deepest gratitude to my main supervisor Anna Flyman Mattson, who has been more than patient and who has kept encouraging me throughout the process of this thesis, as well as for my presentation of the study at the World Congress of Applied Linguistics, AILA2021. Thank you for always making time and being available for discussion and guidance, and for your contagious passion for grammar and language acquisition that is beyond this world.

A special thank you is directed to my co-supervisor Joost van de Weijer for conducting the statistical analysis and taking time to take on the enormous task of explaining statistics to me. You are a brilliant and brave man in doing so.

I am deeply grateful to, and would like to send an enormous thank you to Bram Bulté for invaluable input and discussions on complexity. Another special thank you is directed towards Alex Housen for valuable input regarding the exploration of the morphology of the noun phrase.

To my sister Anna Rasmusson, thank you for still liking me and for all the proofreading I put you trough over the years. I would also like to thank Lotta Johnsson for intense grammatical discussions and encouragement, to Sara Skjutare for Excel support, and to Anders Agebjörn for helping me grasp some of the statistical terminology.

A huge thank you to my mother and my friends for love and support.

Last but not least, I would like to thank all the friends I have met at SOL over the years.

To my father Hans-Åke Rasmusson

(1945-2020)

Table of contents

List of figures

List of tables

Abbreviations

CAF	Complexity, Accuracy, Fluency
CEFR	Common European Framework of Reference for Languages
CDST	Complex Dynamic Systems Theory
Dep	Dependent
DST	Dynamic Systems Theory
IL	Interlanguage
L1	First language
L2	Second language
MLC	Mean length of clause
MLNP	Mean length of noun phrase
MLNPm	Mean length of noun phrase morph
MNPDep	Dependents per noun phrase head
N	Noun
NP	Noun phrase
NS	Native speaker
NNS	Non-native speaker
PP	Prepositional phrase
PT	Processability Theory
SLA	Second language acquisition
TL	Target language
WNPCx	Weighted noun phrase complexity

Chapter 1 Introduction

Complexity has, for quite some time, been of great interest in many different sciences, such as biology, physics, psychology and philosophy, just to mention a few. Complexity has also been frequently studied in the language sciences and has an extended history within second language acquisition (SLA) research. However, the notion of complexity is a complex phenomenon, and one that has been intensely debated, as well as vaguely defined and operationalized. To say the least, complexity does not come problem free. An example can be found within the CEFR, *Common European Framework of Reference for Languages* (2001), used for teaching, learning and assessment of languages. Here the notion of complexity is frequently employed along with its counterpart simple/basic, e.g. complex language, complex material, and simple syntax, but what it actually refers to is not clearly defined (Gyllstad et al. 2014). It is not only within the documents of the CEFR that *complexity* is a notion difficult to define. The construct of linguistic complexity can be examined on different levels, e.g. on an abstract theoretical level, on an observational level, and on an operational level (Bulté & Housen 2012). The distinction between these levels is important to make. On a theoretical level, it has to be established what complexity is, on an observational level, how it is or can be manifested in the L2 performance, and on an operational level, how the manifestation can be captured or quantified (Bulté & Housen 2012:27). Within SLA research, the *complexity*, *accuracy* and *fluency* (CAF) constructs have been debated by scholars throughout the years, mainly due to the large variety of ways in which the constructs have been operationalized, and also the rather vague definitions of the constructs in numerous studies (Bulté & Housen 2012). Another concern has been that studies may employ several measures that measure the same phenomena (Norris & Ortega 2009). These issues have led to confusion, terminological as well as conceptual, and have also led to difficulties in interpreting and comparing results across studies (e.g. Norris & Ortega 2009; Pallotti 2009; Bulté & Housen 2012, 2014; Bulté 2013). Bulté and Housen (2015) state the need for longitudinal studies over more extensive periods of time to establish whether linguistic complexity develops over time. This type of study has been rare within the field of L2 complexity. In an investigation of forty studies conducted within in the period of 1995–2008, Bulté and Housen (2012) showed that subordination is the most popular measure to gauge complexity. However, Ortega (2012) as well as Bulté and Housen (2015) provided evidence of this measure not being fully adequate when measuring complexity in the language production of more advanced learners. Norris and Ortega (2009) suggested that complexity should be

measured multidimensionally, i.e. not only through subordination, but also through coordination at lower levels and through phrasal complexity at advanced levels. This suggestion is supported by the results of Bartning et al. (2015), who examined linguistic complexity at the phrasal level in French. Bulté and Housen (2012, 2014, 2015 and 2018) as well as Bulté (2013), and Spoelman and Verspoor (2010) all took multidimensionality into consideration. Empirical studies of phrasal complexity are needed to measure and capture all dimensions of L2 complexity and the development thereof. Each of the sub-dimensions of the complexity construct can, at least in principle, be measured and/or described independently (Bulté 2013). The fact that structural complexity studies commonly employ measures gauging complexity through subordination, and the scarceness of longitudinal studies and studies on phrasal complexity, suggest that there are many voids still to be filled. To the knowledge of the author, no previous CAF-studies on phrasal complexity regarding development of the noun phrase in written L2 Swedish have so far been presented. Studies are for the most part conducted on L2 English (e.g. Bulté 2013; Bulté & Housen 2014, 2015, 2018; Bulté & Roothooft 2020). However, Bartning et al. (2015), as well as De Clercq and Housen (2017), investigated complexity at the phrasal level in L2 French. Since most studies are conducted on L2 English, it is not surprising that studies on the morphological complexity of the noun phrase are even more scarce. The English noun phrase is rather morphologically poor in comparison to other languages. This might suggest a need to investigate the morphology of the noun phrase complexity as well in languages with a more rich and complex morphology than English. Spoelman and Verspoor (2010) examined noun phrase complexity and word complexity in L2 production of the morphologically rich Finnish language. They studied the morphological complexity, not within the noun phrase, but rather at the word level. De Clercq and Housen (2019) studied the development of morphological complexity cross-linguistically in French and English by investigating verb morphology. In light of this, it could be relevant to investigate the morphology within the noun phrase in Swedish, perhaps especially in regard to the double definiteness.

The present study will make a contribution to begin to fill a void in complexity studies, both on a general level of L2 studies as well as for studies of L2 Swedish, by investigating complexity at the phrasal level, i.e. in this case noun phrase complexity of L2 Swedish. This will be done by investigating the development of noun phrase complexity over time, employing three correlating complexity measures, previously tested on L2 English, as well as a fourth measure – a morphologically directed complexity measure in a longitudinal study of written L2 Swedish

production. The doctoral dissertation of Bulté (2013), as well as the taxonomy of the complexity construct presented in Bulté and Housen (2012), form the basis of the data analysis. The taxonomy, which will be presented in more detail in Chapter 2, is intended to serve as a descriptive-analytic framework for analyses of L2 complexity. The taxonomy is thought to assist researchers in being more specific in what they actually investigate when they state that they investigate L2 complexity.

1.1 Aim and research questions

Bulté and Housen (2014) call for the need to empirically establish if linguistic complexity increases over time, rather than to assume this is the case. The aim of the present study is to establish whether or not a development of the complexity of the noun phrase, i.e. an increase in noun phrase elaboration, can be observed in written L2 Swedish.

To provide an insight to the development of noun phrase complexity in written L2 Swedish production, three phrasal complexity measures employed in Bulté (2013), namely two length measures: MLNP and MNPDep and a weighted measure: WNPCx, will be applied. In Bulté (2013), a positive and strong correlation was found between these measures. The three measures are, however, not tested on L2 Swedish.[1] If the measures also correlate in the present study, it would suggest that the measures are applicable also when investigating L2 Swedish, and that they might be applied to other L2s in general. The phrasal complexity measures employed by Bulté are mainly tested on L2 English. Since Swedish is a morphologically richer language in terms of the noun phrase, it is motivated to also investigate the morphology of the noun phrase. Hence, a fourth measure is constructed and included in the present study, namely a length measure of phrasal complexity through morphology. This measure is referred to as MLNPm. The analysis of the ten L2 learners' written production, and the results of the calculations of the four measures are intended to present an insight to the development of the phrasal complexity both at individual and group levels. Thus, the aim of the thesis will be achieved through the following three research questions:

RQ1. Do the three previously tested, and strongly correlating, measures for L2 English also correlate for L2 Swedish, i.e. do they measure the complexity of the noun phrase in other second languages than English?

[1] MLNP, or at least versions of this measure has been tested in other studies, but then mainly of L2 English.

RQ2. Would a measure that takes the morphology of the noun phrase into consideration present a different picture of noun phrase complexity than the more commonly employed MLNP when gauging complexity in morphologically richer languages, such as Swedish?

RQ3. Is there a development of the complexity of the noun phrase in written L2 Swedish at individual and group level?

The following chapter introduces the theoretical background and previous research related to SLA development, the CAF constructs and complexity, as well as the Swedish noun phrase. The methodology and material are described in the third chapter. In the fourth chapter the results of the analysis are presented and discussed in relation to the research questions and previous findings. Chapter 5 summarizes the findings and suggests directions for future research.

Chapter 2 Theoretical background

2.1 Theories of SLA

In the 1970s, Larry Selinker coined the term *interlanguage* (IL) referring to a unique language system which is different from both the first language (L1) and the target language (TL), but draws on both. This system can be influenced by input as well as by internal processing, and is rather dynamic, i.e. it changes when the learner learns a new rule, deletes a rule, or in any other way decides to restructure his/her system (Selinker 1972). The learner creates new hypotheses regarding the structure of the TL continuously and through these hypotheses develops new grammatical rules. These changes in the interlanguage are not always correct, but are nonetheless signs of L2 development. The interlanguage is both systematic and varied; systematic in that the interlanguage is formed by the rules at hand at that moment, but also that the learner, when it comes to the morphosyntax, follows identifiable stages, and varied in how fast theses stages are obtained and the result thereof. Grammatical developmental stages have since been frequently examined and empirically tested. In the 1970s, a number of researchers performed what came to be known as *the morpheme studies*, e.g. Dulay and Burt (1973, 1974), Bailey, Madden and Krashen (1974) and Larsen-Freeman (1975). These empirical cross-sectional studies examined if a universal acquisition sequence, independent of the learner's L1, could be found in L2 learners, and these studies reflect how developmental stages were first looked upon. The morpheme studies were the starting point of studies on acquisition sequences, but they have, however, also been subject to criticism. The studies have been criticised for comparing morphemes with different functions and for assuming that accuracy reflects the developmental order.

2.1.1 Different approaches to L2 development

One group of researchers argue that the developmental stages are not predictable, but rather a consequence of the individual learner's prerequisites and the context of that learner (e.g. Spoelman & Verspoor 2010; Verspoor, Lowie & Dijk 2008; Verspoor, de Bot & Lowie 2011). However, other researchers, such as Pienemann (e.g. 1998, 2005, 2015), Pienemann and Håkansson (1999) and Baten and Håkansson (2014), are of the opinion that they are predictable and develop in a specific order. One theory within this field of research is the substantially empirically tested processability theory (Pienemann 1998, 2005). Other theories that share the

idea of development being predictable are e.g. the concept-oriented approach, dating back to von Stutterheim and Klein (1987), and the Input Processing (IP) Theory, introduced by VanPatten (1993). The processability theory (PT) is frequently applied within the field of SLA research. PT is a psycholinguistic theory with its origin in cognitive linguistics. The theory is concerned with language development and aims at explaining developmental phenomena of the dynamics of interlanguage. The aim of the processability theory is to establish the order in which the procedural skills are developed in the L2 learner (Pienemann and Håkansson, 1999:386).

The original version (Pienemann 1998) focused on the so-called *developmental problem*, i.e. why language learners follow universal stages of acquisition. The extended version (2005) also address the so-called *logical problem*, i.e. how do learners learn what is not represented in the input they receive. The stages of PT follow a hierarchical order, where a structure of an earlier stage has to be processable to the learner before s/he can process structures at a later stage. Therefore, the development should follow a linear pattern. However, there does occur some variation in the learner language, explained within PT by means of *The Hypothesis Space* (Pienemann 1998). The variation of the production of structures can only occur within the processing stage at which the learner is at that specific moment in time.

Processing Determinism (O'Grady 2015) is another theory where the processing is considered to be the driving force behind language development. The course of development is considered to be a product of processing pressures; what is less costly to process is learnt before structures that require more processing. O'Grady exemplifies it e.g. by the development sequence of the sentential negation of English in children where the sentential negation is initially expressed by *no*, to be followed by preverbal negation, and finally by a negation co-occurring with an auxiliary verb.[2] According to O'Grady (2015:21) "development is uniform, where processing cost is relevant".

As mentioned in the beginning of 2.1.1, other researchers argue that learner language development is based on the prerequisites and context of the individual learner. Researchers within the DST approach apply this view on L2 development. DST, i.e. the Dynamic Systems Theory was first developed as a theory of physics and theoretical mathematics to account for emergent properties of complex systems in nature, and later applied to developmental

[2] O'Grady (2015:7) for more extensive examples.

psychology by Thelen and Smith (1994). Hence, it is not a specific theory for SLA, but rather a general theory explaining change over time in complex systems consisting of variables which interrelate and affect each other (Verspoor & Behrens 2011). Kees de Bot (2017) proposed the label Complex Dynamic Systems Theory (CDST) as the general label since Complex Theory (CT) and DST refer to more or less the same circumstance. Both DST and CDST are found in the SLA literature. Within the DST approach to second language development (SLD), language is viewed as a dynamic and complex system, and one of the basic characteristics of dynamic systems is non-linearity in development (e.g. Verspoor, de Bot and Lowie 2011). The DST approach to SLA is fairly new and can be traced to Larsen-Freeman (1997). When applied to SLA, DST views language development as non-linear and individual, and hence DST researchers also do not agree with development being incremental, stage-bound and universal. The DST approach has not been empirically tested to the degree that PT has, but there are researcher carrying out empirical studies and more research within this approach is on its way. DST researchers claim to find support for language development not being linear, at least not on the level of individual learners. Verspoor, Lowie and Dijk (2008) found that language develops non-linearly and also that there is a dynamic interrelationship to be found between different components of language. These findings were supported by Spoelman and Verspoor (2010).

DST researchers such as de Bot, Verspoor, Spoelman, and Lowie, just to mention a few, search for a comprehensive theory of SLA, and argue that DST can be seen as "a comprehensive theory that can unify and make relevant a number of different 'middle level' theories on Second Language Acquisition (SLA)", given that these theories are in line with DST principle (de Bot et al. 2013:200). Particularly, theories that can fall under the umbrella terms 'usage-based' or 'emergentist' approaches or theories, seem to be compatible with DST thinking. However, Ortega (2013:104–105) points out that how long it will take before we will see the fruit of these theories is difficult to predict.

In the introductory chapter of Verspoor, de Bot and Lowie (2011:2), de Bot et al. discuss the ultimate test of a theory, which would lie in the ability to generate "powerful and testable hypotheses", and continue that whether DST would meet this requirement is yet to be seen because "prediction is not what the dynamic approach is after". Pienemann (2015:140–141) critics the DST approach proposed by de Bot, Lowie and Verspoor, in that it would not have the status of a theory since it cannot be falsified. Baten and Håkansson (2014) stand critical to

the DST claim that development is non-linear. They performed a theoretical and methodological study on PT and DST, and found clear stages in the development of subordinate clauses.

It is quite clear that there are different views on L2 development and whether L2 development follows linear or non-linear developmental trajectories. From an applied language assessment point of view, it is of considerable interest to be able to predict development. To be able to predict development is of great help to language teachers when assessing the proficiency level of learners. PT provides this in many cases, but when considering the noun phrase, one would not expect to see a development since it belongs to the same level. The CAF measures, which are a good complement to usage-based models, differ from PT in that instead of following a predictable linear development, they provide information on the complexity of the learner language at a specific point in time (Flyman Mattsson, 2017).

2.2 Language proficiency and language development in the context of SLA

Language proficiency and language development are two notions, which are intimately connected within an SLA context. L2 proficiency refers to "a learner's skill in using the L2" (Ellis, R. 2015:976). Proficiency can be contrasted with competence, which refers to the knowledge of the L2 internalized by the learner, i.e. the implicit or explicit abstract knowledge of a language. The learner's ability to apply this knowledge in various tasks is what is referred to as proficiency (Ellis, R. 2015). Competence, which is mainly declarative, is hence included in language proficiency, which also incorporates a procedural component (Bulté 2013). While language proficiency refers to a property of learners at a specific point in time, and therefore is a concept which is static and synchronic, development refers to changes in learners' proficiency over a timespan, and is a dynamic and diachronic concept (e.g. Bulté 2013). The developmental changes can be of varying size, and move towards or away from a set goal, such as the proficiency of the norm, e.g. of a native speaker. Generally, development refers to some purely linguistic component of proficiency, such as grammar or phonology. Developmental stages are often decided on the rationale of some specific criteria (Ortega 2012), while terms used to define proficiency levels are more general, such as *beginner*, *intermediate*, or *advanced*.

As in the case of the notion complexity, development has been interpreted and applied in various ways. Furthermore, it has often been interpreted as change in a certain direction, often affiliated with an increase, growth, progress etc, but any kind of change that takes place over

time could pertain to development, in principle (Bulté 2013). Granfeldt and Ågren (2013:29) tentatively define second language development as "the progressive growth of one or more aspects of the interlanguage system (phonology, morphology, syntax, etc.)". Proficiency and development in second language acquisition are by some (e.g. Pienemann & Johnston 1987; Pienemann & Mackey 1993) considered as two separate theoretical constructs, while others, such as Ellis (2008) are of the opinion that perhaps proficiency levels and developmental levels could be matched. Granfeldt and Ågren (2013) investigated the relationship between second language proficiency and second language development empirically in written (L3) French. Their study examined a possible relationship between CEFR, which is employed to assess L2 proficiency, and Pienemann's (1998; 2005) processability theory (PT). The overall correlation they found between the CEFR ratings and the PT analysis was considered to be strong, but they also conclude that the correlation cannot be held as evidence of a relationship between the two theoretical constructs. Moreover, they also found uneven profiles where learners were stronger in their communicative proficiency than in their morphosyntactic development, or the other way around.

L2 production is often used to manifest the more abstract constructs proficiency and development since it is observable, concrete and possible to analyse. However, it is important to consider the fact that what a learner knows and what a learner actually produces is not always isomorphic, at least not when only one sample of the L2 production is analysed (Bulté 2013). The type of learner we are dealing with can also be important to consider in regard to what a learner knows and what a learner produces. Norrby and Håkansson (2007) present four different types of learners: *the Risk-taker*, *the Careful and thorough*, *the Recycler*, and *the Achiever*. The Risk-taker is a daring learner, who is likely to write at a level at which s/he would write in the L1 often resulting in many formal mistakes. The more careful learner produces language with very few mistakes, but still with great syntactic variety. The Recycler uses familiar structures, which lends few challenges and not much progression, while the Achiever is characterised by clear signs of progression. If no development in complexity can be detected, it may not have to do with the learner language, but it could be due to the type of learner. Furthermore, Bulté also issues that it is important to discriminate between L2 proficiency and L2 development as theoretical constructs, L2 production as behavioural constructs, and measures used to assess L2 production as statistical constructs. The manifestation of L2 proficiency and L2 development does not give a comprehensive picture of the level of either of the two constructs at any given

point in time, but rather a snapshot of the proficiency of a learner and the development at specific times.

2.3 Complexity, Accuracy, and Fluency (CAF)

The Complexity, Accuracy, and Fluency (CAF) research traces back to the 1970s, when L2 researchers were looking for metrics of complexity and accuracy to gauge L2 proficiency reliably (e.g. Larsen-Freeman 1978), and when a distinction between the fluency of L2 speech and the accuracy of the L2 usage was made when investigating the communicative proficiency in L2 classrooms (e.g. Brumfit 1979, 1984; Hammerly 1990). The CAF constructs, i.e. *Complexity*, *Accuracy*, and *Fluency*, were first brought together in a proficiency model introduced by Skehan (1996, 1998b) in the 1990s, to which the third dimension, *complexity*, was added. At this time, the three principal proficiency dimensions also received their traditional working definitions (see. Housen & Kuiken 2009; Housen et al. 2012; Bulté 2013):

> *Complexity*: "the extent to which the language produced in performing a task is elaborate and varied" (Ellis, R. 2003:340).
>
> *Accuracy*: "a learner's ability to produce target-like and error-free language" (cf. Ellis, R. 2008, Ellis & Barkhuizen 2005; Skehan 1998b).
>
> *Fluency*: "the ability to process the L2 with 'native-like rapidity'" (Lennon 1990:390), or "the extent to which the language produced in performing a task manifests pausing, hesitation, or reformulation" (Ellis, R. 2003:342).

CAF can be employed to describe the performance of both native speakers (NS) and first language (L1) learners, as well as second language (L2) learners (Pallotti 2009). Through the years, CAF has been used as a descriptor for assessment of language learners' oral and written performance, but also as indicators of their proficiency, and for gauging progress in language learning (Housen & Kuiken 2009). Pallotti (2009) has, in order to be able to determine if the language is adequate for reaching the goals of communication, suggested a complement to the CAF constructs, namely *Adequacy*, which he suggests can be seen both as a dimension of its own, independent from CAF, as well as a means of interpreting CAF measures.

2.3.1 Complexity, Accuracy, and Fluency in SLA research

The research on the CAF constructs is of current interest. The CAF constructs have been investigated as dependent variables, e.g. in studies on effects of age, instruction, and tasks, among other factors. In studies of SLA, the constructs' influence on the L2 performance or proficiency has also been investigated as independent variables (see. e.g. Housen & Kuiken 2009; Baten & Håkansson 2015). Even though a lot of research on CAF is carried out, the constructs lend researchers many challenges. As mentioned previously, some of these challenges concern defining the constructs and operationalizing them in a consistent way, and how they interact with each other. Larsen-Freeman (2009) suggests that when investigating the constructs as separate components there is a risk of neglecting the interaction between the three of them. Housen, Kuiken and Vedder (2012), on the other hand, point out that there is empirical evidence suggesting the three constructs are both central and separate parts of L2 production, and hence also, by extension, of L2 proficiency. Even though they, according to Housen et al. (2012), are separate dimensions, this does not rule out that they do interact in L2 production and development. Furthermore, they continue that evidence from studies such as Larsen-Freeman (2006), Verspoor et al. (2008) and Spoelman & Verspoor (2010), indicate that "complexity, accuracy and fluency do not develop collinearly in SLA, that they interact in intricate ways, and that this interaction is sometimes mutually supportive and sometimes competitive." Skehan (1998a) suggested that complexity, accuracy and fluency are distinct and draw on different systems. Fluency draws on a memory-based system, while accuracy and complexity draw on the learner's rule-based system, and hence require syntactic processing. Complexity is also distinguished from accuracy in that complexity is affected by risk-taking, whereas accuracy is a reflection of attempts to keep control of the resources available to the learner and to avoid errors (Ellis, R. 2015). Skehan (1998b) is of the opinion that learners have a limited processing capacity, which leads to that one of the three will be prioritised. Trade-off will then be noticeable, e.g. if accuracy is focused, then either fluency or complexity, or both, will be affected. If the learner happens to be a "risk taker", s/he might produce more complex structures at the expense of both accuracy and fluency. Complexity and accuracy are in competition according to Skehan, something which Robinson (e.g. 2001, 2003, and 2011) does not agree with. Robinson suggests that structural accuracy and functional complexity are closely linked, so if the complexity of a task, i.e. the cognitive complexity, increases this would lead to both greater complexity, i.e. linguistic complexity, as well as greater accuracy. Both theories intend to explain the effect of the cognitive complexity of a task on complexity,

accuracy, and fluency in L2 production. However, support for Skehan's theory has been greater than the support for Robinson's theory (Ellis, R. 2015).

It should be mentioned that the constructs of CAF do not make up a theory or framework, but are rather dimensions used to describe the speaker or learner's performance (Pallotti 2009). The dimensions have been held to indicate the changes taking place in the underlying L2 system, e.g. the internalisation of L2 elements, modification of L2 knowledge, consolidation and proceduralisation of L2 knowledge (see. Housen et al. 2012:3), and can therefore be considered as a means of explaining these phenomena within cognitive language acquisition theory (Flyman Mattsson 2017).

To assess CAF, researchers generally adopt quantitative measures. These measures are, according to Bulté (2013:47–48), "designed to provide an objective representation of the degree of complexity, accuracy and fluency in any given text, and by extension of the level of development of proficiency of the L2 learner that produced the text". Complexity measures will be discussed in detail in 3.3.2.

2.3.2 Complexity in SLA research

Complexity is the construct that has received the most attention in SLA studies, but it is also the most problematic one of the CAF triad, both in regard to theoretical as well as to methodological concerns. Complexity, in current research, has been investigated either as a dependent or an independent variable. As the former, the learner's L2 performance or proficiency in terms of complexity is measured to provide information about the effect of some other variable such as age. As the latter, complexity is the variable, whose influence on some aspect of L2 performance and proficiency is under investigation, e.g. research on the effects of instruction in SLA (see. Bulté & Housen 2012).

As mentioned previously, there has been great differences and also contradictions in results of studies where complexity has been investigated as a dependent and/or as an independent variable. Bulté and Housen (2012) suggest that different ways to define and operationalise the construct are the reasons behind these differences and contradictions. In many studies complexity is not defined at all, or it is defined vaguely, or in circular terms. To illustrate this, three definitions of complexity are presented:

> "[complexity is the] use of more challenging and difficult language…
> Complexity is the extent to which learners produce elaborated language"
> (Ellis & Barkhuizen 2005:139)

> "Grammatical and lexical complexity mean that a wide variety of both basic and sophisticated structures and words are available to the learner"
> (Wolfe-Quintero, Inagaki & Kim 1998:69, 101)

> "Complexity refers to … the complexity of the underlying interlanguage system developed"
> (Skehan 2003:8)

Norris and Ortega (2009), Bulté and Housen (2012), Pallotti (2015), among others, have called for a more unified and specified definition of complexity, as well as in terms of operationalizing complexity. Bulté and Housen (2012:23) present a taxonomy of the complexity construct, see Figure 1. There are two approaches to complexity, i.e. *relative complexity* (also referred to as cognitive complexity, or difficulty) and *absolute complexity* (or simply just complexity). In the relative approach, complexity is defined in terms of what is difficult to the language user, especially in terms of cognitive efforts or resources needed. The relative complexity is further divided into *subjective determinants*, which are learner-dependent, and *objective determinants*, which are learner-independent. As can be seen in Figure 1, the learner-independent factors include the absolute or inherent complexity, i.e. the relative complexity notion is broader than the absolute complexity notion. In the absolute approach, complexity is defined in objective, quantitative terms, i.e. following Bulté and Housen (2012:24) "the *number* of discrete components that a language feature or a language system consists of, and the *number* of connections between the different components".

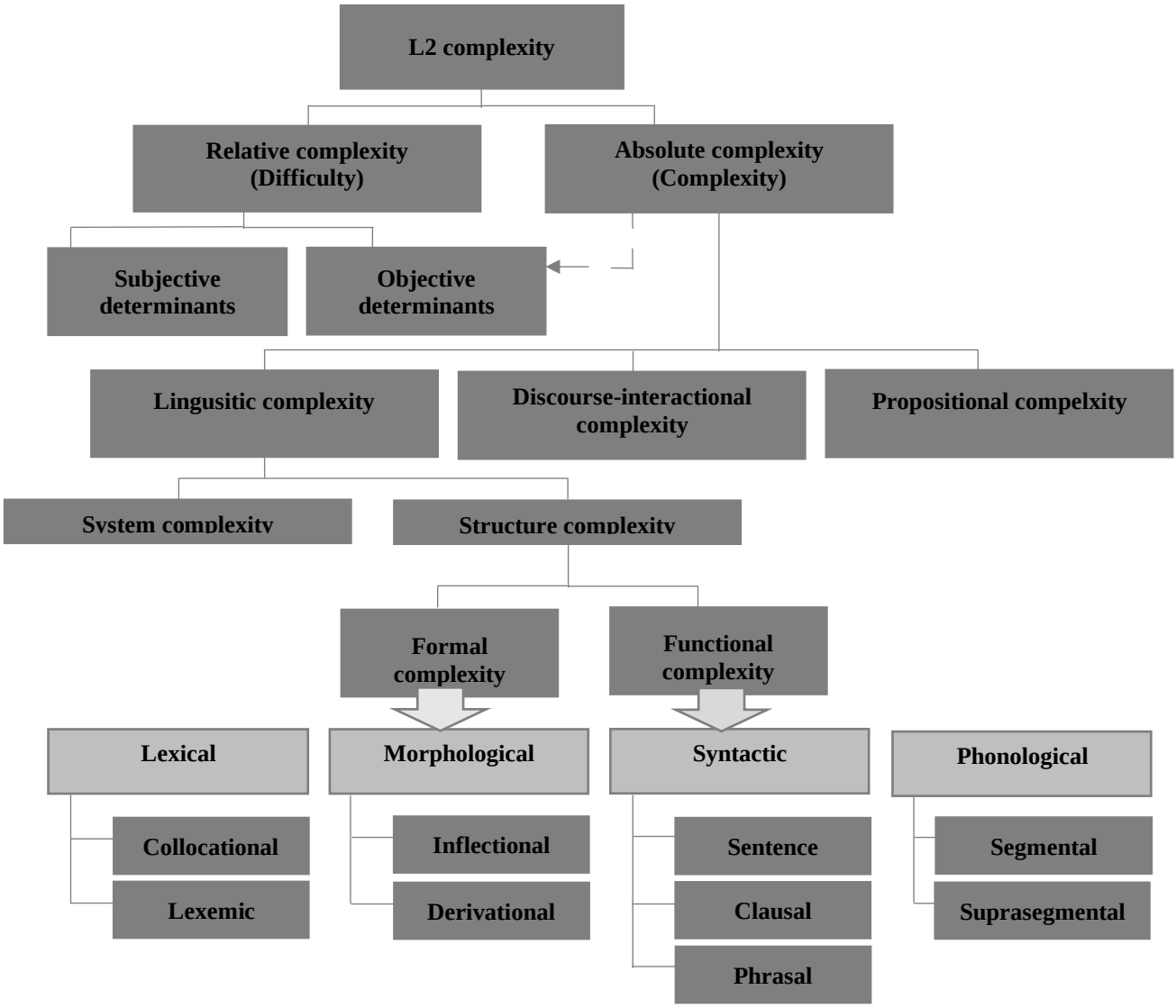

Figure 1. Taxonomy of the complexity constructs.

Based on Bulté & Housen (2012:23)

Absolute complexity, henceforth complexity, consist of three components: *propositional complexity*, *discourse-interactional complexity*, and *linguistic complexity*. The term linguistic complexity is commonly used interchangeably with cognitive complexity in the SLA literature. Housen et al. (2012) stress the importance of distinguishing them since linguistic complexity is an essential component of cognitive complexity, but it does not correspond with it. In the taxonomy, linguistic complexity is split into *system complexity* and *structure complexity*, the former is considered as a dynamic property of the L2 system of the learner on a larger scale (global complexity), while the latter is looked upon as a stable property of what constitutes the learner's L2 system (local complexity), such as individual linguistic items, structures or rules (Bulté & Housen 2012). Structure complexity is further broken down into *formal complexity*

and *functional complexity* of an L2 feature. These can in turn be broken down into *lexical* complexity, i.e. collocational and lexemic, *morphological complexity*, i.e. inflectional and derivational morphology, *syntactic complexity*, i.e. *sentence*, *clausal* and *phrasal complexity*, and *phonological complexity*, i.e. segmental and suprasegmental (see Figure 1). *Functional complexity* refers to "the number of meanings and functions of a linguistic structure and to the degree of transparency, or multiplicity, of the mapping between the form and meanings/functions of a linguistic feature." (Bulté & Housen 2012:25; Bulté 2013:61–63). Structures, where there is a one-to-one mapping of meaning and form are considered less functionally complex. *Formal complexity* has been defined in a variety of ways[3], the number of components included within a structure itself, "in terms of the number of operations to be applied on a base structure to arrive at the target structure", but formal complexity has also been argued to have to do "with the dependency distance between a form and its closest head or dependent" (Bulté 2013:62).

The sub-dimensions of linguistic complexity presented in the taxonomy can according to Bulté (2013:62) be assessed across various language domains, i.e. lexis, morphology, syntax, and phonology, as well as their subdomains, such as for example inflectional and derivational morphological complexity, or sentential, clausal, and phrasal syntactic complexity.

2.4 Measuring complexity

Bulté and Housen (2012) examined how complexity has been measured in a collection of 40 empirical L2 studies published between 1995 and 2008[4], and found that most studies employ general measures, which tap complexity constructs that are of a global and overarching nature, such as e.g. the mean length of T-unit, C-unit, or AS-unit. Specific measures of complexity, such as e.g. number of relative clauses per T-unit, are less commonly used. In their survey, they also found that in most of the studies claiming to measure L2 complexity, only a scarce number of measures were calculated, i.e. twenty-two of the forty studies only employed one or two measures. Bulté and Housen (2012) suggested that there not being enough adequate computational tools for automatically measuring complexity, along with the number of hours and work that have to be put into manual computation, could be the reason behind this. Another

[3] Bulté and Housen 2012:25 for a more exhaustive list of definitions.
[4] Bulté and Housen (2012:30-33) for the list of studies and measures.

concern when measuring complexity in these studies is that the same measures are used in most of them (Bulté and Housen 2012). This means that only a small portion of linguistic complexity is measured, namely lexical diversity and/or syntactic sentential complexity through subordination. Syntactic complexity[5] can e.g. also be measured by assessing the overall syntactic complexity, targeting sentential complexity in terms of compositionality, amount of coordination, and subsentential (clausal or phrasal) complexity. This is also discussed by Norris and Ortega (2009).

2.4.1 Measuring phrasal complexity

According to Bulté (2013:79) the typical aim of syntactic complexity measures is to quantify at least one of the following: "range of syntactic structures, length of unit, degree of structural complexity (sophistication) of specific syntactic structures, and amount and type of syntactic linking through coordination, subordination and embedding". Norris and Ortega (2009) assert that elaboration at the phrasal level, such as nominalisation and other processes at the phrasal level, is something that learners at later stages of L2 development lean more on than clausal subordination. Therefore, it should be of interest to measure complexity at this subsentential level. For this purpose, Norris and Ortega (2009) suggested the mean length of clause (MLC), i.e. the number of words per clause, as a measure of phrasal complexity. The assumption here is that the number of possible phrases are limited in a clause, which implies that an increased clause length would reflect an increase in phrase length, due to the head being modified (Bulté & Housen 2012; Bulté 2013). However, Bulté and Housen (2012) and Bulté (2013) mean that MLC is not to be considered a 'pure' measure of phrasal complexity since the length of a clause can be increased through other means than expansion at the phrasal level, such as expansions at the clausal level by, e.g. the addition of adjuncts. Furthermore, they remark that the length of the clause is determined by the definition and operationalisation of a clause.

2.5 Previous research on phrasal complexity and development

Bulté (2013) performed a longitudinal study on a dataset of ten Dutch L2 learners of English, with the aims to contribute to the definition and operationalisation of L2 complexity, to look into the development over time in written L2 English, and to investigate DST approaches to L2 complexity. The dataset contained eleven measurement points over 19 months. Both syntactic

[5] Bulté (2013:80-86) for a brief discussion of the different subdivisions of syntactic complexity measures, problems and alternative measures.

and lexical complexity were investigated, but only the results of the syntactic analysis, more specifically the analysis of the phrasal complexity, will be considered here.

Bulté and Housen (2012) and Bulté (2013) suggested measures of intra-phrasal complexity, such as dependents per noun phrase head (MNPDep) and the number of words per noun phrase (MLNP). The latter was also employed in Bulté and Housen (2018). Bulté (2013) also presented a third measure as an alternative to the two length measures - WNPCx, which is a weighted measure. The WNPCx was created in order to come around two concerns when calculating complexity, i.e. that when counting the number of words of an NP (MLNP) the same weight is given to all NP constituents creating problems particularly concerning calculating the inherent composition of embedded clauses which can be more or less infinite in length. When counting the number of dependents per phrasal head (MNPDep) the same weight is still rewarded to e.g. articles and embedded clauses. However, after a strong and significant correlation was found between the three measures only MLNP was chosen for the continued study since these measures are considered to measure the same underlying syntactic complexity construct.

Bulté (2013) states that an overall increase over time was found for almost all of the syntactic complexity measures examined in the study. However, the overall increase was in most cases not smooth. In individual scores, a great deal of variation can be observed. For the three phrasal complexity measures: MLNP, MNPDep and WNPCx, a clear upward trend is detected. The scores at times 1–4 are consistently lower than those of times 7–11, with only one exception. The group scores seem to be irregular, i.e. there are sudden drops in data collection point 8 and 11 (2013:162–3). On the individual level, there is great variation both across and within learners, and therefore Bulté notes that finding any trends of individual developmental trajectories is difficult. Furthermore, the variable ‘time’ was found to have a significant effect on MLNP (as well as on MNPDep and WNPCx). The NPs produced by the learners of the study became longer with time.

Two learners were examined more closely to display the interindividual variation, learner 109 and 120. The former appeared to have more or less no progress throughout the study, and scores appeared to be random. The lowest score appeared at the last data collection point and the highest at time 6. Learner 109 shows great variation in the scores over time. For 120, however, the scores appeared to increase (broadly speaking) throughout the study, i.e. the scores of times

1–6 are consistently lower than those of times 7–11. However, there is no clear developmental pattern within the two sets of scores (2013:248–49).

To conclude on the development on MLNP, over time the scores for individual learners appears to contain a great deal of variability across as well as within learners. Most learners show an increase in MLNP scores over time, but in a very inconsistent way (2013:275–6).

Bulté and Housen (2018) studied variation and variability, both within and across learners using ten participants from the same dataset examined in Bulté (2013). They confirm a clear upward trend for the MLNP throughout their study on syntactic complexity, but they also state that the non-linear character of the trajectory with local peaks and drops of this measure is common. Also in this study the group scores are higher in the later data collections points, i.e. in this case times 5–11, and sudden drops in times 8 and 11 are shown. Again, the interindividual variation is great and clear trends of developmental trajectories are hard to find (Bulté & Housen 2018:155).

The observed individual pathways show great variability between learners and the results are interpreted as indications that "different learners follow different developmental paths when it comes to changes in L2 complexity over time", which is speculated to possibly be due to differences between learners and learning profiles such as discussed by Norrby and Håkansson (2007). Some learners in the study seem to have a gradual development, while others present variability and also sudden shifts. Bulté and Housen lend a word of caution when drawing conclusions based on single-case studies, but also when applying conclusions of group studies on individual learners. These findings are suggested to support the DST claim made by Verspoor, Lowie, Chan and Vahtrick (2017) that one cannot generalise from the individual to the group and vice versa, later also supported by Bulté and Housen (2020). Therefore it is recommended to look at both the individual developmental trajectories and the group score to fully understand the L2 development.

Bulté and Housen's studies are not the only studies including phrasal complexity. In the following paragraphs a selection of phrasal complexity studies will be presented in brief. Lu (2011) compared 14 syntactic complexity measures commonly used in research on written L2 development in a study on L2 English. For phrasal complexity MLC, as discussed previously,

not a pure measure of phrasal complexity according to Bulté and Housen (2012), the number of complex nominals per clause (CN/C), and the number of complex nominals per T-unit were investigated. CN/C and CN/T showed significant differences between three and two adjacent levels respectively. Several patterns of development for the ten measures that displayed significant between-level differences. Seven measures displayed significant positive changes from lower to higher levels, among these are MLC, CN/C, and CN/T. MLC and CN/C progressed linearly across all four levels. CN/T also increased linearly across the first three, but displayed an insignificant decline from the third to the fourth level. Lu found the results to indicate "stronger discriminative power of the complex nominal and coordinated phrase measures than most other measures" (2011:57) , and continues to suggest that "this finding suggest that complexity at the phrasal level deserves closer attention".

De Clercq and Housen (2017) measured the length of the NP in a similar way to the MLNP (Bulté & Housen 2012; Bulté 2013), but called the measure LenNP, and used it to investigate and compare the development of syntactic elaboration in oral narratives of L2 French and L2 English. De Clercq and Housen operationalised the NP by only including NPs with a common noun as head which were not embedded in another NP. They found statistically significant differences between the L2 languages for LenNP. Between the learner groups in both languages no statistically significant differences were found. The NS French benchmark group scored significantly higher than the learner groups, leading the authors to conclude that "the learners may have mastered the basic NP structure but the development at a more advanced level might involve further NP complexification" (2017:329), all in line with previous results where phrasal elaboration is preferred in written French. The English learners generally presented more syntactically complex narratives than the French learners, while the trend tended to be the reversed for the NS benchmark groups.

Kuiken and Vedder (2019) compared L2 learners and NS writers of Dutch, Italian and Spanish. For NP complexity they used two measures, number of post-modifying noun phrases per 100 words and the mean length of post-modifying noun phrases. They found that the highest number of post-modifying NPs were acquired by Dutch learners. The highest scores for the length of post-modifying NPs were obtained by the Spanish learners. However, the authors could not discern any clear developmental patterns from their results. Their study does indicate that the

process of gradual complexification in written L2 production vary across proficiency levels, across languages, and between L1 and L2.

Verspoor et al. (2008), measured the length of the NP in L2 English and found development of the NP to increase over time. However, only slightly and with noticeable fluctuations. There was a general trend upward, but they report on a fair amount of variability, and claim that developmental trajectories are unpredictable. All the measures examined show a great deal of variability and development is not linear, showing moments of progress and regress.

Spoelman & Verspoor (2010) studied intra-individual variability in accuracy and complexity measures in Finnish L2 writing, starting at absolute beginner's level, in a longitudinal study. Word, NP and sentence constructions were studied for developmental patterns of complexity. NP complexity was calculated using an NP length ratio, and how frequent the different NPs, i.e. one, two, or three or more word NPs, were used. The authors concluded that accuracy and complexity measures do not remain constant over time and are not collinear, are multivariate and dynamic and therefore should be examined across the full developmental trajectory, and are characterised by peaks and regressions, as well as by progress and backsliding. Therefore they argue that L2 complexity is clearly non-linear. All the complexity measures increased over time, but Spoelman and Verspoor (2010:551) conclude that "intra-individual variability behaves according to the principles of Dynamic Systems".

Bartning et al. (2015) too used an NP ratio, and based their scoring principle on Ravid and Berman (2010). They investigated pre- and post-modification in the noun phrase in very advanced L2 French speakers and NS. They found that native speakers (NS) produced a larger quantity of complex NPs than non-native speakers (NNS). In almost all categories, NS produced higher frequencies, although not significantly so. NS produced higher proportions of complex NPs than NNS at level 3 and 4 of the scoring principle applied. The difference for level 4 is significant, as well as when levels 3 and 4 are cumulated. NNS use a higher proportion of complex NPs at level 1 and 2; the difference is significant when the levels are cumulated. Bartning et al. claim that phrasal complexity is indeed a measure of advanced level proficiency in NNS, worth investigating, and the study answers to the call of Norris and Ortega (2009) for more research on phrasal complexity in advanced stages of L2 development.

A note of caution is that none of the complexity measures recommended in the SLA literature are completely problem free, as mentioned e.g. in the discussion on MLC as well as regarding MLNP. Also, the studies mentioned here more or less all have different foci and some are cross-sectional and some are longitudinal.

So far, no studies seem to have investigated the complexity of NPs through morphology. Most studies on L2 complexity have been carried out with English as the target language, and since English has little inflectional morphology this does not come as a surprise. For English, the number of possibilities to gauge development of inflectional morphology are limited: one can either count the number of inflected verb forms, or the number of inflected forms/grammatical word class may be analysed. Due to the sparse morphology of English, there tends to be an increase at earlier stages of L2 development, but a ceiling effect is soon reached (Bulté 2013, Brezina & Pallotti 2019). Inflectional morphology is an area where cross-linguistic differences can become relevant for the measurement of complexity in different L2s (Bulté 2013). The results of Brezina and Pallotti's (2019) study on morphological complexity in written English and Italian L2 production confirm these differences. De Clercq and Housen (2019), investigated the inflectionally richer French language and the inflectionally poorer English language, in terms of morphological complexity in L2 production. Their study indicates that morphological complexity increases in L2 French more continuously than in L2 English. These studies both concern morphological complexity and are carried out on verb morphology. However, it is possible that the inflectional morphology of the NP should be included when investigating phrasal complexity, and even more so when investigating NP complexity in L2 Swedish. The results of Brezina and Pallotti (2019), and De Clercq and Housen (2019) suggest that it could be of interest to investigate the complexity of the NP through the means of inflectional morphology in Swedish, and also to examine if perhaps a measure of phrasal complexity, that also takes this into consideration, might capture the complexity within the NP (in Swedish) in a more detailed manner. Furthermore, as De Clercq and Housen (2019) mention, morphological development has, in several studies e.g. been established to be "a crucial aspect of L2 development, of continual importance to the learning process". Bulté and Roothooft (2020) employed the morphological complexity index, MCI (Brezina and Pallotti 2019), in their investigation on L2 complexity and L2 proficiency levels in spoken L2 English, and found the overall difference between proficiency levels to be significant. The difference is for the most part found between the lowest level and all the other levels combined. An observed increase in morphological richness was found between the higher levels.

2.6 The Swedish Noun Phrase

The Swedish noun phrase (NP) is structured around a head, which can be a noun, a pronoun, or a proper noun. The head can make up the whole phrase by itself, or it can be accompanied by prenominal and/or postnominal dependents.

The most frequently appearing prenominal dependents are articles, demonstratives, possessives, quantifiers and adjectives, as exemplified in (1)–(5) with the NPs within square brackets and the head marked in bold:

(1)	[en **bil]**, [ett **hus]**	articles *en, ett*
	(a car, a house)	
(2)	[den här **bilen]**, [detta **huset]**	demonstrative pronouns: *den här, detta*
	(this car, this house)	
(3)	[min **bil**], [mitt **hus**], [hans **bil**], [hans **hus**]	possessives: *min, mitt, hans*[6]
	(my car, my house, his car, his house)	
(4)	[alla **bilar**], [många **hus**]	quantifiers: *alla, många*
	(all cars, many houses)	
(5)	[röda **bilar**], [vackra **hus**]	adjectives: *röda, vackra*
	(red cars, beautiful houses)	

In the presence of prenominal dependents the nominal head can be missing, such as in the examples from The Swedish Academy Grammar (SAG *3* Fraser: Teleman et al. 1999:13) below:

(6)	Han har frågat ***många***.
	(He has asked many.)
(7)	Du får tala med *någon **annan** i föreningen*.
	(You will have to talk to someone else in the association.)
(8)	Här ligger *två kroniskt **sjuka***.
	(Here lies two chronically ill.)

The most common post-modifiers are prepositional phrases (PP) and relative clauses (SAG 3 Fraser: Teleman et al. 1999:14) such as in (9) and (10):

[6] There is agreement for the possessive pronouns in the genitive case *min* 'my' and *mitt* 'my' and the heads *bil* 'car' and *hus* 'house', but not for *hans* 'his)

(9) *flickan från Värmland*
(the girl from Värmland)

(10) *flickan som är från Värmland*
(the girl who is from Värmland)

Nominal phrases headed by nouns can have a more complex structure than those headed by pronouns or proper nouns, i.e. the possibility for pre- and post-modification is greater for nouns than for pronouns (SAG 3 Fraser: Teleman et al. 1999).

Swedish NPs can be definite and indefinite, with dependents agreeing with the nominal head in gender, number and definiteness. Phrases containing prenominal modifiers are moreover characterised by the so-called *double definiteness*, which is shown in the following examples.

(11)	**den** glad<u>e</u> mann**en**	(**the** happy man)
(12)	**den** glad<u>a</u> kvinna**n**	(**the** happy woman)
(13)	**det** stor<u>a</u> hus**et**	(**the** big house)
(14)	**de** stor<u>a</u> hus**en**	(**the** big houses)

As can be seen in example (11)–14), definiteness is marked by the definite article (den/det/de) and a definite suffix on the noun, both marked in bold. There is also agreement between the pre-modifying adjective and the noun. In the definite NP, the agreement is marked on the suffix of the adjective, underlined in the examples above. Considering the NP, English and Swedish differ in regard to the inflectional morphology. In English, the definiteness is marked only on the article (marked in bold), such as in: ***the*** *red car*, and the noun is not marked for definiteness, only for plural with the plural *–s*, as in: *the red car**s***, while Swedish employs double definiteness as presented above.

The definiteness of the Swedish NP has been investigated quite extensively, also from an L2 perspective. However, these studies tend to be concerned with accuracy and/or relative complexity (difficulty). Axelsson (1994) investigated the definiteness of the NP in Swedish in regard to difficultness, i.e. what is easily acquired and what is more difficult to acquire, and found three stages for the acquisition of NPs and definiteness. In her doctoral dissertation, Nyqvist (2013) also looked at what is difficult and what is more easily acquired, when it comes to species and article use. Lehtonen (2015) examined reference categories and articles in L2 Swedish and L2 English in reference to the CEFR scale. Also this study falls under the accuracy

and/or difficulty dimension. Complexity and accuracy are two independent dimensions, and of which the analyses of complexity should be conducted independently of accuracy. A recent addition to the research on double definiteness, is the doctoral dissertation of Agebjörn (2021). Agebjörn studied L2 learners (beginners and advanced learners) with Russian as L1 and the development of definiteness.

Chapter 3 Method

As discussed previously there are different ways to view development. Development can occur linearly, i.e. in patterns of increasing complexity, or non-linearly, where the development does not follow a linear pattern. The aim of this thesis is to establish if a development of noun phrase complexity can be observed in written L2 Swedish, regardless of whether said development is linear or non-linear. Three research questions will aid in achieving the aim. RQ1 and RQ2 concern the complexity measures, which are further discussed and exemplified in 3.3.2, while RQ3 concern development. To measure complexity and thereby establish if a development can be observed, data from the ASU corpus is employed. The ASU corpus and the material is presented in 3.1 and 3.2, and the guidelines for the analysis of the data is presented in 3.3

3.1 The ASU Corpus

The ASU corpus (*Andraspråkets StrukturUtveckling*, 'the Structural Development of the Second language') is a longitudinal oral and written text corpus of adult learners of Swedish with a corresponding native speaker part (Hammarberg 2010). The corpus is accessible online at Språkbanken ('the Swedish Language Bank') for research and educational purposes, but requires an approved login. Since the material was collected during 1990–93 (learner part) and 1998 (native speaker part), the corpus has not only been converted into a more modern format, but it has also been connected to ITG, which is a user interface handled by Språkbanken at Gothenburg University (Hammarberg 2013). The ITG provides ways to search, analyse and edit the data[7]. In the ASU corpus, the language of individual learners is documented longitudinally, which lends the possibility to study the language of each learner, and to compare a learner with him/herself over time. Hence, it is possible to detect the development of a specific learner throughout the data, as well as on a group level. The corpus is designed to make it possible to compare different dimensions, such as: longitudinally across acquisitional stages, between individuals, between learners and native speakers, between speech and writing, and between genres (Hammarberg 2010).

All learners are anonymised in the corpus and no information is presented within the present study, therefore no further ethical considerations are considered. Furthermore, the ASU-corpus

[7] For a guide on how to use the ASU corpus and the ITG, see Hammarberg and Olsson (2010).

is partly funded by the Swedish Research Council and follows their rules and guidelines for research.

The learner part consists of data from ten informants, three females and seven males, who were all students in a preparatory course in Swedish for foreign students at Stockholm University. The learners were roughly considered to be semi-formal learners since they lived in Sweden and received their linguistic input both from the preparatory course and from the surrounding linguistic environment. Furthermore, they are considered to be qualified learners and fast learners, i.e. they all have prior experience of foreign language acquisition, along with a strong instrumental motivation for learning Swedish in order to proceed with their studies within their respective specialisations, and they advanced, within the timeframe of one to two academic years, from beginner level to the level required for university studies conducted in Swedish.

Apart from the previously stated characteristics, the learners are stated to be homogenous in aspects regarding:

- Age – young adults (19–28 years, median 20,5 years)
- Socioeconomical background – middleclass
- Prior residency within a Swedish speaking linguistic environment: 10 days to 6 months. Two learners had some informal input through contact with Swedes, which can be noticed in the corpus, but they were nevertheless considered to belong in a beginner's course .
- Course progression: all the participants were placed in the same class during the first academic year, with the same teacher, material and schedule.
- Prior L2 knowledge: all had some degree of knowledge of English, which agreed with the required level of English for university studies.
- Field of studies: none were enrolled in language studies apart from the preparatory course. The different fields were economics, medicine, technology, and film.

In other aspects, there was some variation among the learners, such as in aspects regarding:

- L1: Chinese (Mandarin and Shanghai), German, English, Greek, Polish, Portuguese, and Spanish. One participant, had rudimentary knowledge in an indigenous language of Mozambique[8].

[8] For a more detailed description of each participant see Hammarberg 2010:7

- Varying L2 knowledge (apart from English) based on self-evaluation (advanced, intermediate, and low) in: Japanese, French, Swahili, Russian, German, Italian, Spanish and Portuguese[9].
- Cultural background
- The rate of progress and obtained proficiency

Each learner is identified by a capital letter and a number, e.g. C1, and each text is identified by *person + medium + session + essay 1 or 2*, e.g. C1S071, i.e. person C1, Written (S = skriftligt), session 7, essay 1.

The written data of the learner part contains 220 essays in total, i.e. twenty-two per informant, produced in eleven sessions. A total of around 50 000 word tokens. The written data of the native speakers contains 70 essays, i.e. ten essays per informant produced at five sessions. A total of around 25 000 word tokens.

The data of the native speaker part of the corpus were collected in 1998, i.e. after the learner corpus was completed and edited. The native speaker part contains data from seven native speakers (NS) of Swedish, who were enrolled in undergraduate studies at Stockholm University. The aim was to include informants who were equivalent to the informants of the learner part. All NS, four females and three males, were born and raised in Sweden, and Swedish was their only L1. They were between 20–29 years old (median 23 years), had L2 knowledge of 2–4 languages with varying proficiency levels (English included), and their field of study varied within philosophy, literature, art history, and social anthropology.

3.2 Material

3.2.1 Material – learner part

An L2 learner is constrained by his or her linguistic proficiency in the L2 (Cumming 1989) and for an L2 learner to be able to write s/he has to have the required vocabulary, know how to apply grammatical rules, and how to combine words into e.g. clauses and phrases. Therefore, the written production of an L2 learner holds evidence regarding the learner's level of his or her overall linguistic L2 development and proficiency (Bulté 2013:47). A vast amount of corpus-based studies, such as e.g. Biber 1988, have established that linguistic differences

[9] See. Hammarberg 2010:7

between written and spoken texts are present and that they extend also across register and genres (Bulté 2013:46). Furthermore, written texts are inclined to contain large numbers of noun phrases, and elaborated noun phrases (Bulté 2013:46). Therefore, the written material of the ASU corpus is chosen for the present study.

The material, i.e. the informants' production, was collected during writing and recording sessions, which were not part of the preparatory course they were enrolled in. The written material was collected during writing sessions in groups held in a classroom. Two essays on given topics per informant were produced at each session. The participants were given a topic and a heading, and in some cases a short instruction and/or a set of pictures was provided. The essays were written by hand, and according to the research team they were produced in a spontaneous manner without pondering, or extensive reviews (Hammarberg 2010:8). The tasks include: *picture series*, *narrative*, *description*, *discussion* and *other* (*'övrigt'*), and each task type contains several different topics, some of which were repeated in later sessions. The descriptive task, which is the material used in the present study, contains four essays produced during sessions: S1, S4, S7, S11, where S7 and S11 describe the same topic. Session 11 is a follow up after the participants had finished their L2 courses, and hence no longer had any instruction in L2 Swedish. As shown in Table 1, sessions S1 and S4 were collected at different occasions in the latter part of 1990, i.e. on the 12th September and the 14th or 19th November, session S7 on 18th March 1991, and finally session S11 was collected on the 30th April in 1993, apart from informant Q2 who completed this essay on 15th May. Q2 is also the only participant who did not pass the test to meet the qualification requirements for university studies in Swedish, later known as the TISUS-test, within the two first years.

Table 1. Topics of the descriptive task

Session	**Topic**	**Topic translated into English**
S1 September -90	En gäststudent i Stockholm	A guest student in Stockholm
S4 November -90	Universitetet i Frescati	The University in Frescati
S7 March -91	Familjesidan i en daglig tidning	The family page of a daily newspaper
S11 April -93 ([10])	Familjesidan i en daglig tidning	The family page of a daily newspaper

After Hammarberg (2010:12)

[10] Informant Q2 completed this essay on 15th May.

In order to investigate a potential development of NP complexity over time, the texts are required to contain NPs and have the potential to yield elaborated NPs as well. For the present study, descriptive texts where considered suitable to meet this objective. It may be that other types of text would be equally or perhaps even more suitable for yielding NPs, and it would of course be possible to include texts from different genres, but due to time and space limitations, and the fact that it is not the task that is under investigation but rather the development of the NP complexity, only one type of text is included.

3.2.2 Material – native speaker part

The written material of the native speaker part was also collected in a classroom during group sessions. Two essays on given topics were produced at each session, and the procedure was the same as for the L2 learners. The five writing-session were held one week apart. Task types and topics were similar, or identical to those of the learner part. The native speaker part only includes one essay from the descriptive task, on the topic '*Universitetet i Frescati'* (the University in Frescati). The native speaker part is included in order to be able to compare the learners to the native speakers. An NP can only be elaborated to a certain degree and therefore the native speaker part is analysed to be applied as a kind of norm for the comparison.

3.3 Analysis

The analyses and calculations were carried out manually by the author and an experienced rater. The reason behind the choice to carry out the analyses and calculations manually is not that there are no tools available for measuring complexity, but rather that the tools at hand also come with some disadvantages, such as they only provide the final score of the different measures, and not the underlying analyses. Bulté and Housen (2014:48) point out that the algorithms used in complexity measuring tools, such as e.g. *Coh Metrix* (Graesser et al. 2004) and *L2 Syntactic complexity analyser* (Lu 2010) are still too rigid in order for them to fully identify, segment and parse the production of L2 learners in an accurate manner, and not produce any measurement noise. Hence, despite the value of labour-saving, and the obvious concern regarding reliability created by human analysis, complexity measuring tools are not employed in the present study. However, according to Bulté and Housen (2014), it is possible to reduce the subjectivity of manual coding by creating clear guidelines for coding, which are presented in Tables 2–4, and

to discuss cases that are ambiguous. Another obvious way to reduce subjectivity is to include more than one rater for inter-rater agreement. For the present study, this was implemented.

Instead of the use of any computational tools, the material was segmented in Excel spreadsheets, following the analyses of Bulté (2013) and Bulté and Housen (2014), with the exception that the definition of a linguistic clause employed in Bulté and Housen (2014:48) was not used in this study. Since focus is on the phrasal level alone, and not on clausal and/or sentential levels, it was not considered necessary to divide the material into clauses where a verb and a subject are present. For this study, the material was divided into clause elements. The Excel spreadsheet was adapted for the analysis and calculation of all four measures.

In section 3.3.1, definitions and restrictions concerning the noun phrase and the analyses thereof will be presented, but first the definition of a *word* will be discussed, i.e. what is to be considered a word within the frame of the present study. Flyman Mattsson and Håkansson (2021:90) list some guidelines as to define a word: identical repetitions are not to be counted if they bear no stylistic value, compounds written apart should be counted as one word (**hus tak* = *hustak*), and words from English or any L1 or other L2(s) should not be counted as a word in Swedish. However, the present study makes an exception to this, namely that non-Swedish words are included if the word is used as if it would have been a Swedish word, e.g. *mina snälla *klassmates* (my friendly classmates). This is not considered to have affected the outcome of the measuring of morphological complexity within the NP in the data of the present study since the words included are, as in the example with **klassmates*, used as if they were Swedish words and there are very few cases. It would, nevertheless, be recommended not to included them in line with Flyman Mattsson and Håkansson (2021), especially when the morphology is taken into account. It could potentially result in missing out on a morpheme in regard to definiteness. Furthermore, words functioning as a unit, such as *med hjälp av* (by means of), *på grund av* (due to), etc. could, since they as a unit play a particular role in Swedish be considered as one word. The same holds true for the abbreviations of such units, e.g. *p.g.a.* – *på grund av* (due to). Spelling mistakes are not considered here. In the study, proper names are also considered a unit and counted as one word, e.g. *Stockholms universitet*, and *Yngve Gamlin*.

3.3.1 The noun phrase

In the following section, definitions and restrictions of the NP within the frame of this thesis will be discussed. The present study follows the analysis of Bulté (2013), presented below, in

as much detail as possible in order to compare results, but with an exception in relation to NPs headed by proper nouns and pronouns, which is discussed shortly.

An NP can be made more complex through:

- coordination, e.g. *tables and chairs*
- pre- and post-modification, i.e. embedding of clauses and phrases, see (15) embedment of phrase and (16) embedding of clauses below.

When measuring NP complexity in Bulté (2013), the coordinated NP (the compound NP in Bulté's terminology) is considered to be more complex due to there being a certain degree of integration. However, there is no dependency relationship for the coordinated NP as there is for embedded clauses or NPs. Furthermore, if they are regarded as two separate NPs and heads, the conjunction does not really belong to any of the NPs, but rather to the coordinated NP, i.e. the superordinate NP (see e.g. Johannessen 1998). Coordinated NPs, such as *tables and chairs*, are treated as single constituents that are occupying one slot in the clause's argument structure. The present study also follows this line of reasoning.

NPs can be embedded in a PP, which in turn can be embedded in an NP. In (15) below, the PP *i huset* contains the embedded NP *huset*, which in its turn contains another PP *vid sjön*, in which the NP *sjön* is embedded. The two NPs and their heads are marked with square brackets and bold respectively.

(15) *i* [***huset*** *vid* [***sjön***]]
in the house by the lake

Clauses, such as complement and relative clauses, can also be embedded in NPs, in which more NPs can be embedded. In (16) below, a relative clause is embedded in the superordinated NP *kvinnan*. The relative clause contains two NPs embedded in PPs, as described in (15) above. In total (16) contains three NPs.

(16) [***kvinnan*** *som bor i* [***huset*** *vid* [***sjön***]]]
the woman who lives in the house by the lake

Bulté (2013) and Bulté and Housen (2014) exclude both NPs headed by pronouns and proper nouns in their lexical analysis, due to them not allowing for pre- and post-dependents in English to the same extent as NPs headed by a noun. However, Bulté (2013:118) does include proper nouns in the syntactic analysis. As stated in 2.6 the possibility to take attributes is greater for nouns than for proper nouns and pronouns (SAG 3 Fraser: Teleman et al. 1999). Therefore, NPs headed by pronouns or proper nouns are excluded in the syntactic analysis in the present study. They are however, counted as words if part of another NP, but when heading the NP the NP is not analysed.

(17) [*en gäststudent* [*i* ***Stockholms universitet***]] from C1S011
a guest student of Stockholm University

In example (17), the NP headed by *gäststudent* contains a proper noun *Stockholms universitet*, which is included in the wordcount of the superordinate NP. Furthermore, the proper noun is counted as one word, so the NP contains four words in total. The proper noun *Stockholms universitet* could potentially be considered to include NP complexity, but the possibility that it is learnt as a lexical unit cannot be excluded.

The exclusion of proper nouns (and pronouns) could lead to the risk of missing out on fairly complex NPs though, such as:

(18) *Lilla* ***Anna*** [*som bor i* ***huset*** *vid* ***sjön***].
Little Anna who lives in the house by the lake.

(19) ***Hon*** [*som bor i* ***huset*** *vid* ***sjön***.]
She who lives in the house by the lake.

In (18) the proper noun, *Anna*, is the head of the NP, which also contains a pre-dependent and a post-dependent, which in turn contains two NPs, *huset* (the house) and the subordinated NP *sjön* (the lake). In (19), the pronoun *hon* (she) is the head of the NP and the NP also contains the same post-dependent as in (18) above. As mentioned previously, the possibility to take attributes are greater for nouns than for proper names and pronouns. Therefore, phrases with proper names and pronouns as head of the NP are not included in the analysis even though there may be some complexity in these phrases. For example, in (19) it is not really possible to elaborate *hon* with any pre-modification. Cases like these did not occur in the data.

NPs in PPs such as *at* [*my office*], are included in the analysis, but the preposition is not counted as part of the NP. Coordinated NPs such as [***tables** and **chairs***] are analysed as containing two separate heads (marked in bold), and thereby also as two separate phrases. The conjunction, however, is in line with Bulté (2013), only counted once as discussed previously. For NPs containing embedded clauses, all words of the clause are counted as words of the NP, e.g. *university students*, [*who come from many different countries*] = seven words in total: six words in the clause. Compound words, such as *university students* are written as one word in Swedish, and hence counted as one word. Even though English compounds are written apart, they are counted as one word in Bulté (2013: 127–28), and therefore the total amount of words in the example NP are seven, and not eight. NPs can also be embedded in another NP. In this case, the NPs are counted and analysed separately, but when calculating the length of the superordinated phrase, their constituent words are included. Therefore, the superordinated NP of [*the sisters of* [*my grandmother*]] contains five words, and the subordinated NP in the PP *of* [*my grandmother*] contains two words. This scenario is the same as that stated for the NP embedded in a subordinate clause, functioning as a dependent of an NP (Bulté & Housen 2014). Every word in addition to the head is described as dependents of the NP, i.e. determiners, and pre- and post-modifiers, are all counted as dependents in the measures. Another example of NP embedded in NP is provided in the following example (20), where the NP in genitive *gamla människors* is embedded in the NP *problem*.

(20) [*gamla* ***människors problem***] from C2S071
old people's problems

Development of morphological complexity can not only be gauged through inflectional morphology, i.e. morphemes that add grammatical information to a word, as discussed previously, but also by measures of derivational morphology, which also considers lexical aspects, such as forming new words. Since the study focuses on grammatical development, and the fact that "derivational morphology belongs as much to the lexis as to the morphology of a language system" (Bulté 2013:87), only the inflectional morphology of the NP is studied.

The elaborateness of the inflectional system of the language under investigation sets the potential to gauge and calculate the inflectional complexity. In contrast to English, Swedish is a richer inflectional language. Therefore, the L2 learner also faces a larger range of inflectional

affixes, and there might be a risk of leaving out complexity of Swedish NPs. In the following, the criteria applied to the analysis of the morphological aspects of the NP complexity are presented.

Only concatenative morphemes are counted. This is why verbs such as *lek-te* (play-ed), where the past tense is marked by a suffix, are included in the analysis while the verb *sprang* (ran), where the past tense is only marked by ablaut, is not included.

Nouns

- definite suffixes — *katt-en* (the cat), *katt-er-na* (the cats)
  hus-et (the house), *hus-en* (the houses)
- plural suffixes (not the zero pl.[11]) — *katt-er* (cats), *katt-er-na* (the cats)
- genitive suffix — *-s*, *katten-s* (the cat's)

Adjectives

- comparative suffixes: — *röd* (red) – *röd-are* (more red)
- the definite suffixes: — *den röd-a bilen* (the red car),
  de röd-a bilarna (the red cars)
- plural suffixes: — *röd-a bilar* (red cars)
- neuter suffixes — **rött* vs. *brun-t*

* Neuter suffixes (t-word) are only counted if they are attached to the base form and not if it is replacing the common suffix (n- word). Hence *rött* (red) is counted as one morpheme, while *brun-t* is counted as two morphemes. As for pronouns, such as *min* (min bil - my car) and *mitt* (mitt hus - my house), where *min* is the base form and *mitt* is the neuter suffix, the neuter form is also counted as one morpheme. This is since it is impossible to know if the neuter suffix is learnt as a lexical unit or not when it replaces the base form.

[11] Nouns such as e.g. *träd* 'tree', *ett träd* 'one tree' – *många träd* 'many trees', i.e. the same form for both the singular and the plural.

Since the NPs can contain embedded clauses containing other phrases, such as verb phrases (VP), and all words of the embedded clauses are counted in the MLNP, the inflectional morphology of verbs will also have to be considered.

Verbs

- suffixes of the present tense e.g. *-ar, -er*
- suffixes of the past tense e.g. *-de, -te*
- suffixes of the supine e.g. *-at, -it*
- passive suffixes *loppet sprang-s* (the race was run)
- suffixes of the past participle e.g. *spel-ad*/*spel-at*/*spel-ade*

Adverbs are counted as one morpheme since they are not inflected (marked in bold) as in: *de **allmänt** intresserade* (the generally interested).

3.3.2 The measures

Measures of syntactic complexity at the phrasal level are still scarce in number since it is of fairly new interest to measure phrasal complexity in L1 and L2 complexity research (Bulté & Housen 2015). For phrasal complexity, the mean length of noun phrase has been employed in recent studies (cf. Bartning et al. 2015; Bulté 2013; Bulté & Housen 2014, 2015, 2018; 2020; Bulté and Roothooft 2020; De Clercq & Housen 2017; Lahmann et al. 2019; Spoelman & Verspoor 2010; Verspoor et al. 2008).

The measures employed in the present study are two length measures (MLNP and MNPDep) and one weighted measure (WNPCx) previously employed in Bulté (2013) when measuring the development of complexity in L2 English. For the present study another length measure, MLNPm, was added in order to attempt to gauge NP complexity in regard to the morphology within the NP. The measures are described in more detail below.

MLNP, i.e. the number of words per noun phrase, is a complexity measure of length, where the number of words per noun phrase is divided by the number of noun phrases:

$$\frac{\#Words}{\#NPs}$$

(21) *efter* [*en angenäm* ***vandring*** *över* [*vidsträckta* ***vidder***]] from NS.Z4[12]
(after a pleasant hike over widespread open spaces)

Words: 6 + 2 = 8
NPs (marked with square brackets, the head in bold): 2
MLNP: 8/2 = 4

The preposition *efter* 'after' is not counted as a word of the NP, which therefore includes six words. In total, there are two NPs in (21), the superordinated NP with *vandring* 'hike' as the head of the NP, which in turn contains an embedded NP with *vidder* 'open spaces' as the head of the embedded NP.

An alternative complexity measure of length, presented in Bulté (2013), is MNPDep, i.e. the number of dependents per noun phrase head, where the noun phrase dependents are divided by the number of noun phrase heads. Within the frame of this study, everything within the noun phrase in addition to the head is considered a dependent, e.g. an article is a dependent, as well as a subordinate clause.

#NP dependents
#NP heads

(22) *efter* [*en angenäm* ***vandring*** *över* [*vidsträckta* ***vidder***]]
(after a pleasant hike over widespread open spaces)

NP dependents: 3 + 1 = 4
NP heads (in bold): 2
MNPDep: 4/2 = 2

In example (22) above, the superordinated NP contains three dependents, namely, two pre-dependents, i.e. the indefinite article *en* 'a' and the adjective *angenäm* 'pleasant', and one post-modifier, i.e. the PP *över vidsträckta vidder* 'over widespread open spaces'. The NP embedded in the PP includes one pre-modifier in the shape of an adjective *vidsträckta* 'widespread'. More details of the calculation of MLNP and MNPDep is shown below in Table 2.

[12] Native speaker in the ASU corpus

Table 2. Calculation of phrasal complexity measures (MLNP and MNPDep)

Text sample	**NP**	**words**	**dependents**
*[Min nya **bil**]* 1 NP with 2 pre-dependents	1	3 (mean = 3)	2 (mean = 2)
*[**stugor** med [många vackra **fönster**]]* 1 NP with a post-modifier, a PP with an embedded NP, with 2 pre-dependents	2	5 + 3 = 8 (mean = 4)	1 + 2 = 3 (mean 1,5)
*på [en lång **stig** med [**träd** och stora gröna **buskar**]]* 1 NP with 2 pre-modifiers and 1 post-modifier embedded in a PP, this NP contains a compound NP with 2 heads, one without a dependent and one with 2 dependents	3	9 + 5 = 14 (mean ≈ 4,7)	3 + 0 + 2 = 5 (mean ≈ 1,7)
*[den långa **vägen** som gick till [den gamla **skolan**]]* 1 NP with 2 pre-dependents and 1 post-modifier, an embedded clause, containing a PP with an embedded NP, with 2 pre-dependents	2	9 + 3 = 12 (mean = 6)	3 + 2 = 5 (mean = 2,5)

After Bulté (2013:127)

The third measure suggested by Bulté (2013), WNPCx – weighted noun phrase complexity, is a measure where the type of dependent is considered. As mentioned in Chapter 2, this measure came about as an attempt to overcome two problems regarding the previous measures. The first one being that when counting the number of words of an NP, all constituents are rewarded the same weight without regard of their linguistic nature (Bulté 2013:127). Since embedded clauses can more or less be infinite in length, and the make-up of such clauses will decide the length of the NPs, this can cause problems when calculating phrasal complexity. The second problem according to Bulté, is that even if counting the number of dependents instead of words per phrasal head overcomes the just mentioned problem with the inherent make-up of embedded clauses when calculating the phrasal complexity, the same weight would be rewarded an article as an embedded clause. Bulté is hesitant that it would do justice to the inherent phrasal complexity, and suggested the WNPCx in which different weight is rewarded to NP dependents as well as to NPs that are structurally different. Table 3 displays how the weight is distributed in WNPCx.

Table 3. NP weight

Weighted NP complexity measure		
All immediate NP dependent	[*blåa* ***bilar***[13]]	1 mark
NP in PP	*under* [***huset***]	0,5 extra mark
Compound NP	[***träd*** *och* ***buskar***]	1 extra mark
NP embedded in NP	[***vägen*** *till* [*den gamla* ***skolan***]]	2 extra marks
Embedded clause	[***vägen*** *som går dit*]	3 extra marks

After Bulté (2013:127)

Bulté (2013) present a formula for calculating WNPCx, shown below. For the formula the number of dependents are summed up and if the NP contains any of the constructs presented in Table 3 these are rewarded marks accordingly, i.e. 0,5 mark times the number of NP in PP, 1 mark times the number of coordinated NPs (or compound NPs in Bulté 2013), 2 marks times the number of NP in NP, and 3 marks times the number of clauses on NP. The marks are summed up and divided by the number of NP heads. How to calculate weighted NP complexity is further presented in Table 4.

$$\frac{\#NP\ dependents + (0{,}5*\#NP\ in\ PP) + (1*\#Compound\ NPs) + (2*\#NP\ in\ NP) + (3*\#clause\ in\ NP)}{\#NP\ heads}$$

Table 4. Calculation of weighted NP complexity

Text sample	**Weighted NP complexity**
Min nya ***bil*** 1 NP with 2 pre-dependents	2 dependents, 1 nominal head ⇨ 2/1 = 2
*[****stugor*** *med [många vackra* ***fönster****]]* 1 NP with a post-modifier, a PP with an embedded NP, with 2 pre-dependents	3 dependents, 1 PP, 1 embedded NP ⇨ (3 + 0,5 + 2)/2 = 2,75
på [en lång ***stig*** *med [****träd*** *och stora gröna* ***buskar****]]* 1 NP with 2 pre-modifiers and 1 post-modifier embedded in a PP, this NP contains a compound NP with 2 heads, one without dependents and one with 2 dependents	5 dependents, 2 PPs, 1 compound NP, 1 embedded NP ⇨ (5 + (2x0,5) + 1 + 2)/3 = 3
[den långa ***vägen*** *som gick till [den gamla* ***skolan****]]* 1 NP with 2 pre-dependents and 1 post-modifier, an embedded clause, containing a PP with an embedded NP, with 2 pre-dependents	5 dependents, 1 PP, 1 embedded NP, 1 embedded clause ⇨ (5 + 0,5 + 2 + 3)/2 = 5,25

After Bulté (2013:128)

[13] Translation: *blåa bilar* 'blue cars', *under huset* 'under the house', *träd och buskar* 'trees and bushes', *vägen till den gamla skolan* 'the road to the old school', *vägen som går dit*, 'the road that leads there'.

Example (23) contains NP in PP and NP in NP, and example (24) contains NP in PP, an embedded clause and NP in NP. A more detailed description of how the marks are distributed along with a description of how the measure is calculated follows below.

(23) *efter* [*en angenäm* ***vandring*** *över* [*vidsträckta* ***vidder***]]
(after a pleasant hike over widespread open spaces)

(4 NP dependents + (2x0,5 NP in PP) + (2 NP in NP[14])/2 NP heads = 3,5

In example (22) previously, it has already been established that there are four dependents, which yields 4 marks. It also contains two NPs embedded in PPs, which yields 1 mark (2x0,5), and one NP embedded in another NP, which yields an additional 2 marks. In total, 7 marks. This score is divided by the number of heads, which in this case is 2, i.e. 7/2 = 3,5.

(24) [*mycket* ***kunskap*** *som är samlat på* [*ett* ***ställe***]]
(much knowledge which is gathered in one place) from NS.Z2[15]

(3 NP dependents + (0,5 NP in PP) + (2 NP in NP) + (3 embedded clause))/2 NP heads = 4,25

Example (24) contains 3 dependents (3 marks), one NP embedded in a PP (0,5 marks), one embedded NP (2 marks), and an embedded clause (3 marks). In total 8,5 marks, which are divided by the number of heads, in this case 2, i.e. 8,5/2 = 4,25.

The measure of phrasal complexity through morphology, MLNPm, i.e. the number of morphemes per noun phrase, added in the present study. The motivation behind MLNPm is that if a learner were to first produce noun phrases as in (25) and then elaborate the phrase as in (26), then the progression of the morphology would be possible to detect. MLNP would not detect this progression. It would in both examples be nine words, while for MLNPm there are nine morphemes in (25) and fourteen in (26).

(25) han klippa gräs och ge till häst i hage
(he cut grass and give to horse in paddock)

[14] Note that 2 refers to two marks for NP in NP and then all the marks are divided by the two NP heads.
[15] Native speaker in the ASU corpus

(26) han klipper gräset och ger till hästarna i hagen
(he cuts the grass and gives to the horses in the paddock)

MLNPm is also a length measure, where the number of morphemes per noun phrase is divided by the number of noun phrases:

Number of morphemes
NP heads

(27) *efter* [*en angenäm* ***vandring*** *över* [*vidsträckt-a* ***vidd-er***]]
(after a pleasant hike over widespread open spaces)

Morphemes: 8 + 4 = 12
NPs (marked with square brackets, the head in bold): 2
MLNPm: 12/2 = 6

In (27) there are eight morphemes in the superordinated NP with *vandring* as head, and four in the NP in NP with *vidder* as head.

In the following examples, some additional considerations regarding the analysis and calculation of noun phrase complexity are presented and explained. Examples (28) and (29) both have in common that the number of dependents is decided by how many potential errors the learner could make. Only the surface structure is analysed.

(28) *de* ***raka*** *och* ***långa*** *vägarna*
(the straight and long roads)

raka and *långa* are counted as two dependents since it is possible for the learner to make two mistakes concerning the attributive agreement.

(29) *de långa* ***stigarna*** *och* ***vägarna***
(the long paths and roads)

In example (29) there are two heads, i.e. *stigarna* and *vägarna* and two pre-dependents, i.e. *de* and *långa*, and even though these dependents describe both heads they are only counted once

since only the surface structure is analysed within the present study. The learner can only make one mistake concerning the attributive agreement.

(30) *de stora **husen** som ligger vid **sjön** och som ligger mer avskilt*
(the big houses that are by the lake and that are more secluded)

The superordinate NP *husen* in example (30), is post-modified by two separate relative clauses, which yield two different dependents. The calculation of the MNPDep and the WNPCx would hence be performed as follows:

NP dependents: *de*, *stora*, *som ligger vid sjön*, and *som ligger mer avskilt* = 4 dependents
NP heads: ***husen***, ***sjön*** = 2 heads
MNPDep: 4/2 = 2
WNPCx: (4 + 0,5 + 2 + (3x2 clause in NP[16]))/2 NP heads = 6,25

(31) *varje svensk **tidning** både rikstäckande och lokal.* from P1S111
(every Swedish newspaper both nationwide and local)

Rikstäckande and *lokal* are counted as two (post-modifying) dependents for the same reason as in (29) above, i.e. it would be possible for the learner to make two mistakes concerning the agreement. *Rikstäckande* is a participle, and in this case it does not change form if the head *tidning* would be in plural as the adjective *lokal* would do, i.e. *lokal-a*. However, since participles to a great extent inflect and have the same syntactic function as adjectives they are treated and counted as such (SAG 2 Ord: Teleman et al. 1999:582).

MNPDep: 4/1 = 4
WNPCx: 4/1 = 4

(32) *meddelander om föda eller döda* from C1S071
(messages about born and dead)
(33). *meddelanden av föda och […] av döda* from C1S071
(messages of born and of dead)

[16] Note that 3 in 3x2 refers to the number of marks which is rewarded an embedded clause and 2 the number of embedded clauses.

In example (32) the preposition *om* is shared by *föda* and *döda*, which belong to the same coordinated NP. The coordinated NP is considered as one dependent.

MNPDep: 1/3 = 0,33
WNPCx: (1 + 0,5 + 1 + 2)/3 = 1,5

In the following example (33), the superordinate NP *meddelanden* is followed by two separate PPs and hence they are two separate dependents.

MNPDep: 2/3 = 0,67
WNPCx: (2 + (2x0,5) + 1 + 2)/3 = 2

Had the subordinate NP of *meddelanden* in turn contained another embedded NP then the learner would have been rewarded another 2 marks for NP in NP, but when the superordinate NP contains two, or more embedded NPs on the same level as in example (33) then the learner is not rewarded more than 2 marks for all of them. This is because the NP does not become more complex, and if the superordinated NP contains coordinated NPs then the learner already is rewarded 1 mark for the compound, as discussed previously.

Chapter 4 Results and discussion

This chapter presents the results of the analysis presented in Chapter 3 and discusses them in relation to the research questions, as well as to previous research. First the results regarding whether or not the measures do correlate also for L2 Swedish are presented. Secondly, the results of the MLNPm and whether this measure can present a different picture than MLNP are presented. Lastly, the results regarding development of the complexity of the NP over time are presented, both at individual and group level, as well as how the L2 learners performed in relation to the benchmark L1 speakers. The results are then discussed in relation to the research questions and previous studies in 4.4.

4.1 Correlation between the measures in L2 Swedish

In order to see if there is a correlation between the measures employed in the present study, a pairs plot was created, Figure 2. All four measures were included, and the pairwise relationships between the variables in the scatterplots displayed a very strong correlation between all four measures.

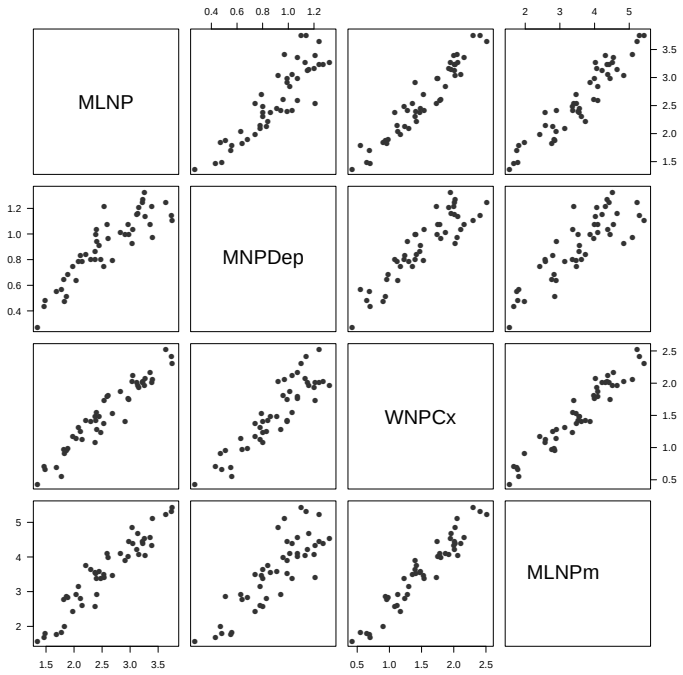

Figure 2. Pairs plot - correlation of all four measures

In the pairs plot, which allows for a view of both the distribution of single variables and the relationship between variables, the names of the variables are displayed along the diagonal

boxes and all other boxes show a scatterplot of the relationship between each of the pairwise combinations of MLNP, MNPDep, WNPCx and MLNPm. The top right box of the matrix shows a scatterplot of values for MLNP and MLNPm. The box below displays the same for MNPDep and MLNPm, and so on.

4.2 MLNP and MLNPm group scores

The patterns of MLNP and MLNPm, displayed in Figure 3, follow more or less the same paths. Figure 3 shows that the scores for MLNPm are higher than those of MLNP, which follows naturally from that the learners are using more inflected words.

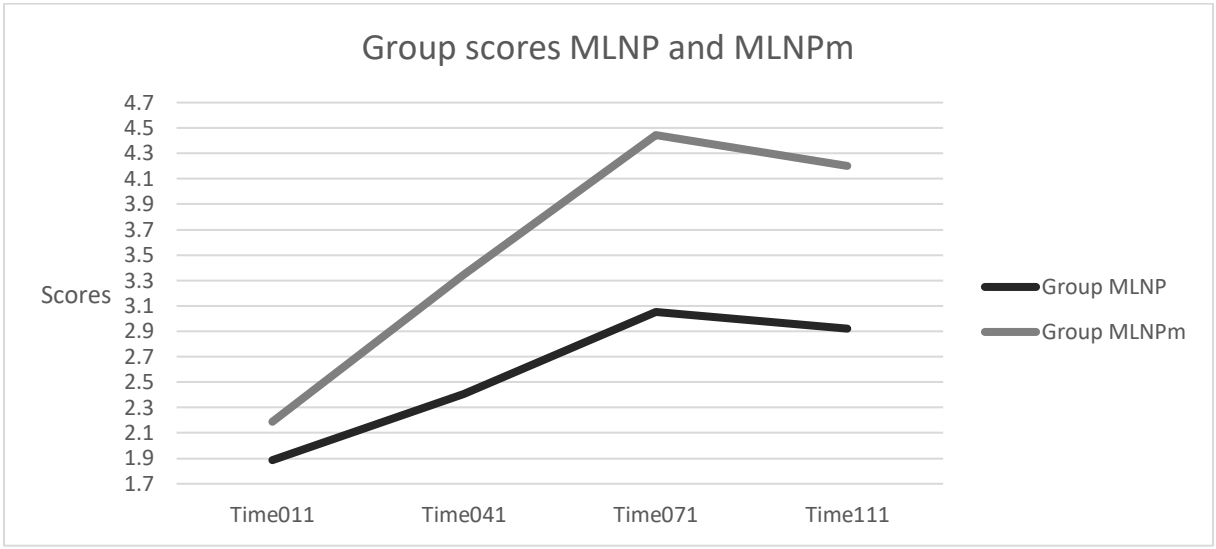

Figure 3. MLNP and MLNPm scores for all measurement points

4.3 Development of phrasal complexity - individual and group scores

The results regarding the correlation between the measures, as well as the results related to RQ2, i.e. if MLNPm can provide a different picture than MLNP, have been presented above. Hence, it is time to examine the outcome of the analysis of the measures presented in Bulté (2013), i.e. MLNP, MNPDep, and WNPCx, as well as the recently added MLNPm, in terms of

development. First, the results of the MLNP are presented. As we can see, C4, G3, P1, Q1 and S1 follow an altogether linear pattern. C1, C2 and Q2 follow the same pattern except for the last measurement time (Time111), where they provide a lower complexity score than at the previous measurement point, and G2 also follows a linear pattern from Time011 to Time071, but has a marginally lower score (-0.02 points) at Time111 than at Time041, while E2 displays a more irregular and non-linear pattern. E2 has a higher score at the first measurement point,

Time011, than at the last measurement point, Time111, and has the highest score at the second measurement point, Time041, followed by the second highest score at the third measurement point, Time071. E2 has the highest score of all the learners at Time011, but is also the learner who deviates most from the other participants in three of the four measures. See Figure 9 in Appendix 1 for E2 performance in all four measures.

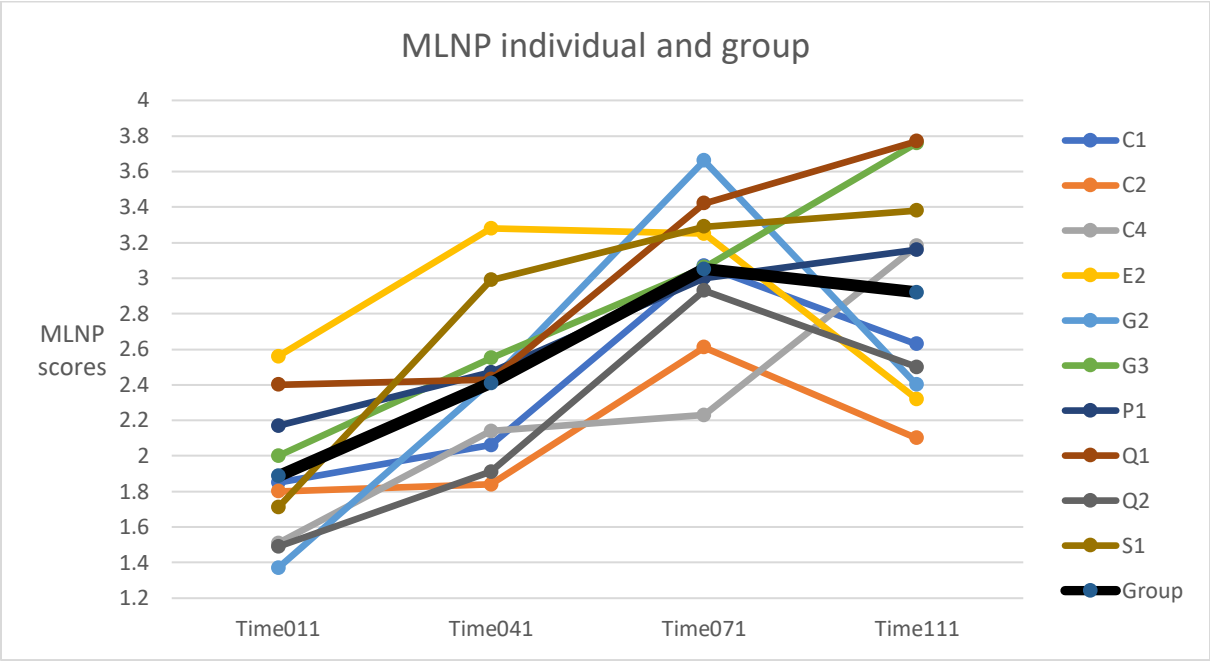

Figure 4. Individual pathways and group mean for MLNP

At group level, overall linear tendencies can be discerned for MLNP, at least between Time011 and Time071. There is a decline from the third to the last measurement point for five of the learners.

For MNPDep fewer than half of the participants (C4, G3, P1 and Q1) do follow an altogether linear pattern. S1, who follows a linear pattern for all the other measures, has a marginally lower score (0.06 points) at Time111 than at Time071 for MNPDep. C1, C2 and Q2 also follow a linear pattern up until the last measurement point, see Figure 5. G2 does, however, follow a linear pattern from the first to the third measurement point, but has a lower score for Time111 than s/he has for Time041, just as s/he displayed for MLNP, however, the difference is more noticeable for MNPDep. See Figure 10 in Appendix 1 for G2 performance in all measures. E2, again follows a more irregular pattern with greater fluctuations than the other learners. The

lowest score is found at Time111, the second lowest at Time011, the highest at Time041, and the second highest at Time071.

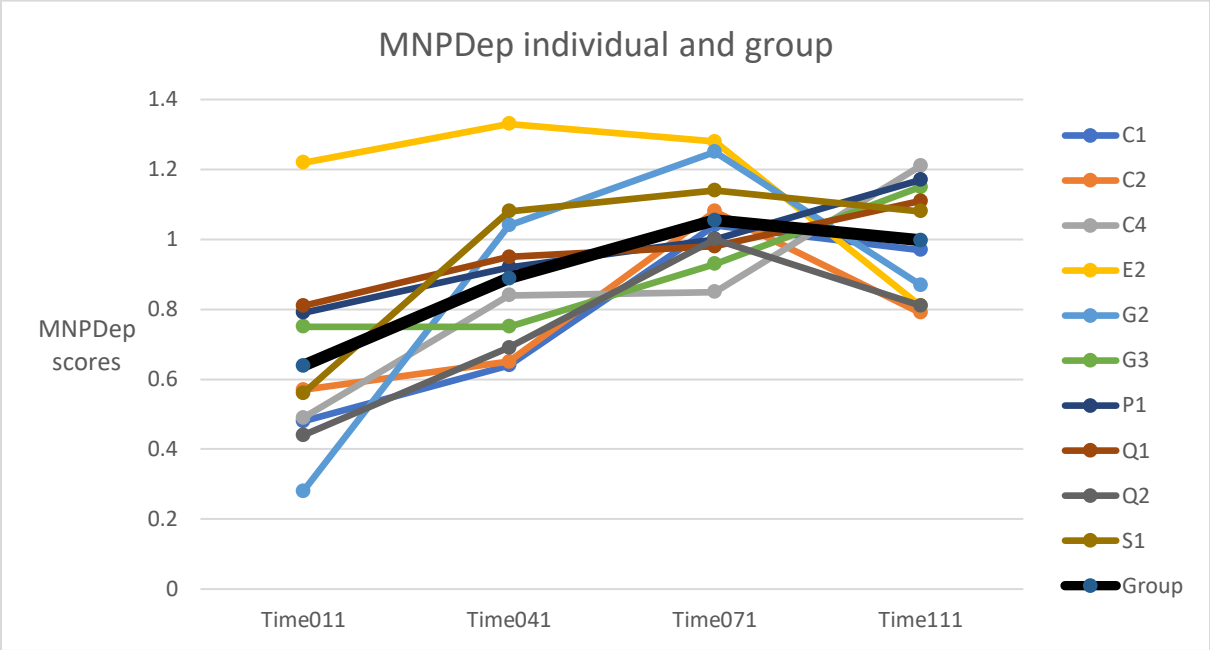

Figure 5. Individual pathways and group mean for MNPDep

Also for MNPDep, there is an overall tendency for a linear development of the complexity of the NP, at least from Time011 to Time071. There is a decline from the third to the last measurement point for six of the learners.

For the weighted complexity measure, WNPCx, five out of ten participants (C4, G3, P1, Q1 and S1) follow a linear pattern, the other five follow a linear pattern until the last measurement point, see Figure 6. Once again G2 also follows a linear pattern up until Time071 but s/he has a slightly lower score (-0.07 points) at the last measurement point than at the second measurement point. For WNPCx, E2 does follow a linear pattern from Time011 to Time071, but s/he has a lower score at Time111 than during all the previous measurement points.

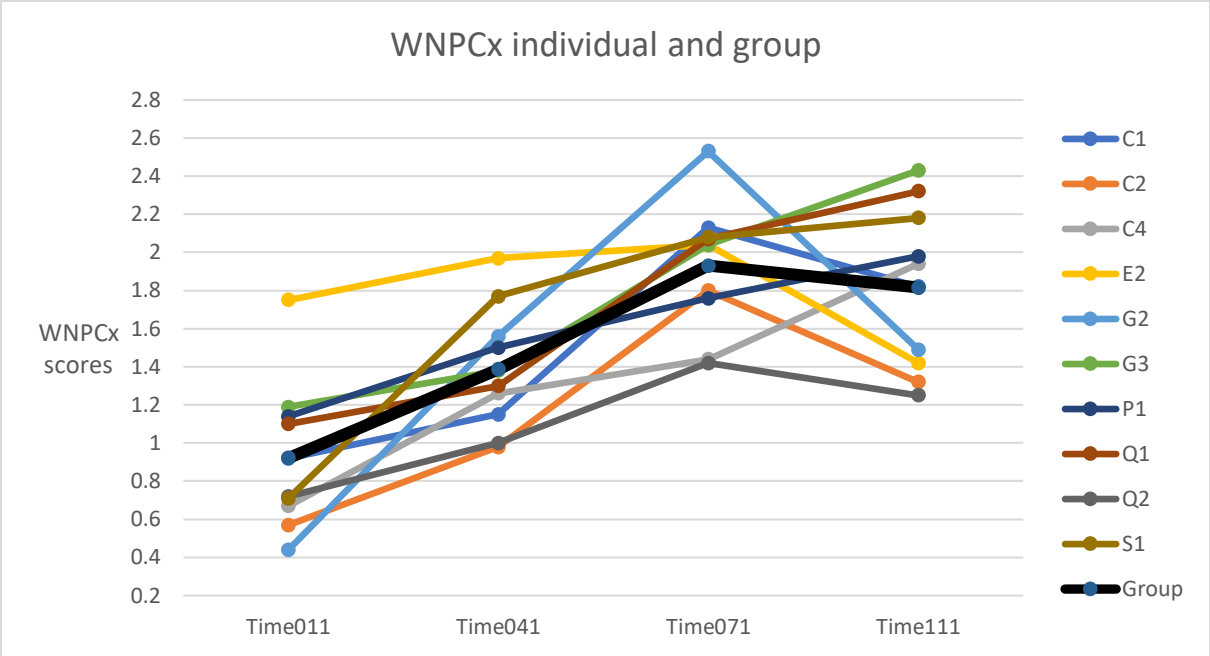

Figure 6. Individual pathways and group mean for WNPCx

Also for WNPCx, there is an overall tendency for a linear development of the complexity of the NP, at least from Time011 to Time071. There is a decline from the third to the last measurement point for five of the learners.

Turning to MLNPm, half of the participants, (C4, G3, P1, Q1 and S1) follow a linear pattern, while four (C1, C2, G2 and Q2) do so from the first to the third measurement point, but have lower complexity score for the last measurement point, see Figure 7. Only one learner, E2, follows an altogether non-linear pattern with fluctuating scores.

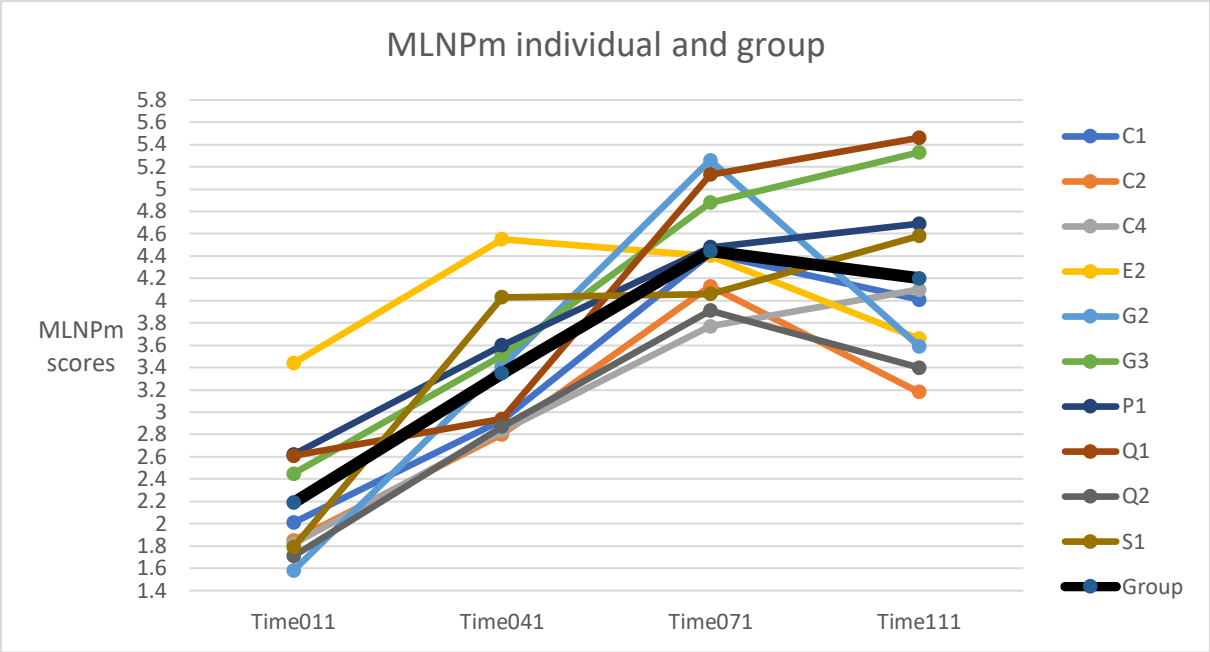

Figure 7. Individual pathways and group mean for MLNPm

On the group level, there is a straight linear development from Time011 to Time071 and a decline from the third to the last measurement point.

For both MLNP and MLNPm C4, G3, P1, Q1 and S1 follow a linear pattern, C1, C2, G2 and Q2 follow a linear pattern from Time011[17] to Time071, as displayed in Figures 4 and 7, and E2 follows non-linear pattern. For MLNP, G2 has a marginally lower score (-0.02 points) at Time111 than at Time041. However, this difference is very small and G2 still follows a linear pattern through Time011 to Time071 for both measures. E2 displays irregular individual pathways. This holds true for both measures. See Figure 9 and 10 in Appendix 1 for E2 and G2 respectively.

A *t*-test comparing L2 at Time011 and Time111 was carried out for the four measures. For MLNP $t(9)=5.12$; $p<0.001$, for MNPDep $t(9)=3.70$; $p<0.001$, for WNPCx $t(9)=5.50$; $p<0.001$, and for MLNPm $t(9)=7.85$; $p<0.001$, i.e. the difference between Time011 and Time111 was significant for all four measures.

[17] As noted in 3.2.1, material from S1, S4, S7, and S11 was selected for the present study. At these sessions, two texts from each informant were collected, and the last number of the measurement point refers to if it is text 1 or 2 of the data collection point. Hence, Time011 means that it is the first session (01), and 1 means that the text analysed is the one referred to in the corpus as text 1 of that session. Time011 is also the first measurement point in the present study. The second measurement point is Time041, the third is Time071, and Time111 is the fourth and last measurement point.

Figure 8 shows the curvilinear relationship between complexity scores and time. For all four measures a significant curvilinear effect of time on NP complexity is found. As can be noticed, both in Figure 8, as well as in the diagrams of all four measures, there is a tendency of a decline for about half of the learners at Time111 for all measures. This decline will be discussed further in 4.4. However, there is a general trend of development of the noun phrase complexity over time.

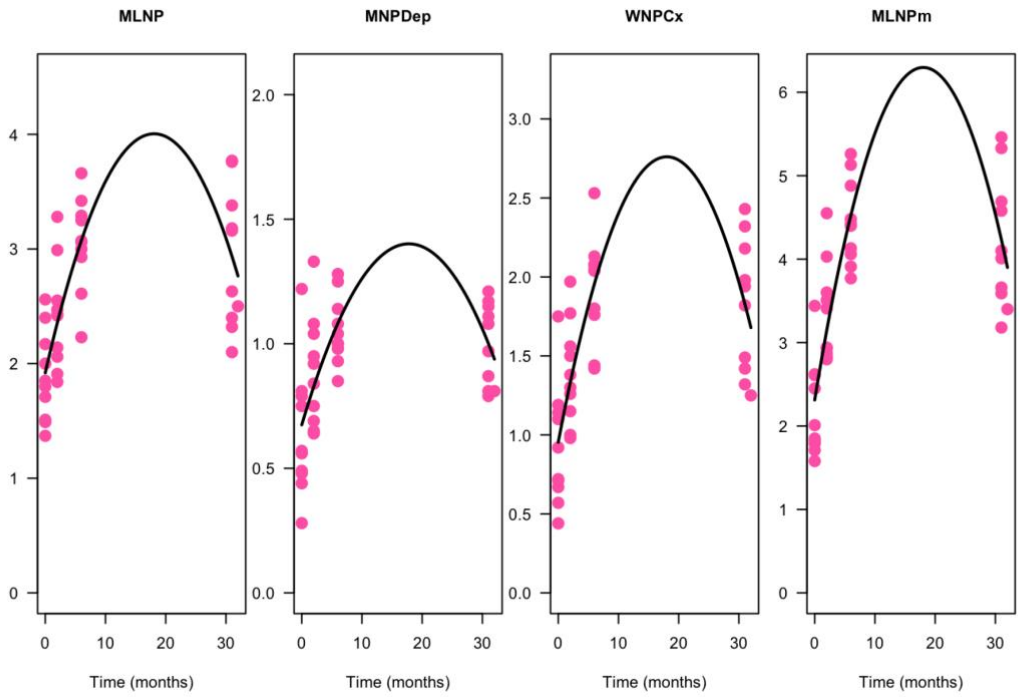

Figure 8. Curvilinear effect of time on NP complexity

4.3.2 L2 and L1

As mentioned in 3.2.2, native speaker data were included since a noun phrase can only be elaborated to a certain degree, and hence the L1 speakers are considered as a kind of benchmark. For a comparison of L2 and L1, two *t*-tests were performed for each measure. The first is a *t*-test comparing L1 to L2 at the first measurement point, i.e. Time011, the second *t*-test compares L1 to L2 at the last measurement point, i.e. Time111. The comparison in the first was significant for all four measures (i.e. the difference between L1 and L2 in complexity scores was significant for this test): for MLNP $t(10.58)=3.91$; $p<0.001$, for MNPDep $t(13.20)=2.76$; $p<0.050$, for WNPCx $t(12.58)=3.97$; $p<0.001$, for MLNPm $t(13.16)=5.94$; $p<0.001$, while the comparison of L1 to L2 at Time111 was not significant for any of the measures. This would suggest that there is a development of the noun phrase complexity for the L2 learners over time.

4.4 Discussion

In this section the findings related to the three research questions, repeated below, are discussed in light of the previous research and theoretical background presented in Chapter 2. Beginning by focusing on the first research question and the strong correlation between the measures presented in 4.1, continuing by discussing the results concerning if MLNPm would provide a different picture of noun phrase complexity than MLNP, and lastly the results concerning the third research question are addressed.

- Do the three previously tested and strongly correlating measures for L2 English also correlate for L2 Swedish, i.e. do they measure the complexity of the noun phrase in other second languages than English?
- Would a measure that takes the morphology of the noun phrase into consideration present a different picture of noun phrase complexity than the more commonly employed MLNP when gauging complexity in morphologically richer languages, such as Swedish?
- Is there a development of the complexity of the noun phrase in written L2 Swedish at individual and group level?

The very strong correlation for all four measures, presented in Figure 2, is in line with the results for MLNP, MNPDep, and WNPCx in Bulté (2013). This correlation would verify that the measures do gauge noun phrase complexity and suggest that the measures are applicable when investigating noun phrase complexity in L2 Swedish as well. The fourth measure MLNPm, added in this study, correlates very strongly with the previous three measures. The strong correlation among the measures can be seen as a verification that they all measure the same component of complexity (Bulté 2013:208), and in that case MLNPm would also be suitable for measuring noun phrase complexity, at least in L2 Swedish. Since the measures all correlate and measure the same thing it would be wise to decide on which measures to move forward with in order not to use several measures that measure the same phenomena, as discussed by Norris and Ortega (2009). Of the three previously tested measures, MLNP is by far the least time consuming and most frequently employed, but as discussed by Bulté (2013), it gives the same weight to all dependents. On the other hand, WNPCx is much more complicated and time-consuming, and places higher demands on the analysis. It could potentially provide different results depending on the rater and the decisions made as to what to include, and how to analyse

specific grammatical constructions. Which phenomena is of interest would of course determine which measure to employ. The present study does, however, not include any exclusion of measures, but rather an investigation as to if they can be used to detect a development of noun phrase complexity in other L2s than L2 English, in this case L2 Swedish, and they do seem to do just that.

Research on morphological complexity within the frame of CAF, has so far been scarce, at least when considering noun phrase complexity. As mentioned previously, MLNPm correlates strongly with the other three measures and can therefore be considered to be used to measure noun phrase complexity. In this study MLNPm did not provide a different picture than that of the MLNP, but it is however likely that the measure would show differences in the development of noun phrase complexity in learners that have not come as far in using inflectional morphemes as the learners in this study had. Figure 3 displays consistently higher scores for MLNPm than for MLNP. This could perhaps lead one to conclude that MLNPm provides a different picture, but since there is an increase in both measures, it is more likely that the higher scores for MLNPm is due to the learners using more inflected words. A measure that takes the morphology into consideration could potentially reflect complexity which would be excluded using the MLNP, but perhaps the MLNPm should be used in less advanced learners. Furthermore, MLNPm would need to be further researched and tested on more L2s, and L2s with different morphology from English and Swedish. It would be interesting to test MLNPm on languages with richer morphology of the NP than English in order to get more interesting results than the ceiling effect often found in morphological complexity studies on English.

MLNPm is somewhat more time consuming than MLNP since one has to consider what type of morpheme the learner uses and so on. MLNP is perhaps more straight forward in regard to definitions of a word than what to include and not to include for an analysis of the morphology of the NP. If it is possible to calculate morphemes/phrase automatically then the time consuming factor of MLNPm would be eliminated and the measure could be an interesting measure to apply on L2 production and the assessment thereof. As mentioned previously, perhaps a measure for less advanced learners than those of the present study. Teachers need easy and non-time consuming tools to aid them in the assessment of L2 production. However, MLNPm would need more research before it would be aiding teachers.

The third research question concerns if there is a development over time of the noun phrase complexity in the written L2 production of the ten learners analysed, both at individual and group level. The *t*-tests performed on the first and last measurement point provide evidence suggesting that there is a development of NP complexity over time, and the results of the *t*-tests are, as mentioned in 4.3, significant. The result of the *t*-test carried out on L1 and the last measurement point of the L2 production was not significant and the result could suggest that the L2 learners have moved closer to the L1 norm, also indicating a development. The diagrams of both the individual pathways, as well as the group scores, also suggest that the noun phrase complexity develops over time. The significant effect of time on noun phrase complexity shown in Figure 8 corroborates the findings for MLNP, MNPDep and WNPCx in Bulté (2013). An overall increase was found in Bulté (2013), where measures displayed a clear upward trend at group level from beginning to end, however not a smooth trajectory on an individual level. An overall increase for MLNP was also found in Bulté and Housen (2018). The results of the present study also correspond to the results of these studies. In his study, Bulté had eleven data points, while the present study, due to space and time limitations, only has four. If the present study had more data points, there would perhaps also be more variation at the individual level with more peaks and drops. This is a mere speculation at this point, but it would be something to keep in mind for future research.

Results from previous studies on L2 complexity, e.g. Spoelman & Verspoor (2010) and Bulté and Housen (2018), have suggested that development occurs in non-linear manners, rather than following predictable linear trajectories, at least at an individual level. For all learners but E2, there is at least a linear increase in the complexity scores of all four measures from Time011 to Time071. E2 has a decline from Time041 to Time071 for all measures except WNPCx, where E2 has an increase of 0,07. For MLNP and MNPDep, the decline is even smaller, varying from 0,03-0,05. For MLNPm the decline is 0,15. E2 has lower scores for Time111 than s/he had at earlier measurement times. Perhaps could the fact that E2 had the highest scores of all at the first measurement point play a significant role in that E2's development does not appear as clearly as for the others after Time041.

Five of the participants follow an altogether linear increase in complexity scores, with the one exception of S1 who had a lower score at Time111 than at Time071 for MNPDep. However, the decline for the other five learners, which is quite noticeable in Figure 8, as well as in the diagrams of each measure, cannot be ignored. The written production collected at Time071 and

Time111 both concern the same topic: the family page of a daily newspaper, which could lead one to expect the learners to perform better due to some familiarity with the topic. This fact cannot be completely overseen when considering the results of those learners who do not have a decline in complexity scores from Time071 to Time111, but it does not need to be the sole reason behind their linear development either. But, as stated by Bulté (2013:209), it is more or less impossible "to completely eliminate and consequently rule out task effects (or learning effects when a task is repeated)". On the other hand, the written production of Time071 was collected a little more than two years earlier than the data for Time111, which also were collected after the L2 Swedish course was completed and the participants no longer had any L2 instructions. Bulté and Housen (2020) discuss that the variation across learners can be due to many different aspects not related to the developmental process per se, such as the already mentioned familiarity with topic, but also mood, fatigue, stylistic choice and measurement noise. It is also possible that the learners lean more towards using subordination than noun phrase elaboration for the task at hand. Verspoor and colleagues (see e.g. Verspoor et al. 2008 or Spoelman & Verspoor 2010) refer to this phenomena as a "competitive relationship" between noun phrase complexity and sentence complexity (e.g. subordination). Their results imply that development at one level comes at the cost of the other. Whether this is the case in the present study as well is not possible to conclude since only noun phrase complexity is investigated here, but it could in theory be held as an explanation as to why there is a decline in the noun phrase for some learners. For future research the measures applied here should be tested multidimensionally, i.e. together with measures of both coordination and subordination. Then it would be possible to investigate a potential competitive relationship between noun phrase complexity and sentence complexity. The results of MLNP presented in Bulté (2013) also displayed a decline[18] at the last measurement point, in his case both at group level, as well as for almost all the individual pathways, with the exception of one learner. In his study the participants still had instruction during the last measurement points, unlike in the present study. Bulté speculated that the decline in noun phrase complexity in his study could have been due to that the learners were asked to write somewhat different texts for the sessions concerned, making them possibly rely more on subordination than on noun phrase elaboration (B. Bulté, personal communication, 10 May 2021). Since the texts from Time071 and Time111 address

[18] In Bulté (2013:162–63), the decline in the last measurement point was also found in the two other noun phrase complexity measures examined, i.e. MNPDep and WNPCx.

the same topic and both texts are descriptive in nature, this would not be likely to be the case here. However, it could be that some learners in the present study were more inclined to use e.g. subordination at the cost of NPs. The same decline as in Bulté (2013) was also visible in the development of group and individual scores for the MLNP in Bulté and Housen (2018), which is based on the same data as Bulté (2013). In a recent study, Bulté and Housen (2020) found some evidence of a competitive growth between the two syntactic dimensions discussed above, and came to the conclusion that different complexity dimensions do not develop in parallel. To what extent the just mentioned factors had any impact on the outcome is difficult to conclude within the scope of this study, but they are important to bear in mind for future studies.

Bulté (2013) carried out a case study on two of the participants' MLNP scores: learner 120 had a, broadly speaking, increase throughout the study, while learner 109 performed more or less in the same way as E2. It is not stated whether 109, like E2, had the highest score of all the learners at the first measurement point. However, learner 109 also had the lowest score at the last measurement point, the second lowest at the first time, the highest score came somewhere in the middle of the data (at time 6 of 11 measurement points), and the second highest at the second last time. This non-linear pattern of individual learners, such as E2S and 109, supports the claim of DST researchers that on an individual level it is difficult to detect any clear developmental trajectories.

Spoelman and Verspoor (2010) recommend focusing on the individual developmental patterns and the variability displayed in them. From a DST perspective, it would be expected to find a high degree of variability between learners, but also more fluctuations, high peaks and sudden spurs within the individual developmental trajectories as discussed by e.g. Kuiken and Vedder (2019). However, the results of the present study do not provide any strong evidence of a non-linear development of noun phrase complexity, or the amount of variability on the individual level as would have been expected in relation to previous studies. There is some variability in the inter-individual development. Again, it could be speculated that the present study may not include enough data points to really detect such variation. As mentioned previously, DST studies suggest that there is a lot of variability within individual development trajectories, but a linear developmental trajectory can be present at group level. The present study presents, like Bulté (2013), Bulté and Housen (2018) and Verspoor et al. (2008), an upward trend on group

level over time. Bulté and Housen speculated that variability in learners could possibly be linked to the type of learners and learning profiles such as those discussed by Norrby and Håkansson (2007), and also discussed in Flyman Mattsson (2017) in terms of an individualised analysis of language where PT-level and level of complexity as well as accuracy are analysed.

The L1 group was, as mentioned earlier, included to function as a kind of benchmark or norm since the noun phrase cannot be elaborated unlimited. The *t*-test of L1 and Time111 was not significant. This could suggest that the L2 learners are getting closer to the L1 norm at the last measurement point, which would be fairly reasonable since there is a development of noun phrase complexity along the course of time.

Chapter 5 Conclusions

The thesis focused on the development of noun phrase complexity in L2 Swedish written production, an area in need of exploration. For the most part linguistic complexity research has been carried out on L2 English and more so on the development of complexity at the sentential level through subordination. This thesis set out to begin filling the void in complexity studies of L2 Swedish by investigating phrasal complexity, more specifically noun phrase complexity. The aim of the study was to establish whether or not a development of the complexity of the noun phrase in written L2 Swedish, i.e. an increase in noun phrase elaboration, can be observed. To provide an insight to said development, three phrasal complexity measures employed in Bulté (2013), namely two length measures: MLNP and MNPDep and a weighted measure: WNPCx, were applied. The measures strongly correlate for L2 English, but had not been tested on L2 Swedish. Therefore, the study began by testing MLNP, MNPDep and WNPCx for L2 Swedish. For this study a fourth measure MLNPm (morphemes/NP heads) was added to investigate if a measure taking the morphology into consideration would present a different picture than MLNP. Hence, the study continued by testing the MLNPm. Finally, by the application of the three previously (on L2 English) tested phrasal complexity measures, MLNP, MNPDep, and WNPCx, as well as a recently added measure MLNPm, noun phrase complexity was gauged in order to investigate if a development of the complexity could be detected at individual and group level.

All four measures correlated strongly, as the three previously tested measures also did in Bulté (2013), and they were considered to be applicable when measuring noun phrase complexity in L2 Swedish. The MLNPm correlated with the three previously tested measures, and was therefore also considered to measure noun phrase complexity in L2 Swedish. As for if MLNPm would present a different picture than MLNP, the measure did not. It seems as if the learners had come too far in their use of inflected morphemes, and no difference in the development of the two measures could be found. It is more likely that MLNPm could be a measure for earlier stages of NP complexity. More research on the morphological complexity of the noun phrase is needed, both on different L2s as well as L2s with a different morphology, i.e. preferably languages with a richer morphology of the noun phrase than English to avoid an early ceiling effect, as discussed by e.g. Bulté (2013). MLNP is less time-consuming to calculate than MLNPm, and time is valuable both to researchers and teachers. A way to calculate MLNPm

automatically and correctly would be desirable in order for the measure to be valuable and useful for teachers in assessing L2 development. However, that was beyond the scope of the study.

A development of noun phrase complexity was found for all learners, however it was not linear for all learners and they display some inter-individual variability. Nevertheless, linear or non-linear, a development of noun phrase complexity in the written production of the L2 Swedish learners of the study was detected. A decline from the third to the last measurement point was found for half of the learners, and possible reasons for this were discussed. One being that they no longer had any L2 instructions at the last session which was collected a little over two years after the third session. Another being what Verspoor and colleagues (see e.g. Verspoor et al. 2008, or Spoelman & Verspoor 2010) refer to as a competitive relationship between subordination and NP elaboration. Bulté (2013) also had a decline in his data and speculated that it could possibly be due to the learners producing different types of texts, which could affect whether they were more inclined to opt for subordination than elaboration of the NP. Since Bulté also detected a decline in his data, the lack of instruction of the learners in the present study does perhaps not hold as a plausible explanation.

The need to further investigate noun phrase complexity and the morphological complexity of the noun phrase has already been dealt with above. But as mentioned in the very beginning of this thesis, complexity should be measured multidimensionally in order to account for all aspects of the development of complexity. Therefore, the measures employed here should be employed in multidimensional studies in order to see if there is any competitive relationship between e.g. noun phrase complexity and sentence complexity, but also to examine not just noun phrase complexity but all aspects of structure complexity. If such a task was to be undertaken it would be suggested to not divide the material into clause elements as was done when investigating only noun phrase complexity within the frame of this study. Topic and task should also be accounted for in order to prevent the risk of task effects and/or learning effects if repeating the same task. It would be interesting to see the results of a multidimensional study, that has focused on eliminating as much as possible of a task effect, as well as on ruling out any possible effect of the learners no longer having any L2 instruction and see if there still would be such a decline as the one found in the present study.

References

Agebjörn, A. (2021). Learning of Definiteness by Belarusian Students of Swedish as a Foreign Language. Doctoral dissertation. Göteborg: Göteborg University.

Axelsson, M. (1994). Noun phrase development in Swedish as a second language. A study of adult learners acquiring definiteness and the semantics and morphology of adjectives. Doctoral dissertation. Stockholm: Stockholm University.

Bartning, I., Arvidsson, K. & Forsberg Lundell, F. (2015). Complexity at the phrasal level in spoken L1 and very advanced L2 French. *LIA Language, Interaction and Acquisition* 6(2), 181–201,

Baten, K. & Håkansson, G. (2015). The Development of Subordinate Clauses in German and Swedish as L2s. A Theoretical and Methodological Comparison. *Studies in Second Language Acquisition*, 37, 517–547.

Bailey, N., Madden, C. & Krashen, S. (1974). Is there a 'natural sequence' in adult second language learning? *Language Learning*, 24, 235–243

Biber, D. (1988). *Variation across speech and writing.* Cambridge: Cambridge University Press.

Brezina, V. & Pallotti, G. (2019). Morphological complexity in written L2 texts. *Second Language Research*, Vol 35(1), 99–119.

Brumfit, C. J. (1979). Communicative language teaching: An educational perspective. In: Brumfit, C. J. & Johnson, K. (eds.). *The Communicative Approach to Language Teaching*, 183–191. Oxford: Oxford University Press.

Brumfit, C. J. (1984). *Communicative methodology in language teaching*. Cambridge: Cambridge University Press.

Bulté, B. (2013). The development of complexity in second language acquisition. A dynamic systems approach. Unpublished doctoral dissertation, University of Brussels (VUB), Brussels, Belgium.

Bulté, B. & Housen, A. (2012). Defining and operationalising L2 complexity. In: Housen, A., Kuiken, F. & Vedder, I. (eds.) *Dimensions of L2 Performance and Proficiency. Complexity, Accuracy and Fluency in SLA.* Amsterdam: John Benjamins Publishing CO.

Bulté, B. & Housen, A. (2014). Conceptualizing and measuring short-term changes in L2 writing complexity. *Journal of Second Language Writing*, 26, 42–65.

Bulté, B. & Housen, A. (2015). Evaluating short-term changes in L2 complexity development. *Círculo de Lingüística Aplicado a la Comunicacíon* 63, 42–76.

Bulté, B. & Housen, A. (2018). Syntactic complexity in L2 writing: Individual pathways and emerging group trends. *International Journal of Applied Linguistics* 28, 147–164.

Bulté, B. & Housen, A. (2020). A DUB-inspired Case Study of Multidimensional L2 Complexity Development: Competing or Connecting Growers? In W. Lowie, M. Michel, A. Rousse-Malpat, M. Keijzer, & R. Steinkrauss (Eds.), : *Usage-based dynamics in second language development: In celebration of Marjolijn Verspoor* (pp. 50–86. Bristol, UK: Multilingual Matters.

Bulté, B. & Roothooft, H. (2020) Investigating the interrelationship between rated L2 proficiency and linguistic complexity in L2 speech. *System*, 91.

Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment.* Cambridge: Cambridge University Press.

Cumming, A. (1989). “Writing Expertise and Second Language Proficiency”. *Language Learning* 39(1), 81–135.
de Bot, K., Lowie, W., Thorne, S. L., & Verspoor, M. (2013). Dynamic Systems Theory as a comprehensive theory of second language development. In: del Pilar García Mayo, M., Junkal Gutierrez Mangado, M., & Martínez Adrían, M. (eds.) *Contemporary Approaches to Second Language Acquisition*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
de Bot, K., Lowie, W., Thorne, S.L. & Verspoor, M. (2013). Dynamic Systems Theory as a comprehensive theory of second language development. In: Garcia Mayo, MP., Gutiérrez Mangado, MJ. & Martínez-Adrian, M. (eds.) *Contemporary Approaches to Second Language Acquisition*. Amsterdam: John Benjamins.
de Bot, K. (2017). Complexity Theory and Dynamic Systems Theory. In: Ortega, L. & Han, Z. (eds.) *Complexity Theory and language Development: In celebration of Diane Larsen-Freeman*. Amsterdam: John Benjamins.
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, 2, 315–334.
De Clercq, B. & Housen, A. (2019). The development of morphological complexity: A cross-linguistic study of L2 French and English. *Second Language Research*, Vol 35, 1, 71–97.
Dulay, H. & Burt, M. (1973). Should we teach children syntax? *Language Learning*, 23, 245–258.
Dulay, H. & Burt, M. (1974). Natural sequences in child second language acquisition. *Language Learning*, 24, 37–53.
Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.
Ellis, R. (2008). *The study of second language acquisition* (2nd ed.) Oxford: Oxford University Press.
Ellis, R. (2015). *The Study of Second Language Acquisition*. (2nd edition). Oxford: Oxford University Press.
Ellis, R. & Barkhuizen, G. (2005). *Analysing Learner Language*. Oxford: Oxford University Press.
Flyman Mattsson, A. (2017). *Svenska som andraspråk på språkvetenskaplig grund*. Lund: Studentlitteratur.
Flyman Mattsson, A. & Håkansson, G. (2021). *Bedömning av svenska som andraspråk. En analysmodell baserad på grammatiska utvecklingsstadier*. [Assessment of Swedish as a second language. An analysis model based on grammatical developmental stages]. (2nd ed.) Lund: Studentlitteratur.
Graesser, A., McNamara, D, Louwerse, M. & Cai, Z. (2004). ”Coh-Metrix: Analysis of text on cohesion and language”. *Behavior Research Methods* 36(2), 193–202.
Granfeldt, J. & Ågren, M. (2013). Stages of processability and levels of proficiency in the Common European Framework of Reference for Language. The case of L3 French. In: Flyman Mattsson, A. & Norrby, C. (eds.) *Language Acquisition and Use in Multilingual Contexts*. Travaux de l'Institut de Linguistique de Lund; Vol. 52. Lund: Media-Tryck, Lund University.
Gyllstad, H., Granfeldt, J., Bernardini, P. & Källkvist, M. (2014). Linguistic correlates to communicative proficiency levels of the CEFR. The case of syntactic complexity in written L2 English, L3 French and L4 Italian. *EUROSLA Yearbook*, 14, 1–30.
Hammarberg, B. (2010). Introduktion till ASU-korpusen – en longitudinell muntlig och skriftlig textkorpus av vuxna inlärares svenska med en motsvarande del från infödda svenskar. [Introduction to the ASU-Corpus, – a Longitudinal Oral and Written Text Corpus of Adult Learners of Swedish with a Corresponding Part from Native Swedes]: Institutionen för lingvistik, Stockholms Universitet.
Hammarberg, B. (2013). Andraspråksforskning med ASU-korpusen. In: *Nordland - Nordisk tidskrift for andrespråksforskning*, 8(1):7–33.

Hammarberg, B & Olsson, L-J. (2010). Arbeta med ASU-korpusen. Partiell bruksanvisning till ITG:s användargränssnitt. [Working with the ASU-corpus. Partial instruction to the ITG user interface]. Stockholms universitet, Institutionen för lingvistik.

Hammerly, H. (1991). *Fluency and Accuracy: Toward Balance in Language Teaching and Learning.* Clevedon: Multilingual Matters.

Housen, A. & Kuiken, F. (2009). Complexity, Accuracy, and fluency in Second Language Acquisition. *Applied Linguistics*, Vol. 30, Issue 4, 461–473.

Housen, A., Kuiken, F. & Vedder, I. (2012). Complexity, accuracy and fluency: Definitions, measurements and research. In: Housen, A., Kuiken, F. & Vedder, I. (eds.) *Dimensions of L2 Performance and Proficiency. Complexity, Accuracy and Fluency in SLA*. Amsterdam: John Benjamins Publishing CO.

Johannessen, J.B., (1998). *Coordination*. Oxford Studies in Comparative Syntax. New York: Oxford University Press.

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.

Lahmann, C., Steinkrauss, R. & Schmid, M.S. (2019). Measuring linguistic complexity in long-term L2 speakers of English and L1 attriters of German. *International Journal of Applied Linguistics*. 29, 173–191.

Larsen-Freeman, D. (1975). The acquisition of grammatical morphemes by adult ESL students. *TESOL Quarterly*, 9, 409–430.

Larsen-Freeman, D. (1978). An ESL index of development. *TESOL Quarterly*, 12(4), 439–448.

Larsen-Freeman, D. (1997). Chaos/complexity science and second language acquisition. *Applied Linguistics* 18(2), 141–165.

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.

Larsen-Freeman, D. (2009). Adjusting expectations: The study of complexity, accuracy and fluency in second language acquisition. *Applied Linguistics*, 30(4), 579–589.

Lehtonen, P. (2015). Referens och artikelbruk på olika CEFR-nivåer hos finskspråkiga universitetsstudenter - en jämförelse mellan svenska och engelska. [Reference and article use at different CEFR-levels in Finnish speaking university students - a comparison between Swedish and English]. Dissertation Jyväskylä University, Jyväskylä, Finland.

Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning*, 40(3), 387-417.

Lu, X. (2010). "Automatic analysis of syntactic complexity in second language writing". *International Journal of Corpus Linguistics* 15(4), 474–496.

Lu, X. (2011). "A Corpus-Based Evaluation of Syntactic Complexity Measures as Indices of College-Level ESL Writers' Language Development". *TESOL Quarterly* 45(1), 36–62.

Norrby, C. & Håkansson, G. (2007). "The interaction of complexity and grammatical processability: The case of Swedish as a foreign language". *IRAL* 45(1), 45–68.

Norris, J.M. & Ortega, L. (2009). Towards an Organic Approach to Investigating CAF in Instructed SLA: The Case of Complexity. *Applied Linguistics*, Vol. 30, Issue 4, 555–578.

Nyqvist, E-L. (2013). Species och artikelbruk i finskspråkiga grundskoleelevers inlärarsvenska. En longitudinell undersökning i årskurserna 7–9. [Definiteness and article use in Finnish speaking primary students Swedish

learner language]. Doctoral dissertation Åbo Akademi University. Åbo, Finland.

O'Grady, W. (2015). Processing Determinism. *Language Learning*, 65:1, pp. 6–32.

Ortega, L. (2012). Interlanguage complexity: A construct in search of theoretical renewal. In B. Kortmann, & B Szmrecsanyi (Eds.), *Linguistic Complexity: Second Language Acquisition, Indigenization, Contact* (pp. 127–55. Berlin: de Gryter.

Ortega, L. (2013). Understanding Second Language Acquisition. New York: Routledge.

Pallotti, G. (2009). CAF: Defining, Refining and Differentiating Constructs. *Applied Linguistics*, Vol. 30, Issue 4, 590-–601.

Pallotti, G. (2015). A simple view of linguistic complexity. *Second Language Research*, 31, (1), 117–134.

Pienemann, M. & Johnston, M. (1987). Factors influencing the development of language proficiency. In D. Nunan (ed.), *Applying Second Language Research*, Adelaide: NCRC. 45–142.

Pienemann, M. & Mackey, A. (1993). An empirical study of children´s ESL development and *Rapid Profile*. In P. McKay (ed.), *ESL development: Language and literacy in Schools*. Vol. 2. 115–259. Canberra: National Language and Literacy Institute of Australia.

Pienemann, M. (1998). *Language processing and second language development: Processability Theory,* Studies in Bilingualism 15. Amsterdam: Johan Benjamins.

Pienemann, M. (ed.). (2005). *Cross-linguistic aspects of Processability Theory*, Studies in Bilingualism 30. Amsterdam: John Benjamins.

Pienemann, M. & Håkansson, G. (1999). A unified approach toward the development of Swedish as L2. A processability account. *Studies in second language acquisition* 21, 383–420.

Ravid, D. & Berman, R. (2010). Developing noun phrase complexity at school age: A text-embedded cross-linguistic analysis. *First Language*,30, 3–26.

Robinson, P. (ed.) (2001). *Cognition and Second Language Instruction*. Cambridge: Cambridge University Press.

Robinson, P. (2003). Attention and memory during SLA. In Doughty, C. & Long, M (eds.) *Handbook of Second Language Acquisition*. Malden, Mass.: Blackwell.

Robinson, P. (ed.) (2011). *Second Language Task Complexity: Researching the Cognition Hypothesis of Language Learning and Performance*. Vol. 2, Amsterdam/Philadelphia: John Benjamins Publishing.

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics*, 10, 209–231.

Skehan, P. (1996). Second language acquisition research and task-based instruction. In Willis, D. & Willis, J. (eds.) *Challenge and Change in Language Teaching*. London: Heinemann.

Skehan, P. (1998a). Task-based instruction. *Annual Review of Applied Linguistics* 18, 268–86.

Skehan, P. (1998b). A cognitive approach to Language Learning. Oxford: Oxford University.

Skehan, P. (2003). Task based instruction. *Language Teaching* 36(1), 1–14.

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, 532–553.

Teleman, U., Hellberg, S. & Andersson, E. (1999). 1st edition, 4th print. *Svenska Akademiens Grammatik*: *2 Ord.* Stockholm: Svenska Akademien.

Teleman, U., Hellberg, S. & Andersson, E. (1999). 1st edition, 4th print. *Svenska Akademiens Grammatik*: *3 Fraser*. Stockholm: Svenska Akademien.

Thelen, E. & Smith, L. (1994). *A Dynamic Systems Approach to the development of cognition and action*.

Cambridge, MA: The MIT Press.

VanPatten, B. (1993). Grammar instruction for the acquisition rich classroom. Foreign Language Annals, 26, 433–450.

Verspoor, M., Lowie, W. & van Dijk, M. (2008). ”Variability in Second Language Development From a Dynamic Systems Perspective”. *Modern Language Journal*, 92(2), 214–231.

Verspoor, M., de Bot, K. & Lowie, W. (2011). *A Dynamic Approach to Second Language Development Methods and Techniques*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Verspoor, M. & Behrens, H. (2011). Dynamic Systems Theory and a usage-based approach to Second Language Development. In: Verspoor, M., de Bot, K. & Lowie, W. (eds.) *A Dynamic Approach to Second Language Development Methods and Techniques*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Verspoor, M. Lowie, W, Chan, H. & Vahtrick, L. (2017).Linguistic complexity in second language development: Variability and variation at advanced stages. *Recherches en didactique des langues et des cultures*, *14*(1).

von Stutterheim, C. & Klein, W. (1987). A concept-oriented approach to second language studies. In C. W. Pfaff (Ed.), *First and second language acquisition processes* (pp. 191–205). Cambridge, MA: Newbury.

Wolfe-Quintero, K., Inagaki, S. & Kim, , H-Y. (1998). *Second language development in writing: Measures of fluency, accuracy, and complexity*. Honolulu, HI: University of Hawaii, Second Language Teaching & Curriculum Center.

Appendix 1 Learners E2 and G2

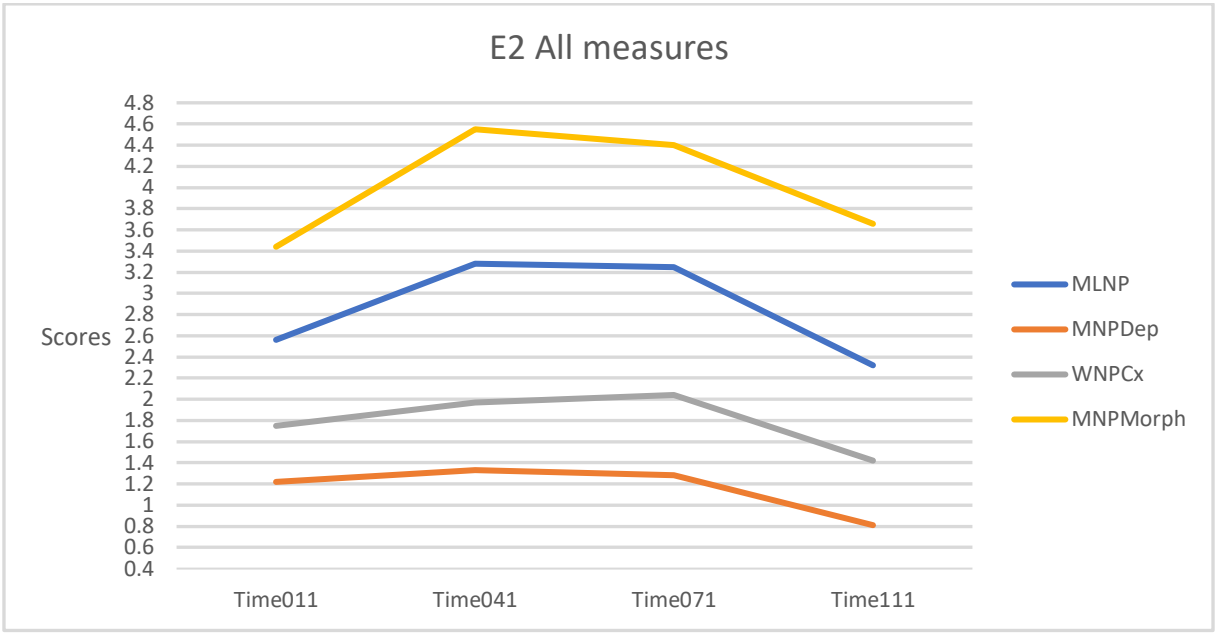

Figure 9. E2 all measures

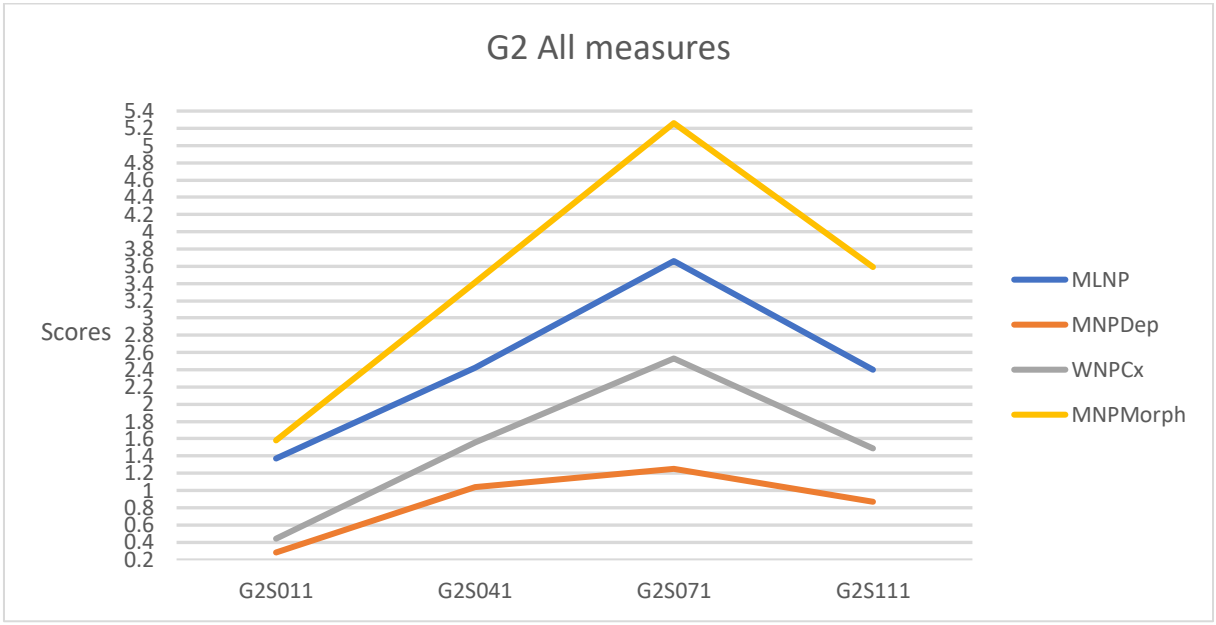

Figure 10. G2 all measures

Appendix 2 Group scores all four measures

Figure 11. Group scores for all measures