JYVÄSKYLÄ STUDIES IN HUMANITIES
233

Corinne Tilma

The dynamics of foreign versus second language development in Finnish writing

JYVÄSKYLÄN YLIOPISTO

JYVÄSKYLÄ STUDIES IN HUMANITIES 233

Corinne Tilma

The dynamics of foreign versus second language development in Finnish writing

Esitetään Jyväskylän yliopiston humanistisen tiedekunnan suostumuksella
julkisesti tarkastettavaksi Fennicum-rakennuksen salissa F205
marraskuun 28. päivänä 2014 kello 12.

Academic dissertation to be publicly discussed, by permission of
the Faculty of Humanities of the University of Jyväskylä,
in building Fennicum, auditorium F205, on November 28, 2014 at 12 o'clock noon.

UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2014

The dynamics of foreign versus second language development in Finnish writing

JYVÄSKYLÄ STUDIES IN HUMANITIES 233

Corinne Tilma

The dynamics of foreign versus second language development in Finnish writing

UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2014

Editors
Maisa Martin
Department of Languages, Applied linguistics, University of Jyväskylä
Pekka Olsbo, Timo Hautala
Publishing Unit, University Library of Jyväskylä

URN:ISBN: 978-951-39-5869-5
ISBN 978-951-39-5869-5 (PDF)
ISSN 1459-4331

ISBN 978-951-39-5868-8 (nid.)
ISSN 1459-4323

Jyväskylä University Printing House, Jyväskylä 2014

ABSTRACT

Tilma, Corinne
The dynamics of foreign versus second language development in Finnish writing.
Jyväskylä: University of Jyväskylä; Groningen: University of Groningen, 2014, 208 p.
(Jyväskylä Studies in Humanities,
ISSN 1459-4323; 233)
ISBN 978-951-39-5868-8 (nid.)
ISBN 978-951-39-5869-5 (PDF)

This longitudinal study from a dynamic perspective explores the development of beginner learners Finnish as a foreign (FL; four learners in the Netherlands; focus on grammar) and as a second (L2; four learners in Finland; focus on meaning) language. The study explores syntactic and morphological complexity and accuracy measures (free response written data; nine months) and looks for differences in outcomes between the two groups of learners (overall averages) and for differences in developmental patterns and interactions over time between measures (one focal learner per group).

The group study shows similarities in syntactic and morphological complexity (increase) during the time span for the FL an L2 groups but differences in case use and some related complexity and accuracy measures.

The focal learner study shows differences in measures related to cases between the focal L2 learner and the L2 group but many similarities between the two focal learners. However, the FL learner shows steeper peaks (use of the 12 other cases, case form errors), indicating a greater degree of development. Besides, the FL learner shows more complexity and accuracy early on; the focus on grammar helps to discover forms and use them from the beginning. For the L2 learner the process takes longer but once the forms are used, this is done relatively more correctly.

Furthermore, the focal learner study shows several similarities in interactions over time between measures, which may be cautiously regarded as general developmental patterns for learners for Finnish. Besides, it shows several different interactions between measures, which supports the idea of different learning trajectories of case use; the FL learner does not learn complexity and case uses simultaneously while the L2 learner learns words and their use more holistically in context. These differences may be very cautiously regarded as due to the language learning context or the types of instruction.

Keywords: dynamic systems theory, Finnish, second/foreign language

Author's address Corinne Tilma
Department of Languages, Finnish as second and foreign language
P.O.Box 35, FI-40014 University of Jyväskylä, Finland
corinne.c.tilma@student.jyu.fi
Department of Languages, Applied linguistics
Postbus 716, 9700 AS Groningen, the Netherlands
c.tilma@rug.nl

Supervisors Prof.dr. M. Martin
Department of Languages
University of Jyväskylä
Dr. M. H. Verspoor
Department of Languages
University of Groningen

Reviewers Prof.dr. J. Kalliokoski
Prof.dr. H. Dufva
Prof.dr. C.M. de Glopper
Prof.dr. A. Housen

Opponents Prof.dr. J. Kalliokoski
Prof.dr. A. Housen

rijksuniversiteit
groningen

The dynamics of foreign versus second language development in Finnish writing

Proefschrift

ter verkrijging van de graad van doctor aan de
Rijksuniversiteit Groningen
op gezag van de
rector magnificus prof. dr. E. Sterken
en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op

donderdag 30 oktober 2014 om 16.15 uur

door

Corinne Tilma

geboren op 2 augustus 1966
te Leeuwarden

Promotor

Prof. dr. C.L.J. de Bot

Copromotor

Dr. M.H. Verspoor

Dr. M. Martin

Beoordelingscommissie

Prof. dr. H. Dufva

Prof. dr. J. Kalliokoski

Prof. dr. A. Housen

Prof. dr. C.M. de Glopper

PREFACE

Now that most of the words of the book are on paper, it has become time to put the names of the people who helped to make this dissertation possible, on paper as well. Without doubt I will forget to name everybody who was there along the way and I apologize for that.

A big THANK YOU to:

-My supervisors and copromotores Maisa Martin and Marjolijn Verspoor. Maisa and the University of Jyväskylä made it possible to conduct a research from a dynamic point of view in Finland and Maisa's confidence in me and in the research have been of great importance for me all along the way. Her door was always open for me and when it was not possible to see each other face to face because of the distance, we mailed or skyped; the advantages of modern times! Maisa's open-minded attitude was a great encouragement for me to bring out the best. In Groningen, during the Master year of Applied Linguistics, the everlasting enthusiasm of Marjolijn had already inspired me to become a researcher. Then, from the very beginning of the research project, Marjolijn stood aside of me and she was always there to help me and listen to my worries. Every time when I could not see the wood for the trees, she gave helpful suggestions to refine the research and to make the research as clear as possible.

-My promotor Kees de Bot for his helpful comments on the dissertation and his confidence in it and in me.

-The four reviewers of the dissertation Jyrki Kalliokoski, Alex Housen, Hannele Dufva and Kees de Glopper for their helpful feedback. This made it possible to refine the research even more.

-Mik van Es for the statistical assistance and helpful advice on the use of the data. With his expertise he made an important contribution to the study.

-Lili Ahonen, my very enthusiastic lecturer of Finnish in Groningen. She inspired me to use my interest in the Finnish grammar for higher goals and encouraged me to do research on the Finnish language in Finland. Moreover, she gave me access to the written data that was needed for the research.

-The teachers in Jyväskylä who made it possible for me to find participants for the study and the master student who checked the participants' written texts.

-The members of the support groups in Jyväskylä and in Groningen and my paranymphs for their helpful feedback on the research. Learning and getting help from people who are in the same situation is great and very stimulating.

-The eight participants of the study in Jyväskylä and in Groningen. They all gave their permission to use their (sometimes very personal) writings for the research; without them this dissertation would not have existed.

-My family and friends for their patience, their time and their never ending support and confidence in me.

Jyväskylä/Groningen, September 2014
Corinne Tilma

FIGURES

TABLES

CONTENTS

0 INTRODUCTION

Already in the nineteenth century the differences between the Indo-European language group, to which most of European languages belong and the Finno-Ugrian language group, to which the Finnish language belongs, were described. However, only quite recently the extent to which the grammars of two language groups differ, in combination with the role of linguistic transparency in the process of learning difficult grammatical rules, is outlined more intensively. This was done among others by Spoelman and Verspoor (2010) in a study from a dynamic perspective: quite a new point of view in second language acquisition that finds its roots in Dynamic Systems Theory. The current study also takes this dynamic perspective on Finnish because it lends itself so well to such analyses with its intricate morphological features.

Dynamic perspective

For some time now a group of usage-based theories have considered language to be a complex dynamic system in itself, consisting of a number of subsystems that interact constantly (De Bot et al., 2005). These theories are all compatible with a Dynamic Systems Theory approach, a rather new theory in applied linguistics, first inspired by Larsen-Freeman who explained the importance of seeing language as a complex adaptive system in 1997 (De Bot et al., 2007). The past few years have shown the rapid development of this theory and nowadays it is an important component of second language acquisition research. The Dynamic Systems Theory serves as theoretical background for the dynamic approach that is used in the current study to explore "systems that change through forces" (Verspoor et al., 2011: 9).

In a dynamic perspective on language and language development, patterns that emerge from language use are studied. Language development is assumed to be a nonlinear process and differences in the degree of variability are assumed to give insight into this process. The constant fluctuations of language are considered as information; the differences in the degree of variability provide a window into the developmental processes. In a dynamic

approach individual differences and developmental processes are important. The assumption is that all factors involved such as cultural, social, psychological, cognitive and linguistic continuously interact and that variability is needed for each learner to select the language forms they need to develop their language (Thelen & Smith, 1994). Variability is seen as a sign of change and development in the language system; low variability in a subsystem indicates that the system is relatively organized and stable, high variability indicates that the system is changing and reorganizing until it settles again at the next developmental stage (Verspoor et al., 2008). By looking at data in this way, useful information about changes in various subsystems is detected (Van Geert & Van Dijk, 2002) and new ways of understanding second language development are discovered (Larsen-Freeman, 2002). Theories of second language development "describe[s] and explain[s] the development and use of more than one language in individuals" (Verspoor et al., 2011: 3). They make use of the term *development* (instead of the more commonly used acquisition) because linguistic skills can grow and decline. Besides, they claim that development and use are not distinguished as separate entities in language and they claim that development is an ongoing process. In the context of a language-learning activity, this means that intrinsic dynamics of the learner as well as the external resources are included; the study of the learners and the context are not seen as two different phenomena but they are studied together (Verspoor et al., 2011).

Recently, researchers have shown an increased interest in language development from a dynamic perspective (Spoelman & Verspoor, 2010; Caspi, 2010; Verspoor, Schmid & Xu, 2012; Rousse-Malpat & Verspoor, 2012; Bulté 2013) and questions have been raised about (among others) the syntactic and morphological development of language learners in different stages of learning a second language. Other important issues are the types of instruction (can learners detect intricate morphological rules on their own from exposure to the language or is explicit instruction needed also for the patterns to be discovered?) and immersion or not (is immersion in the language of influence on complexity and accuracy?).

The study

In the current study two groups of learners (four learners each) are compared: a group of foreign language (FL) learners and a group of second language (L2) learners.

The first group, whose data we had already permission for to use before our study took off, the FL learners of Finnish, consisted of students from the University of Groningen in the Netherlands, with Dutch as first language. They learned the language during the bachelor study 'Finno-Ugric languages and cultures', a full time study of three years, with six contact hours per week on language learning. For these learners the Finnish language was a foreign language; the acquisition took place in a country in which Finnish is not the first

language. During their study the students were taught about the differences between the Dutch and Finnish language, and there was a heavy focus on the grammatical rules; they were seen as learners who treat language primarily as an object of study and for whom the focus is on several specific forms, in order to learn and understand them. This focus on grammar was a rather common and traditional choice because in this situation, there was no immersion in the Finnish language in any way. It involved knowing rules by heart, which might be an advantage in some languages (MacWhinney, 1997). However, it is uncertain whether the focus on grammar would learners help to express themselves in the Finnish language, because also frequency of exposure plays an important role in the development of a foreign/second language (Bybee, 2008). Moreover, research shows that some systems in a foreign language are extremely difficult to acquire for learners who have to deal with a first language that does not have a similar system (MacWhinney, 2008), which was the case for Dutch learners of Finnish. For them input would not become intake when they would fail to see features which would have to be processed in a different way from their first language or when they did not notice subtle cues (Ellis, 2007). The intuition for the role of such cues in the foreign language is based on frequency of occurrence of cues and saliency in the input (De Bot et al., 2005), though the failing process of input becoming intake even exists when the aspects are highly frequent in the ambient first language (Ellis, 2007).

With these facts in mind, the question arose whether there would be a difference between the language development of these learners of Finnish (as a foreign language) and those who learn the language as a second language. Therefore, the four Dutch learners were compared to a group of L2 learners; four students who took a Finnish language course at the University of Jyväskylä while they lived in Finland for at least one academic year. The instruction of the course focused mainly on meaning and communication. This focus on meaning in this course was also a common choice because most of the learners had a different first language and as these learners mainly needed to communicate while living in Finland, the focus was much more on meaning and interaction. Besides, during the learning process these learners were exposed to the Finnish language every day.

Aims of the study

This study has two aims. To date only very few longitudinal studies have been conducted on the difference between the learning of a foreign language with explicit grammar instruction and the learning of a second language through mainly input and exposure. More specifically, to date there have been no studies which compare differences in the language development of beginner learners of Finnish as foreign language and as second language in free writing and no research has been found that surveyed the exploration of development over time of beginner learners of Finnish as foreign and second language yet. This taken together seems an opportunity to accomplish new aims. The few

studies with a dynamic focus have been carried out in a small number of areas and the research to date has tended to focus especially on the development of advanced learners of English as foreign language. In the current study, the development of the learners in the two different settings was used to compare the overall averages of the two groups in order to find similarities and differences in the learning processes in the two contexts and in their development longitudinally. In this way, useful information about stages and processes of learning Finnish as foreign and as second language could be found. Therefore, the first aim of the study is to explore the development of beginner learners of Finnish as foreign language with less exposure to the language and more focus on explicit grammar rules and learners of Finnish as a second language with much more exposure and much less focus on grammar rules. Moreover, no research to date has studied the syntactic and morphological development over time of (in this case Finnish as) a foreign or second language and has described the different roles of various subsystems of the language and has subsequently shown how these systems are related. Therefore, the second aim of the study is to find out exactly how various subsystems interact.

All data were coded for a great number of syntactic and morphological complexity and accuracy phenomena. To be able to explore language development at both an inter and intra individual level with all its variability and variation, the study focuses on the effect of the frequency and types of exposure to the Finnish language. In addition, the study focuses on the impact of instruction on the individual development of complexity and accuracy in writing in the Finnish language. Complexity is operationalized as sentence complexity, accuracy as accuracy in case use and form. The variability analyses will give insight into the number and ratio of correct case uses and forms.

In short, the main questions addressed in this study are:

1. How do several syntactic and morphological measures develop and interact with each other over time?
2. Are there differences between the FL and the L2 groups?

Outline of the study

This study gives insight in the relationship between syntactic and morphological development and consists of ten chapters, organized in the following way:

Chapter 1 gives an outline of the theoretical dimensions of the current study and reviews its theoretical background from a dynamic perspective. In line with this, the chapter presents the basic characteristics of Dynamic Systems Theory after which the chapter presents several general questions which will serve as the basis for the current study.

Chapter 2 explains why Finnish can be seen as very different for learners from the Indo-European language family. Moreover, the chapter gives an overview of differences and similarities of learning a language with the focus on grammar and with the focus on meaning. Furthermore, the chapter shows

the Finnish grammar issues that are used in the study in detail after which it states study's research questions and hypotheses.

Chapter 3 presents the design of the study, the participants of the two groups, the way the data have been collected and the operationalizations of the measures in detail. Moreover, it presents the teaching materials and the conditions on which the data is produced. Finally, the chapter explains the study's statistical methods, analyses, examples and some pitfalls in research from a dynamic perspective.

Chapter 4 examines the data at the group level for all coded measures on complexity and accuracy in order to find measures in which the groups may differ in general development. With this information it explores some data of two individual learners as well. Besides, it presents the measures to use in the remainder of the study.

Chapter 5 explores the development and interactions of three syntactic measures (more complex sentences, average sentence length and clause length in morphemes). Besides, it examines the most informative measure to represent general sentence complexity in Finnish.

Chapter 6 examines another three syntactic complexity measures (total use of cases, use of nominative and use of the 12 other cases (than nominative, genitive and partitive).

Chapter 7 presents the morphological complexity measure average word length in morphemes. Besides, the chapter explores the development and interactions of the average clause and word length in morphemes and use of the past and perfect tense. The possible task relatedness of the latter measure is explored as well.

Chapter 8 introduces the accuracy measure CAR (=case accuracy ratio) use and form errors. This measure was found to have a significant difference between the FL and L2 groups. Moreover, it examines the development and interactions of the syntactic accuracy measure CAR use errors and the morphological accuracy measure CAR form errors.

Chapter 9 examines the development and interactions of five measures: CAR use and CAR form errors, average clause and word length in morphemes and complex tense use.

Chapter 10 summarizes the findings at the group and the individual level, links them to the literature and interprets them. Besides, the chapter shows the relevance of the current study for the second language development in general. Finally, it presents suggestions for further research.

1 RESEARCH FROM A DYNAMIC PERSPECTIVE

In this chapter we will present an outline of the dynamic approach, the theoretical framework of the current study by providing an overview of the basic characteristics of Dynamic Systems Theory (DST) and discussing several studies, done from a dynamic perspective. We will also present the general questions that will serve as the basis for this study.

1.1 Dynamic systems theory and L2 development

The first section presents the roots of a dynamic approach, the theoretical framework of the current study.

> Things taken together are whole and not whole, something being brought together and brought apart, in tune and out of tune; out of all things there comes a unity, and out of a unity all things. (Heraclitus, n.d.: frag. 10)

McKirahan (2003) explained Heraclitus' holistic insights as follows:

> The world is a single dynamic whole made up of many things related to one another in various ways. We need to understand both the many and the one: how the one world works and how the many things in it work as well, and to do so involves understanding that the many things are interrelated in many and unexpected ways, and understanding that they work together, not each on its own, and how they do so; likewise it involves understanding how the world is a unity composed of many parts and how each part contributes to the whole. (McKirahan, 2003: 12)

Even though Heraclitus already implied 2500 years ago that the world was a complex dynamic system, it was only during the past few decades that a holistic way of looking at things became acknowledged in many different research fields (Waldrop, 1992; Thelen & Smith, 1998). In 1994, the starting point of an advanced theory about development of cognition (motor skills) was DST (Thelen & Smith, 1994). DST was eventually (in the field of theoretical

mathematics) developed in order to model development of systems which are complex (Verspoor et al., 2011, ch.1). In the renewed version of the theory that was suggested to be applicable for early cognitive development as well, Thelen and Smith united the advances of the theory within the field of neuroscience and neural development. With this, DST was introduced in the field of human development. They showed the importance of describing the route (followed by an organism) from one state to another; there is not an instant form of the end-state of an organism and the focus is on the process of change (Thelen & Smith, 1994).

Van Geert (1991, 1994) outlined the importance of DST for L1 acquisition (early language development). Besides, he showed the value of looking at language learning as a dynamic system and moreover, contributed useful methods from a dynamic perspective in language acquisition. After Van Geert had created a framework in order to be able to interpret language growth as a process of growth under limited resources and had introduced a dynamic growth model, Larsen-Freeman (1997) showed the importance of this progressive insight in her article on complexity and second [1] language acquisition. She emphasized the "similarities among complex nonlinear systems occurring in nature and language and language acquisition" (Larsen-Freeman, 1997: 142) and suggested that DST might be useful to L2 acquisition as well, because language is a complex, non-linear system that contains subsystems (e.g. morphology and syntax). Moreover, the interdependency of the subsystems should consequently imply a reaction in one or more of the subsystems if one should change (Larsen-Freeman, 1997). When on the other hand the isolated parts are studied, the outcomes would not comprise the interaction between the subsystems. Larsen-Freeman concluded:

> … the behavior of the whole emerges out of the interaction of the subsystems. Thus, describing each subsystem tells us about the subsystems, it does not do justice to the whole of language. (Larsen-Freeman, 1997: 149)

Already in 1991, Larsen-Freeman and Long had pointed out that the process of learning items in linguistics always goes on and that mastering a single item is only a partial indication of development and in 1997, Larsen-Freeman strengthened these thoughts for language acquisition. Moreover, she then stated that "an unstable system is not a contradiction in terms" (Larsen-Freeman, 1997: 156) if a complex dynamic system is seen as constantly being instable. She also argued that the interaction that comes to the surface is only a very small part of all changes that take place; in other words, there would be even more instability. Larsen-Freeman emphasized the importance of details while keeping the whole in mind and the difficulty of "studying the whole of second language acquisition" (Larsen-Freeman, 1997: 159). She so cleared the way in L2 acquisition for focusing on individual aspects and at the same time

---

[1] When in this chapter the term L2 is used in other studies, referring to (results of) studies other than the current study, it may mean foreign language (FL: learn a language in a country in which the language is not the L1) as well.

showing consideration for the complexity of the whole and DST could account for this, looking at ''ever interacting measures, non-linear behavior and sometimes unpredictable outcomes'' (De Bot et al., 2007: 7); DST is able to go into details of, and to gain useful information on the relationship between levels of patterns that can appear in a language in general and that can be on the surface of the individual level as well (Verspoor et al., 2011). In DST these two levels of development are looked at from different perspectives: from above (looking at global structures, explained as "similarities across subjects" (Van Dijk, 2004: 129)) and from below (looking at details, explained as "behavior that is variable, fluid, and highly context dependent" (Van Dijk, 2004: 129)). These details are seen as indicators for variability: "an important internal developmental characteristic" (Van Dijk, 2004: 129). Therefore, variability is valued as a substantial part of language development.

1.2 Language development from a dynamic perspective

Now that the roots of the dynamic approach have been explored, we will show how language development is looked at from a dynamic perspective.

In recent years, the relevance of a dynamic approach to L2 acquisition has been proved by several studies from which it has been shown that "language, language acquisition, and language attrition are much more intricate, complex, and even unpredictable than a linear position would allow"(De Bot et al. 2007: 7).

In their case study on free writing, Verspoor et al. (2004) recognized the difference between the common and the new approach and explained that they looked at variation (*inter*: between the individual and his environment) per se in order to understand the individual's process of development instead of trying to show general tendencies by averaging the variation. This different starting-point was also ascertained by Van Geert (2008), who stated that the process of development needs not to be looked at in the usual way if it is intended to be really understood. He concluded that the way to do so is ''through investigating associations between variables across populations'' (Van Geert, 2008: 179). In line with this, he claimed that knowledge of dynamic systems can contribute significantly to the process of understanding development. What also helped to understand development was the way variability (*intra*: within the individual) was looked at in a dynamic approach; as another important factor for development (Verspoor et al., 2008). To conclude, research from a dynamic perspective brought a coherent approach in L2 development because both cognitive and social aspects of language development were taken into consideration (De Bot et al., 2007).

For a long time only few publications reviewed the dynamic approach in research in L2 acquisition (De Bot et al., 2007). One of them came from Herdina

and Jessner (2002), who found that variability in the multilingual system changes over time, implying that it plays a crucial role in development. Also Van Geert investigated several topics on measurements of measures from a dynamic perspective (Van Geert & Steenbeek, 2005), suggested a technique "that incorporates variability in the analysis of the shape of developmental change" (Van Dijk & Van Geert, 2007: 7) and explored difficulties in model building (Van Geert, 2008).

1.3 Language as a dynamic system

In this section we present the basic characteristics of dynamic systems and within each quality, its role for the current study.

1.3.1 Initial state

This subsection presents the first quality of a dynamic system.

Larsen-Freeman presumed the presence of universal principals and limitations in the use of certain types of constructions in human's language (Larsen-Freeman, 1997). She labeled them as initial conditions, which were supposed to exist in every human language (De Bot et al., 2005). Such an initial state is important for the development of a dynamic system and the underlying butterfly effect can be understood as the essence of chaos; from its starting-point, a system can develop in all possible directions because every small difference in the beginner state will have its effect on the development of the system (De Bot et al., 2007). However, it is impossible to trace all possible information that might affect the development of a system. Moreover, there is no information beforehand about the relevancy of initial conditions for this development (Verspoor et al., 2011). Besides, the sensitive dependence on initial conditions is strongly related to the interconnectedness of systems and to the non-linearity of systems, which are both responsible for limitations as well (Verspoor et al., 2011).

For the current study this means that there is no certainty at all about similar outcomes (Verspoor et al., 2011) even though the learners of Finnish as FL or L2 seem to have several comparable conditions.

1.3.2 System

This subsection presents the two qualities that are covered by the word system. A system is always playing more roles at the same time because the many subsystems in the system language are also systems on themselves (Verspoor et al., 2011).

1.3.2.1 Interconnectness

d'Anglejan and Renaud (1985) claimed that it is inevitable that learner measures overlap and interact with others. Moreover, they stated that factors can only be measured correctly when they are not isolated from other factors (Larsen-Freeman, 1997): in other words in complete interconnectedness. Complete interconnectedness stands for an interrelation between all measures in a dynamic system; all parts within the system are connected to each other; within a system changes in one measure will always affect all. Even though measures may be more or less connected to each other, the change in one measure will have the same effect on the other consequently changed connected measure. This process of change is ongoing and therefore determination of the outcomes of the process of development is complicated. Some complex issues arise when the dynamic point of view comes into sight; firstly, it needs to be outlined whether it is possible at all to examine a total interconnected system. Secondly, if so, the correct way to study developmental changes in a system that influence other systems and are influenced by other systems at the same time needs to be explored (Verspoor et al., 2011). In recent years an increasing amount of research done from a dynamic perspective is seen (see 1.4), so several results are available now. Moreover, the results of the current study may point out the significance and the relevancy of the measures that were included, to give insight in the expectations beforehand and the actual findings in the study.

In the current study we explore how syntactic and morphological development of learners of Finnish as an FL and L2 influence each other. Change in one system will not only influence the other system but all parts of the systems it depends on (e.g. Finnish as FL or L2, L1, learner, human being). To sketch this scale of system complexity, in literature the term nested is used (Briggs & Peat, 1989; De Bot et al., 2005; Caspi, 2010); systems are nested and always part of a larger system. This means that, eventually, they are part of the universe as most complex dynamical system of all (Briggs & Peat, 1989; De Bot et al., 2007).

Figure 1.1 shows interconnectedness; the connections between systems are to be understood as lines with a certain value. With the change in length of one connection, the lengths of all connections change.

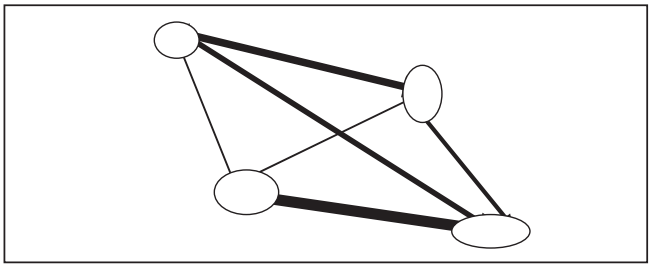

Figure 1.1: Lines with different stroke widths, representing the interconnectedness between measures within a system.

1.3.2.2 Dependency

In a dynamic approach every developing system is dependent of resources (Verspoor et al., 2011). These resources can be internal, inside the system (like memory and talent (Van Geert, 2008)) and external, coming from outside the system (like the language used by the environment and material resources (Verspoor et al., 2011)). These internal and external resources are not standing by themselves; they are interrelated in the sense that they interact. For L2 development from a dynamic perspective this means that a language-learning activity always happens within a context, consisting of resources from within and outside the language learner (Verspoor et al., 2011). The possible relationship between measures that have "a meaningful relationship to each other" (Verspoor et al., 2011: 86) are the following:

Supportive: Subsystems[2] develop in unison because they support each other.

Competitive: Subsystems develop in alternating patterns (when one goes up the other goes down) because they compete with each other.

Conditional: A minimal level of one subsystem is a necessary precondition for another subsystem to develop, also referred to as a precursor interaction. (Verspoor et al., 2011: 86)

Moreover, the resources are limited. The phenomenon of limitation is called carrying capacity: the inevitable limitations that incidentally imply the limitation in growth as well (De Bot et al., 2007). Carrying capacity is explained as "the maximal attainable level of a particular systemic component given the resources available for its growth" (Caspi, 2010: 13). It is therefore understood as the state of knowledge that a subject can get at a particular moment with a particular structure of resources (Van Geert, 1994). From a dynamic approach it is obvious that dependence on internal and external resources exists within language as a system and all its subsystems.

In the current study we explore the external informational resource influence of instruction for learners of Finnish as FL or L2.

Figure 1.2 shows dependence of a system; the highest number of possible connections is to be understood as the highest number of possible lines between systems (though we are aware of the fact that such an image is impossible to represent).

---

2 Also called growers.

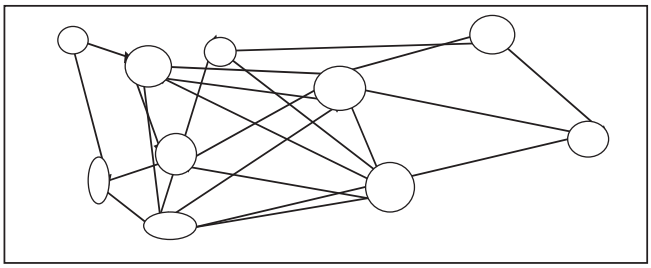

Figure 1.2: Connections, representing the dependency of systems on internal and external resources.

1.3.3 Dynamic

This subsection shows the four qualities that are covered by the word dynamic. This term implies changes in a field or system that are produced by internal or external energy.

1.3.3.1 Change

The process of change was put by Larsen-Freeman as follows:

> As I am writing this article I am contributing to the changing frequencies of word use in this article and to those in the whole system of English. (Larsen-Freeman, 1997: 151)

The previous sentence outlined the total essence of change. In systems, changes take place constantly and because of the complete interconnectedness and dependence on internal and external resources, they take place in the whole system. In this process the state of being of the initial state (see 1.3.1) is not important because change takes place anyway and always (De Bot et al. 2007). Because of this constant change in systems, they develop. This development takes place through interaction of the system with the environment, serving as input and through reorganization of the system itself, after which the system becomes the result of internal changes (Verspoor et al., 2011). After this change, settlement in a new temporary (attractor) state takes place: the state a system prefers to be in at a certain point in time (Verspoor et al., 2011). This state is rather stable with little internal variation (De Bot et al., 2005) and can only be caused by developments within the system (Verspoor et al., 2011). When a system has settled in an attractor state, its state of being is called fossilization: stability in development (Verspoor et al., 2011). Sooner or later a critical point will be reached, after which the system changes again and settles in another attractor state; this process is ongoing. Dependent of the degree of the attraction, a certain amount of energy "is needed to make the system move on to another attractor state" (De Bot et al., 2007: 8). It is not clear why a system settles in a state (which at that moment becomes the attractor state) and why the system does not settle in the opposite (repellor) state; both the attractor state and the repellor state consist of "dynamically interacting forces" (De Bot et al.,

2007: 18). Moreover, the directions into which a system will go are unpredictable. In other words, beforehand we cannot know into which state a system will settle and what the attractor state will be (Verspoor et al., 2011).

With the learner of Finnish as FL or L2 as the system and the Finnish language instructions as input, the current study explores the influence of instruction (focus on grammar and on meaning) on the learners. In order to trace development it is important to explore how the learners change through reorganization within themselves and through interaction with the instruction.

Figure 1.3 shows the attractor and repellor state.

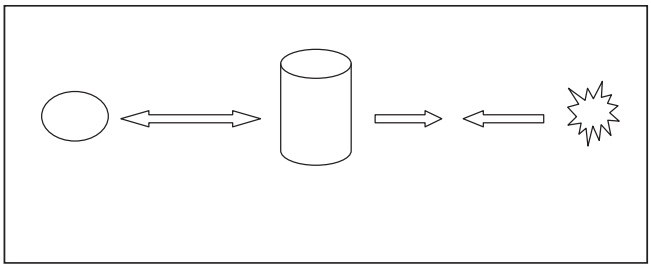

Figure 1.3: A system with its attractor (left) and repellor state (right).

1.3.3.2 Iteration

"The repeated application of the same procedure over and over again" (Verspoor et al., 2011: 16) is another dynamic quality, denoted as iteration. Iteration implies that the following step is never exactly the same as the previous step and this means that every step is a representation of its own history (Verspoor et al., 2011). Like in the interaction between dancers and their environment, every next step differs from the previous and the following one.

In the current study we look at syntactic and morphological development in free writing, among others to find how tense use changes in time; every time when the learner of Finnish as FL or L2 uses tenses or when he notices tense use in his environment, the current representation of the tenses changes.

1.3.3.3 Non-linearity

The relation of cause and effect is non-linear, which makes the applicability of the dynamic approach complex (Larsen-Freeman, 1997; Verspoor et al., 2011). Therefore, computers are an indispensable tool in dynamic language research, as already predicted by Larsen-Freeman in 1997.

The current study investigates how the development of learning Finnish as FL or L2 takes place by exploring non-linearity of changes (peaks and dips) in the learners' syntactic and morphological systems.

Figure 1.4 shows the linearity and non-linearity of cause and effect.

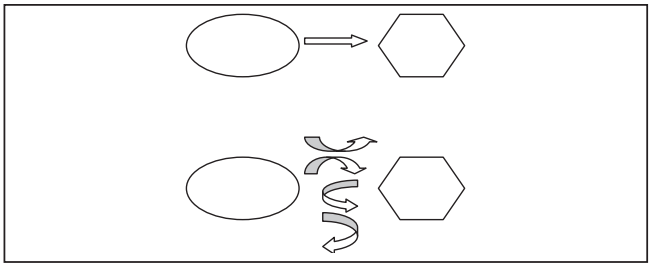

Figure 1.4: The linearity of cause and effect (upper side) and their non-linearity (lower part).

1.3.3.4 Emergence

To study a complex system, understanding the behavior of groups by showing general tendencies (i.e. averaging out variation) is not the proper way (Larsen-Freeman, 1997) because research from a dynamic perspective treats emergent properties as "a functional characteristic of the system" (Verspoor et al., 2011: 17). It assumes that a complex system is more than the sum of its components and that it depends on all parts in order to function well (Larsen-Freeman, 1997; Verspoor et al., 2011). The general learning processes will give the learners tools in order to recognize patterns in the language and to generalize the rules and cues (De Bot et al., 2005). The phenomenon of emergent properties was pointed it out as follows by Verspoor et al.:

> The potential to become ice is not an inherent characteristic of water, but an emergent property that comes out in specific conditions. (Verspoor et al., 2011: 17)

The current study explores the emergence of language rules and cues of the Finnish language as FL or L2 in the learning process by investigating use and form errors in case markings, because errors are seen as useful information.

1.4 Previous findings

This section describes a number of studies from a dynamic perspective and their contributions to the field in general. Besides, it presents these studies' specific findings; they lead to several general questions that will serve as the base for the current study.

1.4.1 Language development with a dynamic approach: a brief history

In 2004, researching variability was still in its infancy. Van Dijk (2004) argued that variability is the result of flexibility and adaptability of systems to the environment and as both the source and indicator of a specific moment in the development. The aim of her study on language development was to contribute "to the methodology

by proposing various simple but powerful techniques for describing and analyzing variability and ambiguity" (Van Dijk, 2004: 8). By using these techniques, intra-individual variability in time serial data could be visualized and outlined. It is noted that in exploring a developmental process, random variability should not to be taken for developmental variability. A solution for this was found in the moving min-max graph, a technique to visualize intra-individual growth and variability (Van Dijk, 2004) (see chapter 3). A min-max graph shows "local variability peaks that precede developmental jumps" (Caspi, 2010: 27) and points out the amount of variation in their relation over time. Van Dijk applied a dynamic perspective to L1 development and several useful techniques were developed that will be applied in the current study.

Also in 2004, Verspoor et al. explored L2 development from a dynamic perspective. Longitudinal studies on L2 development had been done before, but none of them had focused on the variability specifically. In this pilot study on free writing samples, the aim was to find out whether the amount of variability was as large as to be expected from a dynamic perspective. Six weeks of data appeared to be too little to draw any conclusions, but the fact that Verspoor et al. (2004) were the first to apply a dynamic approach to L2 development makes this study important. Moreover, as suggested, the current thesis will also use free writing samples of beginning FL and L2 learners, but over a longer period of time:

> …further research will have to show whether subjects show signs of development during a longer period of free writing or whether some more external force in the form of a more explicit task assignment, instruction or feedback is needed to bring about change… (Verspoor et al., 2004: 421)

Verspoor et al. (2008) examined the longitudinal data of an advanced learner of English as an L2 (academic texts) on the development of several lexical and syntactic measures and their interactions. The study showed the non-linearity in patterns and illustrated that looking at variability as a source of information is relevant. General increase over time was found, but the study also showed that development is nonlinear and that the interaction of subsystems is dynamic.

Spoelman and Verspoor (2010) looked at the development of Finnish as an L2 from a dynamic perspective over a period of three years, exploring both accuracy and complexity at the word, noun phrase and sentence level. They found some types of case errors to decrease faster than others. Moreover, meaningful interactions between complexity measures over time were found, but not between accuracy and complexity measures. For the current study in particular Spoelman and Verspoor (2010) is important because their data are very similar to our data; the same free writing tasks have been used, but in the current study two different groups of Finnish language learners are compared.

Caspi (2010) presented a longitudinal case study of four advanced learners of English as an L2 on both vocabulary and writing development from a DST perspective. After a number of variability analyses, Caspi ran computer

simulations on her own data to test a number of hypotheses. One of the main findings that concern the current study is the fact that in their writing development, the learners first develop in lexical complexity, then lexical accuracy, then syntactic complexity and then syntactic accuracy.

Verspoor et al. (2011) is an edited book in which the methods and techniques, used to visualize variability and test for meaningful jumps, were described in great detail and these will form the basis for the current study.

Taking a dynamic usage based approach Verspoor et al. (2012) looked at the development of a great number of subsystems in a cross-sectional study of Dutch L2 learners of English. They found that at different levels of proficiency different subsystems seem to develop more. In line with Caspi, first lexical measures seem to develop especially at the lower level, then the syntactic subsystems at the higher levels and finally again the lexical measures increase. The study showed that different subsystems of the language develop at different stages of development. It also showed that learners at the lowest levels showed the most variation. As learners become more proficient they show less variation, indicating that they converge more. One of their suggestions was to compare results from a cross-sectional study with the results of a longitudinal study and that is exactly what the current study aims to do.

Bulté (2013) traced ten L2 learners of English over 19 months on a number of complexity measures and found that, although there were group trends, not a single learner showed the same development as the group average on any measure. He found variability and variation in all subsystems in all learners. Moreover Bulté found that no single, specific measure was a good indicator of development. However, when four rather broad measures at both the syntactic and lexical subsystems were combined, the general pattern of development could be shown. The importance of this study is that we cannot create an average learner by taking group averages.

Murakami (2013) explored the effect of L1 in the order of acquisition of morphemes of learners of English as an L2 with seven different L1 backgrounds in a huge corpus study, both cross-sectionally and semi-longitudinally. What is relevant for the current study is that he found a large L1 effect. If a particular morpheme did not occur in the L1 (such as the article) then it was unlikely that the learner would achieve a 90% accuracy rate for that morpheme. He also found that some morphemes were more sensitive to L1 effect (especially the article) than other morphemes (such as third person singular or plural –s). Like Housen had already shown in 2002 and Bulté in 2013, Murakami also found a great deal of variation among learners, even in homogeneous groups. Besides, he found that no single learner behaved like the group trend. What is important for the current study is that two groups of learners will be compared, one homogeneous group with the same L1 (Dutch) and the other heterogeneous group with learners from four different L1's (Portuguese, Japanese, Bahasa Indonesian and German). The differences found may therefore be partly due to L1 effect. However, for the detailed longitudinal study the data of the German learner was used as to minimize the differences in L1 effect.

1.4.2 Specific findings leading to general questions

Van Dijk (2004) found that early language development is a rather variable process. Moreover, she showed that there was a relation between variability and discontinuity at the moment of the first peak or dip in the data, but she found that the discontinuity may consist of different types and will depend of the combination of criteria for each participant. Verspoor et al. (2012) in their cross-sectional study have shown that beginners show more variation than more advanced learners. Murakami (2013) has also found in his semi-longitudinal study that beginners showed more variability than the more proficient learners. Because the learners in the present study are absolute beginners, the assumption is that there will be a great deal of variation among the learners and variability within each learner, which diminish over time. But in line with Van Dijk, some learners may show clear signs of discontinuity (peaks or dips that are statistically significant). Therefore, in the current study when the visualized data suggest a clear peak or dip, these will be tested for significance.

A large amount of non-linear development, variation and changing relationships among the measures was found by Verspoor et al. (2012) especially for learners at the lowest levels, whereas more proficient learners show less variation. This leads to the question what the amount of variation will be for the beginner FL and L2 learners in the current study.

Because of the suggestion of the possibility of different outcomes with different types of instruction in the study of Verspoor et al. (2004), this dissertation explores differences in syntactic and morphological complexity and accuracy measures between two groups, instructed differently in the process of learning and FL or L2; this leads to the question whether the type of instruction leads to differences in the development of complexity and accuracy measures.

Spoelman and Verspoor (2010) found that insight in the developmental process is possible through looking at degrees of variability and that the interactions of several researched complexity measures changed over time but they did not find any interactions between accuracy and complexity measures. Besides, they found that error rates decreased quickly, with the exception of four cases. Moreover, they found word and sentence complexity as well as word and noun phrase complexity to develop in the same way, whereas the noun phrase and sentence complexity appeared to be competitive growers. Spoelman and Verspoor concluded that the intra-individual variability acted in the way it was expected according to a dynamic approach and their study confirmed the presumption that L2 development is non-linear. This leads to the questions whether the FL and L2 learners in the current study do show interactions between accuracy and complexity measures and whether the complexity and accuracy measures show similar development.

Also Verspoor et al. (2004) looked at several aspects to find indicators for change and differences in the development of the subjects. A great deal of inter- and intra-individual variation was detected, but at the same time no signs of

development were found during the period of collecting the data. Since then, several studies have looked at development over a longer period of time in single case studies (Verspoor et al., 2008; Spoelman & Verspoor, 2010) and Bulté (2013) has looked at ten individuals over a period of 19 months. All these longitudinal studies found clear signs of development, despite high degrees of variability and variation. This leads to the question whether the FL and L2 learners in the current study show clear signs of development in the course of one academic year and if so, in which subsystems of complexity and accuracy development occurs. Moreover, the current study will examine if there are clear differences between the FL and L2 groups.

Verspoor et al. (2008) found that more complex constructions were used during the three years of study and they concluded that the finite verb ratio (=average number of words per finite verb) could give interesting insight into complexity, e.g. the measure reflected longer noun phrases than the measures simple and complex sentences. At the start the data showed competition between the simple and compound sentences and the finite verb ratio after which similar increases for complex sentences and the finite verb ratio were shown, the latter relation being supportive from that moment on. This leads to the question about how Finnish as FL and L2 develops and what measures may show this development best. As Finnish is an agglutinative language it may well be that rather than using the finite verb ratio based on words, a finite ratio based on morphemes is a better measure. Once the better measure is established it can be used for the remainder of the study.

Verspoor et al. (2004) showed that their beginners used mainly the simple present and past tense and that the perfect tense emerged later in a rather unpredictable way. Moreover, they showed that in the short time span of six weeks the tenses did not seem to become more complex. In their cross-sectional study, however, with a greater range of proficiency levels, Verspoor et al. (2012) showed that the use of the present tense steadily declined and the use of the past tense steadily inclined across proficiency levels. This was also the case for general measures like all errors combined, and the use of the sentence types (from simple to more complex sentence types), the tenses (from present tense to other tenses) and the accuracy measures (fewer errors over time) but not for specific constructions in which, on the contrary, they found non-linear development, variation and changing relationships. This leads to the question whether the FL and L2 learners show similar tendencies in both the broad and specific measures.

Spoelman and Verspoor (2010) researched development in the Finnish language in a single case study and they found that the accuracy in case use increased quite rapidly, except for two particular ones: the genitive and partitive case. This leads to the question whether the four FL learners with the same L1 and learning context improve in a similar manner, with rather abrupt increases in accuracy. It is also interesting to see if learners in the L2 context have similar trajectories.

Verspoor et al. (2008) found that the development of the language system showed alternations in the focus on lexicon and syntax, as well as competitive interaction between the researched measures. In addition to this, Caspi (2010) found that first words are made more complex after which they are used more accurately, after which the syntax shows respectively more complexity and accuracy. Besides, Caspi showed that variability in the used data was "a meaningful manifestation of internal systemic dynamics" (Caspi, 2010: 173). This leads to the question whether the data of the FL and L2 learners show similarities in the sequence of complexity and accuracy measures.

1.4.3 The current study from a dynamic perspective

On the development of individuals, variation is expected (Housen, 2002; Bulté, 2013; Murakami, 2013), which may be due to individual trajectories, to the L1 and to the learning context. The purpose of the current study is not to find causes but to explore individual trajectories. We will explore how different subsystems in the FL and the L2 develop and interact over time in the data of the participants of the current study. The specific research questions and hypotheses will be presented at the end of chapter 2.

Summary

This chapter has shown that the past twenty years have seen increasingly rapid advances with regards to the place of research from a dynamic perspective in the field of L2 development. When Thelen and Smith had renewed the version of DST in 1994 and the process of change became the point of view, the door to application in L2 development was opened by Larsen-Freeman in 1997. Clearly, within language development, DST took its place in usage based theories, but even though these theories were compatible with DST in quite a few ways, none of them combined all different kinds of aspects of the developmental process in language learning the way DST did. Van Geert put the theory into practice in L1 development and later in L2 development as well.

Moreover, the chapter has illustrated that language is interpreted as a dynamic system and exists of several qualities: interconnectedness, dependence (on internal and external resources), change, iteration, non-linearity (in development) and emergent properties. Also their role in the current study has been shown.

Finally, the application of the theoretical interpretation has been outlined through eight studies (Van Dijk, 2004; Verspoor et al., 2004; Verspoor et al., 2008; Spoelman & Verspoor, 2010; Caspi, 2010; Verspoor et al., 2012; Bulté, 2013; Murakami, 2013) and one book (Verspoor et al., 2011) that have all played a role in the progression of a dynamic approach in (second) language development. Also their specific findings, leading to general questions that guide the current study were discussed and it was pointed out that this study, done from a dynamic perspective, explores individual trajectories over time.

2 DEVELOPMENT IN THE FINNISH LANGUAGE

The first aim of the current chapter is to describe several aspects that deal with learning Finnish as an FL/L2: learning a different language, using different types of instruction and the influence of the language learning environment. The second is to explore those aspects of the Finnish language system that have to be learned at the morphological and syntactic and morphological levels. The final goal is to present the research questions and hypotheses of the study.

2.1 Learning the Finnish language

This section investigates learning Finnish as a foreign or second language as it is a very different language for learners with other language backgrounds and its relation with the role of instruction and the role of the language learning environment.

2.1.1 Learning a very different language

More than 95 percent of the European people speak a language that belongs to the Indo-European language family (Ahonen, 2006). The Indo-European language family does not include the Finnish language; Finnish is a member of the Finno-Ugrian language family (Karlsson, 1999). The Finnish language contains several issues and constructions which learners with an Indo-European language background are totally not familiar with. Learning the Finnish language is for them learning a very different language. However, Dahl (2008) found that the only way in which the Finnish language is different from Indo-European languages is that Finnish is an agglutinative language. Despite of Dahl's insights, MacWhinney (1976) described the Hungarian language, also a member of the Finno-Ugrian languages, as a non-Indo-European language with agglutination, suffixes and vowel harmony and especially different from Indo-European languages in the areas of morphology and syntax (for studies on

the acquisition of Finnish morphology and syntax, see e.g. Martin, 1995; Kaivapalu, 2005; Kajander, 2013; Seilonen, 2013; Honko, 2013). MacWhinney (1997) found that Indo-European learners of Hungarian as second or foreign language (henceforth referred to as an L2) need a lot of work before they are able to understand the rules of the Hungarian grammar. Moreover, they make a lot of errors in the choice of conjugations because of the extreme extent to which the grammar of the Hungarian language differs from the Indo-European languages. He suggested that for adult learners the efficiency of acquisition could increase when they should focus on "particularly difficult parts of a grammatical system" (MacWhinney, 1997: 129).

Also Schot-Saikku (1990) described difficulties for Indo-European learners of Finnish as a second language. She found that the partitive is a problematic issue. Also Spoelman and Verspoor (2010) found that the use of the partitive (next to the genitive and accusative) in the Finnish language is a highly complex and difficult issue for such learners (see also 1.4.3). The problems in the use of the partitive, genitive and accusative singular were not solved during the longitudinal study and "the difficulty [of the use of the partitive, genitive and accusative in the Finnish language] lies in a lack of consistency and/or frequency of relevant instances to entrench the instances" (Spoelman & Verspoor, 2010: 10). They concluded that a lack of rule transparency is a logical result of the complex rules for the linguistic structures of the three cases in singular. The fact that the learner kept on having difficulties with these particular cases supported their insight that "complex and opaque rules are the most difficult to acquire" (Spoelman & Verspoor, 2010: 13). This strengthened the perceptions of DeKeyser (2005) as well; he concluded that a lack of linguistic transparency plays an important role in the process of learning difficult grammatical rules in second language acquisition. In line with this, Ringbom (1987) pointed out that the L1 is of importance in second language acquisition, after he had found that learning a different language is much more difficult than learning a similar language. Finally, also MacWhinney (2008) showed that some systems in a different second language are extremely difficult to achieve for learners who have to deal with an L1 that does not have a similar system, something that was also shown by Spoelman (2013) for Dutch and German learners of Finnish and by Kaivapalu (2005) for Russian learners of Finnish in comparison to Estonian learners.

2.1.2 The focus of instruction

Apparently, in the case of opaque constructions in a second language, a correct pattern may not be found by the learner on his/her own and to extract them explicit instruction may be needed (de Bot et al., 2005). However, Schoonen et al. (2003) found that the English writing proficiency correlated more with the L1 (Dutch) writing proficiency than with the English linguistic knowledge or the accessibility of it. Nevertheless, discerning conceptual distinctions that do not exist in the L1 may be a difficult issue for an adult second language learner (Williams, 2003). He concluded that the effect of explicit instruction depends on

the extent to which the first and second language are related, but that it might in any case be helpful in the situation of learning a very different language. This conclusion was emphasized by Dijkstra (2003) who found a correlation of explicit rule knowledge and written abilities for beginner learners (though not of a very different language: German as FL, English as L1). Also Larsen-Freeman (1997) saw the value of explicit instruction in the fact that it facilitates intake, even though she realized that grammatical items may not be mastered directly. On the contrary, Krashen (1982) found that "people do not learn a language by talking about it (i.e. studying the rules of grammar) but by experiencing enough meaningful input and communication within that language" (De Bot et al., 2005: 45). However, Lightbown (2000) found that merely communicative practice is insufficient to let learners become proficient and accurate in the second language as well.

Schmidt's so called noticing hypothesis showed that noticing is "the necessary and sufficient condition for the conversion of input to intake for learning" (Schmidt, 1994: 17), i.e. noticing grammatical structures guarantees incorporation into the developing language of the learners (Schmidt, 1990). This implies that, even though the effect of giving explicit instruction would be that the salience of target language forms in the input could be increased in a useful way, the process of acquiring target forms would only take place if they were really noticed. This is what Leow (2001) stated as well; he found that learners who were aware of the morphological forms they had to deal with, were significantly more able to recognize them and to produce them in their writing after exposure. He rejected the hypothesis that language learning can take place without some degree of consciousness (Nassaji & Fotos, 2004). Still, just noticing may not be sufficient for input to become intake. The learner should consciously have to notice the gap between the first and second language; in that case, the noticed item could be reflected on, even though Schmidt and Frota (1986) concluded, that learners often continue making the same errors as they did before the noticing. Nevertheless, even when second language learners should receive explicit instructions on grammar patterns, the correct application of the rules often fails in communicative tasks (Green & Hecht, 1992). The latter was also noticed by Haley and Rentz (2002), who found that acquisition of a form, noticed by the learner or introduced through instruction, is delayed when the communicative need to use it is not present.

To date a number of studies have investigated so called Form-Focused Instruction in second language acquisition. This term refers to planned as well as incidental instructional activity, both intended to induce language learners to pay attention to linguistic form (Ellis, 2001). Form-Focused Instruction can therefore be seen as a cover term for a variety of terms that is seen in the current literature; it includes both traditional (Focus on Forms) and communicative (Focus on Form) approaches to teaching forms (Long, 1991; Ellis, 2001), outlined in the parts coming up next. Both phenomena represent the focuses of applied instruction in classroom situations and both want the learners to be able to communicate. Norris and Ortega (2000) concluded that instruction that involves

an explicit focus on grammar has a positive and durable effect for both the focus on Form and on Forms. However, Norris and Ortega themselves say that there may be a bias in these studies as participants are usually tested explicitly on the items that were taught, often in closed tests, and not in free response data, which would show implicit knowledge better. The current study makes use of such free response data to avoid the possible bias of the explicit instruction in controlled tests.

2.1.2.1 The focus on grammar

Traditional learning can be sketched through the acquisition metaphor, in which the learner is regarded as a "container, absorbing new information" (De Bot et al., 2005: 106). The traditional approaches focus on "the language as a system and on the role of instruction" (De Bot et al., 2005: 108), and their main purpose is to accumulate individual language elements, e.g. case markings (Haley & Rentz, 2002).

When in language learning the focus is on grammar, as the traditional notion of teaching, the primary purpose of an activity is to learn a preselected form (Long, 1991; Ellis, 2001; Haley & Rentz, 2002). Learners, seen as students instead of language users, treat language primarily as a study object and focus their attention intensively on some specific elements of grammar which are isolated from the context and the communicative activity, in order to learn the specific grammatical rule (Ellis, 2001; Haley & Rentz, 2002; Nassaji & Fotos, 2004).

The fact that to date a lot of different methods have been used to get data (Norris & Ortega, 2000) has an impact on all kinds of variables that should show the possible effectiveness of the focus on grammar (like on "learners' developmental stage, the structure being taught, the instructional context and the instructional materials'' (Ellis, 1994: 12) and implies different findings. However, in all studies that have been done so far with the purpose to find effects of instruction with the focus on grammar, the explicit focus on forms appears to be effective in promoting language learning and does not change the "natural processes of acquisition" (Ellis, 1994: 12).

2.1.2.2 The focus on meaning

Communicative learning can be sketched through the participation metaphor, in which the learner is regarded as "a person who becomes part of an L2 community" (De Bot et al., 2005: 106). The communicative approaches focus on the learner as "the activist in the learning process" (De Bot et al., 2005: 108). The participation metaphor complements the acquisition metaphor (Pavlenko & Lantolf, 2000).

Within the focus on meaning two different forms can be distinguished: planned focus on meaning and incidental focus on meaning. In both forms lessons show an emphasized focus on meaning or communication (Long, 1991) and the instructions for these subdivisions are different in just one respect. The

instruction concerning the enriched input in the *planned* focus on meaning is intensive; this means that learners are invited to focus primarily on meaning and have the opportunity to attend to a single, preselected form many times (Ellis, 2001). *Incidental* focus on meaning also involves primary attention to meaning but here the instruction is extensive; a lot of linguistic forms can be paid attention to. Incidental focus on meaning distributes attention to a wide range of forms that have not been preselected (Ellis, 2001).

Furthermore, incidental focus on meaning is split up in pre-emptive and reactive focus on meaning, which both can arise when there is a communication or a form problem. In *pre-emptive* focus on meaning either the teacher or a learner takes a time out from a communicative activity to initiate attention to a form that is perceived to be problematic, whereas *reactive* focus on meaning consists of the negative feedback that teachers provide in response to actual or perceived errors that learners make (Ellis, 2001) (see figure 2.2).

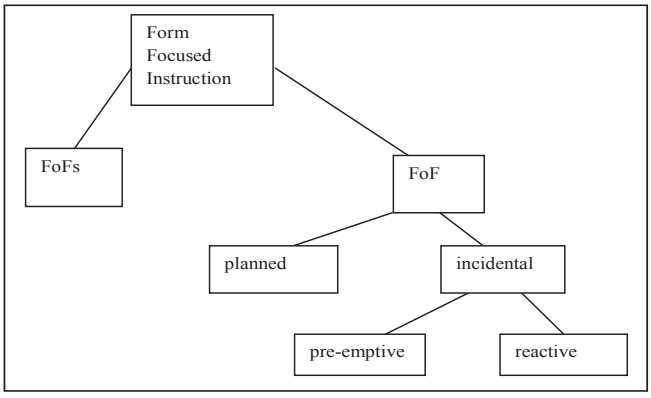

Figure 2.1: Different types of instruction (Ellis, 2001).

2.1.3 The language learning environment

For some decades now immersion programs in language exist, in order to give the learner the opportunity to pay attention to meaning and communication. However, even though the importance of the role of immersion in second language development has been recognized, the interaction that takes place should always be looked at as "a scene in which potential learning can take place" (De Bot et al., 2005: 184).

Several studies have been conducted in order to explore the influence of stay abroad on second language proficiency. In order to be able to find substantial changes in the syntactic complexity of writing in the second language in research on influence of a year abroad, Ortega (2003) chose a one year observation period of college-level instruction. In her longitudinal research, Ortega found that the language learning context (the language to be learned was the first language or not) influences second language proficiency and syntactic complexity in writing (Ortega, 2003). In comparable research on oral segments, such as Collentine (2004), this influence was found as well; the

learners in an environment in which the language to be learned was the first language acquired better narrative abilities. Moreover, they were able to produce a more semantically dense language. On the contrary, the learners in an environment in which the language to be learned was not the first language, developed more discrete grammatical and lexical features (Collentine, 2004). Furthermore, Coleman (1997) explored many investigations on the influence of living abroad on language proficiency. He came to several conclusions, among other things that the fact of living abroad leads to greater language proficiency. Moreover he found that living abroad leads to greater syntactic complexity and improved linguistic skills (among other things in speaking and aural comprehension) and to a reduction in errors and some morpho-syntactic progress. However, the individual variation of the influence of living abroad on language proficiency was considerable and the impact of living abroad on sociolinguistic norms appeared to be larger than its influence on the grammatical skills of the learner (Coleman, 1997). Finally, Ellis (1994) found that there might be a link between formal instruction and natural exposure, and that this combination might help the development of greater second language proficiency.

Summarizing, we may say that Finnish may be a difficult language to learn by speakers from other language groups. There are mixed and complex results about the role of instruction and immersion. Instruction seems to be needed to notice forms the learner may not notice him or herself. However, instruction may also lead to delay in use if there is no communicative need for the form. Finally, a large metalinguistic study (Norris & Ortega, 2000) has found the effect of instruction to be positive on the whole, but with the reservation that the testing may have been biased to the role of instruction. Immersion studies claim that the learning process of the first and a second language are quite similar, especially as far as the order or sequence of acquisition is concerned. As the current study deals with a comparison of two small groups over time, one in an FL setting with grammar taught very explicitly (FoFs) and one in an L2 setting with a focus on meaning (FoF) and hardly any attention to grammar rules, the next section will deal with the forms that need to be acquired in Finnish. In the comparison of the two groups the study will look especially at complexity and accuracy measures.

2.2 The Finnish language as learning target

This section presents several issues on Finnish syntax and morphology with examples and the syntactic and morphological complexity and accuracy measures that will be used in the study.

2.2.1 Sentence patterns

In this subsection we describe several sentence patterns. After the descriptions we present the measures that will be used in the current study.

2.2.1.1 Sentence types and structure

As many Indo-European languages, Finnish makes use of simple, compound, complex and compound-complex sentences (as defined by Verspoor & Sauter, 2000). Below, these sentence types are presented in several examples with finite verbs in bold. A simple sentence (example 1) contains one finite verb and a compound sentence (example 2) contains at least two main clauses that both have their own subject and finite verb. A complex sentence contains at least one subordinate (i.e. an adverbial or a relative clause) (examples 3 and 4) and the compound-complex sentence (example 5) is a combination of the qualities of a compound and a complex sentence. The parts in bold (in the Finnish examples) are the discussed issues in question.

(1) Minun äitini **on** sihteeri. *'My mother is a secreatary.'*
(2) Minä **asun** Suomessa ja minä **olen** saksalainen. *'I live in Finland and I am German.'*
(3) Jos minä **olen** masentunut, minä **menen** kävelylle. *'If I am depressed, I go for a walk.'*
(4) Illalla minä **opiskelen** suomea, koska minä **olen** työssä päivällä. *'In the evening I study Finnish, because I am at work through the day.'*
(5) Kun minä **menin** kouluun, minä **kävin** mummolla ja äitini **oli** töissä. *'When I went to school, I went to grandmother and my mother was at work.'*

Even though the Finnish sentence structure shows rather free word order of the major grammatical functions subject(S) – verb(V) – object(O), it is the basic word order (Vilkuna, 1989; Newson & Maunula, 2006). To start with, this word order occurs in a Finnish active, affirmative sentence that has the aim to express a neutral meaning. In such a sentence, the direct object (example 6), the clauses (examples 7 and 8) and the predicate nominal (example 9) are put at the end of a sentence. SVO is also used in a question that starts with a pronoun (example 10) and in a negative sentence (example 11). In a question that starts with a verb, the VSO sequence appears (example 12). Compound, complex and compound-complex sentences maintain SVO in the part of the sentence after the conjunction (examples 13, 14 and 15).

(6) Mies (S) ostaa (V) **takin** (O). *'The man buys the jacket.'*
(7) Antti (S) tulee (V) **tästä ravintolasta** (X[3]). *'Antti comes out of this restaurant.'*
(8) Minä (S) luen (V) tämän kirjan (O) **ensi kesänä** (X). *'I will read this book next summer.'*
(9) Hän (S) on (V) **terve** (X). *'He is healthy.'*

[3] X: other than the object.

(10) **Kuka** (S) osti (V) juomat (O)**?** *'Who bought the drinks?'*
(11) Sinä (S) **et tunne** (V) häntä (O). *'You do not know him.'*
(12) Tunnet**ko** (V) sinä (S) hänet (O)**?** *'Do you know him*?'
(13) Hän on neljä vuotta **ja** hän on suloinen. *'She is four years old and she is cute.'*
(14) Minä en usein lue kirjallisuutta, **koska** minulla ei ole aikaa. *'I do not read literature, because I do not have time.'*
(15) Minä olen kova lukemaan **ja** minä luulen **että** kukaan ei ole liian vanha oppimaan kirjallisuutta. *'I love to read and I think that nobody is too old to study literature.'*

The current study will explore the use of simple, compound, complex and compound-complex sentences. These are all regarded as general syntactic complexity measures, measured "by any length-based metric with a potentially multiple-clausal unit of production in the denominator" (Norris & Ortega, 2009: 561). The measures will be examined separately in chapter 4. The way we use these measures is not standard in second language acquisition research; only few studies did this before (Spoelman & Verspoor (2010); Verspoor et al. (2012)). These measures have the advantage of providing "a comprehensive overview of the types of sentences contained in a language sample" (Bulté, 2013: 79), though detailed information concerning the sentence composition can not be found this way.

The more complex sentence types (compound, complex and compound-complex) are the first of three syntactic complexity measures to be explored in chapter 5. Because through the addition of all kind of elements (adjectives, subordinate clauses etc.) a sentence can be lengthened, such measures are regarded as general complexity measures (Norris & Ortega, 2009). As the participants of the current study are beginner learners of Finnish but at the same time all adults (and therefore advanced in their L1), we want to find how they develop in terms of sentence complexity: do they write more simple sentences in the beginning and more complex sentences later on, like Verspoor et al. (2012) found?

2.2.1.2 Tenses

The four tenses of the Finnish language are presented in the examples 16 to 19. In the examples, the morphemes of the tenses are shown by slashes and no slashes are used to show morphemes between words. The present and past are simple tenses, whereas the perfect and the pluperfect both are compound tenses; they consist of at least two words (Karlsson, 1999).

(16) Present:	minä sano/n	*'I say'*
(17) Past:	minä sano/i/n	*'I said'*
(18) Perfect:	minä ole/n sano/nut	*'I have said'*
(19) Pluperfect:	minä ol/i/n sano/nut	*'I had said'*

The active present tense is used when an action takes place in non-past time. This means that the moment of the action falls together with the moment of the utterance (example 20) (Hahmo & Liebe, 1998). This can also be the case in a negative sentence: expressed by an inflected verb (example 21) in congruence with the subject and followed by the inflectional stem of the main verb (Karlsson, 1999; VISK, 2004: §108). The active present tense is also used when the action is done in future time (example 22) and when something happens on a regular basis (example 23) or when the utterance expresses an action done at this moment by a not further specified person or group (example 24) (Hahmo & Liebe, 1998; Karlsson, 1999). These kind of utterances are done in the passive voice: the passive voice is used when it is unclear or not relevant who the actor is; the action is more important than the actor (Karlsson, 1999; Ahonen, 2006).The active present tense is also used to express a general truth (example 25) (Karlsson, 1999). Finally, Finnish uses the passive present tense (without ending) in combination with the third person singular of the inflected auxiliary verb for a negation in the passive form (example 26).

(20) Minä **olen** ulkona. *'I am outside.'*
(21) Minä **en tunne** hanta. *'I do not know him.'*
(22) **Haluatko** sinä maitoa huomenna? *'Do you want some milk tomorrow?'*
(23) Minä **harrastan** tanssimista. *'My hobby is dancing.'*
(24) **Puhutaan** suomea! *'Let's speak Finnish!'*
(25) Strutsi **on** eläin. *'An ostrich is an animal.'*
(26) Häntä **ei tunneta**. *'He is not known.'*

The past tense is used in Finnish when an action already took place before the moment of utterance. This is also the case when there is no connection between the action and the present time (examples 27 and 28). The latter holds for the passive voice as well (example 29) (Hahmo & Liebe, 1998; Karlsson, 1999; Ahonen, 2006). The past tense is also used in combination with expressions of time to utter facts that happened in the past (example 30).

(27) Mitä sinä **teit**? *'What did you do?'*
(28) Minä **en kertonut** Pekalle mitään. *'I did not tell Pekka anything.'*
(29) **Tul|iin** Amsterdamiin. *'We arrived in Amsterdam.'*
(30) Viime viikolla he **antoivat** minulle veitsen. *'Last week they gave me a knife.'*

The perfect tense is used when an action was done in the past but still influences the moment of utterance or continues in it (example 31). Besides, the perfect tense is used to express an action in the past of which the exact time is not relevant (example 32) (Karlsson, 1999; Ahonen, 2006).

The pluperfect is used when an action took place before a certain moment in time in the past (example 33) (Hahmo & Liebe, 1998; Karlsson, 1999; Ahonen, 2006). When the (plu)perfect tense is used in negation, the past

participle in congruence with the subject is used in addition of a form of the auxiliary verb *olla* (*to be*) (example 34).

(31) Minä **olen asunut** Helsingissä kaksi vuotta. *'I have lived in Helsinki for two years.'*
(31) Mari **on ostanut** tietokoneen. *'Mari has bought a computer.'*
(33) Äiti **ei ollut nähnyt** häntä ennen kuin hän tuli Tukholmaan. *'Mother had not seen him before he came to Stockholm.'*
(34) He **eivät ole kertoneet** mitään. *'They have not told anything.'*

The complex tense use (past and (plu)perfect) serves as a sign of greater complexity (Bulté, 2013) and are therefore used as one of the two morphological complexity measures in chapters 7 and 9.

2.2.1.3 Finnish cases

Until about the 19th century, most of the books on Finnish grammar were based on Latin categories (Martin, 1995) because they were intended for people with an Indo-European language background (Vihonen, 1978). This implies that in grammar books all cases were derived from the nominative form (Wiik, 1988). However, the application of the Indo-European case system on the Finno-Ugrian languages was questioned regularly, continuing even in the 20th century (among others by Von Farkas, 1956).

The fact that the Latin tradition was taken as point of view for classification of the cases led to misunderstandings in the comprehension of the Finnish case system, especially in the understanding of the use of cases for the direct object; where the languages from the Uralic language family do not distinguish a separate subject and direct object case, Indo-European languages do so, with confusion as a consequence. Grünthal (1941) can serve as an example for this: he used the Finnish word objekti as the equivalent of the German Akkusativ in a text that he had translated (Hasselblatt, 1998). In 1997, Volodin invented the term Indo-European spectacles for the process of following the Latin tradition in the description of the Finnish cases (Volodin, 1997), i.e. treating the accusative case as being of the same order as the direct object (Kiparsky, 2005). In the current study the choices of Iso Suomen Kielioppi **(**VISK, 2004[4]) are followed, among others on the treatment of the cases for the direct object: from Finnish perspective.

The Finnish language has fifteen cases in total (VISK, 2004: §81), all involved in this study. The classification of the Finnish case system is presented in figure 2.1.

[4] VISK: Verkkoversio Iso Suomen Kielioppi (Internet version of the Large Grammar of Finnish, published on the internet in 2004).

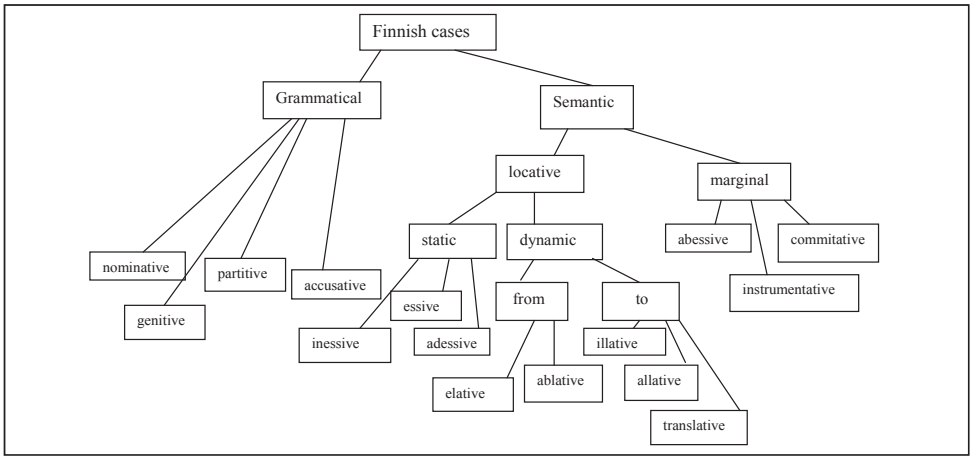

Figure 2.2: Classification of the 15 Finnish cases.

The cases can be classified in grammatical (or structural) and semantic cases. The grammatical cases are composed of the nominative, genitive, partitive and accusative[5]. The semantic cases consist of the other eleven cases and are divided in two subsystems: locative cases and marginal cases. Locative cases can be subdivided in different ways. The first way to do this is from a directional point of view[6]. In this approach, the locative cases are subdivided in two parts, static and dynamic cases. The word static is meant as a way of being (an existence) which is in English denoted by *in*, *on* and *at* (Martin et al., 2010). The static locatives include the inessive, essive and adessive. The word dynamic is defined as "a direction of the movement" (Martin et al., 2010: 64), in English denoted by *from* and *to* (Martin et al., 2010). The cases for the dynamic expression *from* are the ablative and elative, the cases which express *to* are the illative, allative and translative. The locatives can also be subdivided from the point of view of quality; this classification was not used in figure 2.2. In this way an expression of "the nature of the relationships" (Martin et al., 2010: 64) is represented by dividing the locatives into three kinds. The first one is internal, in English denoted by the expressions *in*, *into*, *from* and *out of* and in Finnish connoted by the illative, inessive and elative. The second one is external, in English denoted by *on*, *onto*, *off* and *from* and in Finnish connoted by the allative, adessive and ablative. The third one is general, in English denoted by *being* and *becoming*, in Finnish connoted by the inessive and translative. Finally, the three marginal cases, which are rare and mainly appear in fixed expressions or grammatical constructions, are the instructive, abessive and comitative (Karlsson, 1999; Ahonen, 2006).

Cases are used in nominals (nouns, adjectives, pronouns and numerals) and in the non-finite verb forms: participles and infinitives (Karlsson, 1999).

---

5 Abbreviations of the cases are, when useful, applied in order to indicate the cases; for an overview: see Appendix A.

6 This classification was used for the locative cases in figure 2.2.

There is one exception though; the accusative case is applied only to personal pronouns and to the pronoun *kuka* (*who*) (VISK, 2004: §81).

What follows are six Finnish language issues that concern cases. For every issue choices are to be made in order to use the correct case.

The first issue, *government,* is used to describe the fact that the Finnish language contains several fixed combinations of verbs, substantives, adjectives, adverbs and adpositions with a certain case to express something specific [7](VISK, 2004: §1225). These fixed combinations as well as chunks (made of parts of speech in combination with a certain fixed case to express a certain fixed meaning) can be understood as contractions in an Indo-European language (VISK, 2004: §1225).

Examples 35 to 37 show examples of fixed combinations with a verb. Example 38 shows a substantive with a certain case marking in another part of speech, whereas example 39 shows idiom with a fixed meaning.

(35) vb.-pron.Minä tykkään häne**stä**. *'I like him.'* → tykätä +elative = to like something/somebody
Minä olen varma tä**stä**. *'I am sure of this.'* → olla varma+elative = to be sure of something/somebody
Minä olen ylpeä häne**stä**. *'I am jealous of him.'* → olla ylpeä +elative = to be proud of something/somebody

(36) vb.-subst. Minä puhun suome**a**. *'I speak Finnish.'* → puhua+partitive = to speak a certain language

(37) vb.-adj. Kahvi tuoksuu hyvä**ltä**. *'The coffee smells good.'* → tuoksua +ablative = to smell in a certain way

(38) subst.-subst.Kiitos korti**sta**. *'Thank you for the postcard.'* → kiitos+elative = to thank somebody for something

(39) Mikä elokuva menee Ritzi**ssä**? *'What's on at the Ritz?' lit.: 'Which movie goes in the Ritz?'* → mennä +inessive = to be (going) on in/at

The second issue, *congruence,* is explained as an agreement in number and case and the phenomenon can take place in phrases and clauses (Karlsson, 1999; VISK, 2004: §1267).

Two types of congruence can be distinguished (VISK, 2004: §1267). The first one is morphological congruence. To this category belong the relationships of the headword with an adjective, being an attributive (examples 40 and 41) and the headword with a predicate nominal (examples 42 and 43). Abbreviations are applied, if useful. For an overview see Appendix A:

(40) **sellainen** (attr. nom.sing.) **kirja** (subst.nom.sing.) *'such a book'*

(41) Hän opiskelee **hyvässä** (attr. iness.sing.) **suomalaisessa** (attr. iness.sing.) **yliopistossa** (subst. iness.sing.). *'He[8] is studying at a good Finnish university.'*

---

7 A combination of words with a locative case, in order to express a direction or a quality, is not included as these are not falling under the umbrella of government.

8 The Finnish pronoun 'hän' can mean 'he' as well as 'she'. In all translations in this study the word 'he' is used, but it could be 'she' as well.

(42) **Mies** (subst. nom.sing.) on **vanha** (pred.nom.nom.sing.). *'The man is old.'*
(43) **Miehet** ovat **vanhoja** (pred.nom.part.pl.). *'The men are old.'*

The second type of congruence is congruence in number and person. This is explained as congruence of a verb with its subject (examples 44 and 45).

(44) **Hän** (subst.nom.sing.) **tulee** (vb.3.sing.). *'He is coming.'*
(45) **Miehet** (subst.nom.pl.) **puhuvat** (vb.3.pl.). *'The men are talking.'*

The third syntactic issue concerns *constructions with a logical and a grammatical subject*. In Finnish, the basic case for the subject is nominative (Ahonen, 2006). However, the possibility of a logical subject exists. This means that one part of speech has the semantic role of subject in the sentence, whereas another part of speech is the actual, grammatical, subject (Ahonen, 2006). The logical and grammatical subjects are present in sentences with a verb of necessity or obligation (example 46), with a verb with a modal meaning (example 47) and with a verb which expresses the emotional condition of the logical subject (example 48) and also in sentences which include the so called have-construction (to be explained separately below). These kinds of verbs only occur in the third person singular (Karlsson, 1999). The case of the logical subject is genitive (example 46) or partitive (examples 47 and 48).

(46) Minu**n** (log.subj.) **täytyy** (gramm.subj.) lähteä. *'I have to leave.'*
(47) Minu**a** (log.subj.) **kiinnostaa** urheilu (gramm.subj.). *'I am interested in sports.'*
(48) Minu**a** (log.subj.) **itkettää**. *'I feel like crying.'*

The have-construction falls under the constructions with a logical and a grammatical subject as well. In Finnish the verb *to have* does not exist (Ahonen, 2006) and the possessive structure in Finnish is therefore realized by the combination of a possessor in the adessive (this is the logical subject in the sentence), the third person singular of the verb *olla* (*on*) and the possessed person or thing (this is the grammatical subject of the sentence) (Karlsson, 1999) (examples 49 and 50). The use of the adessive in order to express possession with a personal pronoun is common (example 51) (Karlsson, 1999). When the possessed item is not transferable from one person to another and when, moreover, there is an intimate connection between the possessor and the possessed item, the inessive is applied (example 52). The nominative is the case of the possessed item when the item is countable; this involves singular as well as plural forms. The case of the possessed object is partitive when the item is not countable (example 53). When a sentence expresses an item which is not possessed (example 54) the case of the item is partitive as well. In the latter condition the partitive is also the case of a countable item (example 55); i.e. in this situation the case of the not possessed item is always partitive (on development of the have-construction: see Kajander (2013)).

(49) **Antilla** (possessor/log.subj.) **on** (3.sing.) **pyörä**.
(poss.th.sing./gramm.subj.). *'Antti has a bicycle.'*
(50) **Äidillä** (possessor/log.subj.) **on** (3.sing.) **silmälasit**[9].
(poss.th.pl./gramm.subj.). *'Mother has glasses.'*
(51) **Meillä** (log.subj.) on uusi tietokone (sing. gramm.subj.). *'We have a new computer.'*
(52) Puu**ssa** (log.subj.) on vihreät lehdet (pl. gramm.subj.). *'The tree has green leaves.'*
(53) Liisalla (log.subj.) on kahvi**a** (gramm.subj.). *'Liisa has coffee.'*
(54) Liisalla (log.subj.) **ei ole** kahvi**a** (gramm.subj.). *'Liisa does not have coffee.'*
(55) Minulla (log.subj.) **ei ole** puhelin**ta** (gramm.subj.). *'I do not have a telephone.'*

The fourth issue is syntactic in nature: Finnish contains several *sentence types with no subject*. In such a sentence the predicate is always in the third person singular (VISK, 2004: §1350). In an active sentence without an explicit present subject a zero-subject can be put at the place of the subject. A zero-subject tells us something about a certain place or condition. It is often located at the beginning of a sentence (example 56) but it can also be preceded by another part of the sentence (example 57) (VISK, 2004: §1350; Ahonen, 2006). Moreover, no subject is present in a meteorological sentence (example 58) or in a sentence with a total adverbial clause (example 59) or an infinitive (example 60) as a subject. Finally, in a sentence in the passive voice (example 61), in a generic sentence (example 62) and in a sentence with an imperative form (example 63), no subject is present either.

(56) **Täällä** on kaunista. *'It is cold in here.'*
(57) Ei **noin** voi sanoa. *'One cannot say it like that.'*
(58) On kaunis **ilma**. *'The weather is nice.'*
(59) On tärkeää, **että sinä tulet**. *'It is important that you will come.'*
(60) On helppo **ymmärtää**, että... *'It is easy to understand, that…'*
(61) Tuolla **myydään** jäätelöä. *'There they sell icecream.'*
(62) Tämän oven **saa** helposti auki. *'One can get this door open easily.'*
(63) **Auta** minua! *'Help me!'*

In Finnish, the case of the *direct object* (fifth issue) in the Finnish language is chosen out of the partitive, nominative, accusative and genitive. The choice for the correct case is sketched as follows:

1] Partitive: the sentence is negative

---

[9] n.b. Compound words in the Finnish language are often made of two or more non-derived nouns (tanssi/teatteri:*'dance-theater'*; rauta/tie/asema: *'rail-way-station'*), which are in Finnish not divided by a space (Karlsson, 1999). Within a compound word the cases of the compounds may not be similar, a phenomenon which might cause confusion in the decision of a case for learners of Finnish as a foreign and second language (silmä/lasi*t*: *'glasses'*; revo*n*/tuli *'northern lights'*).

	for semantic reasons (e.g. the expressed action is irresultative)
	for lexical reasons (e.g. the direct object is not countable)
2] Nominative:	for grammatical reasons (e.g. the sentence is in the passive voice)
3] Accusative:	the direct object is a personal pronoun
4] Genitive:	remaining

The last syntactic issue concerns the *predicate nominal*. The examples 64 and 65 show sentences in which the subject is countable and in singular. In such situations the *predicate nominal* is put in the nominative singular. However, when the subject in not countable and in singular, the predicate nominal is put in the partitive singular (examples 66 and 67). When the subject is countable, in plural and part of a larger group, the predicate nominal is put in the partitive plural (example 68). When the subject is countable, in plural and forms an entirety, the predicate nominal is put in the nominative plural (example 69).

(64) Täma on **sanakirja**. *'This is a dictionary.'*
(65) Hän on **englantilainen**. *'He is an Englishman.'*
(66) Kahvi on hyvä**ä**. *'Coffee is good.'*
(67) Aika on raha**a**. *'Time is money.'*
(68) Me olemme opiskelijo**ita**. *'We are students.'*
(69) Ikkunat ovat likaise**t**. *'The windows are dirty.' (i.e. of this particular building)*

The compound measure that consists of incorrect case use in the just presented phenomena serves as syntactic accuracy measure CAR use errors (4.1.2) in chapters 8 and 9, because these phenomena together represent the use of cases in such a way, that the errors that are made in the cases can really be described to the use of the case instead of another reason (like e.g. the incorrect use of a semantic rule). The measure is also examined in chapter 4, in combination with the incorrect case form.

2.2.2 Issues of morphology

To be able to find interesting measures for the exploration of morphological complexity and accuracy, several morphological issues are described. After the descriptions we present the measures that will be used to answer the research questions.

2.2.2.1 Morphemes

Finnish belongs to the group of synthetic languages. To this group belong languages in which words can consist of many morphemes, which are the smallest meaningful units of language. Besides, within words alternations of stems and suffixes can take place and these alternations, arising through inflections (conjugations and declensions), occur frequently in the Finnish

language (Holmberg & Nikanne, 1993; Lagus et al., 2005). The frequency of the inflections in the language, as the characteristic way of the usage of morphemes, is the reason why the Finnish language is called an agglutinative language.

The process of the segmentation of morphemes in order to trace the meaning of each separate morpheme in each word is needed to understand an utterance in total. For verbs (example 70) as well as for other words (in the case of example 71: noun) holds that the endings (bound morphemes or suffixes, which only appear as part of a larger word) are added to the stems (Karlsson, 1999). The morphemes within words are separated by slashes.

(70) Kysy/**i/t/kö**(vb./past/2.sing./que.cl) sinä paljon/**ko**(que.cl) se maksa/**a**(3.sing.)? *'Did you ask how much it costs?'*
(71) Hotelle/**i/ssa/kin**(subst./pl./iness./que.cl.) o/**n**(3.sing.) sauna. *'Also in the hotels is a sauna.'*

Words can be formed from existing words and stems through derivation: the addition of a derivative suffix to the root of a word or to another stem. There are several kinds of derivation; a deverbal suffix which forms a new nominal (example 72), a denominal suffix which forms a new nominal (example 73), a deverbal suffix which forms a new verb (example 74) and a denominal suffix which forms a new verb (example 75). The derivative suffixes serve thus as part of a word, but because they look like separate suffixes they can cause confusion when a word needs extra morphemes to express more (example 76: case marking).

(72) -jA: ajaa *'to drive'* → aja/**ja** *'driver'*
(73) -stO: kirja *'book'* → kirja/**sto** *'library'*
(74) -skele-: oppia *'to learn'* → opi/**skel**la *'to study'*
(75) -ile- teltta *'tent'* → telttai**l**la *'to go camping'*
(76) kirja/sto/**ssa** *'in the library'*

The huge number of deverbal verbal suffixes is typical for Finnish and makes the language very different from Indo-European languages (Karlsson, 1999). A word can consist of a lot of derivative suffixes which are placed after a root of the word and before the inflectional endings (Karlsson, 1999).

In second language writing, syntactic complexity is often seen as quantification of a certain feature, like sentence or clause length (Ortega, 2003). In the agglutinative language Finnish, the use of more morphemes on word or sentence level serves as sign of greater complexity and therefore (and different from e.g. the finite verb ratio for English, which is not a agglutinative language) the average sentence, clause and word length in morphemes are used, respectively as the second and third syntactic complexity measure and the second morphological complexity measure in chapters 5, 7, and 9. The three measures will also be examined in chapter 4.

2.2.2.2 Case markings[10]

First the Finnish case markings are shown after which we present the phenomenon of consonant gradation.

Table 2.A provides an overview of the Finnish cases and their markings in singular and plural. The endings are separately outlined in the third (sing.) and fifth (pl.) row and the markings (i.e. within words) are found in the second (sing.) and fourth (pl.) row.

Table 2.A: The 15 Finnish cases and their markings (VISK, 2004: §81).

case	sing.	ending	pl.	ending
nom.[11]	tuttu, tuote	–	tutu-t, venee-t	*-t*
gen.	tutu-n, tuottee-n	*-n*	tuttu-j[12]-en, poik-i-en, paperi-en ~[13] papere-i-den, tuotte-i-den ~ tuotte-i-tten, nais-ten, vanho-j-en ~ vanha-in	*-en*, *-den* ~ *-tten*, *-ten*, *-in*
part.	tuttu-a, maa-ta, tuote-tta, toin-ta ~ toime-a, tärkeä-ä ~ tärkeä-tä	*-A*[14], *-(t)tA*	tuttu-j-a, poik-i-a, tuotte-i-ta, palvelu-i-ta ~ palvelu-j-a, fyysiko-i-ta ~ fyysikko-j-a	*-A*, *-tA*
acc.	minu-t	*-t*	meidä-t	*-t*
ess.	tuttu-na, tuottee-na	*-nA*	tuttu-i-na, tuotte-i-na	*-nA*
transl.	tutu-ksi, tuottee-ksi	*-ksi*	tutu-i-ksi, tuotte-i-ksi	*-ksi*
iness.	tutu-ssa, tuottee-ssa	*-ssA*	tutu-i-ssa, tuotte-i-ssa	*-ssA*
elat.	tutu-sta, tuottee-sta	*-stA*	tutu-i-sta, tuotte-i-sta	*-stA*
illat.	tuttu-un, tuottee-seen, maa-han, essee-seen ~ essee-hen	*-V*[15]*n*, *-hVn*, *-seen*	tuttu-i-hin, tuotte-i-siin, poik-i-in, korke-i-siin ~ korke-i-hin	*-hin*, *-siin*, *-in*
adess.	tutu-lla, tuottee-lla	*-llA*	tutu-i-lla, tuotte-i-lla	*-llA*
ablat.	tutu-lta, tuottee-lta	*-ltA*	tutu-i-lta, tuotte-i-lta	*-ltA*
allat.	tutu-lle, tuottee-lle	*-lle*[x]	tutu-i-lle, tuotte-i-lle	*-lle*[x]
abess.	tutu-tta, tuottee-tta	*-ttA*	tutu-i-tta, tuotte-i-tta	*-ttA*
com.			tuttu-i-ne, tuotte-i-ne	*-ine*[(16)]
instr.			tutu-i-n, tuotte-i-n	*-in*

Consonant gradation is about the fact that the consonants k, p and t appear differently in open and closed syllables than the rest of the consonants in

10 In the current study, the expression *case marking* is used: the applied case endings and the changes in the stem of the word are explored.
11 The nominative is the default form of the cases.
12 The letters j and i are used as markings of the plural forms.
13 ~= Two forms possible, this one and the one on the row below.
14 A= a or ä.
15 V=vowel, same as the former vowel in the word.
16 +poss.suff.

Finnish (Hahmo & Liebe, 1998).The examples show that in open syllables the three consonants appear in the strong grade (example 77) and in closed syllables they appear in the weak grade (example 78) (Karlsson, 1999; Ahonen, 2006). When an open syllable changes into a closed syllable or vice versa, alternation through conjugation or declension takes place[17].

(77) ka**k**-**k**u	*'cake'*	ku**p**-**p**i	*'cup'*	ma**t**-**t**o	*'carpet'*
(78) ka-**k**un	*'of the cake'*	ku-**p**issa	'in *the cup*'	ma-**t**olla	*'on the carpet'*

There are two ways of consonant gradation: quantitative and qualitative. We now outline the difference between these two forms.

1] Quantitative consonant gradation:

When in the Finnish language a short vowel is placed before the long consonants -kk-, -pp- and -tt-, they alternate in the short consonants -k-, -p- and -t- (examples 79 to 81) (Karlsson, 1999).

(79) Pe**k**-**k**a - Pe-**k**an	*'Pekka - of Pekka'*
(80) ku**p**-**p**i - ku-**p**in	*'cup - of the cup'*
(81) ha**t**-**t**u - ha-**t**un	*'hat - of the hat'*

2] Qualitative consonant gradation:

In general, short consonants alternate with other consonants. This process is called qualitative consonant gradation. Qualitative consonant gradation takes place in the combinations: -k- (Φ) (example 82)[18], -p- (-v-) (example 83) and –t- (-d-) (example 84)[19].

(82) ha**k**ea – haen	*'to collect - I collect'*
(83) kyl**p**y – kyl**v**yn	*'bath - of the bath'*
(84) löy**t**ää – löy**d**än	*'to find - I find'*

Also dependent on consonant gradation are the combination of letters -uku- (-uvu-), -rke- (-rje), -lke- (-lje-) and –hke- (-hje-) (Hahmo & Liebe, 1998; Karlsson, 1999; Ahonen, 2006).

Finally, assimilation alternation occurs in the combinations -nk- (-ng-) (see example 85), -mp- (-mm-), -nt- (-nn-), -lt- (-ll-) and -rt- (-rr-) (Karlsson, 1999; Ahonen, 2006).

(85) Helsi**nk**i – Helsi**ng**in *'Helsinki - of Helsinki'*

In chapters 8 and 9, the incorrect case form serves as morphological accuracy measure CAR form errors (4.1.2). This measure consists of errors in consonant

---

17 The process of consonant alternation does not comprise the letter combinations -sk-, -sp-, -st-, -ks-, -ps-, -ts-, -tk- and -hk-, even though the letters k, p and t are involved (Hahmo & Liebe, 1998; Ahonen, 2006).

18 We are aware of the fact that this example makes use of a verb. A verb may also have to do with consonant gradation. If in the current study a verb is formed incorrectly, the error is part of the measure total of errors.

19 Idem.

gradation and stem forms, both only in case markings. These issues are subjects of grammar rules and therefore, correct use of the Finnish case markings serves as sign of more accuracy. The measure is also examined in chapter 4, in combination with the incorrect case use.

2.3 Research questions and hypotheses

As must be clear now, the Finnish language makes many intricate distinctions in especially the case system that may be difficult to learn for students with a different type of L1, especially when such distinctions are rather opaque, and the question may be whether such subtle rules can ever be acquired without some formal instruction. The overall goal of the dissertation is to compare two groups of learners, one that is immersed in the language and has had little to none grammatical explanations and one that is taught as an FL with a focus on grammar.

Development will be studied from a dynamic perspective, which holds that patterns emerge from language use. As discussed in chapter 1, language development is assumed to be a non-linear process and tracing the differences in the degree of variability may give insight into this process. The constant fluctuations of language are considered as information; the differences in the degree of variability provide a window into the developmental processes. Individual differences and developmental processes are important. However, there still may general trends to be found in different groups.

The assumption in the current study is that all factors such as the effect of the L1 on learning the L2 (Murakami, 2013; Spoelman, 2013), the context such as a foreign language or second language and the type of instruction (focus on grammar or on meaning) interact. Besides, the assumption is that in the developmental process, variability is needed to develop (Thelen & Smith, 1994). In a study from a dynamic approach, variability is seen as a sign of change and development in the language system; low variability indicates that the system is relatively organized and stable, high variability indicates that the system is changing and reorganizing until it settles again at the next developmental stage (Verspoor et al., 2008). In the context of a language-learning activity this means that both the intrinsic dynamics of the learner and the external resources are included; the study of the learners and the context are not seen as two different phenomena but are studied together (De Bot et al., 2011). The aim of the current study is to find relations between different types of resources.

To summarize, the main goal of the current study is to see whether the developmental trajectories in two different contexts -Finnish as a second language or Finnish as a foreign language- are similar or different. This will be done by exploring the trajectories of 8 students in two conditions, with two focal students who will be discussed in detail. To be able to explore this question, the following specific research questions were formulated:

1. *Which of the complexity and accuracy measures as explored in the current study, show significant differences between the overall averages of the FL and the L2 groups?*

To answer question 1, we will perform a longitudinal group study, mainly to find overall differences between the two groups and select measures that are interesting to focus on at the beginners' level of development. Verspoor et al. (2004) suggested the possibility of different outcomes with different types of instruction and Spoelman and Verspoor (2010) found that for a Dutch learner, studying Finnish in the Netherlands (as a foreign language) with the focus on grammar, error rates decreased quickly, but that there remained difficulties with the correct application of the Finnish genitive and partitive cases. Even though Schoonen et al. (2003) found (for Dutch as L1 and English as L2) the linguistic knowledge and its accessibility were of less influence on writing proficiency than the L1 writing proficiency, the current study (in which all participants have no Finno-Ugrian language background) hypothesizes that the FL learners with the focus on grammar will sooner show a decrease in error rates and an increase in complexity in the writing than the L2 learners with the focus on meaning. In chapter 4, the group measures will be compared. In this chapter we will also select the two focal learners who will be traced in detail. The two learners will be selected on L1 background (Germanic), the number of writings available, and general personality traits. Their first two and last two texts are rated holistically on complexity, idiomaticity and accuracy by a team of expert raters to see if there are general differences between the two learners and subsequently on a number of specific variables that will be used in the remainder of the study.

Those variables that have shown significant differences in the group study will be used for the longitudinal case studies in chapters 5 and 6 on syntactic complexity (respectively on sentence length and on use of the Finnish cases), in chapter 7 on syntactic and morphological complexity measures, in chapter 8 on accuracy in use and form of the Finnish cases and in chapter 9 on the accuracy and complexity measures.

2. *Which quantitative syntactic complexity measure, as explored in the longitudinal data of one FL and one L2 learner in the current study, is the most informative expression of sentence complexity?*

To be able to measure learner outcomes, it is useful to know which measure best captures overall complexity. Sentence length can be a robust measure for Indo European languages and especially for English, which is low on morphology. However, for Finnish, an agglutinative language with words becoming internally more complex, a sentence does not necessarily have to have more words to become more complex (Verspoor et al., 2008; Verspoor et al., 2011, ch. 3) and it was found that the finite verb ratio (words) was a better complexity measure than average sentence length or type of sentence (simple or complex sentences); indirectly it includes noun phrase length and other non-

finite constructions. This measure is also based on word counts. For Finnish, we hypothesized that a finite verb ratio based on morphemes may be a better measure as it will also reflect the internal complexity of a word. Chapter 5 will explore the different complexity measures such as complex sentences types, sentence length in morphemes and clause length in morphemes. It will be argued that the finite verb ratio in morphemes shows the internal complexity at clause level best and will be used in further chapters as a base measure to compare other measures to.

3. *To what extent is the complex tense use related to the writing task?*

Verspoor et al. (2004) and Verspoor et al. (2012) did not look at the topics of the writing tasks, which might be of importance for the use of the tenses. Verspoor et al. (2004) found that the perfect tense was applied later than the present and past tense but in a rather unpredictable way and they concluded that the use of the tenses did not seem to become more complex over time. On the other hand, Verspoor et al. (2012) concluded the opposite; they found that the use of the tenses went from mainly use of the present tense to use of the other tenses.

The current study does look at the subjects of the writing tasks and we hypothesize that the complex tense use is task related and 'unpredictable'; on account of the findings of Verspoor et al. (2004) we expect that a writer will normally use the present tense and that he is influenced by the tense in which the question is put and also by the subject the learner has to write about to use other tenses (see chapter 7). In other words, this chapter tests whether complex tense use is task related and subsequently whether the measure is a good indicator of development.

4. *Which of the syntactic and morphological measures, as explored in the longitudinal data of one FL and one L2 learner in the current study, show significant peaks or dips?*

As Van Dijk et al. (2011) found that a beginning learner of English as a second language showed a significant peak in the use of don't-constructions (often a non-target form) before the non-target use of these forms disappeared, the current study will explore which measures may show significant peaks or dips, which indicate a developmental phase shift. It is expected that different measures will grow at different rates. Some will develop rather steadily and others will show clear phase shift. For example, some frequently used rule based items such as a nominative case form may first show a strong regression to non-target forms before the item is mastered. One question is whether instruction or not may have an effect on the rate of these shifts. It is possible that through instructions, some forms are noticed earlier and may cause some overuse. Therefore, in the current study, all measures that show strong peaks or dips in the variability or min-max graphs will be tested for significance by means of a Monte Carlo analysis.

5. *Which of the syntactic and morphological measures, as explored in the longitudinal data of one FL and one L2 learner in the current study, show support or competition?*

Spoelman and Verspoor (2010) found various supportive and competitive growers in the longitudinal data of an L2 learner of Finnish. In the current study a greater number of measures will be examined and even though there are no previous findings to base the hypotheses on, the following seem reasonable hypotheses:

First, it is hypothesized that most general syntactic complexity measures at the sentence and clause level are supportive. It is expected that when the sentence becomes more complex, it becomes longer in both number of words, at both the sentence and clause level (chapter 5).

Secondly, it is hypothesized that the more specific complexity measures at the morpheme level, the average clause length in morphemes and the average word length in morphemes, are supportive growers because they both involve syntactic measures, albeit at different levels (word versus sentence) (chapter 7). In the same vein, it is hypothesized that the average clause length in morphemes and the complexity measure past and perfect tense are competitive growers because we expect that less use of one will mean more complexity (i.e. on average more morphemes per clause and more use of past and (plu)perfect) for the other. The measures average word length in morphemes and past and perfect tense are expected to be supportive because use of the past or perfect tense implies higher use of morphemes.

Thirdly, based on findings by both Spoelman and Verspoor (2010) and Caspi (2010), it is hypothesized that among syntactic and morphological accuracy measures there is a sequencing (precursor) effect: first CAR (=case accuracy ratio) form errors will decline and then CAR use errors. The reason is that the form, which is visible, is more salient than use, with all kinds of subtle meaning nuances (chapter 8).

Fourthly, though Spoelman and Verspoor (2010) did not find any consistent interactions between their accuracy and complexity measures, several relations were explored in the current study. The expectation is that complexity and accuracy compete, especially early on. As the learner tries new more complex constructions at various levels, the chances are that the learner will first make errors in both form and use, which later may disappear. Therefore the expectation is that several syntactic complexity measures (average clause length in morphemes, average word length in morphemes and use of the past and perfect tense with their morphology) will compete initially with both use and form errors. This means that especially early on both the use and form errors are expected to increase when the different measures become more complex and vice versa. This will be outlined in chapter 9.

Summary

This chapter took off with an exploration of the literature about learning a language from another language family (as Finnish is for Indo-European learners). To such learners, the Finnish language contains several unfamiliar conceptual distinctions and constructions. Related to this, we discussed research that has shown that systems in a different second language are extremely difficult to acquire, after which we pointed out the discussion of talking 'about' or 'in' the language. Subsequently, we described traditional learning, with the focus on the language as a system and on the role of grammar and communicative learning, with the focus on the learner as the acting agent in the learning process and on the role of meaning. The section ended with a review on the influence of residence abroad on second language development. It suggests that the language learning context influences second language proficiency; however, the combination of formal instruction and natural exposure might help the development of greater second language proficiency (Ellis, 1994).

After this we explored several syntactic issues of the Finnish language (sentence types, tenses and cases). Also some issues of morphology, important for Finnish as a synthetic agglutinative language, were depicted (morphemes and case markings).

Finally, the research questions and hypotheses for the current study were presented.

3 DESIGN OF THE STUDY

The current study, quasi-experimental in that it cannot control for all conditions, aims to trace the development of Finnish as a foreign or second language in two groups and two selected individuals over the course of one academic year (nine months). The groups differ in both general condition (non-immersion versus immersion) and teaching approach (focus on grammar and focus on meaning).

The previous chapters reviewed the literature on the role of instruction and the measures that are worth investigating in the Finnish language. Moreover, we chose to take a dynamic perspective on development, which means that the study design should be longitudinal and have dense data. As Caspi points out, such a design "facilitates inspection of development as temporal change, and enables the relation and comparison of such change across various levels of the data" (2010: 23).

This chapter presents the design of the study. After introducing the participants, their background, their learning environments, the data collection, the analyses and the techniques, the operationalizations of all coded measures on complexity and accuracy are described in detail. Syntactic and morphological complexity measures are seen as basic descriptors of FL/L2 language performance and as indicators of FL/L2 language proficiency (Bulté & Housen, 2012). Complexity is measured in order to explore effects of the language learning environment and different types of instruction. Bulté and Housen (2012) argue that three different levels of linguistic complexity need to be distinguished. The first is the theoretical level stating what complexity is; in the current study complexity is operationalized as grammatical complexity at the morphological and syntactic levels. The second is the observational level stating how the grammatical complexity in morphology and syntax is manifested in actual language; in the current study the syntactic complexity is operationalized as sentence and clausal complexity and the morphological complexity as word complexity. The third is the operational level which states how these manifestations can be quantified.

We end by discussing possible pitfalls in taking a dynamic perspective.

3.1 The participants introduced

In general the term second language learning serves as an overall term for learning a language that is different from the L1, so it can refer to a second, third or fourth (and so on) language. However, in the previous chapters we introduced the separate terms foreign and second language and by doing so we followed Kecskes and Papp (2000):

> L2 development and FL development should be considered as two different entities because the underlying mechanisms responsible for the two types of development have more differences than similarities and result in two different types of language production. (Kecskes and Papp, 2000: 13)

The implication for the current study is as follows:

Foreign language: Learning a language (different from the L1) in an environment in which another language than the language to learn is the first language: Finnish as FL in the Netherlands.

Second language: Learning a language (different from the L1) in an environment in which the language to learn is the first language: Finnish as L2 in Finland.

In the current study eight learners are involved, in two conditions: FL and L2. In the following subsections the groups as a whole and, where relevant, the individuals are described in detail.

3.1.1 The FL group

The FL group (see Table 3.A) consists of four university students majoring in Finnish at the University of Groningen, the Netherlands (2006-2007). They all lived in the Netherlands at the time of study; Finnish was for them a foreign language.

The learners in this group are all female. Three of them were born in the Netherlands and one was born in Belgium. For these four learners the Dutch language is the L1. This implies that their L1 belongs to the Indo-European, Germanic, language family and for these learners, Finnish is quite a different language (see chapter 2). During their first year of studying Finnish, all FL learners had the same teacher: a native speaker of Finnish and MA in the Finnish language.

None of the learners has a Finno-Ugrian language background. Three of the learners had some previous knowledge of the Finnish grammar; Kim[20] had read about the Finnish grammar, Annet had lived in Finland for several months

[20] All names of the participants are fictional in order to guarantee their anonymity.

and Cleo corresponded (in English) with native speakers of Finnish about the Finnish grammar.

During the time of the study, the FL learners had six contact hours per week for about ten weeks per semester, and received 10 ECTS[21] per semester if they passed the exam, which they all did. From day one on, all competences (speaking, listening, reading and writing) were practiced weekly: speaking in the classroom, watching videos at home and in the classroom and reading and writing texts at home. For all learners the communicative need was not present in their everyday life and this might suggest that for them acquisition of a form was delayed (Haley & Rentz, 2002). All learners wrote two texts per week during a time span of two semesters, with one interruption of six weeks during the Christmas holidays. The students did the same writing assignments. The current study used the writing texts and started in the first week of September and ended in the last week of May; the time span of the data is nine months. For further detail, see the teaching approaches below.

The learners of the FL group did not know that their data would be used for this study at the time they did their homework. Permission to use the data was given after they had finished the course. The main reason these four students were selected was that they had kept a large number of their written texts during their first year. Admittedly, one of the learners, not the focal learner, had had previous experience with Finnish.

3.1.2 The L2 group

The L2 group (see Table 3.A) consists of four students who learned the Finnish language through immersion in the everyday life and one or more courses of Finnish at the University of Jyväskylä, Finland, with a maximum of nine months (2011-2012). For none of these learners, Finnish was a major; they had come to Finland for one or several years in order to do a bachelor (Andrea and Chiya), taught in English, a master (Bowo) taught in English or post doc research (Clara) in different fields. They had moved to Finland just before the courses and data collection had started, so they lived in Finland at the time of study, but were beginners; learning Finnish was for them learning a second language. They lived in residencies with other foreign students. During the first semester a weekly diary was kept by each participant indicating the amount of exposure and practice they had had during that week. All four L2 learners spent about the same amount of time (outside of class) on listening, reading, speaking and writing in Finnish, of which they spent most time on listening and reading (on average about 20 to 30 minutes per week, increasing in time). They spent on average about 10 to 20 minutes per week (increasing in time) minutes on speaking and writing in Finnish.

---

[21] European Credit Transfer and Accumulation System: a standard in order to compare study attainments and performances of students (higher education; European Union/other collaborating European countries).

None of the students had a Finno-Ugrian language background. This group of learners consists of three females and one male. One of them comes from Germany and has German as L1, one comes from Brazil with Portuguese as L1 (both languages from the Indo-European language family, but one Germanic and the other Romance). The third participant comes from Japan with Japanese as L1 (Japanese language family) and the last one comes from Indonesia and speaks Bahasa Indonesian as L1 (Malaysian Polynesian language family) (Anhava, 1998). Also for these learners, Finnish is a very different language.

All four learners were enrolled in the same Finnish course but with different teachers, all native speakers of Finnish with an MA in the Finnish language. The first two parts of the course, which lasted one semester, were followed by all four learners and the third part of the course was followed only by Andrea and Chiya. The fourth and last part of the course was not followed by any of the participants. None of these students had any foreknowledge of the Finnish language. During the time span of the course, the German learner had contact with a Finnish student ('EOTO partner'), volunteering as a buddy. The L2 learners had eight contact hours per week during about ten weeks per semester, and received 5 ECTS[22] per semester after passing an exam. From day one on, the competences talking, listening and reading were practiced. For all learners the communicative need was often present in their everyday life, which might suggest an effective need to acquire forms (Haley & Rentz, 2002).

The writing was not practiced in class but the researcher asked students to volunteer for the current study. In return for their participation, they were offered feedback on the written products.

The participants wrote one text every week for one academic year on the same topics as the Dutch students had; the writing was an extracurricular activity and fell outside the realm of the normal curriculum of the University education. The writing was done weekly with one interruption of four weeks during the Christmas holidays. The writing started in the second week of August, and the participants wrote their last free writing tasks in the first week of May; for this group the time span of the data is nine months as well.

The L2 learners knew beforehand that their data would be used for this study and gave their permission. They joined the research voluntarily, and received one ECTS per semester for their participation. Originally there were more participants but the current ones were selected because their data set is the most complete in that their data consists of the most material (one academic year). Finally, the other conditions of the learners are quite similar.

---

22 European Credit Transfer and Accumulation System: a standard in order to compare study attainments and performances of students (higher education; European Union/other collaborating European countries).

Table 3.A: Information on the study's eight participants.

group/name	sex	country of origin	L1	L2, L3 etc.
FL				
Kim	F	The Netherlands	Dutch	Eng/Ger/Fr
Sanne	F	Belgium	Dutch	Eng/Ger/Fr
Cleo	F	The Netherlands	Dutch	Eng/Ger
Annet	F	The Netherlands	Dutch	Eng/Ger
L2				
Andrea	F	Germany	German	Eng/Fr/Ita
Chiya	F	Japan	Japanese	Eng
Bowo	M	Indonesia	Indonesian	Eng
Clara	F	Brazil	Portuguese	Eng

3.1.3 The individuals

The groups will be compared in general trends in the data, but for the detailed dynamic analyses, two focal students (Kim and Andrea) were selected. Specifically these learners were selected because like the Dutch students, Andrea has a Germanic language as an L1. Both Dutch and German have case markings (with German actually more so than Dutch). Moreover both Kim and Andrea had the largest number of written texts, and used the largest number of total words, in their respective groups. Moreover, both participants were quite interested in the grammar of the Finnish language and eager to learn and understand grammatical rules. Kim expressed her interest in grammar during class and Andrea in the questions she asked about grammar matters when she sent her written texts. However, to control for instructional condition no explicit rules were given to her and the answers to Andrea's questions were as concise as possible. For a total overview of the data size of the two learners, see table 3.B.

Table 3.B: Overview of Kim and Andrea's written data.

participant	number of written texts	number of written words	average number of words per text	sample mean	missing texts
Kim (FL)	37 (out of 39)	3797	102.6	1,0	9,15
Andrea (L2)	32 (out of 35)	2788	87.1	1,2	8,26,29

3.2 The collection of the data

Even though this study is quasi experimental and not all variables could be controlled, an attempt was made to be as systematic as possible in keeping the instructional conditions as separate as possible and the data gathering and processing as similar as possible.

3.2.1 The teaching approaches

The FL group learned Finnish from the course book *Kuulostaa hyvältä!* (Ahonen, 2006) used in the first year of the study. The course book is written in Dutch, the L1 of the learners, and consists of five parts: a grammar book, a word list in which the meaning of new words is given (with some uses of cases), a video tape with short stories that were acted out by native speakers of Finnish (to be watched as homework), a textbook in which the transcriptions of the stories are put together with the Dutch translation and finally, a book with questions about the (video) text. Besides, the computer is a tool used weekly to record spoken Finnish and for (free) writing tasks. In the teaching approach, the focus lies on grammar (see chapter 2). This implies that the learners consider the language to be the study object; the focus of the course is on some specific elements of grammar and the purpose is to learn the grammatical rules of the Finnish language. In every lesson, several specific elements of grammar are discussed extensively. The goal of this particular course is to deal with all the grammar as it is the only course in the program to deal with it explicitly.

In the FL classroom, all grammar instruction was in Dutch. Also the communication between the teacher and the learners was in Dutch for the bigger part. As the period continued, in small talk the Finnish language was used more.

The L2 group learned Finnish from the coure book *Matkalla jossakin suomessa* (Aalto et al., 2003), used in the first semester of the academic year, i.e. the first two parts of the course. The course book is written partly in English and partly in Finnish, (neither the L1's of the learners) and consists of a book that contains a combination of short stories, questions wordlists, some grammar, and exercises. No videotape is included in the materials. In the teaching approach, the focus lies on meaning. The main focus of the class is to be able to manage in real life and only rarely some planned, intensive instruction is given on grammar. In the third and fourth part of the course no textbook is used; then the lessons are distributed on paper weekly.

In the L2 classroom, the communication between the teacher and the learners was mainly in Finnish; only then some English was used when the learner could not get the point of the conversation. During the time span the English language was used in some short moments of explicit instruction.

3.2.2 The procedures concerning the free writing tasks

Every week during the first year of their study, the FL learners wrote in Finnish and the topics of the free writing tasks were related to the stories on the video. The tasks were part of the homework. This means that the free writing tasks were done on the computer at home and that the learners were allowed and encouraged to use all kinds of sources; they could use a dictionary, the internet etc. In total, the learners had to write 39 texts in the first academic year (see appendix C for the topics) and they had to post the completed tasks on the internet site of the university. In the first 15 lessons, the learners were allowed to write any number of words but from task 16 on they were expected to write about 100 words per task. The texts were corrected by the teacher and the checklist (see appendix B for the consolidated version), that was used for this purpose (i.e. classify and code the incorrectly formed and used words and constructions), was created as a tool for *Kuulostaa hyvältä*. Subsequently, the learners had to revise the incorrectly formed and used words in order to be assessed on it at the end of every semester: the more correctly revised words, sentences and constructions, the higher the grade. The first versions of the written texts, i.e. without any corrections, were used for the data collection for the current study. The learners of the FL group were expected to use the checklist actively when they revised their writing tasks.

During one academic year, the L2 learners wrote a text every week. Because the videos that the FL group had used were not available for the L2 group, some free writing tasks were adapted somewhat so that the L2 group could understand them without reference to the video; the topics and the questions were as similar as possible to the FL tasks, though (see appendix C for an overview of the tenses, moods and forms, as used in the questions). The writing was done at home on the computer and the learners were allowed to use resources. In total the learners had to write 35 free writing tasks. In the first nine tasks the number of words to write was not given beforehand but from free writing tasks 10 on, the learners were asked to write 50-100 words per task (appendix D gives an overview of the actual written words per participant per text). We corrected the texts by replacing words that were incorrectly formed or used or were inappropriate in the context (the researcher is graduated in the Finnish language, though not a native speaker of Finnish). The revision only contained the corrected form of the item. No explicit explanation was given to control for the level of instruction. The learners of the L2 group did not have the checklist. Because two of the learners had reacted positively to the offer of revising the free writing tasks, the free writing tasks of Andrea and Chiya were revised from the start. Later on, the free writing tasks of Clara (from lesson 10 on) and Bowo (from lesson 20 on) were revised weekly as well. However, for the current study only the original texts were used.

3.2.3 Example sentences

In the forthcoming chapters, relevant example sentences or words will be used to illustrate phenomena. They will be formatted as follows. Example 1 shows between the brackets: writer, text number, sentence in the text, sentence length in morphemes, average clause length in morphemes and average word length in morphemes. The bold faced characters point to the feature to be discussed, in this case the use of the past and perfect tense. The English translation is always in italics. For a total overview of used abbreviations etc. in examples, see Appendix A and for an overview of all texts of all participants, see the electronic file of the dissertation.

(1) Kun hän **ol/i kuol/lut** // Eero **perust/i** oma/n arkkitehti/toimisto/on[d.obj]. *'When he had died, Eero founded his own architectural office.'* (Kim-t35s9/mp33-pl.perf2/2mrph-past2mrph-SlMrph1/14-ClMrph2/14(6-8:7.0)-WlMrph8/14(1.75))

In case of errors, the feature to be discussed is underlined. In example 2 the feature to be discussed is the incorrect use of cases, in this example the inflectional error (infl). The consonant gradation in this example is also incorrect, but because it is not relevant to the discussion at hand, it is not underlined.

(2) Lauantai**na** mun täytyy kirjoittaa esseettä[cons.gr] suome**lta**[infl] historia**lta**[infl]. *'On Saturday I have to write an essay about the history of Finland.'* (Andrea- t6s3/mp6(7:3))

If the topic to be discussed concerns the use of cases, the entire word in that case is bold. In example 3 the nominative is bold. If the topic concerns errors in the use of cases, only the case morphemes are boldfaced. In example 4 the last letters are in bold.

(3) **Se** on ilmeisesti **kansanomainen tanssi.** *'Apparently it is a folksy dance.'* (Sanne-t27s4/27(5:3))
(4) Suomi e/i ole todella[lex] helppo**[**pr.nom]. *'Finnish is not very easy.'* (Andrea-t4s1/mp4-WlMrph5/6(1.20))

3.2.4 Processing the data

In chapter 2 the measures to be explored in our learners were discussed and defined. Now, we will present the coding and counting of the data after which we will explain how the measures were calculated.

3.2.4.1 Coding and counting

The free writing tasks of the FL group were first checked by the original teacher, who coded for inaccuracies. However, we double checked the texts for several reasons. First of all, not all codes were used for the current study, so only the inaccuracies that pertained to this study were selected. Moreover, the double check was done to make sure all errors were coded consistently and to make sure no errors had been overlooked.

The texts of the learners of the L2 group were coded by an advanced (master, fifth year of study) student of Finnish, who is a native speaker of Finnish. This student received five ECTS for coding the free writing tasks. For the coding the consolidated version of the checklist was used (appendix B). Before the student started to code, we had carefully explained the procedure of coding and in case of doubt or questions the advanced student contacted us or put marks in the coded texts. Finally, the researcher double checked the coded texts in order to make sure that the process of coding had been done the same way as it was done for the FL group and to look for errors that might have been overlooked.

In addition to the coding of errors, the researcher coded and counted several syntactic and morphological complexity measures such as the number of morphemes and words per sentence.

From each text a random sample of approximately 100 words was selected with a 10 percent deviation (to include whole sentences); i.e. the texts used for the study were between 95 and 105 words in length. Because early on, some of the texts did not consist of 100 words, the numbers of occurrences of the various measures were normalized to 100.

Subsequently, an excel data base was created with all texts and measures. In the Excel data base every feature was counted and filled in per writing sample per learner. All the data of one learner of one year were linked to two overall excel sheets (a version consisting of the real number of words and a version with the normalized data) in order to have an overview of all counted features. The data of the sampled overall files per learner was used to explore and analyze by techniques outlined in section 3.4.

3.2.4.2 Operationalization

Now we will explain how the measures were calculated. There are three kinds of measures: syntactic complexity, morphological complexity and accuracy in the use and form of cases.

Syntactic complexity

For syntactic complexity we looked at types of sentences, sentence and clause length in terms of morphemes and the use of cases. The classification of the four sentence types is based on Verspoor and Sauter (2000).

Simple sentence
Per text all simple sentences (correct and incorrect) are counted. The total number of simple sentences is divided by the total number of sentences per text and multiplied with 100 and presented in percentages.

Compound sentence
Per text all compound sentences (correct and incorrect) are counted. The total number of compound sentences is divided by the total number of sentences per text and multiplied with 100 and presented in percentages.

Complex sentence
Per text all complex sentences (correct and incorrect) are counted. The total number of complex sentences is divided by the total number of sentences per text and multiplied with 100 and presented in percentages.

Compound-complex sentence
Per text all compound-complex sentences (correct and incorrect) are counted. The total number of compound-complex sentences is divided by the total number of sentences and multiplied per text with 100 and presented in percentages.

Compound, complex and compound-complex sentences
The total use of compound, complex and compound-complex sentences represent the total number of correct and incorrectly used and/or formed compound, complex and compound-complex sentences per text, in percentages. Because every sentence is subsumed in one of the four types, the total of these specific sentence types is divided by the total number of sentences and multiplied with 100.

Average sentence length in morphemes
Per text the number of morphemes is divided by the total number of sentences.

Average clause length in morphemes
Per text the number of morphemes is divided by the total of finite verbs, so the measure represents the average number of morphemes per finite verb, after the suggestion of Verspoor and Sauter (2000), of a clause consisting of a (non-)visible subject and a finite verb.

Total use of cases
Per text all cases, singular and plural, are counted. At this point, all correct and incorrectly used and formed cases are included.

Use of the nominative
Per text the total use of the nominative, singular and plural, is counted. Correct and incorrectly used and formed nominatives are included.

Use of the genitive
Per text the total use of the genitive, singular and plural, is counted. Correct and incorrectly used and formed genitives are included.

Use of the partitive
Per text the total use of the partitive, singular and plural, is counted. Correct and incorrectly used and formed partitives are included.

Use of the other cases
Per text the use of the Finnish cases abessive, ablative, accusative, adessive, allative, comitative, elative, essive, illative, inessive, instrumentative and translative, singular and plural, are counted; this measure does not contain the nominative, the genitive and the partitive. All correct and incorrectly used and formed other cases are included.

Morphological complexity

For morphological complexity we looked at word length in terms of morphemes and tenses.

Average word length in morphemes
The total of morphemes used in a text is divided by the total of words used in the text.

Use of the past and perfect tense[23]
The use of the past and perfect tense represents the total number of correct and incorrectly used and/or formed past and (plu)perfect tenses, in singular ad plural form, per text, in percentages[24]. The total use of the past and (plu)perfect tense is divided by the total number of sentences per text and multiplied with 100. It should be noted that a sentence can contain more than one tense.

Syntactic accuracy

For syntactic accuracy we looked at all errors and specifically at errors in the use of cases. Moreover we looked at the incorrect use of words and word order. In addition to total numbers of errors, we also looked at case accuracy ratios (CAR), which is calculated by dividing the number of incorrect forms or uses by the total number of cases used.

---

23 Even though the past tense can be ranked under the morphological complexity measures and the perfect tense under the syntactic measures, the total use of the past and perfect tense is ranked among morphological complexity measures. This choice was based on the fact that for both tenses a change in morphemes takes place, which has more impact than the change of the sentence by putting it in another tense; when a different tense is used, hardly anything changes in a sentence except for the tense.

24 Next to these tenses the present tense could be used. The conditional mood was used sporadically and therefore not taken into account.

Case use errors
Per text the incorrectly used cases are counted. Case use errors consist of incorrectly used cases in the application of government and congruence, incorrectly used cases in a sentence with no subject, incorrectly used cases for the direct object, the grammatical subject and the predicate nominal and finally, inflectional errors.

CAR use errors
The CAR use errors represent the total number of incorrectly used cases, singular and plural, divided by the total number of cases used in a text[25].

Incorrect application of words
Per text the instances of the incorrect application of words is counted. Application is defined as semantic, pragmatic, or idiomatic appropriateness of the word within its context.

Incorrect word order
Per text the instances of incorrect word order is counted.

Morphological accuracy

For morphological accuracy we looked at the form errors in cases.

Case form errors
Per text the incorrectly formed cases are counted.

CAR form errors
The CAR form errors represent the total number of incorrectly formed cases, singular and plural, divided by the total number of cases in a text. The case form errors consist of incorrect consonant gradation and stem form errors[26].

CAR use and form errors
The CAR use and form errors belong to both syntactic and morphological accuracy. The measure represents the total number of incorrectly used (syntax) and formed (morphology) cases, singular and plural, divided by the total number of cases in a text.

---

25 When a case use error does not belong to one of the categories, it is called an inflectional error; another case should have been used in order to express oneself correctly, e.g.: Minä käyn viikonloppu*ssa* [should have been: *-na*] kirjastossa. *'In the weekend I go to the libraby'*.

26 The examples show incorrectly formed case markings when the stem form is formed incorrectly in a case, e.g: romaan*ija* [should have been: *-eita/-eja*] *'novels'*, asi*aita* [should have been: *-oita*] *'things'*, teem*aita* [should have been: *–oja*] *'topics'* and san*aita* [should have been: *–oja*] *'words'*.

Total of errors
The total of errors is both a syntactic and morphological measure. Per text all use and form errors at the syntactic, morphological and lexical level are counted.

Table 3.C gives an overview of all measures.

Table 3.C: All measures in an overview.

Syntax		*Morphology*	
Complexity	**Accuracy**	**Complexity**	**Accuracy**
-Simple sentence -Compound sentence -Complex sentence -Compound-complex sentence -Compound, complex and comp-complex sentences -Ave. sentence length in morphemes -Ave. clause length in morphemes -Total use of cases -Nominative use -Genitive use -Partitive use -Use of the other cases	-Case use errors -CAR use errors -Incorrect application of words -Incorrect word order	-Ave. word length in morphemes -Use of the past and perfect tense	-Case form errors -CAR form errors -CAR use and form errors -Total of errors

3.3 Traditional statistical measures

To compare overall development in the two groups, we will use traditional statistics. Paired samples tests will be run (which make use of samples of matched pairs of similar data from the compared groups or individuals) and the information will be imported into the data processing program SPSS 20.0. The α decision level is determined at <.05.

For the development over time of the individuals in the groups, an independent-samples test will be run (which make use of samples of data from the compared groups) and the information will be imported into the data processing program SPSS 20.0 as well. Also for this purpose, the α decision level is determined at <.05. To visualize the outcomes we will use box plots because they show variability within a defined time frame and give insight in the median and the range of the outcomes (Van Dijk, 2004).

To compare overall development in our two focal learners (Kim and Andrea), the holistic scores of Kim and Andrea's first two and last two writings are tested for differences. For a statistical test on complexity and accuracy measures the last eight texts of their data sets are used. For these purposes, again a paired samples test will be run.

3.4 Analyses from a dynamic perspective

The goal of a dynamic analysis is to see when variables change, how they change (slowly or abruptly) and how various variables may interact (e.g. support or compete). For these types of analyses only individual data can be used as averaging over groups may smooth away the actual variability and interaction of variables that we are interested in. Because individuals may manifest "disproportional amounts of variability when they move from one stage to the next" (Verspoor et al., 2011: 83), the answers to these questions can help to understand the phenomenon development.

In the current dissertation, mainly visualization techniques are used (Van Geert & Van Dijk, 2002). Line graphs are used to discover the general trend of development. Moving min-max graphs are used to visualize possible peaks or dips in development. Monte Carlo analyses are performed to see if a peak or dip in the development is random or not. Correlations and moving correlations give insight in the interaction of measures over time.

3.4.1 Techniques on development

The development of data can be visualized in a *line graph* to visualize general direction over time. A line graph shows raw data and can be made of one or more measures of one or more learners.

Figure 3.1 shows an example of a line graph, representing the development pattern of the average clause length in morphemes use by Kim. Kim's general trend line shows a general increase of the measure.

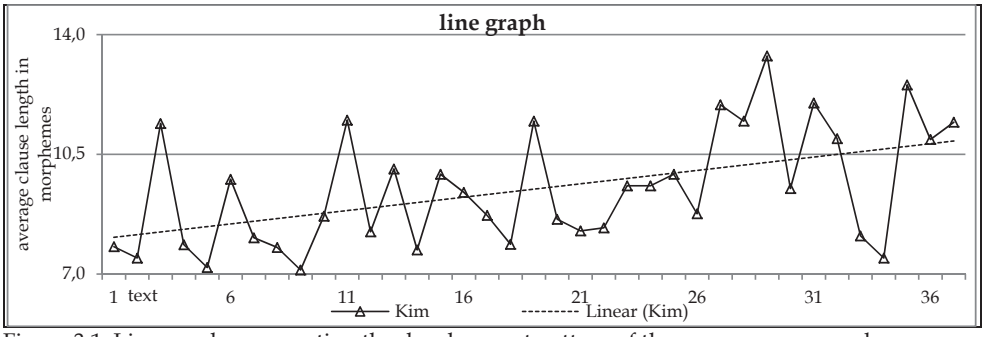

Figure 3.1: Line graph representing the development pattern of the measure average clause length in morphemes by Kim, with linear trend line.

There is always variability in a complex system, but to discover developmental jumps, a min-max graph can be used. In a moving min-max graph the minimum and the maximum scores are represented by two lines and the plain average scores are usually pointed out as dots. (Verspoor et al., 2011, ch.4). The general pattern of variability is highlighted but at the same time the attention is paid to the raw data as well by keeping it visible (Van Geert & Van Dijk, 2002; De Bot et al., 2011). The technique is used to be able to study a developmental trend, while variability around a general trend is outlined by "a score range for each measurement occasion" (Van Geert & Van Dijk, 2002: 353).

Depending on the total number of data points and the degree of generalization one seeks, a window number of data points is decided upon, in our case usually 5. Then the minimum or maximum of the raw data is calculated over the first five data points and the formulae are pulled down so that the min max lines of the subsequent data points windows is shown as follows:

1,2,3,4,5 > 2,3,4,5,6 > 3,4,5,6,7 > 4,5,6,7,8 > and so on.

By using a moving window, fluctuations over time can be traced. The moving window is explained as

> …a timeframe that moves up one position (measurement occasion) each time (the size of the window, e.g., five consecutive data points, 1 month, etc. is called its period). Each window partly overlaps the preceding windows, using all the same measurement occasions minus the first and plus the next. (Van Geert and Van Dijk, 2002: 353)

The min max graph gives insight in the bandwidth of the scores in which the data are presented; a wider bandwidth means a greater amount of variation and vice versa. A moving min-max graph can be of help in visualizing possible transitional phases, e.g. in order to be able to form hypotheses (Verspoor et al., 2011).

Figure 3.2 shows a moving min-max graph of the average clause length in morphemes, again in Kim's data. The trend line illustrates that the average clause length in morphemes grows steadily over time. Moreover, the graph shows that the pattern of widening and narrowing of the bandwidth alternates during the whole period. Overall, the graph shows variability during the whole period, and not yet any sign of stabilization, which is to be expected for a beginner.

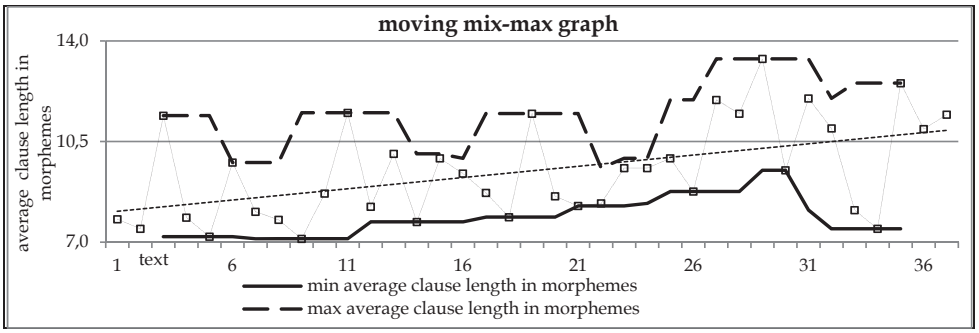

Figure 3.2: Moving min-max graph representing the development pattern of the measure average clause length in morphemes by Kim (window size of 5 data points), with linear trend line.

3.4.2 Technique on simulation

When the score range in a min-max graph shows striking isolated jumps or peaks, a *Monte Carlo analysis* is suggested. A Monte Carlo analysis is a statistical simulation of variability through which we can find out whether isolated peaks are random or not. When the analysis shows that the result is no coincidence, this means that the peak (or dip) is an indicator for growth (Spoelman & Verspoor, 2010; Verspoor et al., 2011). With a Monte Carlo analysis we can also find whether a learner shows more variability in the use of certain measures than another learner. The analyses can be carried out in Excel and used together with the computer program Poptools[27].

In a Monte Carlo analysis the data is used and then reshuffled a number of times (e.g. 5000 times) to see how often the test criterion (e.g. the largest distance found between the original data points) is found again (Verspoor et al. 2011, ch.4). If it is less than 5% chance, then the assumption is that the test criterion was not a random occurrence and therefore developmental. Figure 3.3 shows a developmental peak in Kim's data.

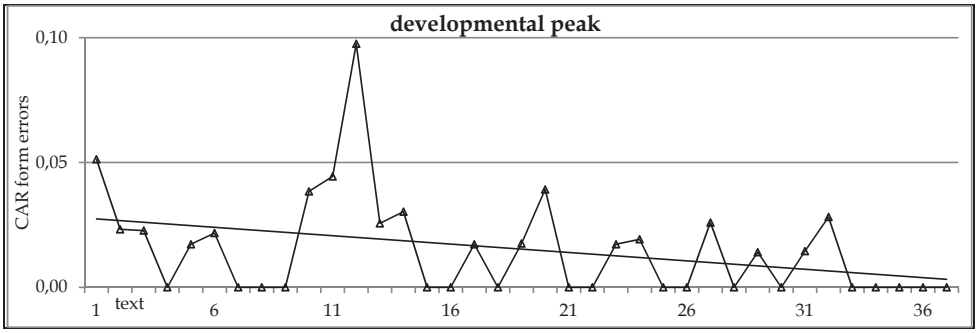

Figure 3.3: The development pattern of the measure CAR form errors by Kim with a developmental peak in measure point 12 (α<.05).

[27] Hood, G. M. (2010) PopTools version 3.2.5. Available on the internet. URL http://www.poptools.org (visited on 10-01-2011).

3.4.3 Techniques on interaction

In order to find interactions between measures over time, a *correlation test* can be done on two or more measures over time. This is done with residual data (Van Geert & Van Dijk, 2002; Verspoor et al., 2011, ch.5), in which the slope is subtracted (Verspoor et al., 2011, ch.4). To calculate the trend for all data values, the sum of the intercept and slope of the entire data series is multiplied by the number of measurements (Verspoor et al., 2011, ch.5). However, the outcome of a correlation test, the *correlation coefficient*, gives only a rather general view on whether the measures are supportive or not. The positive coefficient shows more or less support, the negative coefficient shows more or less competition. It cannot show whether the relation has changed over time. When a correlation coefficient is *very weak* (between .00 and (-).10), the correlated measures do not or hardly influence each other. In case of a *weak* correlation coefficient (between (-).10 and (-).20), the measures have a little influence on each other and when it is *moderately strong* (between (-).20 and (-).40) they have a considerable influence on each other. With a *strong* correlation coefficient (between (-).40 and (-).60) the measures have much influence on each other and finally, in case of a *very strong* coefficient (from (-).60 on), the measures influence each other very strongly; in case of a positive coefficient this will mean that when the one measure goes up, the other one goes up as well or, in case of a very strong negative coefficient, when the one goes up, the other goes down. The calculation of the correlation coefficient for Andrea for the two measures more complex sentence use and average sentence length in morphemes in figure 3.4 shows support: (R=.64). In other words, the measures seem to develop rather synchronously.

Sometimes, a graph will show clear alternations with periods of competition and support, which would point to a trade-off effect for a period of time. In such a case a *moving correlation*, done with residual data, is used to make the changing interaction between the measures more visible (Verspoor et al., 2011). A moving correlation plot "shows temporal changes in the coefficient values in a moving window of several observations" (Caspi, 2010: 30) and is conducted the same way as the moving window of the min-max graph; small sets of values of each grower are used (Verspoor et al., 2011, ch.5). In a moving correlation plot, two measures are represented by one line, revealing the temporal changes in the interaction between two or more measures (Caspi, 2010). The number of data points selected in a moving window depends on the total number of data points. With about 32 data points a window of about five to six data points seems appropriate as it allows enough of a trend without losing sight of more detailed interactions.

Figure 3.4 shows few fluctuations and only one small period of negative correlations for the two measures more complex sentence use and average sentence length in morphemes.

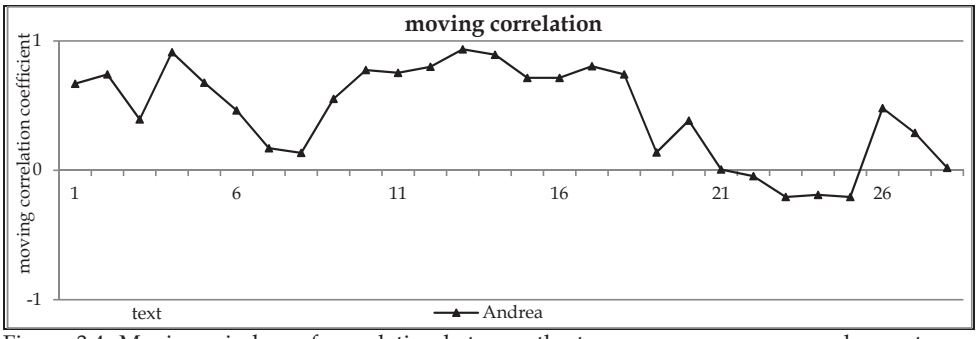

Figure 3.4: Moving window of correlation between the two measures more complex sentence use and average sentence length in morphemes (window size of 5 data points) by Andrea.

3.5 Pitfalls in research from a dynamic perspective

This small section provides three characteristics of research from a dynamic perspective which may, depending on the aim of the study, become at the same time its pitfalls.

Firstly, valuing variability assumes solid sample sizes (Caspi, 2010) and because of the large number of samples, data of only few learners can be investigated; in the current study data of only eight learners will be explored.

Secondly, when we look closely at a limited number of measures, the big picture might be lost; in the current study only 18 measures (for the group study) and seven measures (for the two individuals) are explored.

Finally, in order to be aware of the goal of explaining regularities, Ellis (2007) described the insufficiency of highlighting individual variability. However, the specific aim of the current study is to explore the individual development of the participants and we do not have any intention to look for generalities.

Summary

Kecskes and Papp (2000) suggest that L2 development and FL development are two different entities. They state that the underlying mechanisms, responsible for the two types of development, "have more differences than similarities and result in two different types of language production" (Kecskes & Papp, 2000: 13).

In the forthcoming chapters we will explore this claim by means of two types of analyses. The first is to see if there are overall differences between the two groups of learners (FL and L2). This will be done by averaging over groups using traditional statistics. The second aim is to see if there are differences in the developmental patterns of specific learners.

The FL group consists of four students majoring in Finnish, who all lived in the Netherlands at the time of study and for whom learning Finnish was learning a foreign language. They learned Finnish from an approach that focussed on grammar; the language was the study object and the focus every week was on some specific elements of grammar in order to learn the grammatical rules of the Finnish language. During the first academic year the learners wrote texts every week at home on the computer. The L2 group consists of four students who learned Finnish through a course in Finland. They all lived in Finland at the time of study and learning Finnish was for them learning a second language. The L2 group also took Finnish courses with a strong focus on meaning. However, now and then a preselected form was discussed. The learners volunteered for the current study and wrote a text every week at home on the computer during one academic year. Their tasks were made as similar as possible to those of the FL group.

Even though variables have been controlled for as much as possible in this quasi-experimental study, we do not claim that the groups were completely different in every respect. In addition to their large amount of explicit grammar instruction, the FL group probably had implicit language learning through language exposure by means of the short videos with native speaker actors. The L2 group had daily exposure to Finnish, but we do not know how much they actually interacted in Finnish. Moreover, the students received some grammar instruction, they received feedback on the errors in their writing and they may have looked up explicit grammar rules in grammar books on their own. However, it may be safe to say that the groups are different in the amount of explicit grammar instruction (FL more than L2) and exposure (L2 more than FL).

Then, the chapter gave an overview of the way the syntactic and morphological complexity and accuracy measures were coded, counted and operationalized for the current study. Also the analyses and techniques on simulation and interaction were described. They will be used to examine the overall averages of the FL and the L2 groups, the development of Kim (FL) and Andrea (L2) and the interactions of syntactic and morphological complexity and accuracy measures.

Finally, the chapter gave a review of several pitfalls of a longitudinal case study as this is. In such a study only data of few learners can be investigated. Moreover, the big picture might get lost when a limited number of measures are looked at. Also, we need to be aware of the goal of explaining regularities as Ellis (2007) suggested. We explicitly explained that the specific aim of this study is the exploration of the participants' individual development and that we are not looking for generalities at all.

4 THE FL AND L2 GROUPS COMPARED

In order to fulfill the aim of finding different patterns for the two types of learners with respect to the complexity (operationalized as sentence and word complexity is operationalized as sentence and word complexity) and accuracy measures (operationalized as accuracy in the use and form of the cases), the overall averages of the FL group of learners and the L2 group are traced. In this way we can find measures which show significant differences either in frequency of use of the explored measures or in development over time in the use of more complex language between the FL and the L2 groups. After that the two focal learners are compared to their groups to see to what extent they can be seen as representative of their groups. Finally, a motivation is given for those measures that will be further explored in the longitudinal case studies.

4.1 Results

In chapter 3, participants and procedures in the coding, counting and operationalizations of the different measures are explained. In this section the statistical analyses and results of the overall averages of the measures are given, both in total numbers and in the development over time of the FL and the L2 groups.

The current analyses focus on specific measures, so the broad syntactic measure of the combination of sentence types was not included. Also the tense measure was not explored because of the small numbers of occurrence. In the group analysis also the CAR use and form errors were taken together.

Because the actual number of written words in the texts of the FL and L2 groups showed no significant differences, the variation in the original length of the texts was normalized. From each text a random sample of approximately 100 (from 95 to 105) words containing full sentences were taken. The measures that were taken from texts with fewer than 100 words (which occurred especially at the beginning) were also normalized to 100 words.

For each of the measures the mean number of occurrences is compared. To calculate the mean number of occurrences, only the eleven texts which were

written by all learners of the two groups (texts 10, 11, 12, 13, 14, 20, 21, 23, 25, 26 and 27) were used. For each measure the mean per group (FL and L2) was calculated. A paired samples test was run for each of these measures.

Moreover, for each of these measures the development over time (the correlation between the measure and the number of the text per group) was compared. For each learner the texts were sequentially ordered and a correlation was run between the number of occurrences of a particular measure and the text number. As we did not compare texts to each other in these analyses but correlations, it was possible to use all the texts written by the participants. We compared the correlations between individuals and between groups (2x4). Table 4.A gives an overview of the results.

Table 4.A: The number of uses and the development over time of syntactic and morphological complexity and accuracy for the FL and L2 groups. Bold faced measures indicate significant differences.

	Frequency *(11 texts per participant)*			***Development over time*** *(all texts per participant)*		
	FL	***L2***		***FL***	***L2***	
	R[28](SD)		Sig.	R[29](SD)		Sig.
Simple sentences	.55(.11)	.52(.13)	.35	-.19(.15)	-.36(.17)	.18
Compound sentences	.19(.08)	.17(.05)	.34	-.01(.24)	.12(.13)	.39
Complex sentences	.17(.07)	.25(.18)	.05	.14(.15)	.32(.18)	.18
Comp-complex sentences	.06(.06)	.07(.07)	.68	.20(.18)	.11(.05)	.34
Average sentence length in morphemes	13.6(2.3)	14.5(1.3)	.28	.61(.20)	.49(.12)	.35
Average clause length in morphemes	9.7(1.5)	9.2(1.3)	.14	.52(.20)	.30(.13)	.11
Total use of cases	54.8(3.9)	55.7(5.7)	.57	.26(.19)	-.11(.17)	**.03**
Use of the nominative	22.3(4.4)	26.2(3.5)	**.02**	-.32(.13)	-.18(.13)	.18
Use of the genitive	6.9(3.1)	6.8(2.3)	.94	.53(.09)	.22(.36)	.15
Use of the partitive	13.7(5.4)	10.5(3.3)	.08	-.15(.17)	.04(.21)	.22
Use of the other cases	14.2(4.1)	11.6(3.1)	**.04**	.30(.14)	-.07(.14)	**.01**
Average word length in morphemes	168.1(21.1)	165.2(8.6)	.53	.53(.05)	.06(.18)	**.00**
Case use errors	6.3(1.9)	9.8(3.2)	**.02**	-.12(.20)	.08(.32)	.35
Case form errors	1.2(.75)	1.5(1.2)	.44	-.23(.31)	-.19(.13)	.81
CAR use and form errors	.14(.04)	.20(.05)	**.00**	-.25(.20)	.03(.33)	.21
Total of errors	14.0(3.5)	27.4(4.3)	**.00**	-.20(.23)	.12(.28)	.13
Incorrect application of words	.68(.33)	2.0(.70)	**.00**	-.12(.36)	.14(.08)	.25
Incorrect word order	.34(.27)	1.7(.59)	**.00**	.01(.13)	-.01(.22)	.90

[28] Mean of 11 texts.
[29] Mean of correlation over time.

The FL and L2 group were similar in several developmental measures; overall sentence complexity (types of sentences and average length of sentences) was the same. Both groups showed a similar order of the use of cases and a nominative, genitive and partitive use from the beginning, as had also been found in research on twenty thousand cases, counted in four categories of written Finnish texts (newspapers, magazines, memoirs, reports) (Karlsson, 1983). Besides, the order of the seven most used cases (Karlsson, 1983) was the same for both the FL and L2 groups: nominative, genitive, partitive, inessive, illative, elative, adessive (with only very small differences in frequency of use in the last four cases). However, the groups differed in the number of different types of cases they used and the groups differed in several accuracy measures as well. We will discuss the differences in complexity and accuracy measures separately below.

The FL and L2 groups were different in four complexity measures:

1. The FL group showed more development over time in the total use of cases.
2. The FL group showed a lower use of the nominative.
3. The FL group showed a higher use and more development over time in the use of the other cases.
4. The FL group showed more development over time in the average word length in morphemes.

These findings are complementary. If a nominative is used more frequently, other cases are used less frequently, and the use of cases affects word length in morphemes. The more cases that are used, the more complex not only the noun, but also the words in the sentence may become. First, the FL group developed more in the use of cases, used the nominative less frequently and the other cases more frequently and the L2 group showed less varied choice in cases over time. In other words, the FL group was more successful in acquiring different cases and word complexity. The FL and L2 groups were also found to differ in five accuracy measures:

1. The FL group showed a lower number of case use errors in absolute numbers.
2. The FL group showed a lower number of case use and form errors in terms of the case accuracy ratio.
3. The FL group showed a lower number of total errors.
4. The FL group showed a lower number of word application errors.
5. The FL showed a lower number of word order errors.

Even though the FL group used more cases in absolute numbers, the FL group made fewer errors in their use. In the absolute number of form errors there were no differences. The FL group also made relatively fewer case errors, fewer errors in total, fewer errors in using a word appropriately in the context and fewer errors in word order in the context.

To summarize, the FL group was overall better in the use of cases (complexity) and more accurate in all aspects examined. This answers the first research question: which of the complexity and accuracy measures show significant differences between the overall averages of the FL and the L2 groups? (Please note that the sample is too small to generalize.)

However, we cannot be sure the differences found are due to instruction as the two groups differ in L1 background, which has been shown to have an effect on development in the L2. Murakami (2013) found that if a feature was not found in the L1, there was less accuracy in the L2. Also Spoelman (2013) found that Dutch learners of Finnish made relatively many more errors in cases than Estonian learners, having an L1 with a similar case system as the Finnish language. The FL group consists of three Dutch learners and one Belgium learner, but the L2 group has learners from Brazil, Japan, Indonesia and Germany. As Dutch and German are somewhat similar in the use of cases (although German has more explicit case forms and agreement) the choice was made to compare one learner of each group. In the next few chapters, Kim (FL) and Andrea (L2) will be traced in detail longitudinally. However, to start with, their texts will be judged holistically and data analyzed by means of traditional statistics to see to what extent they are similar or different in a general sense.

4.2 Two learners: holistic and quantitative scores

This section explores two features: the holistic scores on several texts written by the two learners Kim (FL) and Andrea (L2) at the beginning and at the end of the study and the quantitative scores on the texts written in the first half and in the second half of the study.

4.2.1 Holistic scores

Seven experts on Finnish as an L2, all native speakers of Finnish rated two early texts (same topics) and two late texts (same topics) written by the two focal learners. The raters were asked to rate the texts on four criteria: sentence complexity, morphological complexity, accuracy in general and authenticity (idiomaticity). The results of each rater are presented in table 4.B.

Table 4.B: Holistic scores of texts written by Kim and Andrea on sentence complexity (S1) morphological complexity (S2), accuracy in general (S3) and authenticity (S4) by seven raters (1 to 7), scaled from 1 (very low score) to 5 (very high score). The bold faced measures represent the significant difference.

	Kim (FL)										*Andrea (L2)*					
	1	**2**	**3**	**4**	**5**	**6**	**7**	**t**[30]	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**t**
T1S1	3	3	3	2	2	1	3	17	2	3	2	4	1	2	3	17
T1S2	2	3	3	3	1	2	3	17	2	3	3	4	2	2	3	19
T1S3	3	3	4	4	2	2	3	21	2	4	3	4	2	3	3	21
T1S4	3	3	3	3	4	2	3	21	4	3	3	4	1	2	4	21
T2S1[31]	3	4	2	4	4	4	4	**25**	2	2	1	3	1	1	4	**14**
T2S2	3	2	2	3	3	3	4	20	2	3	2	3	1	1	3	15
T2S3	3	1	2	4	3	4	2	19	2	1	1	3	1	2	3	13
T2S4	3	4	2	3	2	4	3	21	1	4	1	3	2	1	3	15
T3S1	4	4	4	5	4	3	5	29	3	3	4	5	3	4	4	26
T3S2	4	4	4	5	4	3	4	28	4	4	4	5	2	3	3	25
T3S3	4	4	3	4	4	2	4	25	4	4	3	5	3	4	4	27
T3S4	4	4	4	4	5	2	3	26	4	3	4	4	3	3	3	24
T4S1	5	4	4	4	5	3	5	30	4	4	4	3	3	3	4	25
T4S2	3	3	3	3	4	2	5	23	3	4	3	4	2	3	4	23
T4S3	3	2	3	3	4	3	2	20	4	4	3	3	3	4	3	24
T4S4	2	4	3	2	4	2	3	20	4	4	4	2	3	3	3	23

Both the raw scores and the statistical analyses show that Kim and Andrea were similar in all developmental measures at the beginning of the study except for sentence complexity in the second text, where Kim scored higher than Andrea. At the end of the study there were no differences between the two learners in the holistic ratings.

4.2.2 Quantitative scores

Although there were no differences in the holistic scores on the texts written by the two writers at the end of the study, the raw numbers in a few complexity and accuracy measures did show some differences towards the end (see table 4.C).

30 Total of all scores in that row.
31 Significantly different: p=.01.

Table 4.C: Raw measures of four texts written by Kim and Andrea.

learners:	Kim	Andrea	Kim	Andrea	Kim	Andrea	Kim	Andrea
texts:	3/3		4/4		34/33		38/34	
average clause length in morphemes	11,4	10,8	7,9	6,6	11,0	10,2	10,9	11,9
average word length in morphemes	1,5	1,8	1,4	1,5	2,0	1,8	1,7	1,7
use complex tenses	0,0	0,0	0,0	1,6	3,0	0,0	4,1	0,0
CAR use errors	0,1	0,0	0,1	0,3	0,0	0,1	0,2	0,1
CAR form errors	0,0	0,0	0,0	0,1	0,0	0,0	0,0	0,1
total use of cases	56,4	50,0	46,2	37,5	53,8	52,9	46,9	49,0
use nominative	26,9	9,7	25,6	23,4	14,4	21,8	16,3	21,0
use other cases	5,1	11,1	3,8	3,1	24,2	14,9	24,5	15,0
case use errors (abs.)	2,6	12,5	5,1	14,1	3,8	5,7	5,1	7,0
CAR use and form errors	0,1	0,4	0,1	0,4	0,1	0,2	0,1	0,1
total of errors	11,5	31,9	11,5	32,8	9,8	26,4	18,4	24,0
incorrect application of words	2,6	1,4	1,3	6,3	1,5	10,3	4,1	6,0
incorrect word order	0,0	1,4	1,3	0,0	0,8	0,0	1,0	1,0

As it was not possible to do a statistical analysis to compare only the two final texts by the writers, the last eight texts of each writer were compared on the use of the other 12 cases (in this case used as an indicator of overall complexity) and the total of errors (in this case used as an indicator of overall accuracy: lexical, syntactic, morphological and word order). A paired samples tests (Table 4.D) shows that there was no difference in the use of other cases, but there was a difference in accuracy. The FL learner Kim made significantly fewer errors.

Table 4.D: The use of other cases and the total of errors in the third (last) part of the texts written by Kim (FL) and Andrea (L2) Bold faced measures represent significant differences.

Measure	***Kim***	***Andrea***	
	R(SD)		Sig.
Use of the other cases	16.5(5.7)	13.7(4.8)	.16
Total of errors	11.9(3.2)	23.9(8.5)	**.001**

Figures 4.1 and 4.2 show box plots of the use of the other cases and the total of errors by the two learners.

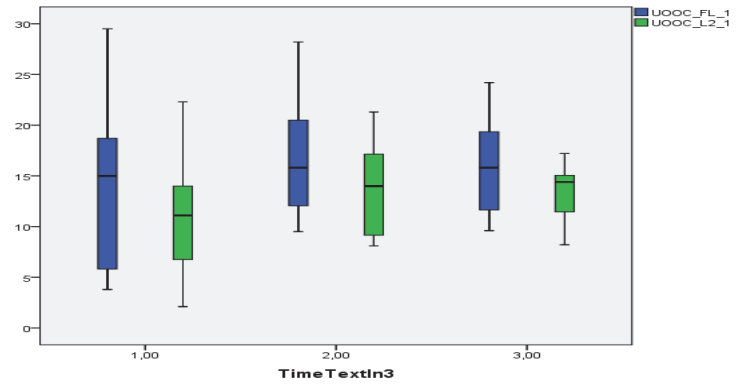

Figure 4.1: The use of other cases by Kim (blue) and Andrea (green) in three periods of time (11-11 and 8 weeks of writing).

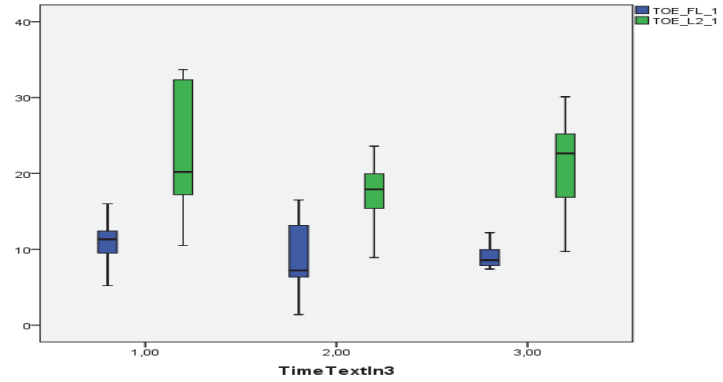

Figure 4.2: The total of errors made by Kim (blue) and Andrea (green) in three periods of time (11-11 and 8 weeks of writing)(p=. 001).

Figure 4.3 shows the actual results of the total of errors by the two learners.

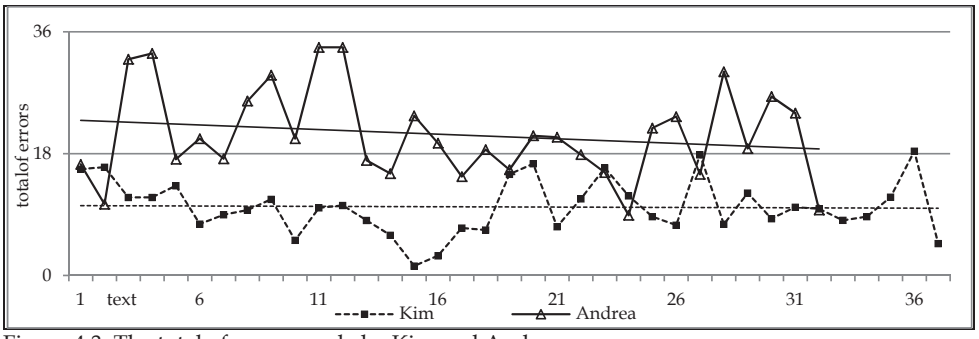

Figure 4.3: The total of errors made by Kim and Andrea.

Because Kim made significantly fewer errors than Andrea, the data were tested for variability as well (Monte Carlo analysis). The Monte Carlo analysis showed a strong effect (p=.001), suggesting that Kim was more variable in making errors than Andrea. Kim's level of errors varied throughout the time span, but especially in the second part of the texts, she made fewer errors and was more variable.

To summarize, the FL and L2 learners were found to be very similar in their writing when evaluated holistically at the end of the study. However, the objective data seemed to suggest some differences. When the last eight texts of the two writers were compared, there was a significant difference in overall accuracy. Kim was more accurate than Andrea. Another interesting finding was that Kim was more variable over time than Andrea in her accuracy.

As the developmental patterns of the FL and L2 learners will be explored longitudinally for all FL and L2 learners, with our two focal learners as specific examples, the next section will discuss our motivation for the measures to be explored.

4.3 Motivation for the measures to explore

One of the aims of this chapter is to determine which measures may be interesting to trace over time. The main differences between the FL and L2 groups all seem to be related to the use of cases. The FL learners use a larger variety of cases and therefore they use the nominative less; moreover they make fewer errors, especially in the use and form of cases. The FL learners also know how to use words appropriately and they make fewer word order errors. These results suggest that a few specific measures are worth exploring in much more detail.

Some general measures to see to what extent they are useful for development in the Finnish language will be explored in chapter 5. First of all, it explores the use of compound, complex and compound-complex sentences, which are often used as quantitative syntactic complexity measures. We intend to find out whether such measures are also suitable for the Finnish agglutinative language in which morphemes play such a large role. For this reason, the chapter will also explore two other syntactic complexity measures: the average sentence length in morphemes and the average clause length in morphemes. Both measure the number of morphemes and therefore both measures are to a certain extent related to the use of cases, which was found to be different between the two groups. One question is if the two focal learners also show such differences. Moreover, we intend to find out whether these two measures can give us more detailed information about sentence complexity in the Finnish language and if so, which of the two can be used as a general broad measure in the current study.

The use of cases (total use of cases, use of the nominative and use of the other cases) will be explored in chapter 6, because significant differences were found in these measures in favor of the FL learners.

The development in the average word length in morphemes is discussed in chapter 7; this measure was also found to be significantly different in favor of the FL group. Also the use of the past and perfect tense are introduced, as this measure gives us the opportunity to see how the two focal learners combine grammatical knowledge of complex tense use with other complexity and accuracy measures. However, before doing so, we will see to what extent the use of the past and perfect tense is related to the writing task.

Several accuracy measures (case use errors, case form errors and CAR use and form errors) are explored in chapter 8, as these were all found to be significantly different in favor of the FL learners. After exploring the totals of these errors, we will also explore them separately to see if the learners develop them in a similar manner.

Finally, the interaction of several complexity and accuracy measures (researched in chapters 5 to 8) are explored in chapter 9, to find out how different subsystems influence each other in the data of the two learners.

Summary

This chapter explored the overall similarities and differences between the FL and the L2 groups. It revealed that the groups were similar in eight complexity measures: simple, compound, complex, compound-complex sentence use, average sentence length and clause length in morphemes, genitive and partitive use. The groups were different in the overall averages of two complexity measures: nominative use and the other cases use. Moreover, when the complexity measures were correlated with time, the groups showed differences in three measures: total use of cases, use of the other cases and average word length in morphemes. As far as accuracy is concerned, there were differences between the two groups in all accuracy measures except case form errors: case use errors, CAR use and form errors, total of errors, incorrect application of words and incorrect word order. However, many differences between the groups could be due to L1 effects rather than context.

Because the main aim of the current study is to explore differences in development of Finnish as an FL or L2 and there has been evidence that the L1 may play a strong role in development, we selected two focal learners whose L1 are as similar as possible to control as much as possible for any potential L1 effect. Holistic scores by seven experienced raters showed no differences between the two learners at the end of the study in sentence complexity, morphological complexity, accuracy in general and authenticity (idiomaticity). However, a statistical analysis of the last eight texts written by the focal learners showed that the L2 learner made more errors than the FL learner. Moreover, a Monte Carlo analysis showed that the FL learner was more variable in her accuracy rates than the L2 learner. For these two learners, type of instruction seems to have had an effect on accuracy. The chapter finished with the motivation for the measures that will be explored in more detail in the coming chapters.

5 SYNTACTIC COMPLEXITY: SENTENCE LENGTH

In this chapter we want to explore the development (including peaks and dips) and interactions of several syntactic complexity measures at the sentence level. Concerning language development of the two learners Kim and Andrea, from here on the assumption is that there is variation among the two learners (they will not develop in exactly the same manner) and that there is variability in each of them as well (especially at early stages they will not show stable patterns). Related to the variability, we assume that there may be developmental peaks (Verspoor et al., 2011), which are a sign of overuse or underuse of certain constructions by the two learners. The assumption is that a greater degree of variability may indicate a greater degree of development for that particular learner.

Moreover, we want to find the most informative measure to represent general sentence complexity in Finnish, with its numerous morphemes. Next to the total of compound, complex and compound-complex sentences, the average number of morphemes per sentence and per clause will be explored. The measure that is the most informative (as a holistic, smooth developmental measure) will then be used in subsequent chapters to compare other developmental trajectories to.

5.1 Development of sentence types

In the group study we found no significant differences in the number of the sentence types (simple, compound, complex and compound-complex), but all different sentence types that occur in the writing samples are examined here, as the sub goal is to find the most informative sentence measure. As the group data showed growth in all sentence types except simple sentences, a similar growth is expected in the two individuals. In all examples in this chapter, the parts in bold in the Finnish sentences are the issues in question. If needed, the issue in bold is clarified in the subsection. All abbreviations in the example

sentences are used to explain the errors and to give the relevant information on the sentence in question and their meaning can be found in the list of abbreviations (see Appendix A).

After illustrating the different sentence types their development will be traced in the two individual learners. Example 1 and 2 show respectively a simple and a compound sentence. Examples 3 and 4 show respectively a complex and a compound-complex sentence. The fact that the word order in Finnish is generally the same in compound and complex sentences (see chapter 2) might be a cause for the rather frequent use of compound, complex and compound-complex sentences from the beginning.

(1) Minulla on kaunis puutarha. *'I have a nice garden.'* (Kim-t2s7/mp2-simple)
(2) Runossa voi sanoa paljon **ja** ei tarvitse käyttää[lex] paljon sanaita[form]. *'In a poem one can say a lot and one does not need to use many words.'* (Andrea-t13s11/mp12-comp)
(3) Minä tykkään pyöräillä aamulla ja illalla kurssin jälkeen, **koska** mä rentoutan[lex] sen läpi. *'I like to cycle in the morning and in the evening after class, because I relax by doing that.'* (Andrea-t5s2/mp5-compl)
(4) **Kun** viikko oli ohi, tulehdus ei ollut ohi, **ja** siksi minä sai uudestaan lääkkeitä. *' When the week was over, the infection was not over and therefore I got medicines again.* (Kim-t28s8/mp26-comp.compl)

Kim

Figures 5.1 and 5.2 show the development of the four sentence types in percentages in Kim's writing; her writing contains more sentence complexity towards the end. The trend lines indicate a decrease of the simple and compound sentences over time (simple sentences a fraction more than the compound sentences). Furthermore, the trend lines show an increase of complex and compound-complex sentences over time (complex sentences more than the compound-complex sentences).

Kim uses all sentence types from the beginning, except for the compound-complex sentence (used from the third measure point). The use of simple and compound sentences throughout continues and shows some peaks. Moreover, several times the two sentence types are not applied at all. At measure point 14, the high percentage of complex sentence use is striking; this text consists of eight sentences, six of which are complex. From measure point 21 on complex sentences are applied in every text. At measure point 10, the use of compound-complex sentence is remarkable. The text consists of five sentences, four of which are compound-complex. The compound-complex sentence type is not applied in all texts; in the first half of the data this happens seven times and in the second half two times.

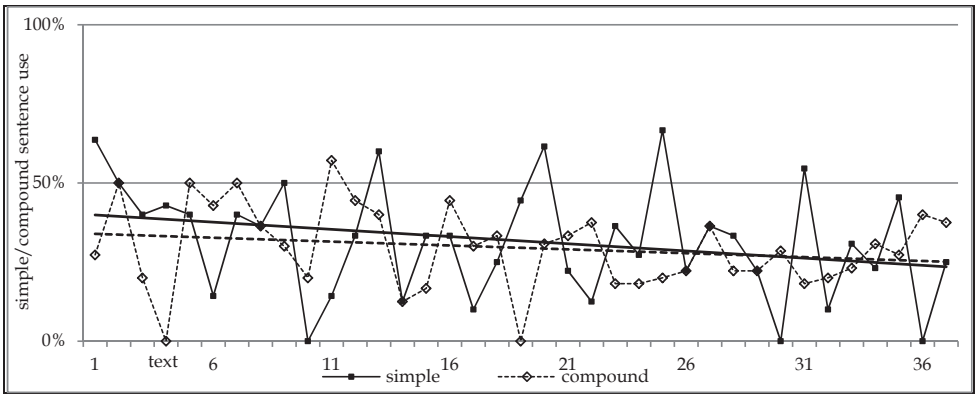

Figure 5.1: Kim's simple and compound sentence use.

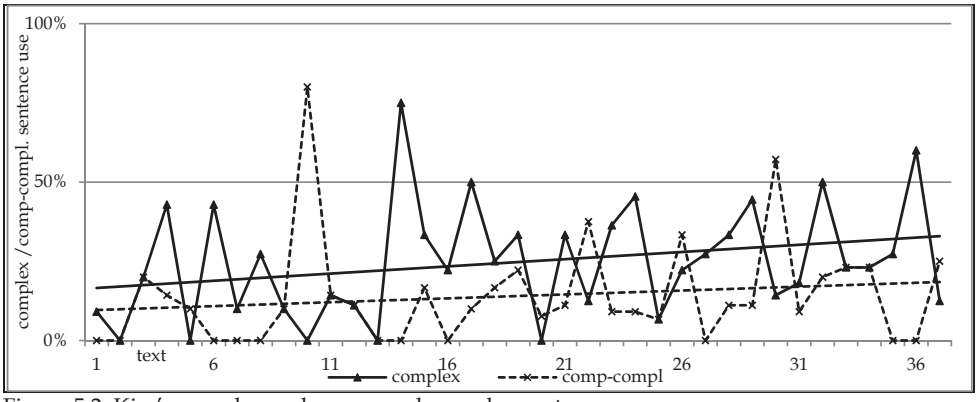

Figure 5.2: Kim's complex and compound-complex sentence use.

Because both the complex and compound-complex sentences show isolated peaks, they were tested for significance using a Monte Carlo analysis. Neither was found to be significant and therefore they are no indicators for growth.

Andrea

Figures 5.3 and 5.4 show the development of the four sentence types in Andrea's writing; Andrea shows more sentence complexity towards the end. The trend line for her simple sentences decreases over time, while the compound sentences stay at the same level throughout. The trend lines of the complex and compound-complex sentences show increase over time (the latter increases less).

Andrea uses simple and the compound sentences from the first measure point, of which the simple sentence is used most frequently. The complex sentence is used from the third and the compound-complex sentence from the sixth measure point. In the last measure point all sentence types are used. The simple sentences show some peaks and are used in all texts. The compound sentences show one large peak (mp17); in this text the compound sentences are applied five times out of eight. The compound sentence is not used in every

text. The complex sentences show some small peaks with two striking peaks in the last six sentences; from measure point 27 (no use of the complex sentence) on, the line goes up to a more than 50 percent use at measure point 31 and 32. In the first half of the data four texts do not contain a complex sentence and in the second half this holds for only one text. The compound-complex sentences show some small peaks. Besides, the compound-complex sentences are not applied in nine measure points in the first half of the data and in eight measure points in the second half; more than half of the written texts by Andrea do not contain any compound-complex sentences.

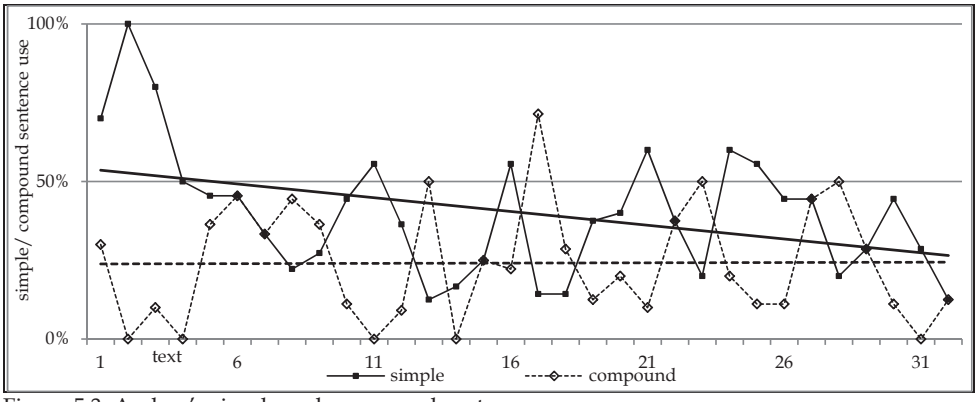

Figure 5.3: Andrea's simple and compound sentence use.

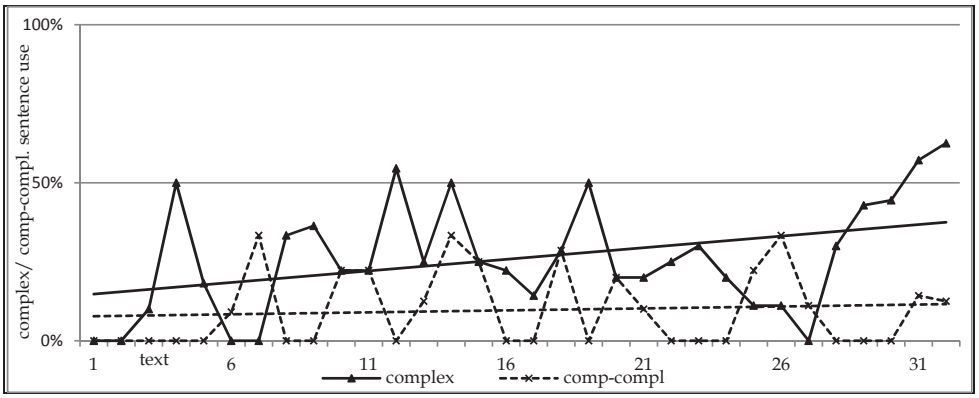

Figure 5.4: Andrea's complex and compound-complex sentence use.

Because the simple sentences show an isolated peak, the measure was tested for significance (Monte Carlo analysis). The analysis did not show significance; this peak is not an indicator for growth.

Kim and Andrea compared

On the whole, Kim and Andrea act quite similarly in sentence use, except for their use in compound sentences. The lines representing the simple sentences illustrate a rather identical pattern over time, even though Andrea's trend line

decreases faster; it takes off much more than Kim's. This might be due to Andrea's high simple sentence use in the first three measure points, specifically the peak in the second one (a 100 percent use of the simple sentence). For both learners the trend lines of the simple sentences end at about the same level. Kim and Andrea's use of compound sentences shows differences; Andrea stays at about the same level throughout, whereas Kim shows a decrease. Kim starts at a higher level than Andrea, who throughout remains around the lowest level of Kim's compound sentence use. The trend lines of the complex sentences are rather similar for both learners and this is also the case for the compound-complex sentences, though for this measure Kim shows more growth than Andrea.

An overall view of both learners' use of all four sentence types shows a shift in sentence use at about data point 26 in which the numbers of complex and compound-complex sentences in the texts exceed the numbers of simple and compound sentences. The variability we find in sentence types is what we would expect from a dynamic perspective. The learners are trying different types out along the way, slightly overusing some at different times. For neither learner there seems to be a clear attractor state at the end with a rather stable mix of sentence types.

5.2 Development of complex sentence types versus average sentence length in morphemes

The previous section showed that the use of all sentence types, except for simple sentences, increase. Therefore to show more holistic development, the number of compound, complex and compound-complex sentences are summed up to see to what extent they relate to sentence length measures expressed in number of morphemes rather than words. In the group study the use of simple, compound, complex and compound-complex sentences as well as the average sentence length in morphemes showed no significant differences. The combination of the more complex sentences was not tested for significance.

Kim

Figure 5.5 shows Kim's more complex sentence use, operationalized as sentences with more than one finite clause per sentence (compound, complex or compound-complex). It illustrates that she uses at least one of each of these in every text. The trend line shows a slight increase.

The more complex sentence use shows quite a variable pattern (with the exception of the first five measure points which together form a rather smooth line) and quite some peaks. The most striking peaks are measure points 10, 30 and 36, in which no simple sentences are used. Also striking are the measure points which show very low use of the three sentence types (mp1, 13, 20 and

25). Looking at the first half of the data, it shows that the more complex sentence use is five times 50 percent or less; in the second half of the data this is three times. In the first part the use is six times 75 percent or more; in the second part this is nine times. Finally, the graph shows a relatively low variable pattern of succession of narrowness and wideness in the bandwidth within the developmental trend of growth; nevertheless and in line with dynamic thinking, the variability stays high during the whole year.

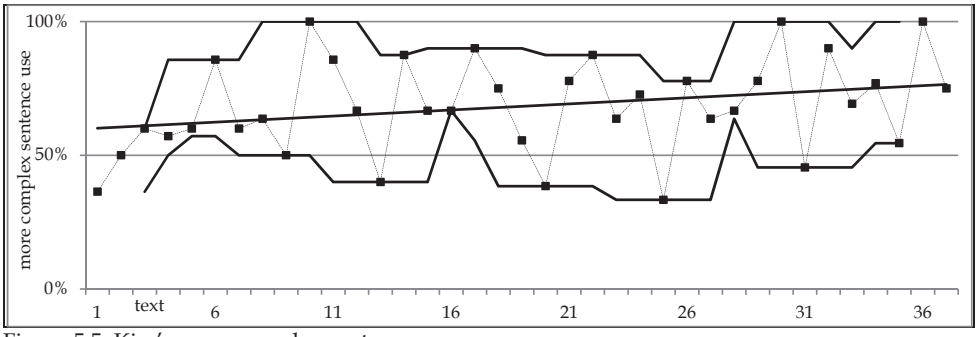

Figure 5.5: Kim's more complex sentence use.

Figure 5.6 illustrates the average sentence length in morphemes in Kim's writing and shows that the measure increases. Looking at the data, the average sentence length in morphemes is 10.6 in the first measure point (the lowest value of all measure points) and 22.9 in the last one.

The data shows some striking peaks and a rather long line of measure points which together form a smooth line upwards (mp26-31). Furthermore, four series of measure points decline (mp6-9, 10-13, 15-18 and 21-24). Finally, the first part of the graph indicates that the bandwidth widens and narrows sequentially, while in the middle of the graph the variability is quite low. At the end the bandwidth broadens to a rather large extend which points at a stage of high variability.

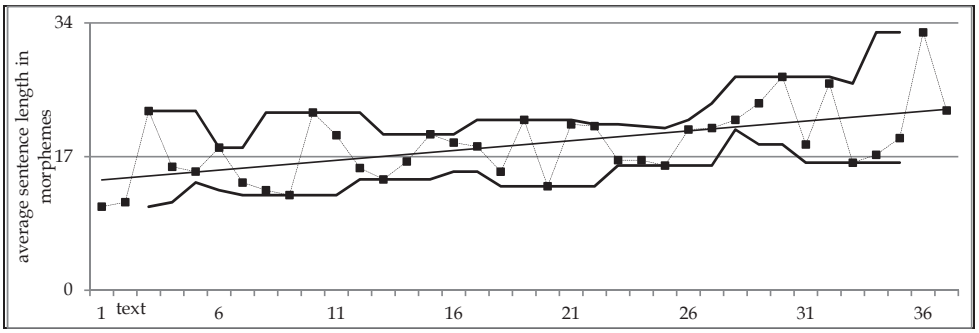

Figure 5.6: Kim's average sentence length in morphemes.

Andrea

Figure 5.7 shows Andrea's more complex sentence use and the trend line illustrates a weak increase. Andrea uses all sentence types in every text except for the second one. Moreover, no measure point indicates a 100 percent use of the compound, complex and compound-complex sentences. This implies that Andrea uses a simple sentence in every text.

The use of the three sentence types shows a pattern of peaks and dips throughout, with a couple of measure points that show a smoother line. The most striking peak is the second measure point. Lastly, the graph shows hardly any variability in the bandwidth within the developmental trend of growth with one exception in the beginning. The overall bandwidth is rather wide during the time span.

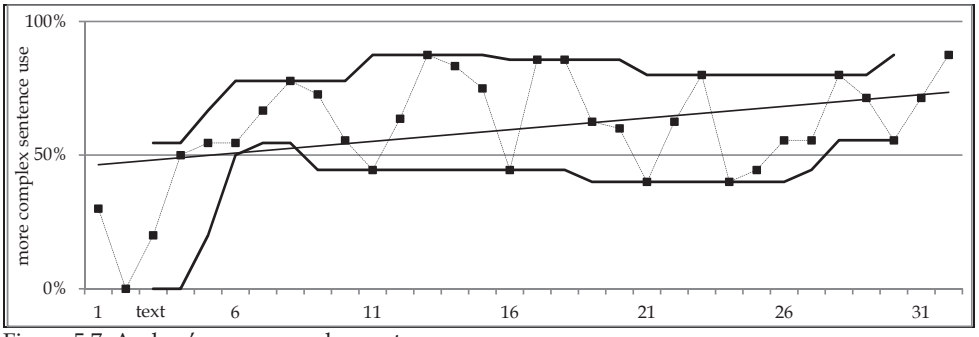

Figure 5.7: Andrea's more complex sentence use.

Figure 5.8 shows the average sentence length in morphemes in the writing of Andrea and shows weak increase. The data show an average sentence length in morphemes of 10.6 in the first measure point and 8.9 in the second measure point (the lowest value). Measure point 31 shows the highest value (23.7).

The pattern of the average sentence length in morphemes shows some striking peaks and two series of measure points show similarities in their pattern; they both form a rather smooth line upwards (mp4-8 and 26-29). To finish, the graph shows that in the first part the bandwidth is rather narrow, after which it gets wider in the second part. In the third part of the graph the bandwidth narrows again before in the last part once again a wider bandwidth seems to arise. Overall, the graph shows relatively low variability.

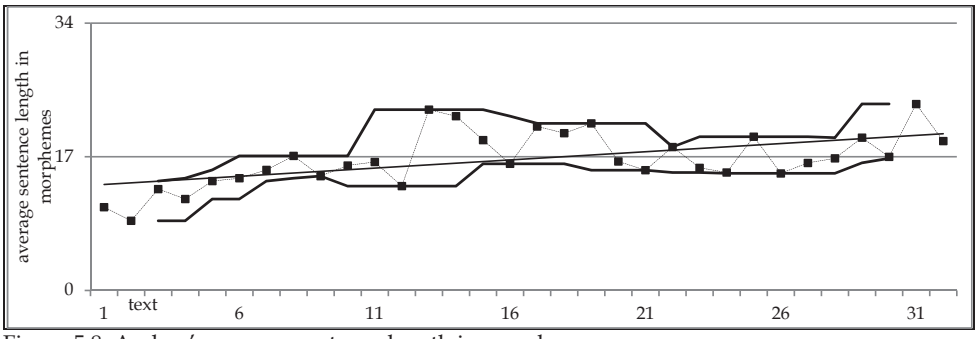

Figure 5.8: Andrea's average sentence length in morphemes.

5.2.1 Interaction between complex sentence types and average sentence length in morphemes

In this subsection, the combination of the two measures, the total of compound, complex and compound-complex sentences and the average sentence length in morphemes, is explored. First Kim and Andrea are discussed separately and then their development is compared. The subsection starts with example sentences.

The examples show the use of the more complex sentences together with the average sentence length in morphemes. Examples 5 and 6 show complex sentences (of two and four finite clauses respectively) and despite of the fact that both sentences belong to the same complexity measure on sentence level, they differ quite a lot in the number of morphemes (11 versus 36). The average sentence length in morphemes (word level) seems therefore to be a more sensitive instrument to use as syntactic complexity measure than the more complex sentences (sentence level). Example 7 seems to underscore this; the sentence consists of a compound-complex sentence with 39 morphemes. All three sentences are examples of more complex sentences, but at the same time they show a lot of variability in the average sentence length of morphemes.

(5) ”Sampo” o/n esine, **joka** teke/e omistaja/n rikaa/ksi[cons.gr]. *'The "Sampo" is a thing, which makes the owner rich.'* (Andrea-t31s8/mp28-compl-SlMrph1/11)

(6) **Kun** minä kuvittele/n, **että** minä saa/n valita itse/lle/ni kesä/paika/n Suome/ssa, minä[infl] e/n tarvitse ajatella kauan, **koska** minu/lla o/n yksi suuri toive, olla kesä/nä[infl] Helsingi/ssä. . *'When I imagine that I get to choose a place to be in the summer for myself, I do not need to think long, because I have one big dream, to be in Helsinki in the summer.'* (Kim-t38s1/mp36-compl-SlMrph1/36)

(7) Ostaminen lahjo/j/a[infl][w.ord] minu/n iso/isä/lle[poss.suff] tai minu/n iso/äidi/lle[poss.suff] o/n vaikea/mmin[part sp], **koska** he/i/llä o/n jo paljon[w.miss], **ja** nyt minä ole/n liian vanha malaa/maa/n[form] taulu/t[d.obj]. *'Buying present for my grandfather*

and grandmother is difficult because they already have a lot, and now I am too old to make a painting.' (Andrea-t12s8/mp11-comp.compl-SlMrph1/39)

Kim

Figure 5.9 shows the more complex sentences and the average sentence length in morphemes in Kim's writing and shows that they both grow. Moreover, they grow quite equally strong. The trend line of the average sentence length in morphemes shows a larger increase than the one of the more complex sentence types and even though the pattern of the average sentence length is not smooth, it shows less variability than the pattern of the more complex sentences.

Overall, the measures show mainly support. Until measure point 15, the supportive correlation between both measures is evident. Measure points 15 to 19 show competition, after which the total of compound, complex and compound-complex sentences and the average sentence length in morphemes seem to be mostly supportive again. Also Kim's correlation coefficients show very strong positive correlations between the measures in both the raw and the residual data (R=.67; R=.65). Appendix E gives a total overview of all correlation coefficients.

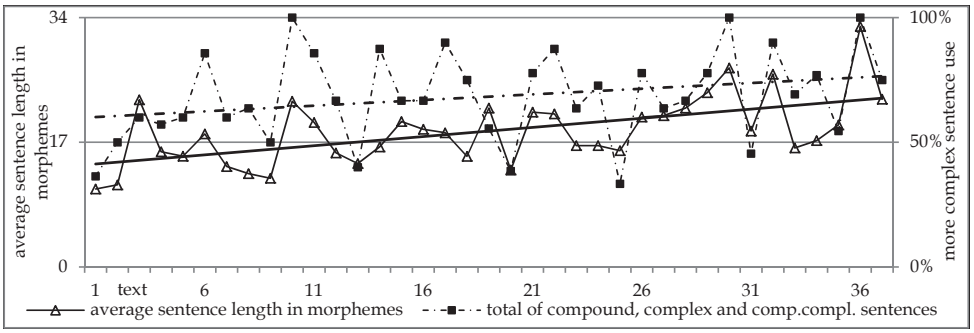

Figure 5.9: Kim's more complex sentence use and average sentence length in morphemes.

Andrea

Figure 5.10 shows the more complex sentences and the average sentence length in morphemes in Andrea's writing and shows growth for both measures, but the trend lines shows slightly smaller increase for the average sentence length in morphemes than for the more complex sentences. Besides, the pattern of the average sentence length in morphemes shows a comparable number of peaks and dips as the pattern of the three sentence types.

Throughout, the supportive relation between the two measures is evident. Several times the more complex sentence use shows a peak where at the same time the average sentence length in morphemes stays rather smooth (mp1-3, 6-10, 17-18, 23 and 28-30). This is not the case the other way around. Also Andrea's correlation coefficients show very strong positive correlations between the measures in both the raw and residual data (R=.71; R=.64).

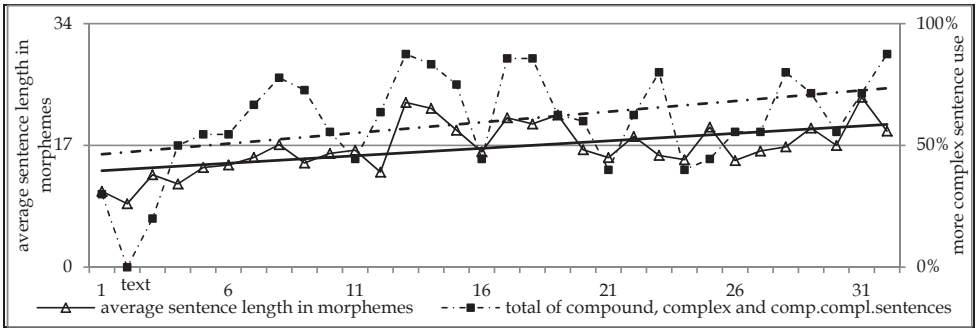

Figure 5.10: Andrea's more complex sentence use and average sentence length in morphemes.

Kim and Andrea compared

Figure 5.11 shows a moving window of correlations of the more complex sentence use and the average sentence length in morphemes. The striking dip in Kim's data (points 14 and 15) is due to the fact that Kim wrote long sentences with words consisting of few morphemes in the related texts; the writing style is quite simple.

We had hypothesized that the more complex sentence use and the average sentence length in morphemes are supportive growers, because we expected that more sentence complexity would imply more morphemes in the sentence. Indeed both learners' correlation coefficients show mainly supportive interactions. This means that a sentence consisting of more morphemes is more likely to be a complex sentence. Both Kim and Andrea wrote more simple sentences with fewer morphemes at the beginning and more complex sentences with more morphemes at the end. The few competitive points in the data suggest that these growers are not supportive all the time.

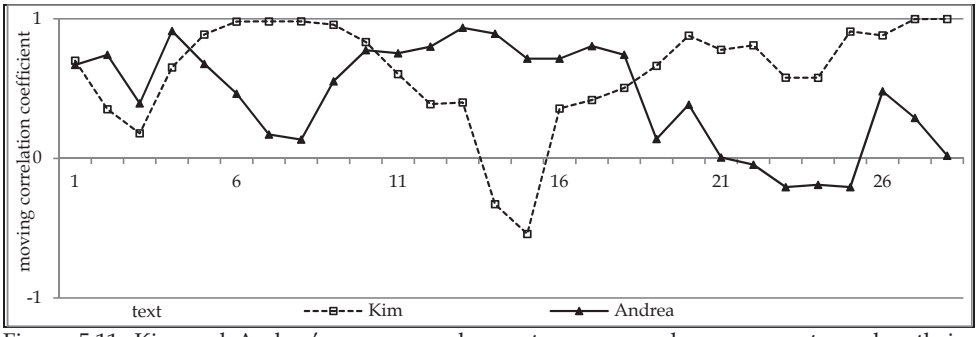

Figure 5.11: Kim and Andrea's more complex sentence use and average sentence length in morphemes and their interaction.

Now that the first two complexity measures have been explored, we will introduce the third one and explore more development and interactions in the next section.

5.3 Development of complex sentence types versus average clause length in morphemes

In the previous section we explored morphemes per sentence. In this section we explore morphemes per clause. One reason to do so is to find out which of these two measures is more informative as a syntactic complexity measure. We counted the finite verbs to get the number of clauses, not distinguishing between coordinate or subordinate clauses. In the group study there was no significant difference in the average clause length in morphemes.

Kim

Figure 5.12 shows the average clause length in morphemes in Kim's writing and illustrates increase. The mean value of the measure is 9.5. The graph shows some striking peaks and two times it shows a rather smooth decline (mp6-9 and 15-18) and also twice it shows a state close to stability (mp20-22 and 23-25). Finally, the first part of the graph shows a rather consistent variable chain of succession of narrowness and wideness in the bandwidth, after which the trend of widening without narrowing too much on the right side seems to prevail. Overall, the graph shows variability in the measure itself and in the bandwidth.

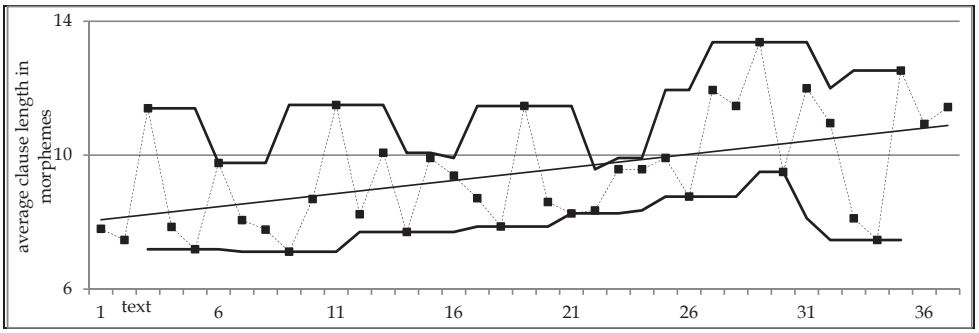

Figure 5.12: Kim's average clause length in morphemes.

Andrea

Figure 5.13 illustrates the average clause length in morphemes in Andrea's data and points a steady increase. The mean value of the measure is 9.1.

The pattern of the average clause length in morphemes shows some striking peaks and two small series of measure points have about equal values (mp9-10 and 17-18). A couple of consecutive measure points show a rather small difference in values (mp1-2, 5-6, 21-22, 24-25 and 29-30). The first and last data points show most variability but overall, the graph shows relatively low variability in the bandwidth within the developmental trend of growth.

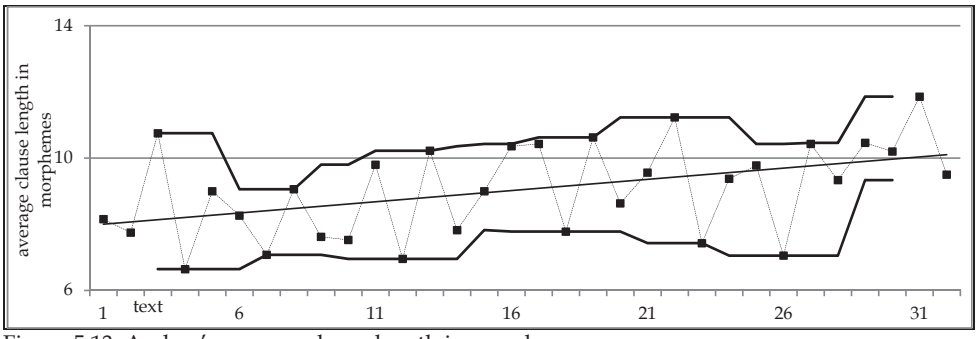

Figure 5.13: Andrea's average clause length in morphemes.

5.3.1 Interaction between complex sentence types and average clause length in morphemes

In this subsection, the more complex sentences and the average clause length in morphemes is explored. Kim and Andrea are viewed separately and then their development is compared. The subsection takes off with example sentences.

The examples illustrate the more complex sentences use together with the average clause length in morphemes. Even though example 8 shows a simple sentence, it is a clause in its totality and is at the same time the average clause length in morphemes. Example 9 shows a compound sentence with two clauses which both consist of only few morphemes. Example 10 shows a sentence with two clauses. Even though examples 8, 9 and 10 all are more complex sentences, the average clause lengths in morphemes differ quite a bit. Therefore the average clause length in morphemes seems to be a more sensitive complexity measure than the more complex sentence use. This appears to be confirmed by example 11, a compound-complex sentence consisting of three clauses which contain seven, eight and nine morphemes. To summarize, all these sentences are more complex sentences, but they show clear differences in the average clause length of morphemes.

(8) Minä osta/n aina soija/juoma/a, joskus kaura/juoma/a ja yleensä soija/jogurtti/a. *'I always buy soya drinks, sometimes oat drink and mostly soya yogurt.'* (Andrea-t3s8/mp3-simple-ClMrph1/16)

(9) Minu/n asunto/ni o/n uusi ja aika iso // **mutta** puu/tarha o/n aika pieni. *'My house is new and pretty big but the garden is pretty small.'* (Kim-t5s2/mp5-comp-ClMrph2/17(10-7:8.5))

(10) Me/i/dän kohde[poss.suff] ol/isi kiva kahvila // **mi/ssä** me voi/mme[tns] lämmitellä ja nauttia kahvi/sta ja munke/i/sta tai pulla/i/sta[form]. *'Our goal would be a nice café where we could warm ourselves and enjoy coffee and doughnuts or coffee bread.'* (Andrea-t33s4/mp30-compl-ClMrph2/26(8-18:13.0))

(11) Me oli/mme käy/neet[tns] Lapi/ssa // **koska** me e/mme rakasta ottaa aurinko/a //**siis** me e/mme[w.miss] kiinostu/neet etelä/Eurooppa/sta[cons.gr]. . *'We had visited Lapland because we do no*

like to take a sunbath, so we are not interested in southern Europe.' (Kim-t22s4/mp20-comp.compl-ClMrph3/19(7-8-9:8.0))

Kim

Figure 5.14 presents the more complex sentences and the average clause length in morphemes in Kim's writing. It shows that both measures grow and that their overall growth develops in quite a similar way. The pattern of the average clause length in morphemes shows a little less variability than the one of the more complex sentence use, but both patterns show a rather large variability.

Overall, the graph shows competition and support between the two measures. Twice, one measure stays at the same level while the other one increases (mp23-24) or declines (mp15-16). Kim's correlation coefficient for the raw data shows a very weak positive correlation between the measures (R=.05) while the residual data show a weak negative correlation (R=-.10).

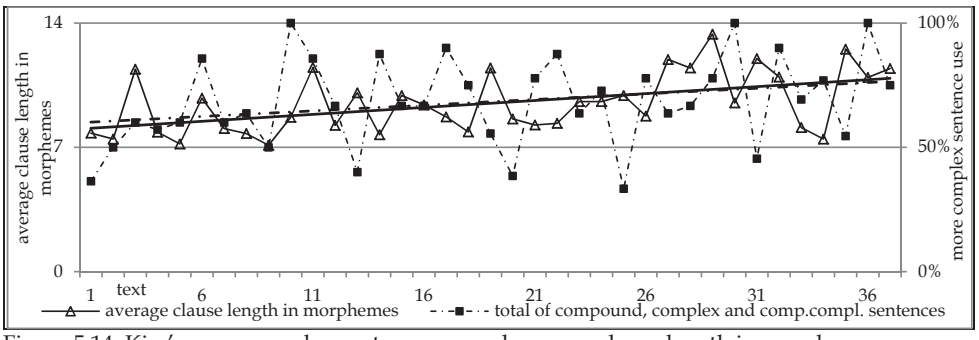

Figure 5.14: Kim's more complex sentence use and average clause length in morphemes.

Andrea

Figure 5.15 shows the more complex sentences and the average clause length in morphemes in Andrea's writing and shows that both measures show growth, even though the increase of the latter is weaker. Both measures show a rather variable pattern but overall, the pattern of the more complex sentence use shows more variability (e.g. mp16-17) as the more complex sentence use shows more rather smooth passages than the average clause length in morphemes (e.g. mp4-6).

Overall, the graph shows large variability in competitive and supportive correlation between both measures. Three times, one of the measures remains at the same level while the other one shows growth (mp 27-28) or declines (mp5-6 and 17-18). Finally, Andrea's correlation coefficient for the raw data show a very weak positive correlation (R=.02) between the measures, whereas the residual data show a weak negative correlation (R=-.19).

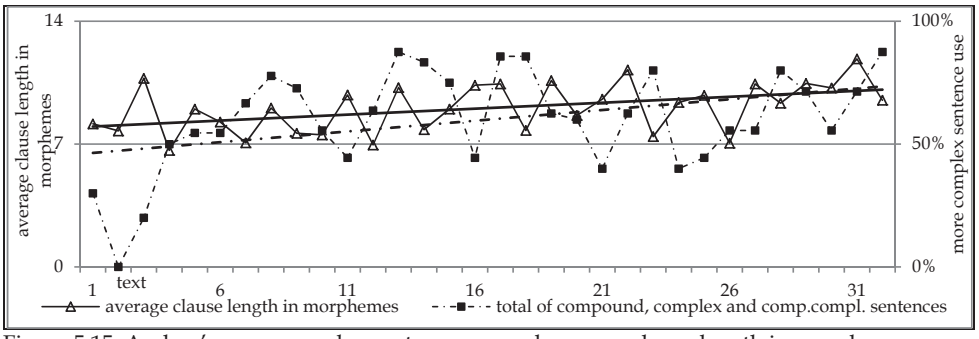

Figure 5.15: Andrea's more complex sentence use and average clause length in morphemes.

Kim and Andrea compared

Figure 5.16 shows a moving window of correlations of the more complex sentences and the average clause length in morphemes. For both learners the relation of the measures changes over time. Kim starts with support, after which the pattern mainly shows competition. Andrea starts with a variable pattern of correlation after which her data mainly shows competition as well.

We had hypothesized that the more complex sentence use and the average clause length in morphemes are supportive growers because our expectation was that the more complex the sentence type becomes, the more morphemes the clause will have. For both learners this is not true (residual data). They both used fewer complex sentences and fewer morphemes in the clauses at the beginning and more complex sentences with less complex clauses or more complex clauses with less complex words in terms of morphemes at the end.

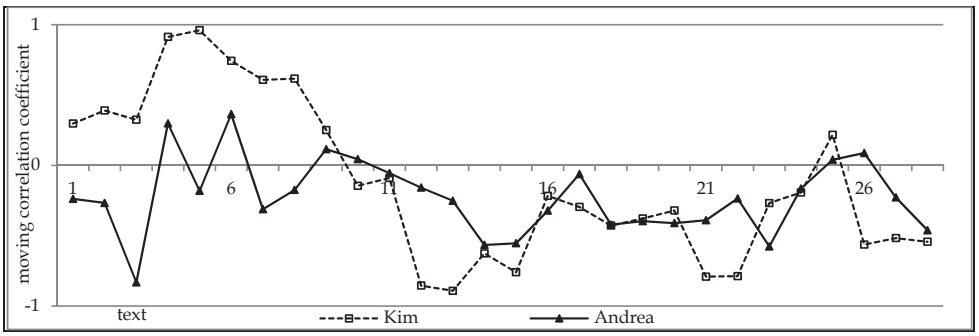

Figure 5.16: Kim and Andrea's more complex sentence use and average clause length in morphemes and their interaction.

Now that the three syntactic complexity measures are explored, we will use the information so far to be able to find the most informative one.

5.4 The most informative measure

In this section we will decide which of the two remaining syntactic complexity measures (average sentence length in morphemes or average clause length in morphemes) we will employ in the next chapters.

In the previous sections we examined two combinations of syntactic complexity measures. We reported that the average sentence length in morphemes (word level) as a general complexity measure should prevail over the more complex sentence types (sentence level). We found strong positive correlations for the measures, as expected. This means that overall the measures support each other. However, we also found that in sentences which all belong to the same complexity measure on sentence level the number of morphemes can differ quite a lot. Therefore it seems that the average sentence length in morphemes is a more sensitive instrument to use. Furthermore, we reported that the average clause length in morphemes (word level) as general complexity measure should prevail over the more complex sentence types. We found that the measures showed moderate negative correlations, which implies that overall the measures compete with each other. This was confirmed by our findings; they showed that in sentences which all belong to the more complex sentence types, the clause lengths in morphemes differ quite a lot. Therefore it seems that also the average clause length in morphemes is a more sensitive instrument to use as syntactic complexity measure. In summary, it means that we can a have a finer-grained view in someone's writing from length in morphemes than from the more complex sentence types.

In the remainder of this section we will give example sentences of the two remaining length measures in morphemes, after which our choice will be motivated.

From the average sentence lengths in morphemes, we cannot know whether we look at a simple sentence with a dense morpheme pattern (like in examples 12 and 13) or at a more complex sentence with a lot of one and two morpheme words or a lot of conjunctions (which can also make a sentence become very long); therefore, we argue that the average sentence lengths in morphemes does not reflect the internal complexity of structures within sentences. On the other hand, the average clause length in morphemes shows the number of morphemes related to one finite verb, irrespective of the length of the total sentence. The examples show sentences in which the numbers of morphemes are not striking when we look at the average sentence length but do when we look at the average clause length in morphemes.

(12) Minä tule/n kutsu/ma/an he/i/dän vappu/pikniki/lle/ni. *'I will invite them to my party on the first of May.'* (Kim-t34s8/mp32-simple-SlMrph1/13- ClMrph1/13)

(13) Minä lainaa/n joskus levy/j/ä kaupunki/kirjasto/sta klassi/sta[infl] musiikki/a[gov][w.miss] kera. *'Sometimes I borrow albums from the library*

with classical music.' (Andrea-t25s4/mp24-simple-SlMrph1/15-ClMrph1/15)

The last step in establishing the most informative expression of sentence complexity is to see how they compare in our focal learners.

Figure 5.17 shows mainly similarities among the measures in Kim's data and an increase (though larger for the average sentence length in morphemes). Figure 5.18 shows fewer similarities in Andrea's data, but still increases for both measures. The average sentence length in morphemes shows a smaller increase than was seen in Kim's data, though. On the other hand, the trend lines of the average clause length in morphemes seem to grow rather similarly for both learners, though the line of Andrea increases slightly less.

Even though both measures show an increase for both learners, the figures show competition between the measures as well; a sentence with a lot of morphemes does not automatically consist of clauses with a lot of morphemes. When there is competition, this could firstly mean that the sentence with a lot of morphemes does not necessarily consist of clauses with a lot of morphemes and secondly that clauses with a lot of morphemes can be found in a relative simple sentences.

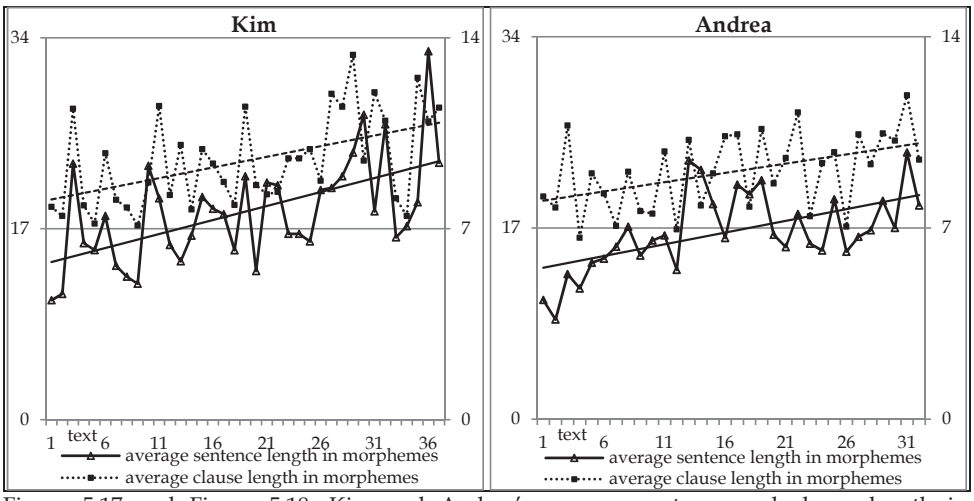

Figure 5.17 and Figure 5.18: Kim and Andrea's average sentence and clause length in morphemes.

To conclude, we hypothesized that the most informative expression of sentence complexity would be the average clause length in morphemes; of the three measures this one deals with the smallest part in a text and looks at morphemes as the smallest semantic unit. In the explorations, the measure indeed showed that it zooms in at a deeper level than the more complex sentences and in the previous part it showed that the measure is more informative than the average sentence length in morphemes as well. In contrast to other length-based measures, clause length is not affected by "variations in the amount of subordination exhibited in production" (Norris & Ortega 2009: 561). We decide

to use the average clause length in morphemes as syntactic complexity measure in the remaining sections of the chapter. This is in line with Norris and Ortega (2009), who claimed that in the average length of all finite clauses, growth can only result from the "addition of pre- or post modification within a phrase or as a result of the use of nominalizations, or the process of reduction of clauses into phrases which help to condense information"; in line with them we regard clause length as a "more narrowly defined source of complexification" (Norris & Ortega 2009: 561). This opinion is very much in line with the mean length of clause discussed by Norris and Ortega (2009) and also with the finite verb ratio used by Verspoor et al. (2008). The difference, of course, is that for Finnish, morpheme rather than word counts are better as words often have a great deal of internal complexity.

5.5 Discussion

This chapter traced the development of four sentence types (simple, compound, complex and compound-complex), the total use of compound, complex and compound-complex sentences, the average sentence length in morphemes and the average clause length in morphemes in the data of the two learners Kim and Andrea.

With respect to the development of the four sentence types, both learners showed a gradual increase in more complex sentence constructions. The similar outcomes of Kim and Andrea are not in line with the idea that type of instruction may have an effect on complexity (Verspoor et al., 2004), nor to the correlation of explicit rule knowledge and written abilities for beginner learners, as found by Dijkstra (2003).

With regard to the development of the more complex sentence patterns versus the average sentence length in morphemes, both measures showed variability during the time span for both learners. This is in line with Bulté (2013) and Murakami (2013), who found variability and variation in all subsystems in all learners. For both Kim and Andrea the correlation coefficient was strongly positive; in their data the two subsystems are supportive growers. Where the data showed competition, it might have been a rather random effect. Remarkable was the difference in Kim and Andrea's patterns of both measures while the correlation coefficients were quite similar. This could be due to variation among learners. Bulté (2013) also found such variation in learners in similar conditions.

With regard to the development of sentence complexity versus the average clause length in morphemes, the latter showed variability during the time span as well. Comparing the two learners in their development, the moving windows of correlation patterns illustrated strikingly similar patterns of peaks and dips. The correlation coefficient for both learners was weakly negative. When the complexity of the sentence changed this could happen in

the opposite direction for the number of morphemes at clause level. The focus on one subsystem seems to be at the expense of the other one.

In addition to tracing the development, the chapter aimed to find the most informative measure to be used in further chapters as a general holistic complexity measure. We found that both the average number of morphemes at the sentence level and clause level (morphemes per finite verb) increased. However, the average number at sentence level grew at the same time as sentence complexity, but at the clause level it competed with sentence complexity, showing that clause level complexity develops separately from sentence type complexity at the early stages. We found that the finite verb ratio for morphemes appears to be the most informative measure for the agglutinative Finnish language. This is in line with Verspoor et al. (2008); in a study applied to English, a language very low on morphology, they found that the finite verb ratio (words) is a better complexity measure than average sentence length or type of sentence (simple or complex).

6 SYNTACTIC COMPLEXITY: USE OF THE FINNISH CASES

The group analyses in chapter 4 showed that the FL group had more development over time than the L2 group in the use of cases: for both the total use of cases and for the use of the other 12 cases. Furthermore, the FL group showed more frequent use of the 12 other cases, whereas the L2 group showed more frequent use of the nominative. However, as the members of the L2 group had different language backgrounds, we decided to explore the development at the individual level for our focal learners with similar L1 language backgrounds.

In every section we will first trace the development of the eight individuals and then compare it to the focal learners. The question is whether we can attribute the differences at group level to the learning condition of these two groups or that L1 background may have played a role.

6.1 Total use of the cases

This section deals with the use of the cases by the two groups and by the two individuals. In the example sentences, the parts in bold are the parts we focus on.

The example sentences illustrate the total use of the cases by the eight individuals of the two groups. Example 1 shows a sentence in which fewer than half of the words contain case markings. Examples 2 and 3 consist of nouns, proper nouns and adjectives with case markings and only one verb. Example 4 illustrates a short sentence with only one case. Example 5 shows that with the application of negation, the sentence automatically contains more verbs and leaves less room for words with case markings. Example 6 shows an enumeration; this implies many case markings. In a construction with the logical subject in the genitive at least two verbs are needed, which implies less room for words with a case marking (at least in the case of no enumeration)

(example 7). Like example 4, also example 8 shows that it is possible to write a sentence with one case.

FL:

(1) On melko vaikea**a** muistaa ja kertoa jo**ta**kin oma**sta** perhee**stä**ni. *'It is rather difficult to remember and to tell something about my own family.'* (Kim-t14s1/mp13(9:4))
(2) **Hän** opiskeli lääketiede**ttä** Turu**n** yliopisto**ssa.** *'He studied medicine at the university of Turku.'* (Sanne-t33s3/mp33(5:4))
(3) Kalevala**n** teksti**ssä rytmi** on **yksi** vahvimm**ista** ominaisuuks**ista.** *'In the text of the Kalevala the rhythm is one of the strongest qualities.'* (Cleo-t25s7/mp16(7:6))
(4) **Minä** en rakastu helposti. *'I do not fall in love easily.'* (Annet-t10s5/mp4(4:1))

L2:

(5) Mieluimmin **mä** menen kirpputori**lle**, koska **tuo**[lex] vaatte**ita**[subj] eivät ole tosi kallis**ta**[pr.nom]. *'I more preferably go to the flea market, because there the clothes are not very expensive.'* (Andrea t9s9/mp8(11:5))
(6) Tänään illa**lla**, **mä** ostin **kuusi** peruna**a**, seitsemän sipul**ia** ja leipä**ä** Prisma**ssa**[infl]. *'Today in the evening I bought six potatoes, seven onions and bread at the Prisma.'* (Chiya t3s1/mp3(11:9))
(7) Minu**n** pitää tuoda paljon olu**tta** ja sima**a**. *'I have to bring a lot of beer and meat.'* (Bowo t33s2/mp29(8:3))
(8) Valitettvasti **mä** en tiedä miten. *'Unfortunately I do not know how.'* (Clara t20s1/mp20(5:1))

6.1.1 Differences between FL and L2

For the total use of the cases on average, the FL group showed a higher correlation with time (r=.26, SD=.19, n=4) than the L2 group (r=-.11, SD=.17, n=4). This difference was significant ($t(6)=2.82$; $p<.05$).

In figures 6.1 and 6.2 the dispersion of the total use of cases of both groups and of all learners individually are shown. In the figures, 35 measure points are shown because the L2 group wrote 35 texts. However, for the statistical analysis the FL group's data consists of 39 texts.

The patterns of both figures look quite similar. They both show considerable dispersion, though the bulk of the L2 group seems to lie more under the halfway line. The left sides of the figures show more use of cases for the L2 group than for the FL group and the right sides of the figures show the opposite: more use of the cases for the FL group. In the FL group the highest total use of case in one text is made by Annet (mp31; 71.1) and in the L2 group by Bowo (mp2; 77.8). Striking is that the lowest total use of cases in one text in the two groups is made by the focal learners of the current study: Kim (mp11; 33.8) and Andrea (mp12; 34.8).

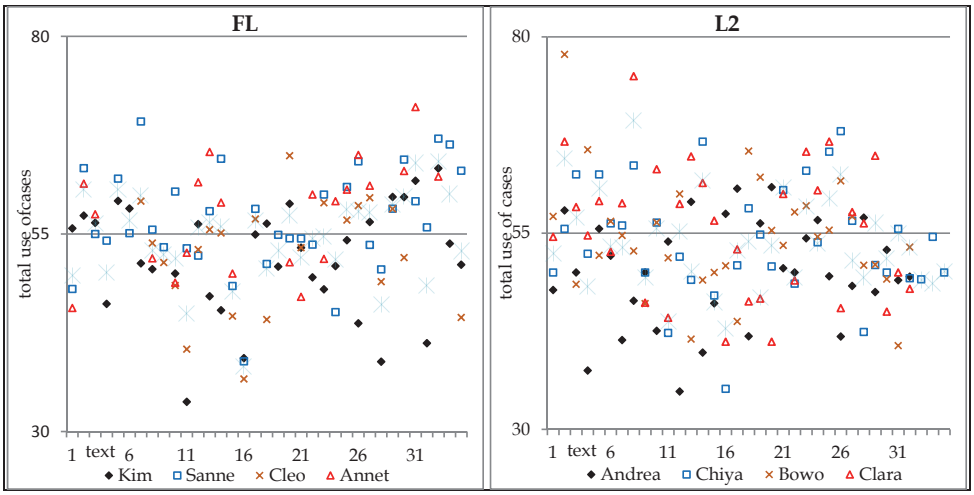

Figure 6.1 and Figure 6.2: FL and L2 participants' total use of cases.

6.1.2 FL, L2, Kim and Andrea compared

Figures 6.3 and 6.4 show the overall averages of the total case use by the FL and L2 groups and by Kim and Andrea. The difference between the figures is striking; both focal learners show a lower use of total use of cases than the overall averages of their groups.

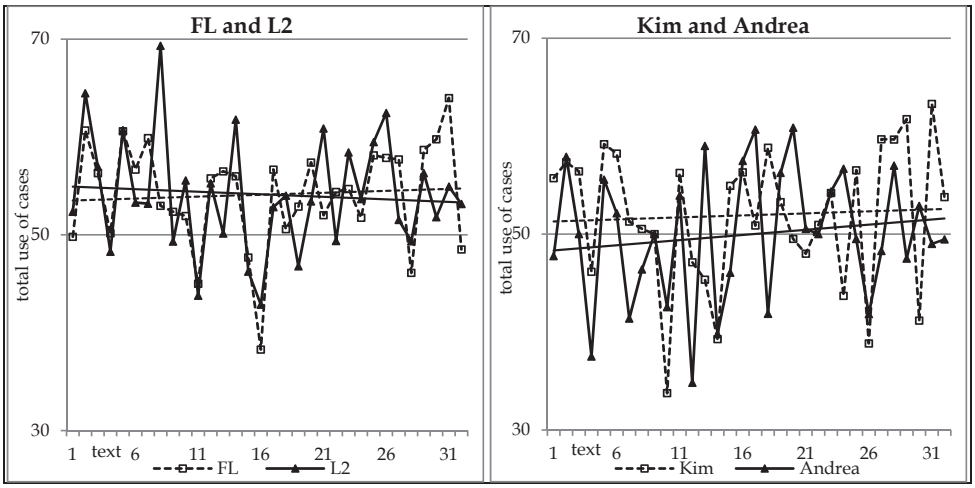

Figure 6.3 and Figure 6.4: Total use of cases by the FL and the L2 groups and by Kim and Andrea.

In figures 6.5 and 6.6 the focal learners are compared to their respective groups. The trend lines of both the FL group and Kim increase slightly, though Kim's shows a little less increase. The fact that her trend line is located at a lower level implies that the other FL participants use more cases during the whole period. Kim's data shows a remarkable dip at measure point 11 and her data is found in

a larger area than her group's data. For the L2 group the trend line declines, whereas Andrea's trend line increases. This implies that the other participants of the L2 group are increasingly using fewer cases during the period. The L2 group's data illustrate a striking peak at measure point 8, whereas Andrea shows two striking dips (mp4 and 13) and the data of Andrea show a larger dip than the L2 group. The graphs that represent the data of the FL group and Kim show 39 measure points. The graphs that represent the data of the L2 group and Andrea show 35 measure points.

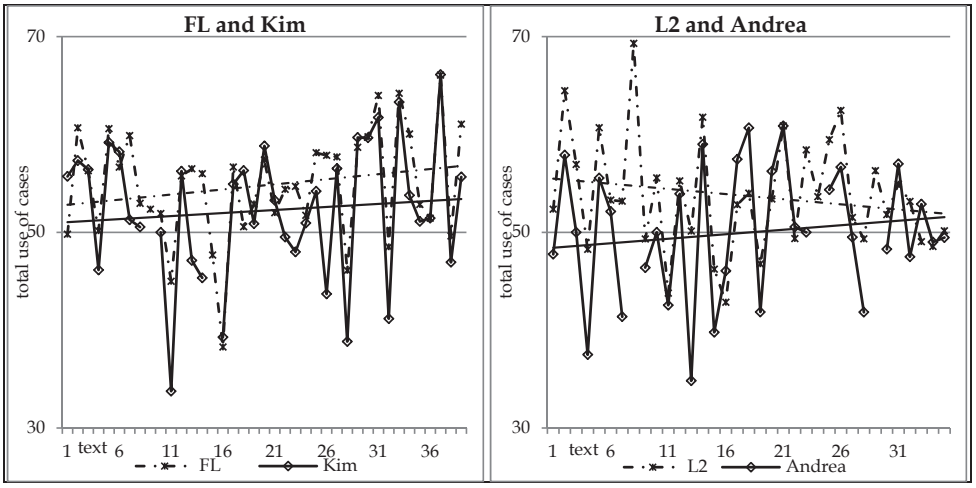

Figure 6.5 and Figure 6.6: Total use of cases by the FL group and Kim and by the L2 group and Andrea.

Now that the total case use is explored we will look at the nominative use.

6.2 Use of the nominative

The nominative is probably the easiest case to be learned as it is the most frequent case in the language. Both groups use the nominative more than any of the other cases, probably because it is also the default case (chapter 2) which learners resort to if they do not know the other cases. In other words, high nominative figures can indicate either correct use (the choice of sentence types in which the nominative is required) or incorrect use (default). Both could be interpreted as aiming at simplicity, so nominative use can mean automatic growth (nominative use in sentences that need a lot of nominatives, as an easy way to prevent incorrect use of cases) or development (incorrect nominative use in sentences that are more complex in use of the cases).

This section analyzes the nominative use by the two groups and by the two individuals and starts with example sentences.

The example sentences illustrate the nominative use (see also chapter 2) by the eight individuals of the two groups. Example 9 shows two possible causes for nominative use in a sentence: the presence of a singular countable subject and the predicate nominal, as do examples 10 and 11. Besides, in the latter the use of adjectives (in congruence with the case of the subject) is a cause for nominative use as well. Example 12 shows a negative sentence in which the presence of a singular countable subject and a predicate nominal usually contains at least two nominatives. In example 13 more than half of the words consist of conjunctions, verbs and adverbs, which leaves little room for words that can contain a case; six times a case is used, four of which the nominative (both subjects in combination with a predicate nominal, for one of which the nominative is used incorrectly. Example 14 shows a sentence with the have-construction; the subject with its adjectives is in nominative. Examples 15 and 16 both show a nominative used as case for the direct object. Both times this was an incorrect choice.

FL:

(9) Minun **asuntoni** on **uusi** ja aika **iso**, mutta **puutarha** on aika **pieni.** *'My house is new and pretty large, but the garden is rather small.'* (Kim-t5s2/mp5(12:5))

(10) **Se** on ilmeisesti **kansanomainen tanssi.** 'Apparently it is a folksy dance.' (Sanne-t27s4/27(5:3))

(11) Minusta **pieni suomalainen runo** on **huvittava.** *'I think the small Finnish poem is funny.'* (Cleo-t13s5/mp7(6:4))

(12) **Se** ei ole **kaupunki.** *'It is not a city.'* (Annet-t8s2/mp4(4:2))

L2:

(13) Mutta valitettavasti **sauna** tässä rakennuksessa on ihan **kylmä**, ja **ilma** on aika koste**a**[pr.nom]. *'But unfortunately the sauna in the building is rather cold, and the air is always humid.'* (Andrea-t18s6/mp17(13:4))

(14) Minulla myös on[w.ord] **pieni heleänvärinen esine**[lex]. *'I also have a small bright colored object.'* (Chiya-t10s5/mp10(6:3))

(15) **Minä** teen usein kotitehtävä**t**[d.obj]. *'I often do homework.'* (Bowo-t16s2/mp14(4:2))

(16) **Mä** näin myös paljon liput[d.obj] ulkona. *'I also saw a lot of flags outside.'* (Clara-t20s5/mp20(6:1))

Chapter 4 showed that the nominative use was significantly different between the two groups and therefore we will now take a further look at the data of the two groups and explore the question whether the focal learners develop the same way as their groups.

6.2.1 Differences between FL and L2

For the average use of the nominative, the L2 group showed a higher frequency (M=26.2, SE=3.5) than the FL group (M=22.3, SE=4.4). This difference was significant ($t(9)=-3.0$; $p<.05$).

Figures 6.7 and 6.8 show the nominative use of both groups and of all learners individually. The graphs look rather similar, but the bulk of data points of the L2 group seems to lie more above the halfway line than it does for the FL group. In the FL group Sanne shows several points of high nominative use. Kim shows a steady average amount of use of the nominative, whereas both Cleo and Annet show a steady rather low use. In the L2 group the measure points of Chiya and Clara show more nominative use than Bowo and Andrea.

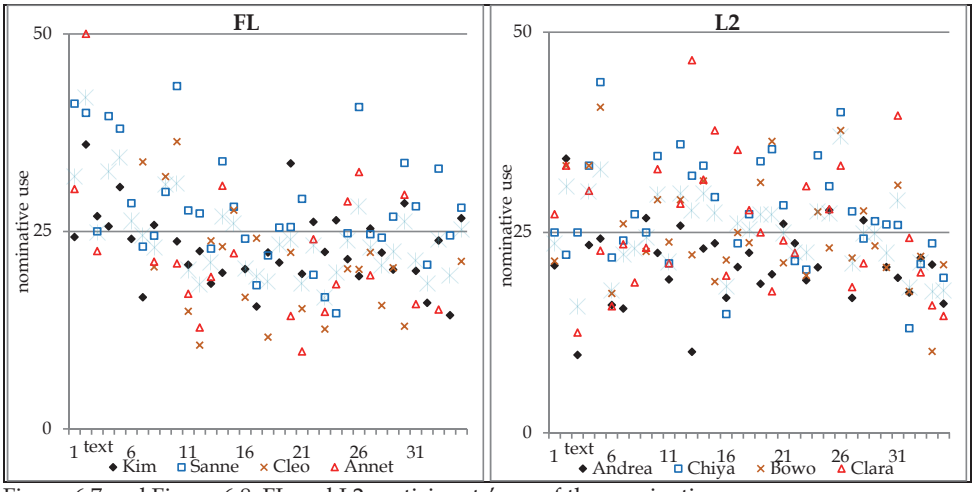

Figure 6.7 and Figure 6.8: FL and L2 participants' use of the nominative.

6.2.2 FL, L2, Kim and Andrea compared

Figures 6.9 and 6.10 show the nominative use by the FL and L2 groups (overall averages) and by Kim and Andrea. The data of the two participants show lower nominative use than the overall averages of the groups.

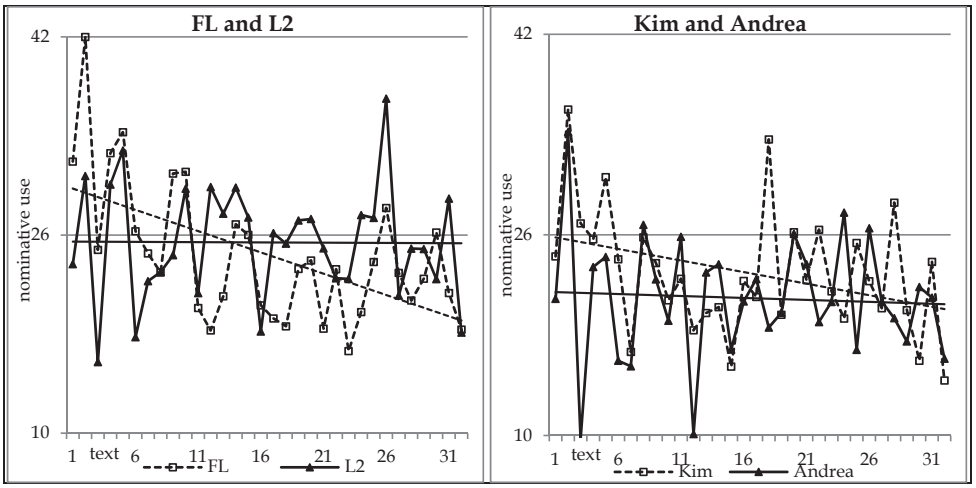

Figure 6.9 and Figure 6.10: Use of the nominative by the FL and the L2 groups and by Kim and Andrea.

In addition, figures 6.11 and 6.12 show the nominative use by the FL group and Kim and by the L2 group and Andrea. The trend lines of both the FL group and Kim decrease but Kim's trend line starts lower and ends higher. This means that the other FL learners, taken together, make more use of the nominative in the beginning but less at the end of the period. Both the FL group and Kim's data show striking peaks at measure points 2 and 18. Kim's data illustrate another peak at measure point 20. Andrea's trend line decreases more gradually and is located at a lower level than the L2 group's trend line. This means that for the other L2 learners, taken together, also the use of the nominative stays about the same during the whole period, but that they use the nominative more than Andrea. Both the L2 group and Andrea's data show a striking dip at measure point 3. Andrea shows another dip at measure point 12.

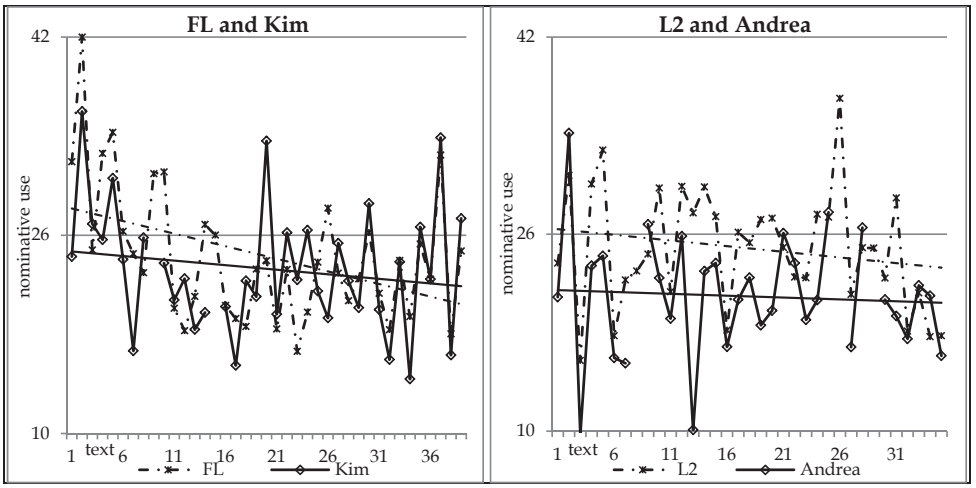

Figure 6.11 and Figure 6.12: Use of the nominative by the FL group and Kim and by the L2 group Andrea.

Because the nominative use shows isolated peaks in the data of both learners, they were tested for significance using a Monte Carlo analysis. Neither for Kim nor for Andrea the analysis showed significance, so they may be considered random peaks.

6.3 Use of the 12 other cases

Because the nominative seems to be overused by learners, other case endings may be used less. Also the genitive and partitive are used rather frequently. The remaining 15 cases are not used frequently enough to analyze separately so they are taken together in our analysis. We may assume that the more different cases learners use, the more complex language they use at the morpheme level. This section analyzes the use of the 12 cases other than the nominative, genitive and partitive by the two groups and by the two individuals. The section starts with example sentences.

The example sentences illustrate the use of the 12 other cases by the individuals of the two groups. Example 17 shows the use of two different cases, of which the first one is used quite often and the second one is used less. Example 18 shows the use of more words with the same case; this happens when a verb needs words with a certain case in order to express something specific (government). Example 19 shows the incorrect use of the essive. When Cleo answered the question of the 14th text (Millainen sinä olit pienenä? (*lit.*: *'How were you as a child?'*), she incorrectly used the case that had been used in the question. Example 20 shows an inflection error. Examples 21 to 24 show sentences with correct and incorrect use (and in 24: one incorrect form) of different cases.

FL:

(17) Lääkäri määräsi minu**lle** lääkkeitä viiko**ksi.** *'The doctor prescribed medication for a week to me.'* (Kim-t28s7/mp26(5:2))

(18) Minä pidän kelttiläise**stä**, irlantilaise**sta** ja skotlantilaise**sta** kansanmusiiki**sta** . *'I like Celtic, Irish and Scottish folk music.'* (Sanne-t25s6/mp25(7:4))

(19) Kun minä olin vielä piene**nä**[pr.nom]… *'When I was still a child...'* (Cleo-t14s1/mp8(5:1))

(20) Si**itä** kerrottiin historia**lla**[infl]. *'About that it was told in the history.'* (Annet-t22s7/mp13(3:2))

L2:

(21) Lauantai**na** mun täytyy kirjoittaa esseettä[cons.gr] suome**lta**[infl] historia**lta**[infl]. *'On Saturday I have to write an essay about the history of Finland.'* (Andrea-t6s3/mp6(7:3))

(22) Piene**nä** minä menin uimaan minun perhen[form][poss.suff] kanssa joka kesä**llä**[infl]. *'As a child I went swimming together with my family every summer.'* (Chiya-t14s9/mp14(9:2))

(23) Joka kesä**nä**[infl] on iso tangofestivaali Seinäjoe**lla.** *'Every summer there is a big tango festival in Seinäjoki.'* (Bowo-t27s3/mp24(6:2))
(24) Paras paikka kesä**llä** on mökki**ssä**[cons.gr] ehkä Mikkeli**ssä.** *'The best place in the summer is at a cottage, perhaps in Mikkeli.'* (Clara-t34s1/mp31(7:3))

What now follows is an exploration of the other 12 cases use because both the frequency of use and the development over time seemed to be significantly different between the two groups (chapter 4). The development of Kim and Andrea is examined as well.

6.3.1 Differences between FL and L2

For the average use of the 12 other cases, the FL group showed a higher frequency (M=14.2, SD=4.1) than the L2 group (M=11.6, SD=3.1). This difference was significant ($t(9)=2.45$; $p<.05$). On average the FL group showed also a higher correlation with time (r=.30, SD=.14, n=4) than the L2 group (r=-.07, SD=.14, n=4). This difference was significant as well ($t(6)=3.87$; $p<.05$).

Figures 6.13 and 6.14 show the use of the other cases of both the FL and L2 groups and within, of all learners individually. The figures look quite similar; the total number of measure points of both groups is concentrated in the area underneath the line halfway of the figure, though the FL data is concentrated on a little higher level. Besides, the FL group shows more data points above the midline than the L2 group. The patterns of the data illustrate quite some variability for both groups and for all learners. All FL learners use about the same number of cases other than nominative, genitive and partitive The FL data show only one striking peak (Sanne, mp7: 43.6). Also the L2 data show that the learners use about the same number of other cases, with the highest use by Clara at measure point 7 (35.3).

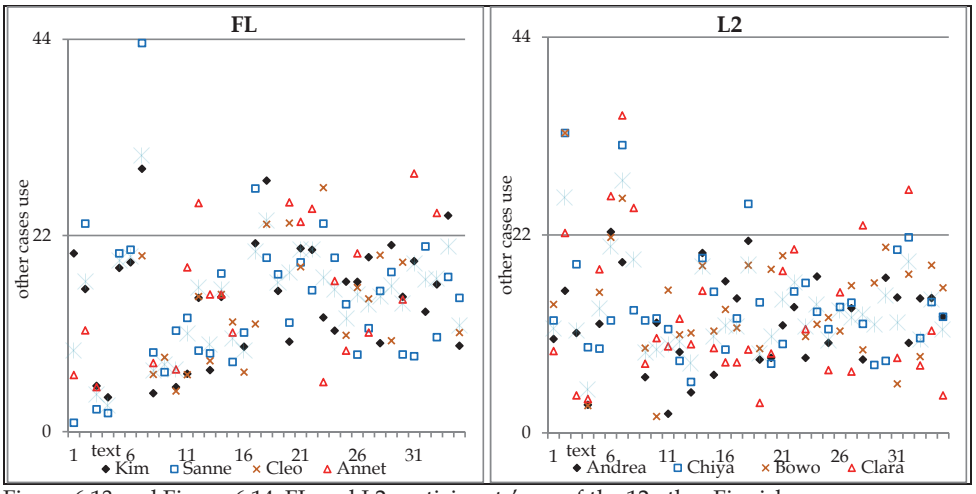

Figure 6.13 and Figure 6.14: FL and L2 participants' use of the 12 other Finnish cases.

6.3.2 FL, L2, Kim and Andrea compared

Figures 6.15 and 6.16 show the 12 other cases use by the FL and L2 groups (overall averages) and by Kim and Andrea. The graphs show similarities but differences as well.

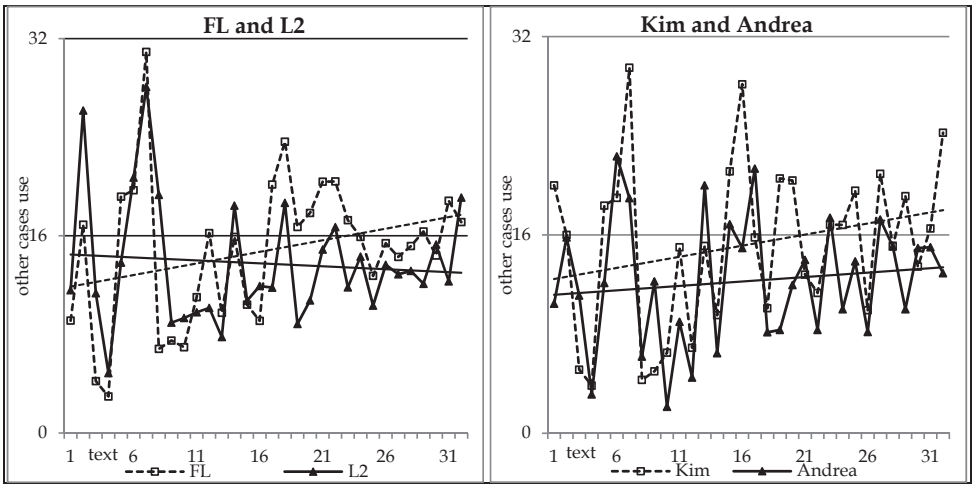

Figure 6.15 and Figure 6.16: Use of the other cases by the FL and the L2 groups and by Kim and Andrea.

In addition, the figures 6.17 and 6.18 show the 12 other cases use by the FL group and Kim and by the L2 group and Andrea. The FL group's trend line starts somewhat lower and end somewhat higher than Kim's trend line, but they both increase; Kim and the other FL learners use the other cases in a similar way during the period. The data of both the FL group and Kim show striking peaks at measure points 7 and 18. The first part of the data shows a lot of variability for both the FL group and Kim but there seems to be more stability in the second half of the period. The L2 group's trend line and Andrea show remarkable differences: slightly decrease versus slightly increase. Besides, the total trend line of Andrea is located at a lower level than the one of the L2 group. This means that the other L2 learners, taken together, use the other cases more at the beginning and about as much as Andrea at the end of the whole period. The L2 group's data show striking peaks at measure points 2 and 7 and one striking dip (mp4: also seen in Andrea's data). Andrea's data show a dip at measure point 11. For both the L2 group and Andrea the first half of the period shows a lot of variability in the other cases use, and for both there seems to be more stability in the second half.

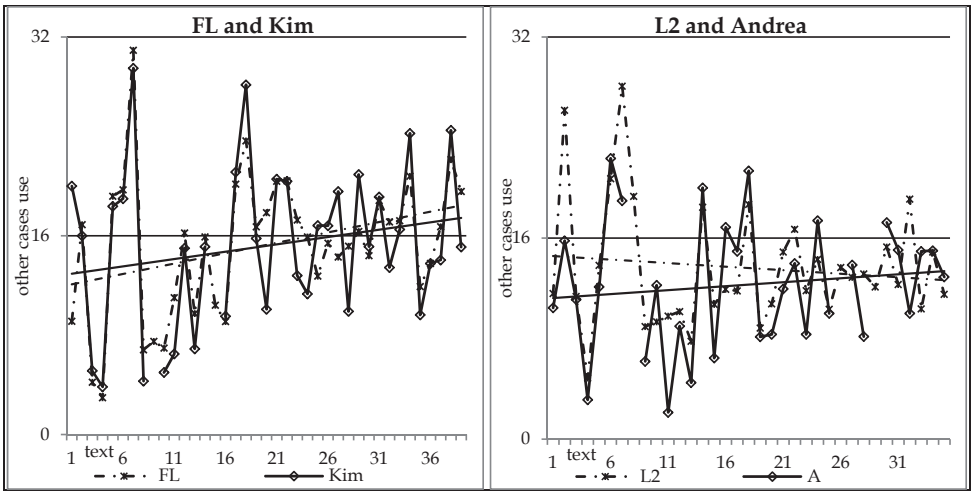

Figure 6.17 and Figure 6.18: Use of the other cases by the FL group and Kim and by the L2 group Andrea.

Because the other cases use shows isolated peaks in Kim's data, it was tested for significance (Monte Carlo analysis). The analysis showed a trend (p=.0574), suggesting that the learner is overusing the cases as a way to acquire them. For Sanne (FL), the analysis showed this trend as well (p=.0254). For both other learners of the FL group the analyses were not significant. Moreover, the data did not show more variability for Kim than for Andrea (Monte Carlo analysis).

Below we will discuss the findings of this chapter.

6.4 Discussion

This chapter traced three syntactic complexity measures: the total use of the cases, the use of the nominative and the use of the 12 other cases (i.e.: with the exception of the nominative, genitive and partitive case). The FL and the L2 groups' overall averages were shown and their comparisons with Kim's and Andrea's outcomes were outlined.

In total use of cases the comparison of the group outcomes with Kim's and Andrea's showed several differences; Kim and Andrea showed lower total use of cases than their respective groups. However, the L2 group showed a decrease but Andrea's trend line increased considerably, which would be in line with the expectation that the L1 plays a role in recognizing and using cases.

The nominative was used less by Kim and Andrea than by their groups at the beginning, but Kim used the nominative more than her group at the end. Andrea's nominative use stayed at a lower level than the L2 group during the time span, again in line with L1 expectations.

The use of other cases was the same for Kim and her group, though Kim started a bit higher and ended a bit lower. Because the peaks in Kim's use of

other cases were analyzed and found to be almost significant, the data of the other FL learners was analyzed, too; Sanne showed significant overuse of the other cases. The fact that two FL learners had significant peaks suggests that for these learners overuse serves as a way to acquire them. Andrea and her group differed quite throughout; Andrea started with fewer other cases use than her group but ended at about the same level; she showed an increase whereas the L2 group showed a decrease. There were no significant peaks found in Andrea's data. Besides the role of L1, Andrea's interest in grammar might have played a role in the outcomes.

With regard to the final outcomes of the two groups discussed in chapter 4, we might have concluded that context and instruction has an effect; however, the L1 effect (Murakami, 2013) seems to minimize these differences as there in the end there are few clear differences between Kim and Andrea). The one effect that might possibly be contributed to instruction is the differences in trajectories. Kim and one other FL learner showed significant peaks of overuse in the use of other cases, whereas Andrea or the L2 learners did not. However, not every FL learner showed these peaks either.

7 SYNTACTIC AND MORPHOLOGICAL COMPLEXITY

Like chapters 5 and 6, this chapter will explore development at both group and individual level of a measure that showed significant differences in chapter 4. Average word length in morphemes showed significant differences in development over time in favor of the FL group. In addition, the development (including peaks and dips) and interactions of several syntactic and morphological complexity measures will be explored: average clause length in morphemes versus average word length in morphemes and use of past and perfect tense. However, before using the past and perfect tenses as complexity measure, the effect of topic and task is examined.

7.1 Average word length in morphemes

The average word length in morphemes has not been used as a complexity measure yet in the previous chapters. Therefore, we will start with the difference between the FL and L2 groups. As in the previous chapters, bolded items in examples point to the issues in question.

The example sentences illustrate the average word length in morphemes used by the two focal learners to illustrate differences between word lengths in number of morphemes rather than characters, which is often used as a complexity measure in English. The point is that the number of characters is not an indication of internal word complexity if operationalized as number of morphemes in the word. Example 1 shows a sentence consisting of four short words (one or two morphemes each) from Kim's last written text; apparently the need to write use words with more morphemes is not always there, not even in the last writing. Example 2 shows only two morphemes: a long adjective with 14 characters with the less commonly used case, the translative (which has previously been counted as part of use of the other cases). Example 3 shows the rather commonly used first person singular past tense of the verb

olla (*to be*), which means three morphemes in four letters. Example 4 illustrates two conjunctions, both consisting of few letters and of one morpheme. Example 5 and 6 show long words (resp. 13 and 12 letters). The former consists of only one morpheme and the latter of three (compound and case marking). Example 7 shows a ten letters word consisting of four morphemes (compound, case markings singular and plural). Example 8 shows a three letter word[32] consisting of two morphemes.

FL:

(1) Nyt o/n tärkeä päivä. *'Now it is an important day.'* (Kim-t39s5/mp37-WlMrph1/2/1/1)
(2) luterilaise/ksi *'lutheran'* (Sanne-t39s6/mp38-WlMrph2)
(3) ol/i/n *'I was'* (Cleo-t35s1/mp22-WlMrph3)
(4) … että…ja… *' that…and…'* (Annet-t33s5/mp21-WlMrph1/1)

L2:

(5) toivottavasti *'hopefully'* (Andrea-t35s6/mp32-WlMrph1)
(6) yli/opisto/sta *'from the university'* (Chiya-t35s2/mp35-WlMrph3)
(7) revo/n/tul/ia *'northern lights'* (Bowo-t35s2/mp32-WlMrph4)
(8) mu/n *'my'* (Clara-t35s1/mp32-WlMrph2)

7.1.1 Differences between FL and L2

For the average word length in morphemes, on average the FL group showed a higher correlation with time (r=.53, SD=.05 n=4) than the L2 group (r=.06, SD=.18, n=4). This difference was significant ($t(6)=5.0; p<.05$).

Figures 7.1 and 7.2 show the overall average word length in morphemes of both groups and of all participants individually. The patterns of the FL and L2 groups look quite different. Though both show considerable dispersion, the bulk of the L2 group is located more under the midline. Throughout, the FL group shows a larger average word length in morphemes. In the FL group the largest average word length in morphemes in one text is realized by Annet (mp31; 2.1) and in the L2 group by Chiya and Bowo (mp2; 2.1). The smallest average word length in morphemes in one text is produced by our focal learners: Kim (FL: mp4; 1.4) and Andrea (L2: mp28; 1.4).

---

32 minu/n->mu/n *'my'*

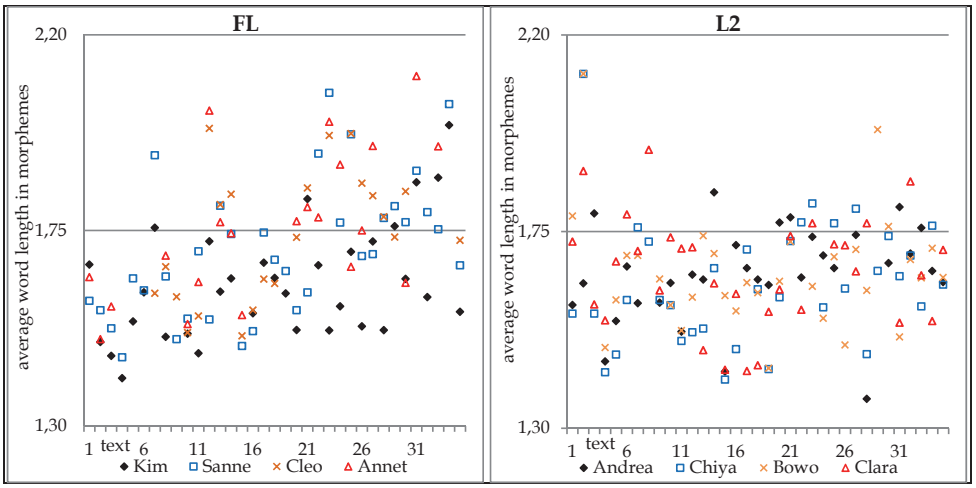

Figure 7.1 and Figure 7.2: FL and L2 participants' average word length in morphemes.

7.1.2 FL, L2, Kim and Andrea compared

Figures 7.3 and 7.4 show the overall average word length in morphemes by the FL and L2 groups and by Kim and Andrea. The similarity in trend of the figures is striking.

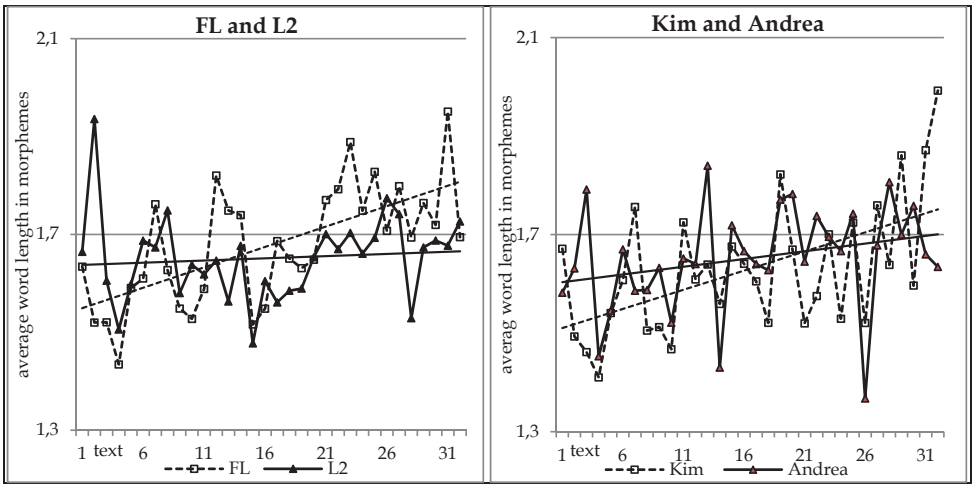

Figure 7.3 and Figure 7.4: Average word length in morphemes by the FL and the L2 groups and by Kim and Andrea.

In addition, figures 7.5 and 7.6 show the average word length in morphemes by the FL group and Kim and by the L2 group and Andrea. The FL group's trend lines and Kim's both show remarkable increase, though the former grows a little faster and is situated at a higher level; Kim uses words with fewer morphemes than the group as a whole. Both the data of the FL group and Kim show a striking peak at measure point 34 and a remarkable dip at measure

point 4. The L2 group's trend line shows hardly any growth, whereas Andrea's trend line increases, starting at a lower level and ending at a higher level than the L2 group. At the beginning Andrea uses words with fewer morphemes than the L2 group and at the end she uses words with more morphemes. The L2 group's data show a striking peak at measure point 2, and Andrea's data at measure point 14. The dips of the group and Andrea are at the same measure points (mp4, 15 and 28). Overall, the complexity at word level shows clear differences in development.

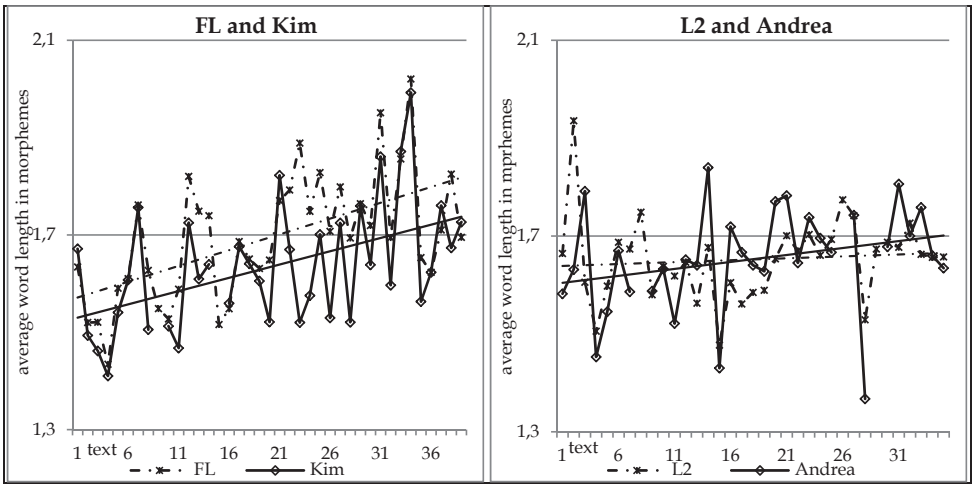

Figure 7.5 and Figure 7.6: Average word length in morphemes by the FL group and Kim and by the L2 group and Andrea.

Because the average word length in morphemes shows an isolated peak in Andrea's data, it was tested for significance (Monte Carlo analysis). The analysis showed no significance, so the peak was no developmental sign of over or underuse.

7.2 Development of average clause length in morphemes versus average word length in morphemes

This section explores the development of average clause length in morphemes and average word length in morphemes. Both measures express complexity at the word level, which encourages the expectation that a growing average word length in morphemes means a growing average clause length in morphemes (consisting of additions of the former measure) and the other way around. However, some examples show that there is no one on one relation between the two measures. Therefore, one of the questions is whether they develop in tandem or whether they compete. As learners use more words per clause, these

words do not necessarily have to be more complex. The other question is whether Kim and Andrea develop in a similar manner in this respect.

Kim

Figure 7.7 shows Kim's average word length in morphemes. In this morphological complexity measure, complexity is regarded as the average number of morphemes per word. The measure shows quite some increase for Kim during the time span.

The pattern of the average word length in morphemes shows quite some variability during the period. The largest variability takes place in the first seven measure points; measure points 1 to 4 show a large decrease, after which the line goes up again from measure point 5 to 7. Texts 12 to 17 show relatively low variability (1.61-1.67). Measure point 19 shows a remarkable peak (1.82, coming from 1.52 in mp18) and a striking dip is seen at measure point 33 (1.56, coming from 1.99 in mp32). Finally, the graph shows a rather variable pattern of succession of narrowness and wideness in the bandwidth within the developmental trend of growth; the variability stays rather high during the whole period.

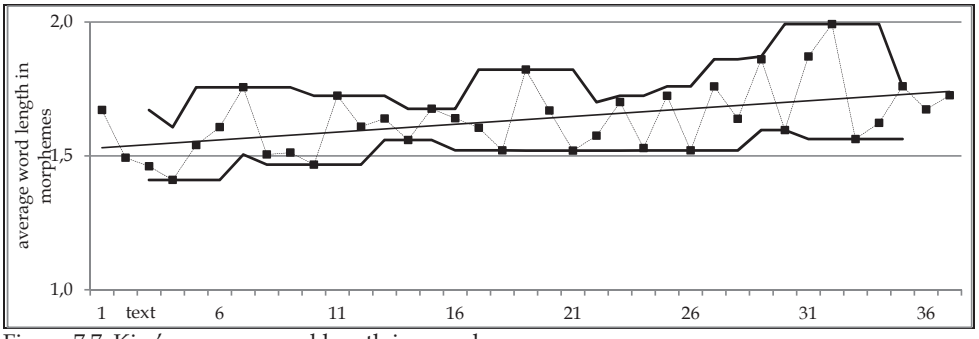

Figure 7.7: Kim's average word length in morphemes.

Andrea

Figure 7.8 illustrates Andrea's average word length in morphemes and shows a slight increase.

The graph shows a rather variable pattern for the average word length in morphemes during the period. The pattern shows variability throughout with some remarkable peaks (mp 3 and 13) and dips (mp 4, 14 and 26); what strikes is that these measure points represent consecutive peaks and dips (with the exception of the dip at measure point 26). Several quite striking sequences of measure points are found, some of which show low variability (mp15-18 and 22-24 and some show rather increase (mp 1-3, 4-6 and 26-28). Finally, narrowness and wideness in the bandwidth alternate as is expected in dynamic thinking. Besides, the variability remains rather high during the period.

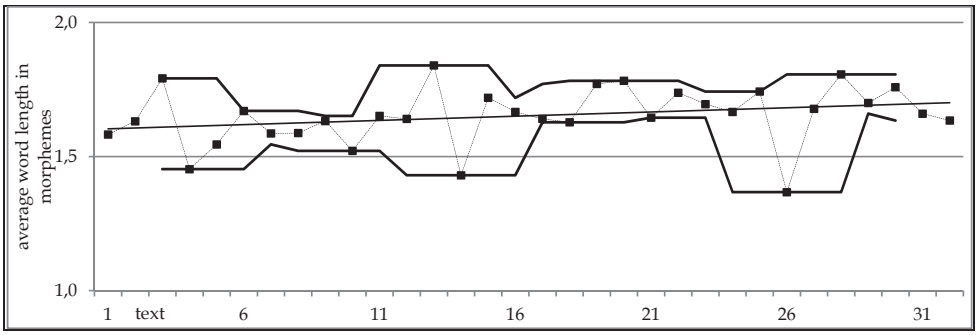

Figure 7.8: Andrea's average word length in morphemes.

7.2.1 Interaction between average clause length in morphemes and average word length in morphemes

In this subsection we look at the average clause length in morphemes and the average word length in morphemes. First we take a look at Kim and Andrea separately and then we compare their development. We start with example sentences.

The example sentences illustrate the fact that the use of the average clause and word length in morphemes do not necessarily have a one-on-one relation. The first two examples present positive correlations of the two measures. Example 9 consists of just one and two-morpheme words in two small clauses. In example 10, both the clauses and words consist of a relatively high number of morphemes. In contrast, example 11 shows some competition between the measures: the clause length in morphemes is a little above Andrea's average (9.1) and at the same time the average word length in morphemes is rather low and much lower than her average (1.65). One possible cause for the low average word length in morphemes is the use of some adverbs (one morpheme) and *oli* (3.sing.past2mrph: a derivative of the verb *olla* (*to be*)). See chapter 2 for the relationship of the headword with a predicate nominal and the possible consequences for the nominative use.

(9) X e/i ole kylä // mutta o/n kaupunki. *'X is not a village but a city.'* (Kim-t2s2/mp2-ClMrph2/9(5-4:4.5)-WlMrph7/9(1.29))

(10) Alu/ssa he ol/i/vat ol/leet vain ystäv/i/ä // ja he ol/i/vat teh/neet kaikenlais/i/a kivo/j/a tavaro/i/ta[lex] yhde/ssä. *'In the beginning they had been just friends and they had done all kinds of nice things together.'* (Kim-t39s2/mp37-ClMrph2/30(12-18:15.0)-WlMrph14/30(2.14))

(11) Käyminen sauna/ssa säännöllisesti[part sp][w.ord] ol/i[tns] luultavasti myös hyvä estäminen[lex]. *'Going to a sauna regularly was probably also a good prevention.'* (Andrea-t28s5/mp26-ClMrph1/10-WlMrph8/10(1.25))

Kim

Figure 7.9 presents the average clause and word length in morphemes in Kim's writing. Looking at the trend lines, it shows that both measures grow and they seem to grow at about the same rate. Both patterns show a great deal of variability. In the first half the patterns look quite similar but in the second half this impression changes slightly; here the measures show more differences in frequency. However, the ups and downs of both measures remain similar during the time span.

Overall, a large majority of measure points show support between the two measures and only twice one measure stays at the same level (mp23-24 and 8-9) while the other one decreases. Kim's correlation coefficients show a very strong positive correlation between the measures for the raw data (R=.66) and a strong positive correlation for the residual data (R=.56).

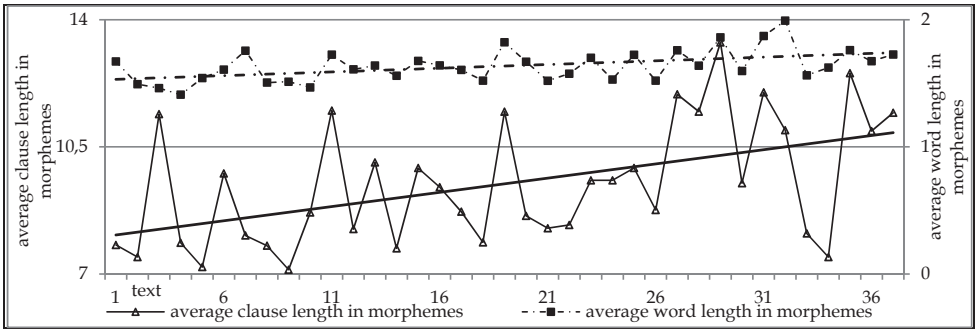

Figure 7.9: Kim's average clause and word length in morphemes.

Andrea

Figure 7.10 shows the average clause length in morphemes and the average word length in morphemes in Andrea's writing and it shows that the measures both grow, though the trend line of the average clause length in morphemes shows a larger increase than the one of the average word length in morphemes. Both measures show a rather large degree of variability and similarities as well. However, the patterns show quite some differences in frequency.

Overall, the graph shows competition and support between the two measures. Once, the average word length in morphemes stays at the same level while the average clause length in morphemes increases (mp7-8). Andrea's correlation coefficients show a very strong positive correlation between the measures for the raw data (R=.61) and a strong positive correlation for the residual data (R=.57).

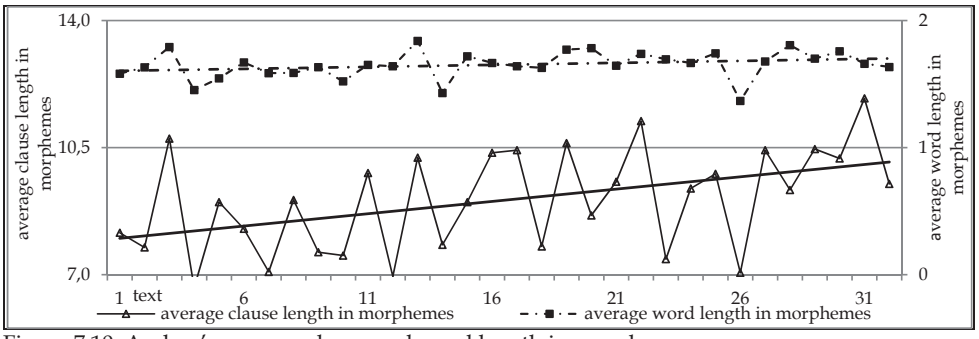

Figure 7.10: Andrea's average clause and word length in morphemes.

Kim and Andrea compared

Figure 7.11 shows a moving window of correlations of residual data. The measures of Kim and Andrea show mainly support and only a few points of competition.

These findings support the hypothesis that the average clause length in morphemes and the average word length in morphemes are supportive growers. Indeed this is true for both Kim and Andrea (residual data); at the beginning they both write shorter clauses and words and at the end longer clauses and words in terms of morphemes. Moreover, overall the moving correlation patterns of Kim and Andrea show quite some similarities.

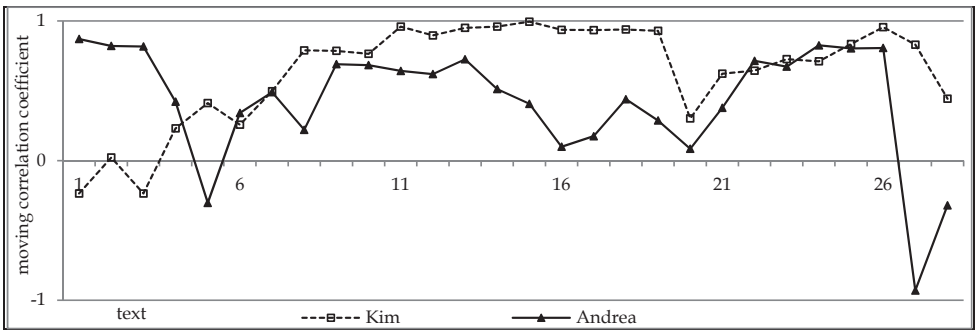

Figure 7.11: Kim and Andrea's average clause and word length in morphemes and their interaction.

7.3 Tenses

One research question was whether the use of tenses could be considered a good developmental measure or whether it shows a strong task effect. Therefore, in the next section we will first look at task effect, before exploring the two morpheme measures in relation to tense use.

There was no significant difference at group level in the use of past and perfect tenses, but in other studies they have shown to be developmental measures (cf. Verspoor et al., 2012). However, as Bulté (2013) has pointed out, some complexity measures may be task induced. Therefore to make sure this is not the case in the current study, we will first explore the use of the tenses in the texts, this time only at the individual level. After this, we will discuss Kim and Andrea separately and we will compare their development. The section takes off with example sentences.

Examples 12 and 13 show sentences written in the tense that was used in the task question (12: present, 13: past). This can be interpreted as a task effect of the use of the tenses. Examples 14 and 15 show sentences which are written in another tense (past) than the question was put in (present). It seems that Andrea wanted to share memories and knowledge with the reader and used the past tense for this purpose. The examples include the task questions.

(12) Q: **Vietetään**kö sinun kotimaassasi juhannusta? *'Do they celebrate Midsummer in your own country?'*
Minä **olen** kotoisin Saksan eteläosasta, ja siellä juhannusta **vietetään**. Minä **en tiedä**, **vietetäänkö** juhannusta kaikissa osissa Saksassa[infl]. *'I am from the southern part of Germany and Midsummer is celebrated there. I do not know whether Midsummer is celebrated in every part of Germany.'* (Andrea-t24s1,2/mp22-present)

(13) Q: Milloin Suomi **tuli** itsenäiseksi? *'When did Finland become independent?'*
Koko ajan Turku **oli** tärkeä kaupunki. Kun Venäjä **voitti** Ruotsin sodassa, Suomi **tuli** Venäjän osa[infl]. *'All time, Turku was an important city. When Russia beat Sweden in the war, Finland became a part of Russia.'* (Kim-t20s3,4/mp18-past]

(14) Q: Mitä sinä **tiedät** suomalaisesta joulusta? *'What do you know about Finnish Christmas?'*
Minun ystäväni **antoi** minulle kirjan suomeksi[w.cl]. *'My friend gave me a book in Finnish.'* (Andrea-t21/s10/mp20-past)

(15) Q: Valitse taiteilija, jonka tauluista sinä **pidät** eniten. *'Choose the artist, whose paintings you like the best.'*
Hän **maalasi**[lex] paljon motiivia[lex] Kalevalasta. *'He painted a lot of motifs from the Kaleva.'* (Andrea-t31s6/mp28-past)

Kim

Figure 7.12 shows the use of the tenses in Kim's writing and shows a decrease of the use of the present tense and an increase for the use of the past and perfect tense, which includes expressions like *olen kiinnostunut* (*I am interested*) and *olen masentunut* (*I am depressed*). The trend line of the past tense grows considerably faster than the trend line of the perfect tense, though. Their trend lines cross at the end.

The perfect tense is not used until measure point 10 and the past tense not until measure point 13; Kim just uses the present tense in the first nine

measure points. Another four times this is the case. In several texts in which the past tense is used most frequently, the present tense is not or hardly used and there is no text that shows a most frequent use of the perfect tense. Fourteen times the present, past and perfect tense are used together.

The use of the present tense shows one striking peak (mp34) and the use of the past and the perfect tenses show some small peaks. In the second half of the period, Kim shows a good mix of use of the tenses. Even though the present tense is used throughout, it is not the main tense anymore and every tense is used. Overall, Kim shows development, even though there may not be an attractor state (see 1.3.3) yet at the end of the year.

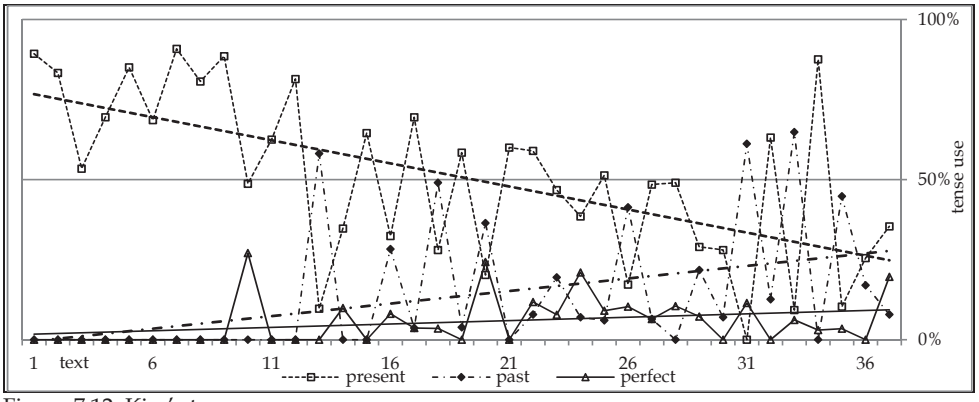

Figure 7.12: Kim's tense use.

Andrea

Figure 7.13 illustrates the use of the tenses in Andrea's writing and shows a fast decrease of the present tense. The past tense shows very little increase and the hardly noticeable trend line of the perfect tense shows some decrease. No trend lines cross in the figure.

In about half of all texts Andrea makes use of just the present tense and in the other half at least one other tense is applied. The perfect tense is used at measure point 4 for the first time and the past tense at measure point 8. Three times, the past tense is used most frequently and there are no texts in which the perfect tense is used most frequently. Three times the present, past and perfect tense are used together (mp12, 24 and 29).

The use of the present tense shows some striking peaks and dips (e.g. measure point 25, in which the present tense was not used at all). The use of the past tense is striking in two points; at measure point 13, the large use of the past tense entails that the other tenses are not or hardly used, but at measure point 20 the present tense is also used frequently. Finally, the perfect tense use shows one small peak (mp10). Andrea does not show a balanced mix of the use of the tenses towards the end; she keeps on using the present tense a lot throughout, while the use of the past tense hardly increases and the perfect tense stays at the same level. Her data show relatively little development.

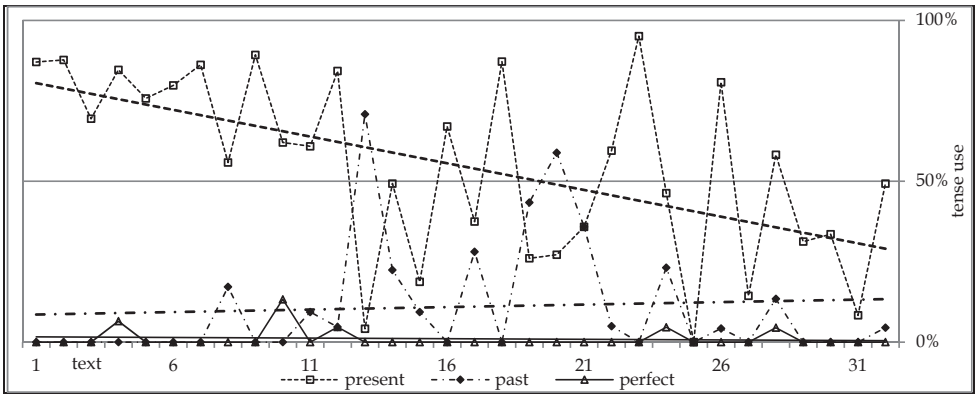

Figure 7.13: Andrea's tense use.

Kim and Andrea compared

In order to find an answer to the question whether task relatedness might play a role in the use of the tenses, the trend lines and the data of Kim and Andrea's tense use and also some example sentences are compared. The trend lines representing the present tense use illustrated a rather identical pattern over time, whereas the trend lines for the past tense looked quite different; where Kim showed a growing use of the past tense, Andrea took off higher after which the line showed hardly any increase anymore (figure 7.12 and 7.13). Furthermore, their use of the perfect tense is quite different; Andrea's trend line decreases, whereas Kim's trend line shows growth.

The similarities and differences in the use of the tenses by both learners are illustrated in table 7.A. It shows that there are only two striking similarities in the use of the tenses; first, the use of just the present tense in the first part of the period and second, the use of the past tense at measure point 13. The latter also holds for measure point 20 for which, however, both participants wrote a different task and which is therefore not interpreted as similar. Finally, the same tasks for Kim (mp15) and Andrea (mp16) show for both a small peak for the use of the present tense with a task question on literature put in the present tense. A similar peak is seen for the use of the past tense, for Kim at measure point 18 and for Andrea at measure point 19, but in this task, the task question for the L2 group was inadvertently put in the present tense, whereas the question for the FL group was put in the past tense.

Table 7.A: Similarities (bold) and differences in Kim and Andrea's tense use.

Tense use	Kim	Andrea
Just present tense	**1-9**, **11**, **12**, **15** and 21	**1-3**, **4-6**, **7**, **9**, **16**, 18, 23, 25 and 30-32
First perfect tense use	10	4
First past tense use	13	8
Past tense used most frequently	**13**, **18**, **20**, 26, 31, 33, 35; (**13**, 31, 33 and 35 hardly/ no present tense)	**13**, **19** and **20**; (**13** hardly present tense)

Our exploration of task effect on the use of tenses shows that there is no clear relation between task and use of tense. Therefore, we may assume that the patterns found are developmental.

To summarize, an overall view at the three tenses' trend lines by both learners illustrates that the use of the present tense decreases but plays an important role throughout. Moreover, the less used past and the even less used perfect tense grow slightly (the latter with the exception of Andrea, who shows a perfect tense use of almost none) but comparing the two learners shows that Kim's good mix of tenses use is a mark of development, whereas Andrea's use of the tenses is a sign of relatively little development. Overall, their outcomes are what would be expected in the use of tenses by beginner learners of an FL/L2.

7.4 Development of average clause length in morphemes versus use of past and perfect tense

In this section we will introduce the use of the past and perfect tense and we will explore the average clause length in morphemes in combination with the use of the past and perfect tense to show their development and their interactions. We will also explore the use of the past and perfect tense to find whether the measure is suitable to use as complexity measure.

Kim

Figure 7.14 shows the past and perfect tense use in Kim's writing and illustrates increase. The data show enormous variability though the first nine measure points show no use of these tenses. The highest frequency of the complex tense use in Kim's data is 90 percent (mp31).

The graph shows several striking peaks. Three times the data show a rather smooth line with little variability (mp22-24, 27-28 and 36-37). The first

part of the data (1-9 excluded) shows a rather stable pattern of mainly wideness in the bandwidth, after which the bandwidth narrows for a short while (mp23-24). Subsequently the trend of widening without narrowing goes on in the second part, though this time on a slightly higher level. Overall, the graph shows variability throughout.

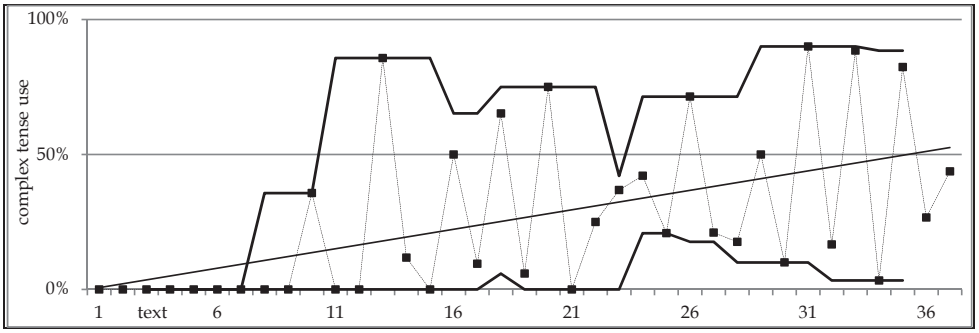

Figure 7.14: Kim's past and perfect tense use.

Andrea

Figure 7.15 shows the use of the past and perfect tense use in Andrea's writing and shows a small increase. The data show considerable variability, though. The largest use of these two tenses is shown in the middle of the period and the most frequent use is found at measure point 13 (94 percent).

The graph shows several striking peaks. In measure points 10-12, the data show a smooth line with little variability. The first half of the graph starts with a narrow bandwidth increasing to large wideness, after which in the middle of the period the bandwidth narrows for a short while (mp16). Also in the second half the bandwidth first widens and then narrows again. Overall, the use of the past and perfect tense remains variable during the whole period.

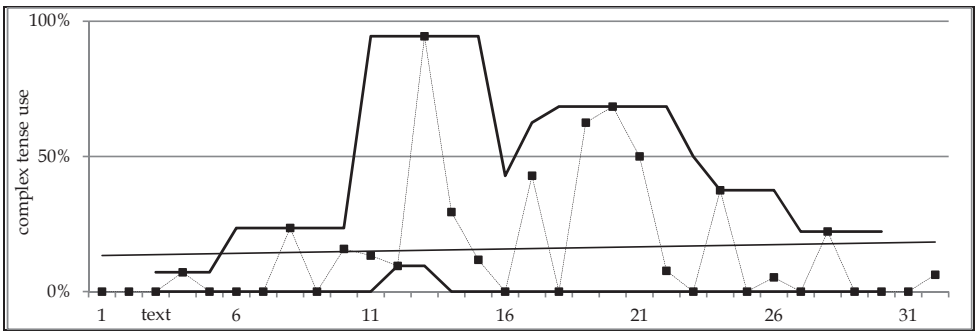

Figure 7.15: Andrea's past and perfect tense use.

Because the use of past and perfect tense shows an isolated peak in Andrea's data, it was tested for significance (Monte Carlo analysis). The analysis showed no significance, so the peak was no indicator for growth.

7.4.1 Interaction between average clause length in morphemes and use of the past and perfect tense

This subsection looks at the combination of the average clause length in morphemes with the past and perfect tense use. First Kim and Andrea are discussed separately and then their development is compared. The subsection starts with example sentences.

The example sentences (including the questions) illustrate the past and perfect tense use in combination with the average clause length in morphemes. Example 16 shows a sentence in the past tense with a small clause length in morphemes, coming from a writing task of which the question was put in the past tense as well. The dates were written in numbers and counted as one morpheme per date. In the question the subject was described in the third person singular (implying two morphemes for the finite verb). Example 17 shows a sentence that is written in the past tense while the question was put in the present tense. In this case the writer is the subject and the sentence is written in the first person singular past tense (implying three morphemes for the finite verb). If the present tense would have been applied, it would have been *tarvitse/n: present2mrph*. Example 18 shows a striking peak in the complex tense use. In this text the learners had to write about their youth, which connotes writing in the past tense whit the probable frequent use of the first person singular (and therefore at least three morphemes). Example 19 illustrates the use of the pluperfect (with a relative large number of morphemes in the clause) and the perfect with subjects in the third person singular. For a total overview of topics of the free writing tasks, see Appendix C.

(16) Q:Kuka **oli** Eliel Saarinen? *'Who was Eliel Saarinen?'*
Eliel Saarinen **el/i** 1873-1950. *'Eliel Saarinen lived from 1873-1950.'* (Kim-t35s3/mp33-past2mrph-ClMrph1/5)

(17) Q: **Tarvitset**ko sinä nyt uuden takin ja uudet kengät? *'Do you now need a new jacket and new shoes?'*
Todellinen[lex] minä **tarvits/i/n** uude/t kengä/t. *'Actually, I needed new shoes.'* (Andrea-t9s1/mp8-past3mrph-ClMrph1/9)

(18) Q: Missä sinä **asuit** lapsena? Millainen **olit** pienenä? *'Where did you live as a child? How were you when you were little?'*
Piene/nä minä **men/i/n** yleensä mielellä/än[poss.suff] koulu/un // koska minä **tapas/i/n** mu/n kavere/j/a[form][poss.suff] siellä // ja minä **tykäs/i/n**[cons.gr] oppia uus/i/a asio/i/ta. *'When I was little, I usually loved to go to school, because I met my friends there and I liked to learn new things.'* (Andrea-t14s5/mp13-past3mrph-past3mrph-past3mrph-ClMrph3/34(11-11-12:11.3))

(19) Q: Kuka **oli** Eliel Saarinen? *'Who was Eliel Saarinen?'*
Kun Eliel **ol/i kuol/lut**_// Eero **perust/i** oma/n arkkitehti/toimisto/on[d.obj]. *'When Eliel had died, Eero founded his own architectural office.'* (Kim-t35s9/mp33-pl.perf2/2mrph-past2mrph-ClMrph2/14(6-8:7.0))

Kim

Figure 7.16 shows the average clause length in morphemes and the use of the past and perfect tense in Kim's writing and shows increase for both measures. The clause length in morphemes grows less and shows less variability than the complex tense use though, but both measures show quite some variability.

Overall, the data show some competition (mp1-9 are put aside, so: mp10-37) and several small periods of supportive correlation (mp13-14, 21-24, 28-32 and 34-37). Kim's correlation coefficient (raw data) shows a positive correlation (R=.22) between the measures, but the very weak negative one for the residual data shows that this support should not be overvalued (R=-.03) (Verspoor et al., 2008).

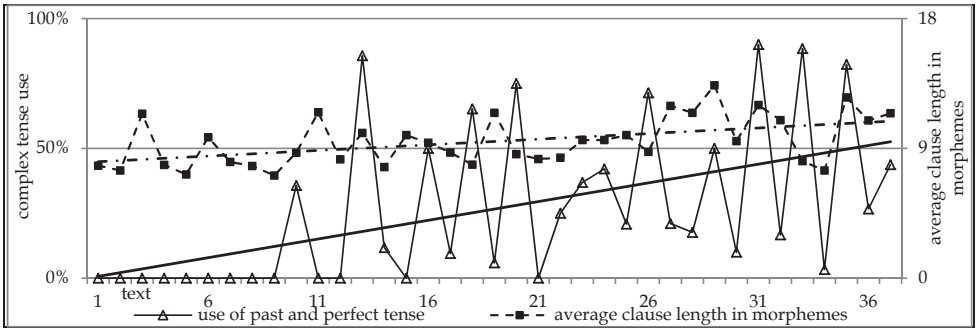

Figure 7.16: Kim's average clause length in morphemes and past and perfect tense use.

Andrea

Figure 7.17 shows the average clause length in morphemes and the use of the past and perfect tense in the writing of Andrea and shows growth for both measures. The measures seem to increase quite similarly. Both patterns show variability, though the pattern of the average clause length in morphemes seems less variable than the pattern of the complex tenses.

Overall, the data show some competition (mp1-3 are put aside, so: mp4-32), and the graph illustrates several small periods of supportive correlation as well (mp8-9, 12-13, 17-19 and 23-24). Andrea's correlation coefficients (raw and residual data) show weak positive correlations (R=.16; R=.15) between the measures.

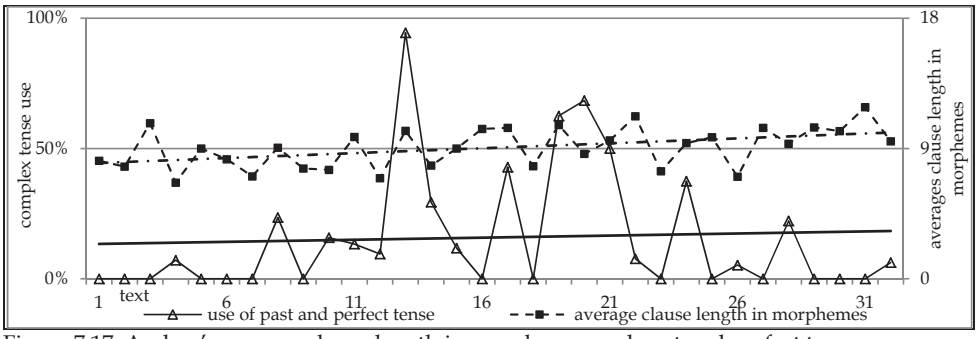

Figure 7.17: Andrea's average clause length in morphemes and past and perfect tense use.

Kim and Andrea compared

Figure 7.18 shows the measures' moving windows of correlations. It shows alternations of negative and positive correlations for both learners, though this is less the case for Andrea than for Kim. Kim and Andrea show relatively different patterns between the measures; Kim starts and ends with support between the measures and alternations in between. Andrea starts and ends with competition between the measures and mainly support in between.

We hypothesized that the average clause length in morphemes and the use of the past and perfect tense are competitive growers because we expected that less use of one would mean more complexity for the other. Kim's data indeed show a negative (but quite weak) correlation and this means very little evidence for the hypothesis. The hypothesis must be rejected for Andrea's data, showing a moderate positive correlation of both measures. The alternations between support and competition imply no actual influence of the measures on each other over time, though. Both learners keep on writing more morphemes per clause with and without using the past and perfect tense.

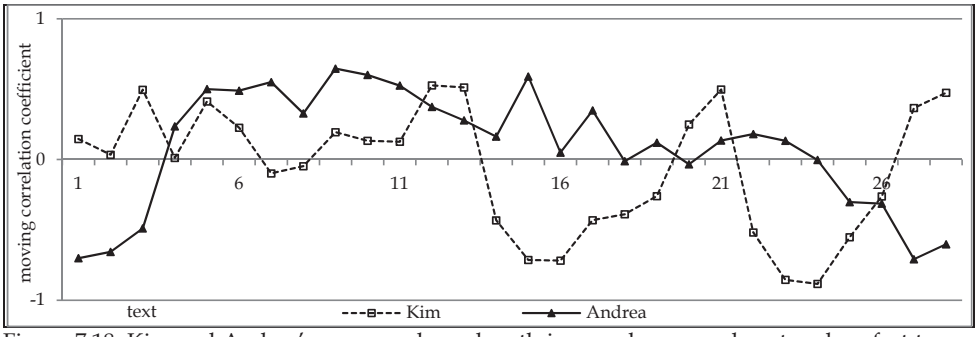

Figure 7.18: Kim and Andrea's average clause length in morphemes and past and perfect tense use and their interaction.

Now that we have explored the combination of the average clause length in morphemes and the complex tense use, we are also interested to find whether the latter does influence the average word length in morphemes or not, as the average clause and word length in morphemes showed large support (see 7.2).

7.5 Development of average word length in morphemes versus use of past and perfect tense

This section examines the development of average word length in morphemes in combination with the use of the past and perfect tense.

7.5.1 Interaction between average word length in morphemes and use of the past and perfect tense

The subsection starts with example sentences. A relevant point is that the number of morphemes per word and the use of more complex tenses are related. Then the development of the more complex tenses in the focal learners are explored separately and compared.

The example sentences show the average word length in morphemes in combination with the past and perfect tense. Example 20 illustrates a sentence in the past tense in plural (first person) and this combination is responsible for a third of the morphemes in this sentence. Example 21 shows the use of a perfect tense in singular (first person) and this combination is responsible for four morphemes (out of ten).

(20) Me **nä/i/mme** jopa[w.ord] poro/n metsä/ssä. *'We even saw a reindeer in the forest.'* (Andrea-t22s4/mp21-past3mrph-WlMrph5/9(1.8))
(21) Minä **ole/n ol/lut** usein suomalaise/ssa sauna/ssa. *'I have often been in a Finnish sauna.'* (Kim-t18s1/mp16-perf2/2mrph-WlMrph6/10(1.7))

Kim

Figure 7.19 illustrates the average word length in morphemes and the past and perfect tense in Kim's writing and shows increase for both measures. However, the past and perfect tense use grows considerably faster than the word length in morphemes. Though both patterns in Kim's data illustrate a variable pattern, the complex tense use shows more variability.

Overall, the measure points show competition and support as well (mp1-9 are put aside, so: mp10-37), the latter also found in several small periods (mp12-14, 21-23, 28-31 and 35-37). Kim's correlation coefficient (raw data) show a weak positive correlation (R=.17) between the measures, while the correlation coefficient of the residual data show a very weak negative correlation (R=-.08).

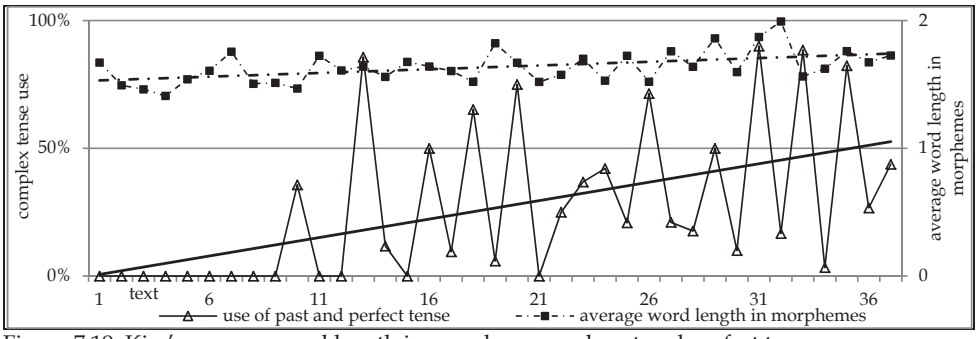

Figure 7.19: Kim's average word length in morphemes and past and perfect tense use.

Andrea

Figure 7.20 illustrates the average word length in morphemes and the complex tenses in Andrea's writing and both measures show a rather equal increase. The data show considerable differences in variability, though.

Overall, next to some measure points that show a low degree of competition (mp1-3 are put aside, so: mp4-32) some small periods of supportive correlation are shown as well (mp12-14, 18-19 and 28-29). Andrea's correlation coefficients of both the raw and the residual data illustrate moderately strong positive correlations (R=.32; R=.31) between the measures.

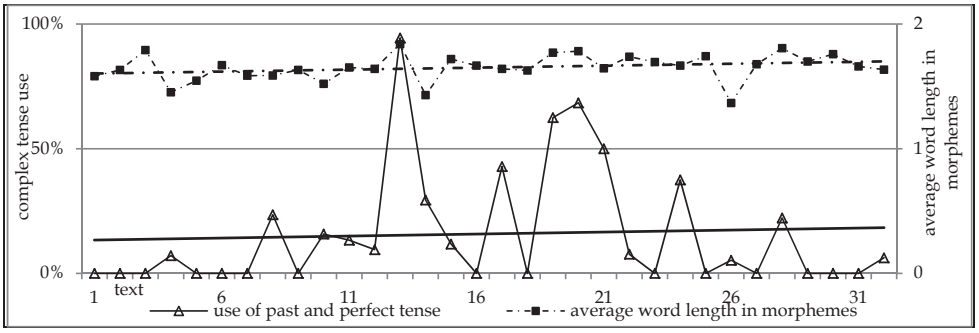

Figure 7.20: Andrea's average word length in morphemes and past and perfect tense use.

Kim and Andrea compared

Figure 7.21 shows a moving window of correlations of the average word length in morphemes and the complex tenses. It illustrates several alternations in support and competition for both participants.

We hypothesized that the measures are supportive growers, because using the past or perfect tense means more morphemes. This is true for the correlation coefficient of Andrea, whose residual data are strongly positive correlated but not for Kim, whose residual data show a negative weak correlation (almost zero). However, again the alternations between support and competition indicate that the measures do not actually influence each other

consistently over time. Both learners keep on writing more morphemes per word with and without using the past and perfect tense.

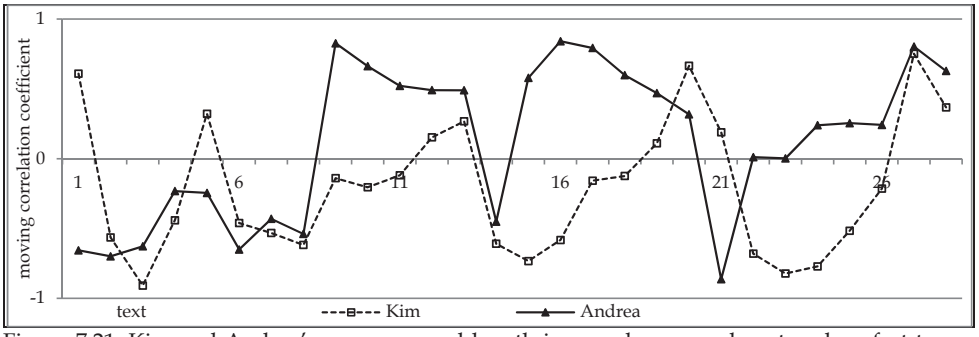

Figure 7.21: Kim and Andrea's average word length in morphemes and past and perfect tense use and their interaction.

In this subsection we found that there is not a one-on-one relation between the use of complex tenses and average word length in morphemes, but no systematic interaction was found either. We will now discuss all findings.

7.6 Discussion

This chapter traced the development of average word length in morphemes at both the group and individual level. Then it explored average clause length in morphemes and past and perfect tense in the two focal learners.

The focal learners were strikingly similar to their respective groups in the development of the average word length in morphemes. Both the FL group and Kim showed rather a large increase, though Kim relatively a bit less. Both the L2 group and Andrea showed very little increase, though for Andrea this was relatively a bit more. Moreover, Andrea started using words with fewer morphemes than the L2 group while at the end she used words with more morphemes.

At the beginning Kim used fewer complex words than Andrea, but at the end she used more. However, at the end there were no significant differences in either the holistic scores or on the quantitative scores in this area.

The interactions between the average clause length versus the average word length in morphemes in the moving windows of correlation showed a strikingly similar pattern of mainly support and only a few isolated points of competition for both learners. For both learners the correlation coefficient was strongly positive; they both wrote shorter clauses and words at the beginning and longer clauses and words (in terms of morphemes) at the end.

The next question was whether the use of tenses could be considered a good developmental measure, but before doing so we tested for task effects on the use of tense and no clear relation was found. The only clear similarities

between the focal learners were the use of mainly the present tense in the first part of the period, which is in line with Verspoor et al. (2012), who noted that beginners use the present tense mainly. The fact that the learners did not show similarities in their tense use seems to demonstrate that overall the past and perfect tense use is not determined exclusively by task. The current study shows increase of complexity in the use of the past and perfect tense for both learners over the 10 month study, which is in line with Verspoor *et al*. (2012), who found tense a strong discriminator across proficiency levels. Therefore, the tense measure seems to be suitable to use as a developmental complexity measure.

In their use of tenses, Kim and Andrea were quite different. At the end, Kim showed a rather balanced mix of tenses, which can be considered as a mark of development, but Andrea did not, which can be seen as a sign of relatively little development. In other words, here we might actually see an effect of context and/or instruction. Apparently, Kim was challenged to use a larger range of tenses and Andrea was not. Whether this is due to explicit instruction and more awareness (focus on form versus on meaning) or on types of exposure and relative frequency in the input we do not know. It might be that in everyday spoken Finnish the present tense is so much more frequent than the other tenses that Andrea did not feel the need to use them. Another related possibility was that her communicative classes and activities did not provide opportunities to use the more complex tenses. Moreover, as the holistic scores on their last writings (chapter 4) did not show any clear differences in this respect, we have to be very cautious in attributing this difference to "more development" and/or "context".

With regard to the development of the average clause length in morphemes versus the use of the past and perfect tense, the moving windows of correlation patterns of the two learners showed quite some differences. Moreover, Kim's correlation coefficient was very weakly negative, whereas Andrea's correlation coefficient was weakly positive. However, support and competition alternated quite a lot which means that there seems to be no strong relation between the measures over time. Both Kim and Andrea kept on writing more morphemes per clause, regardless of using the past or perfect tense.

With regard to the development of the average word length in morphemes versus the use of the past and perfect tense, again the moving windows of correlation patterns of the two learners showed quite some differences. Kim's correlation coefficient was very weakly negative, whereas Andrea's correlation coefficient was moderately strong positive. Also for these measures the alternations in support and competition indicate no strong relation between them over time. Both learners kept on writing more morphemes per word with and without using the past and perfect tense.

8 ACCURACY: USE AND FORM OF THE FINNISH CASES

This chapter further explores the development of accuracy in both use and form of the Finnish cases at group level. The group analysis in chapter 4 had shown that there were significant differences in accuracy; the L2 group clearly made more case use and form errors. As in chapters 6 and 7, the question is whether we can attribute these differences at group level to condition; the FL group received a great deal of explicit instruction on the Finnish case system and the L2 group received hardly any. Because of the different L1 backgrounds in the L2 group, we need to compare the two focal learners from each group with a similar L1 background to their respective groups to see how representative they are for the group and then to each other (Verspoor et al., 2004; Murakami, 2013). Another objective is to explore the dynamic development of cases by examining the trajectories (including peaks and dips) and interactions of the different accuracy measures in the two focal learners.

8.1 Case use and form errors in absolute numbers

The section on absolute numbers in errors serves to obtain a clear picture of the data behind the ratio of correct versus incorrect use and form errors (Case Accuracy Ratio, CAR), which will be discussed extensively below. In the current study we have distinguished between use (appropriate use in the context) and form (appropriate grammatical form).

Figures 8.1 and 8.2 provide an overview of the raw numbers of the incorrectly used cases per group and per participant. All learners show incorrect use of the cases during the whole period of writing, and as expected from a dynamic perspective, they all show different patterns.

For the average use of cases in absolute numbers, the L2 group showed a higher error rate (M=9.8, SE=3.2) than the FL group (M=6.3, SE=1.9). This difference was significant ($t(10)=-2.9$; $p<.05$).

The figures show a few differences. Even though all individuals in both groups make quite some case use errors throughout the whole period, they remain below the midline of 16 case use errors per text most of the time, especially the FL group members. Only once a FL group member (Annet) shows more than 16 case use errors, whereas this happens 13 times in the L2 group. Looking at the two focal learners, we see that Kim has the most case use errors at measure point 10 (8.8) and Andrea at measure point 12 (25.8), which is also the second highest number of case use errors in the L2 group.

Finally, when we compare the general group trends we see that the FL group average starts low, has a large increase noticeable in the bulb, and then a decline and stabilization. The L2 group also starts low, after which it shows an increase without any clear bulb, and a rather stable pattern until the end with a sudden decrease. Overall the trend line of the L2 group is higher than the one of the FL group, as would be expected because of the statistical differences found in chapter 4.

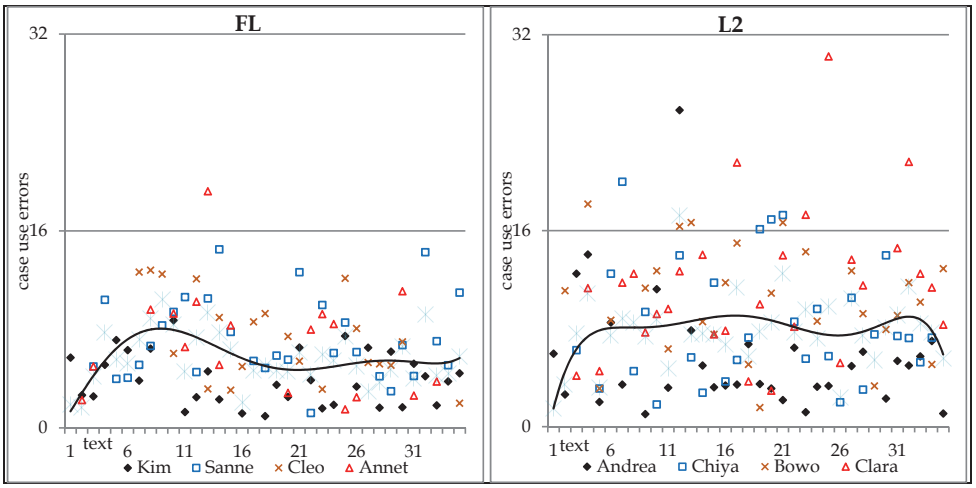

Figure 8.1 and Figure 8.2: FL and L2 participants' case use errors.

Now we turn to form errors. Figures 8.3 and 8.4 compare the number of case form errors per group and per participant. A first glance shows different patterns. Even though both groups show case form errors throughout the whole period and both groups transcend the halfway line of four case form errors per text, the L2 group does so more often: seven times for the FL group and 15 times for the L2 group. In the FL group, Sanne does so most often (three times), and in L2 Clara (five times). In the FL group Cleo (mp21; 5.4) makes most of the case form errors in one text and in the L2 group Andrea (mp13; 7.9). Striking is that Kim does so as well at measure point 13 (4.6). Not all learners have case form errors in all writings: Sanne does not make a case form errors before measure point 6 while Annet and Cleo do not make case form errors after point 24 and 27 respectively. Again, this implies that all learners show variable patterns within the data and compared to the patterns of the other participants.

Finally, the FL group trend starts high after which it decreases quite suddenly and subsequently increases slightly again. This pattern is repeated once more with a small increase at the end. The L2 group trend takes off from a lower point and increases slightly, after which it shows decrease and goes down to a lower level, increases slightly again before it considerably decreases at the end. Overall the trend lines of the groups are located in the same area.

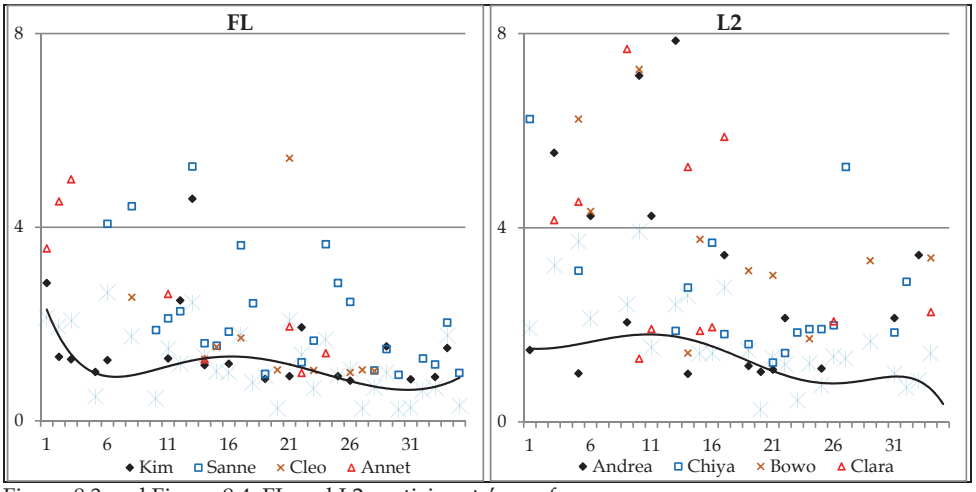

Figure 8.3 and Figure 8.4: FL and L2 participants' case form errors.

The groups do not seem to be very different at the end as far as form errors are concerned, but the absolute numbers give information about the total number of errors made, but not in relation to the number of case used. Therefore, the following section will look at the accuracy ratios of these measures, first taken together and later on separated in use and form.

8.2 CAR use and form errors

The absolute numbers above were given to get an idea of case errors made in the texts; however, we want to know how many of the used cases contain case errors. Therefore, the measure CAR was developed to calculate the ratio of correct versus incorrect uses. After giving example sentences, this section analyzes the CAR use and form errors by the two groups and by the two individuals. As in previous chapters, the bolded parts in examples sentence indicate the issue in question.

All example sentences below (1 to 8) show case use and form errors made by all eight individuals of the two groups, taken from those texts with the highest number of use and form errors per learner (correct cases in italics and incorrect cases in bold).

FL:

(1) *Se* on *ikävää*, mutta *minä* luen nyt hyvin paljon kirjallisuu**ta**[form]. *'It is boring, but I read a lot of literature now.'* (Kim-t13s9/mp12)
(2) *Hän* on auttanut *minua* hyvin ja *minä* haluan kiittää hänet**t**[gov] kai**kk**esta[cons.gr]. *'He helped me a lot and I want to thank him for everything.'* (Sanne-t32s9/mp32)
(3) *Minulla* ei ole lempiruok**a**[subj], *minä* pidän paljon ruoka**a**[gov][lex]. *'I do not have favorite food, I like a lot of food.'* (Cleo-t8s4/mp2)
(4) *Minusta tämän pienen suomalaisen* run**o**[d.obj] ei ole vaikea**ta**[pr.nom] ymmärtää. *'In my opinion this little Finnish poem is not hard to understand.'* (Annet-t13s3/mp8)

L2:

(5) Kun *minä* en askarrele[cons.gr] *minä* tykkään löytää[lex] kirja**t**[d.obj] lahja**na**[infl]. *'When I am not making handcrafts I like to look for books as a present.'* (Andrea-t12s7/mp11)
(6) Sitten[lex], *joka päivä minä* menen[w.ord] yliopisto**on**[infl] *polkupyörällä. 'Then, every day I go to the university by bike.'* (Chiya-t7s3/mp7)
(7) Kun *sää* on *kylmä*, *minä* usein[w.ord] käytä[form] pip**o**[d.obj] ja *talvitakkia. 'When the weather is cold I often wear a head and a winter jacket.'* (Bowo-t10s4/mp8)
(8) Brasilialai**nen**[congr] kirjallisuu**si**lla[form] on paljon portugali**a**lai**nen**[congr][form] *vaikutteita*, mutta jälkeen alu**ssa**[infl] 1700 luvu**lla**[infl] [w.ord], *Brasilia* aloitti luoda[gov] *omaa kirjallisuutta. 'In the Brazilian literature is a lot of Portuguese influence, but after the beginning of the 18th century, Brazil started to create own literature.'* (Clara-t17s5/mp17)

The CAR use and form errors seemed to be different between the two groups (chapter 4) and will be further explored for both the groups and focal learners.

8.2.1 Differences between FL and L2

To calculate the ratio of use and form errors, the total of all case errors (use and form) per text of 100 words was divided by the total number of cases used per text. On average, the L2 group showed a higher frequency of CAR use and form errors (M=.18, SE=.05) than the FL group (M=.13, SE=.05). This difference was significant ($t(34)=-4.97$; $p<.05$).

Figures 8.5 and 8.6 show the CAR use and form errors per individual in each group. The figures look rather different even though there is variation in both groups and variability in each learner. The FL group shows fewer case errors than the L2 group throughout, but this does not hold for every individual. In the FL group the participants show a rather similar pattern with a small decrease in case errors, with the exception of Sanne who shows some striking peaks (mp32; .28 and 35; .19). In the L2 group Andrea makes a lot of case errors in the beginning (though alternating with texts with few errors) after which the number of case errors stabilizes somewhat. Chiya and Bowo both show a rather stable pattern of case errors, which decreases little during the period. Clara shows such a rather stable pattern as well, but also shows some

striking peaks in the middle of the period. In the FL group Annet (mp13; .29) has the highest CAR use and form errors and in the L2 group Clara (mp17; .52) does.

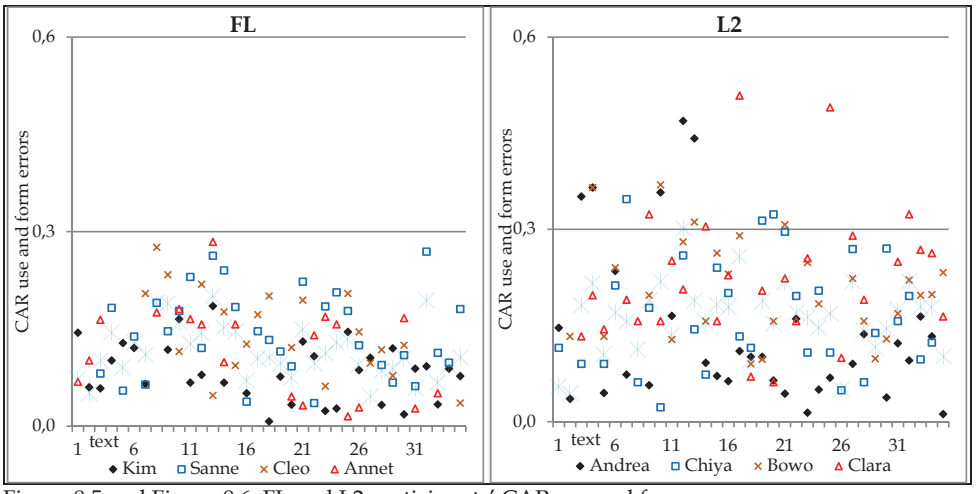

Figure 8.5 and Figure 8.6: FL and L2 participants' CAR use and form errors.

8.2.2 FL, L2, Kim and Andrea compared

Figures 8.7 and 8.8 illustrate the overall averages in CAR use and form errors by the FL and L2 groups and by Kim and Andrea individually. The considerable difference between the figures is striking. Also the differences in patterns within each figure are striking: in 8.7, the increasing trend line for the L2 group, showing more case use and form errors during the time span (while the errors in the FL group diminish) and in 8.8, the huge number of errors made by Andrea in the first half of the period (compared to Kim).

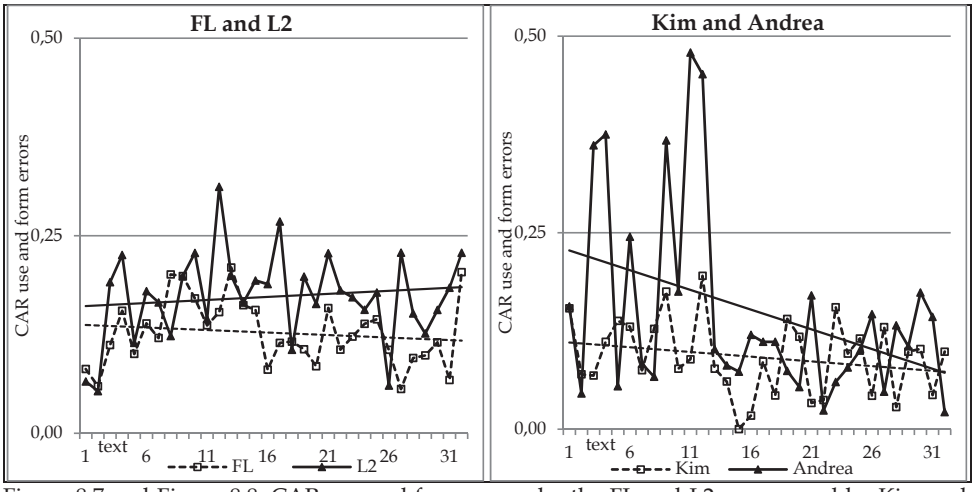

Figure 8.7 and Figure 8.8: CAR use and form errors by the FL and L2 groups and by Kim and Andrea.

In addition, figures 8.9 and 8.10 show the CAR use and form errors by the FL group and Kim and by the L2 group and Andrea. The FL group's trend line decreases slightly and so does Kim's, even though her trend line is located at a lower level. This means that the other FL learners, taken together, make a little more use and form errors during the whole period, but the decrease of errors takes place in the same pattern. Both the patterns of the FL group and Kim are located at the same level, with the brief exception of the end of the first part.

The L2 group's trend line increases slightly, whereas Andrea's decreases considerably; she starts much higher and ends much lower than the group. The fact that Andrea shows a striking number of use and form errors in the cases in the first part of the period means that the other L2 learners, taken together, are increasingly making much more use and form errors in the cases during the period. The trend lines of both participants end at the same level. Overall, Kim and Andrea show a lower pattern at the end than the overall averages of their groups.

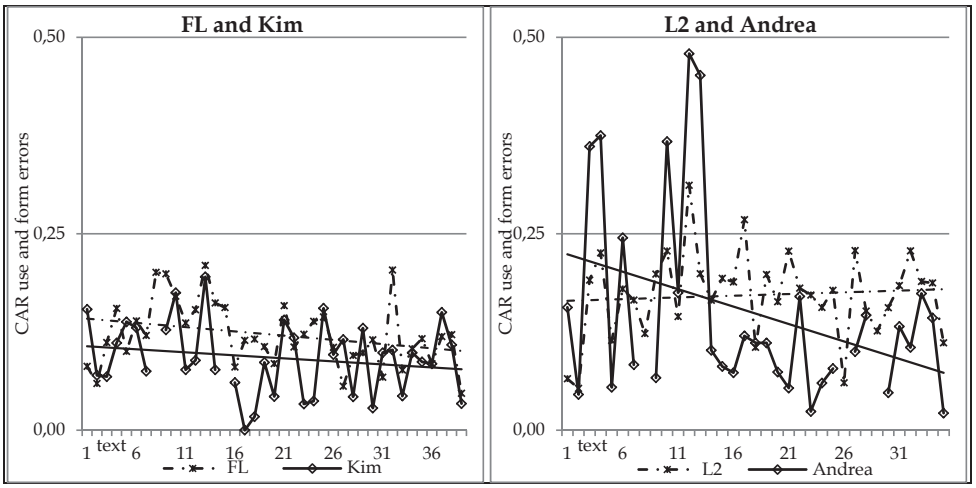

Figure 8.9 and Figure 8.10: CAR use and form errors by the FL group and Kim and by the L2 group and Andrea.

To summarize, there are significant differences in total CAR use and form errors (chapter 4), but if we look at development over time in the figures, we see that these differences could be mainly due to their differences in developmental paths, which is especially evident in the focal learners. The L2 learner makes many more errors early on, but shows a sharp decline after data point 16 and then she becomes very similar to the FL learner at the end. However, the data did not show more variability for Kim than for Andrea (Monte Carlo).

In this section the CAR use and form errors have been taken together. In the next section we will explore the two accuracy measures separately.

8.3 Development of CAR use errors versus CAR form errors

This section explores further the development of accuracy by looking at the CAR use and form errors separately in the two focal learners. The question is whether these types of errors have a similar development and whether the two learners develop in a similar manner.

Kim

We start off with use errors. Figure 8.11 shows a linear trend line for the CAR use errors which stays at about the same level throughout the period, while at the same time the min-max shows a considerable degree of variability, ranging between .000 (mp15) and .175 (mp9).

The CAR use errors show mainly a variable pattern, with some small rather stable periods. Twice the difference in errors in subsequent measure points is less than .002 (mp2-3 and 33-34). The first nine measure points show a rather variable pattern, after which point 10 shows a remarkable dip. From the next point on, the pattern becomes more stable again (mp11-18). After this period the pattern shows a series of considerable variability (mp19-29, with two exceptions in mp20 and 22). From measure point 30 to 34 the data show a brief rather stable period and the pattern ends with variability (mp35-37).

In total 18 measure points show a lower average use than the mean value of the measure (.076 for Kim), 11 measure points of which are found in the first part of the period. Lastly, the graph shows a low variable pattern of succession of narrowness and wideness; the bandwidth widens after measure point 6 to become smaller again at measure point 12. From that point on, the wideness of the bandwidth stays about the same until measure point 31 in which it narrows for a short while to widen again in the last part of the graph.

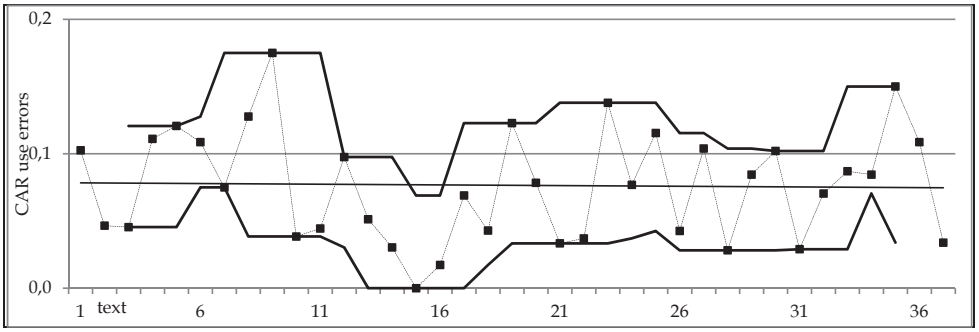

Figure 8.11: Kim's CAR use errors.

Figure 8.12 illustrates Kim's CAR form errors and shows a steady decrease to almost zero at measure point 37, but with variability and a peak at measure point 12 (.098), while there are 17 measure points without errors in form.

The CAR form errors illustrate a rather variable pattern in which measure point 12 shows the most striking peak. Twice the difference in form errors in subsequent measure points is less than .002 (mp2-3 and 23-24; the points of no use not taken into account). Two brief periods (three measure points or more) of no incorrect used cases are shown (mp7-9 and 33-37). The first part shows the most variability .The data show a lower incorrect form of the cases than the mean value of the measure (.015 for Kim) in 20 measure points (seven in the first part of the period, 13 in the second part). To conclude, the pattern of narrowness and wideness is variable, but with the exception of points 11-13, the min-line remains at zero. The bandwidth starts quite wide and narrows after measure point 4, to widen again (mp8-14). After this, the pattern of narrowing and widening again takes place twice (in which the max-line diminishes in height) before the moving min-max graph ends at zero.

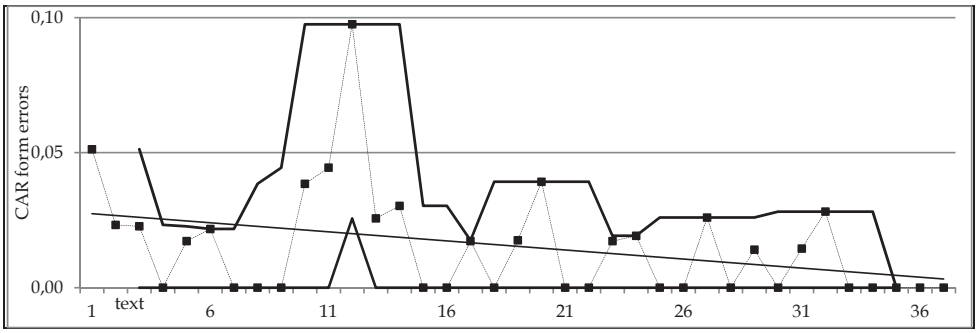

Figure 8.12: Kim's CAR form errors.

Because the CAR form errors show an isolated peak in Kim's data, it was tested for significance (Monte Carlo analysis). The analysis showed a rather strong trend (p=.0710), suggesting that this is a developmental peak. The data of the other FL learners did not show developmental peaks.

Andrea

Figure 8.13 shows the CAR use errors made by Andrea and shows a decrease, but with striking variability. The highest number of incorrect used cases is found at measure point 11 (.479).

The first 13 measure points show substantial variability, with one remarkable peak (mp12). The second part of the period shows a more stable pattern, starting with measure point 14 to the end of the period. The latter period shows a series of patterns with little variability, to end with two measure points which show a large difference in frequency (mp31 (.144) and 32 (.022)).

In total 22 measure points show a lower average use than the mean value of the measure (.116 for Andrea): nine in the first part of the period, 13 in the second part. Finally, the first part of the graph illustrates a pattern of succession of narrowness and wideness; the rather wide bandwidth in the first eight points widens to an enormous large extent in seven measure points (mp 9-15). It

becomes rather small for a short period (mp15-17) before the band width widens again until the end.

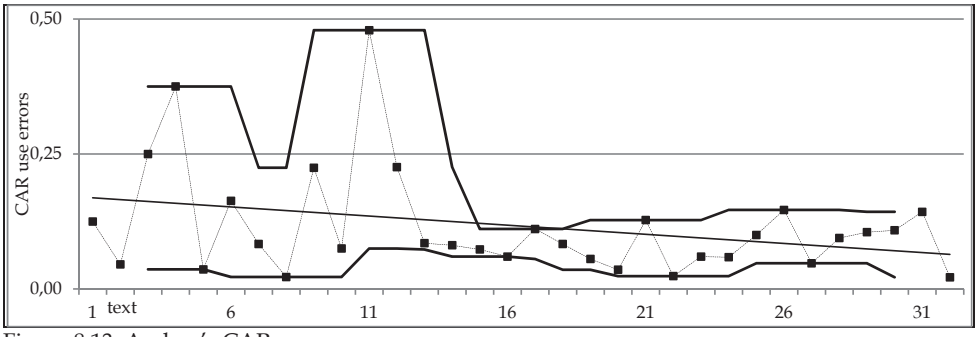

Figure 8.13: Andrea's CAR use errors.

Figure 8.14 illustrates the incorrectly formed cases by Andrea and shows a considerable decrease. The data show that in total 15 measure points do not show any CAR form errors.

In the first part of the period the case form errors show a considerable variable pattern, with measure point 12 as most striking peak, whereas in the second part some small periods of relative stability are shown, i.e. the difference in form errors in subsequent measure points is less than .002 (mp19-20; the points of no use not taken into account). The data show one short period of subsequent measure points with no incorrect applied case forms (mp25-27). The first part of the period shows on the one hand more variability in frequency but on the other hand also more small periods of relative stability than the second part.

In 23 measure points the data show a less frequent appearance of case form errors than the mean value of the measure (.033 for Andrea) (nine in the first part of the period, 14 in the second part). Lastly, the patterns of narrowness and wideness show quite some variability but the total min-line remains at zero. The bandwidth starts quite wide after which it narrows for a very short while (mp6). Subsequently it widens to a larger (mp7-9) and even larger extent (mp10-14). After this the pattern of narrowing sets in, which becomes a pattern of widening again from point 24 on.

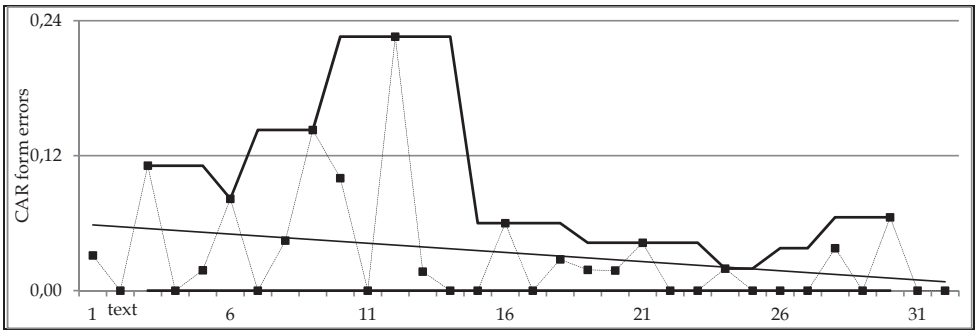

Figure 8.14: Andrea's CAR form errors.

Because the CAR use errors and the CAR form errors show isolated peaks in Andrea's data, they were tested for significance (Monte Carlo analysis). Neither of the analyses showed significance; the peaks are not necessarily indicators for growth, but the first 16 data points are more variable than the last ones, so we may still consider this a shift in DST terms, though we did not find more variability for the first 16 data points than for the second part Andrea's CAR form errors, nor for her CAR use errors.

8.3.1 Interaction between CAR use and CAR form errors

In this subsection we investigate the combination of the two measures CAR use and CAR form errors. After discussing Kim and Andrea separately, we compare their development. The subsection takes off with example sentences.

Examples 9 to 12 show sentences that contain both errors in the use and form of the cases. The case use errors in the example sentences consist of incorrectly used cases for the direct object, inflectional errors, case use errors in the application of congruence and case use errors in the grammatical subject, whereas the case form errors consist of consonant gradation and stem form errors.

(9) …me emme syö **ne**[d.obj] ja me ostamme paljon vihanneks**etia**[form]…
'*…we eat them and we buy a lot of vegetables* ' (Kim-t3s4/mp3)
(10) ...käyttää vaatteita, joi**ta**[congr] sopivat[w.miss] ihmisen ulkonä**k**ö**lle**
[cons.gr][infl], luon**nelle**[form[33]][infl] ja tun**nelle**[form[34]][infl]. '*…to use clothes which fit to the appearance, to the environment and to the feeling.*'
(Andrea-t10s5/mp9)
(11) Joskus muoti antaa uu**tta**[d.obj] idea**ta**[form][d.obj]... '*Sometimes fashion gives new ideas…*' (Andrea-t10s8/mp9)
(12) Runo**ja**[subj] ovat 1900-lu**k**usta[cons.gr]. '*The poems are from the twentieth century.*' (Andrea-t13s3/mp12)

Kim

Figure 8.15 shows Kim's CAR use and CAR form errors and illustrates little different patterns; both measures decrease but the form errors decrease more. The data show rather large variability for both the CAR use and the CAR form errors. In some small periods their patterns seem to look like each other while at other points the patterns show quite some differences.

Overall, the graph shows support and competition between the measures. Nine times the CAR form errors stay at zero while the CAR use errors alter. Kim's correlation coefficients indicates an overall correlation of

---

33 Though the error contains a consonant gradation error as well, it is regarded as form error; the form should be luon**tee**lle.
34 Though the error contains a consonant gradation error as well, it is regarded as form error; the form should be tun**tee**lle.

almost zero (very weak positive) between the CAR use and CAR form errors for both the raw and the residual data (R=.01 and R=.00).

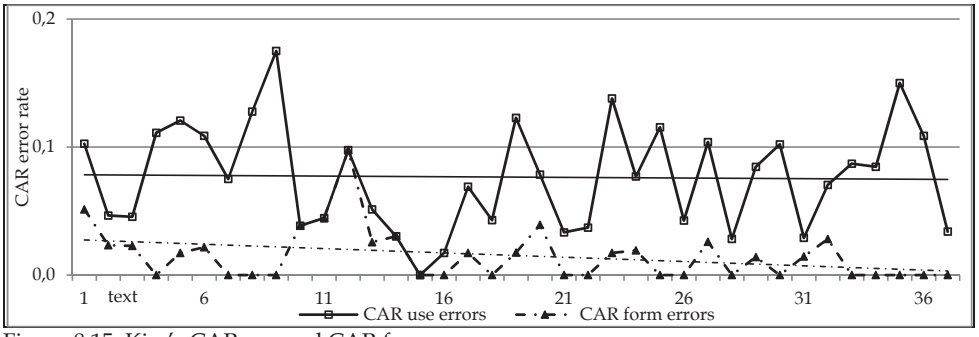

Figure 8.15: Kim's CAR use and CAR form errors.

Andrea

Figure 8.16 shows Andrea's CAR use and CAR form errors and they both illustrate a rather similar decreasing pattern. The data show quite some variability and the patterns look remarkable equivalent at some points.

Overall, the graph shows support and competition between the measures. Five times the CAR form errors stay at zero while the CAR use errors alter. Andrea's correlation coefficient (raw data) showed a moderately strong positive correlation with the (R=.25), while the residual data showed a weak positive correlation (R=.17).

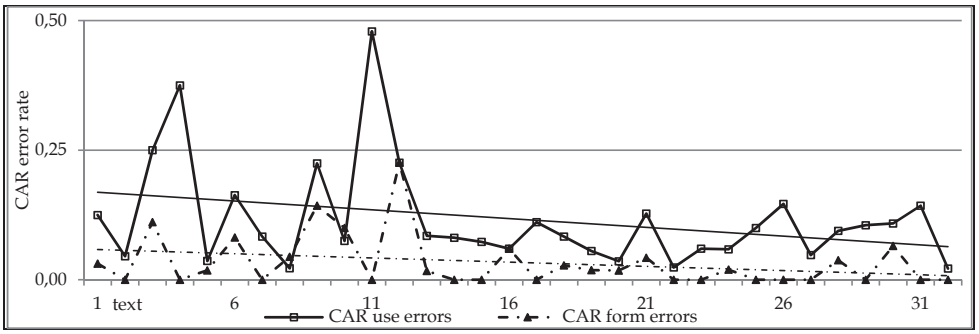

Figure 8.16: Andrea's CAR use and CAR form errors.

Kim and Andrea compared

We looked at a moving correlation between the two measures to see if there are differences between Kim and Andrea in the way the two measures interact.

Figure 8.17 shows a moving window of correlations and shows that positive correlations of the CAR use and CAR form errors prevail over negative correlations for both learners. The patterns in the moving window of correlation differ; Kim shows support in the second part of the period, indicating that the use and form errors decline simultaneously, whereas Andrea shows alternation

between support and competition, indicating that there is no obvious connection between the use and form errors.

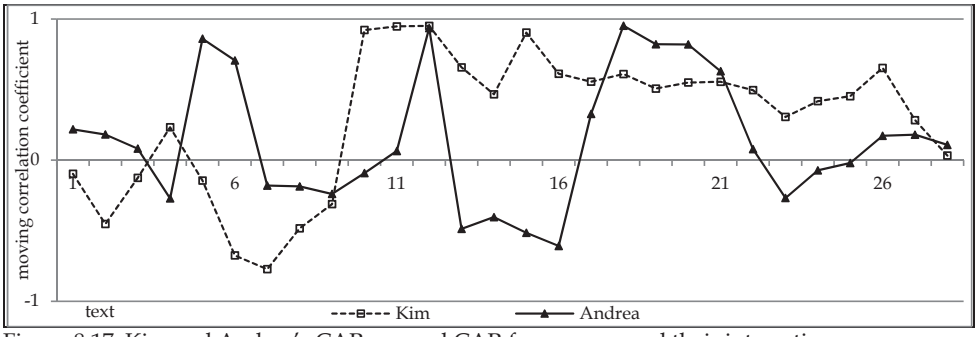

Figure 8.17: Kim and Andrea's CAR use and CAR form errors and their interaction.

In this section we examined the development of the CAR use errors and CAR form errors and of the two learners concerning these measures. Now we will discuss the findings.

8.4 Discussion

The expectation had been that of all measures in the study, especially the CAR use and form errors would show significant differences between the FL and the L2 groups because of instruction: focus on grammar in the FL group and focus on meaning in the L2 group. Indeed the FL group showed fewer case use and form errors (CAR) than the L2 group. However, when we compared these errors in the four texts of the two focal learners with a similar L1 background (quantitative scores, see chapter 4), this clear difference was not there at all.

In case use errors, the absolute numbers also showed a significant difference in statistical terms (chapter 4) and the graphs also clearly showed a higher number of case use errors in the L2 group. The developmental trajectories also look different: In the FL group, the group trend line showed a bulb, suggesting U-shaped behavior (overuse of errors), and a subtle decline in case use errors; this was not the case for the L2 group.

In case form errors, the absolute numbers showed no statistical difference between the groups (chapter 4) and indeed the group trend lines were rather similar.

Compared to their respective groups, the two focal learners showed fewer CAR use and form errors, with the exception of the first part in Andrea's data. Both the FL group and Kim showed a decrease, though Kim made fewer case errors throughout. Andrea also showed a large decrease, whereas the L2 group as a whole made more errors in the use and form of the cases over time. It seems that Andrea, with her L1 background and her interest in grammar, is aware of cases and attempts to use them, but with a great number of errors

early on. This is in line with research from a dynamic perspective in which trial and error is needed to progress.

The isolated peak in Kim's data concerning the CAR form errors was significant (p=.0710), but for Andrea it was not. However, also Andrea's data clearly showed a shift with a sudden decrease in form errors. As far as CAR use errors are concerned, Andrea's data also showed a rather sudden drop, whereas Kim's data did not. Also Andrea's correlation between the two measures was stronger than for Kim. This would suggest that for Andrea the form and use of cases are integrated and learned as one form-meaning mapping, whereas for Kim the form and use of cases are less integrated. She can reason out the forms, but not the uses, which are more subtly meaning based.

The moving windows of correlation patterns showed a strikingly different pattern of peaks and dips for the two learners. Kim showed support in the second part of the period, indicating that the use and form errors in cases declined simultaneously, whereas Andrea showed alternation between support and competition, indicating that there was no obvious connection between the use and form errors in cases. Both Kim's and Andrea's (residual) data showed a positive correlation coefficient, though for Kim this was minimal.

Beforehand we hypothesized that the CAR use and CAR form errors are supportive growers, because we expected that paying attention to the phenomenon of many cases would for both learners be a new experience. Therefore, paying a lot of attention to the right use of a case was thought to imply attention to the correct form of the case and that little attention to the use of a case would also imply little attention to the form of the case. Indeed both Kim's and Andrea's data showed a positive correlation coefficient, though for Kim this was a minimum. However, even though Kim and Andrea's correlation patterns mainly showed differences, they represent the inter and intra variability, as expected in a dynamic way of thinking.

To summarize, the outcomes suggest that there might be an effect of different types of instruction (Verspoor et al., 2004). Although the group outcomes showed significant differences, the analyses showed that in the comparison of the individuals with similar L1's, Kim did make fewer case errors than Andrea at the beginning, but at the end they made about the same number of case errors. It seems that the focus on grammar and the focus on meaning have different effects especially at the beginning of the learning process: the focus on grammar approach leading to fewer case errors early on.

9 ACCURACY AND COMPLEXITY

This chapter aims to explore the relation between complexity and accuracy. For complexity, three measures will be used at different levels: the average clause length in morphemes, which was argued to be the best general syntactic complexity measure in chapter 5, the average word length in morphemes, a good measure of complexity at the word level, mostly because of case markings and the total use of past and perfect tenses, a good measure of verb phrase complexity. The development at each of these three levels will be will compared to the development in accuracy to see to what extent they may compete or support each other.

9.1 Interaction between average clause length in morphemes and CAR use errors

In this section we look at the average clause length in morphemes in combination with CAR use errors. First we will look at Kim and Andrea separately and then we will compare their development. The section starts with example sentences. In all examples in this chapter, the parts in bold in the Finnish sentences are the issues in question.

In the examples the focus is on the average clause length in morphemes together with CAR use errors. Example 1 shows competition between the measures. Both clauses in the sentence contain more morphemes than Kim's mean value of the measure (9.5). Moreover, in both clauses high averages of the clause lengths in morphemes go together with correct case use. Example 2 also shows two clauses. They both contain seven morphemes, which is lower than Andrea's mean value of the measure (9.1).The first one shows competition, the second one support. The first clause contains two case use errors; a low average clause length in morphemes and a large number of case use errors cause competition. The second clause shows correct case use; low average clause length in morphemes with correct case use causes support.

(1) Ja minä ole/n myös kuunnel/lut si/tä pari päivä/ä // ja kerta kerra/lla Värttinä tul/i kaunii/mma/ksi. *'And I also listened to this a couple of days, and every time Värttinä got better.'* (Kim-t25s5/mp23-ClMrph2/22(12-10:11.0))
(2) Ostaminen lahjo**/j/a**[infl][w.ord] voi olla hausk**a**[pr.nom] // jos minu/lla o/n hyvä idea. *'Buying presents can be fun if I have a good idea.'* (Andrea-t12s1/mp11-ClMrph2/14(7-7:7.0))

Kim

Figure 9.1 presents Kim's average clause length in morphemes and CAR use errors and shows a little decrease for the case use errors and an increase for the clause length in morphemes. The data of both measures show rather large variability. At some points the patterns seem to look quite similar but they show some differences as well, especially in frequency.

Overall, the graph shows support and competition between the measures. For both Kim's raw and residual data the correlation coefficients show very weak negative correlations between the measures (both R=-.08).

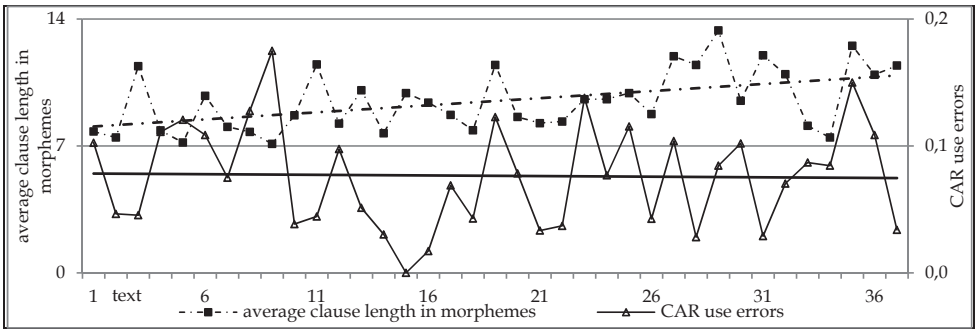

Figure 9.1: Kim's average clause length in morphemes and CAR use errors.

Andrea

Figure 9.2 shows Andrea's average clause length in morphemes and CAR use errors and shows considerable decrease for the use errors in cases. The average clause length in morphemes, on the other hand, shows increase. Both measures show large variability in the data. The patterns show some similarities but differences rule.

Overall, the graph shows support and competition between the measures. Andrea's correlation coefficient (raw data) shows a weak negative correlation between the measures (R=-.16), whereas the residual data shows a very weak negative correlation (R=-.02).

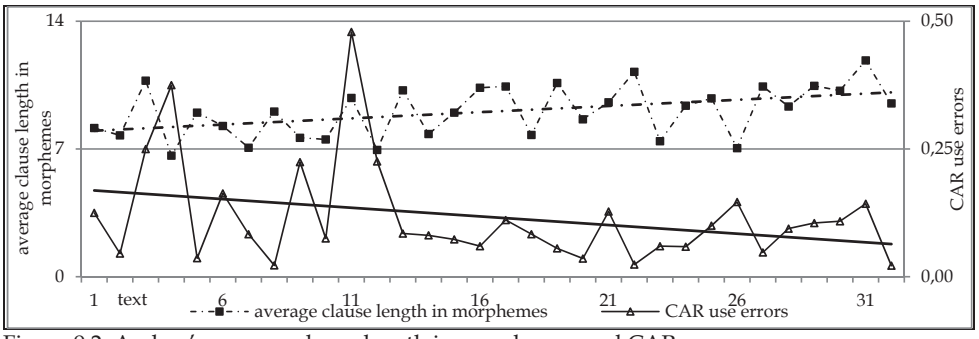

Figure 9.2: Andrea's average clause length in morphemes and CAR use errors.

Kim and Andrea compared

Figure 9.3 illustrates the moving windows of correlations of the measures. For both Kim and Andrea positive and negative correlations alternate. However, the patterns show considerable differences. In the first part of the data, Kim shows competition after which in the second part mainly support between the measures is seen. The latter means that for Kim, writing more complex clauses goes hand in hand with making more case use errors. It seems that Kim does not learn words and their case use simultaneously. Andrea shows alternations in the first part of the data but in the second part there is mainly competition between the measures. The latter means that for Andrea, writing more complex clauses goes hand in hand with making fewer errors. As argued in chapter 8, it seems that Andrea learns the words and their case use more holistically in context.

We had hypothesized that the syntactic complexity measure average clause length in morphemes and the syntactic accuracy measure CAR use errors would be supportive; we expected more case use errors when the average clause length in morphemes should increase and fewer case use errors when the average clause length in morphemes would decrease. Both Kim and Andrea's correlation coefficients (residual data) show that this is not true; the case use errors and the average clause length in morphemes do affect each other negatively, though hardly.

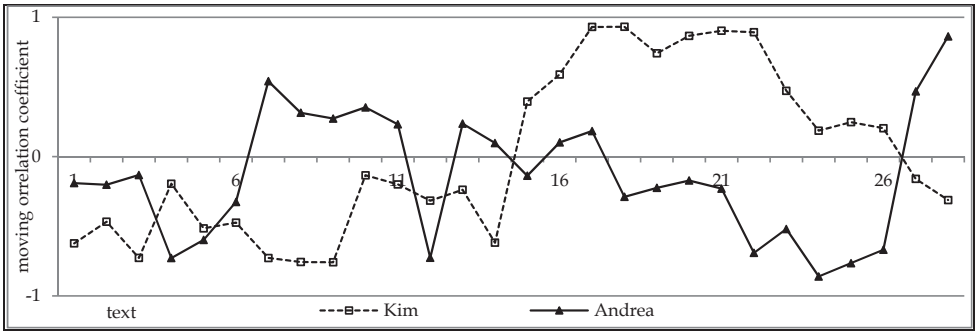

Figure 9.3: Kim and Andrea's average clause length in morphemes and CAR use errors and their interaction.

9.2 Interaction between average clause length in morphemes and CAR form errors

In this section, the average clause length in morphemes is looked at in combination with CAR form errors. Kim and Andrea are looked at separately after which their development is compared. The section starts with example sentences.

The examples show sentences in which both measures are found. The two examples in this section both serve as an example for the hypothesis that the measures would be supportive: the longer/shorter the average clause length in morphemes, the more/fewer CAR form errors. Example 3 shows one clause consisting of many morphemes and three case form errors. The task has a rather difficult topic (poetry) but Andrea seems to feel the urge to tell a lot about it. The topic is important to her and she does not seem to care a lot for the case markings. Example 4 illustrates two short clauses with correct case markings, consisting of fewer morphemes than Kim's mean value of the measure; it seems that she is focused on writing correct case forms.

(3) Romaan**i/**j/a[form][subj] kerto/vat tärke/i/tä asi**a/**i/ta[form], ehkä filosofis/i/a teem**a/i/t**a[form][lex]. *'Novels tell important things, maybe philosophical subjects.'* (Andrea-t13s5/mp12-ClMrph1/18)

(4) Hän autta/a Jarkko/a// vaikka hän e/i tarvitse apu/a. *'She helps Jarkko, though he does not need help.'* (Kim-t36s5/mp34-ClMrph2/13(6-7:6.5))

Kim

Figure 9.4 shows Kim's average clause length in morphemes and CAR form errors and shows considerable decrease for the CAR form errors, while the average clause length in morphemes increases. Both measures show quite large variability and seem to show mainly differences as well.

Overall, the graph shows support and also competition between the two measures. Once, the CAR form errors grow while the average clause length in morphemes remains the same (mp23-24) and nine times the CAR form errors remain at zero while the average clause length in morphemes alters. Kim's correlation coefficient (raw data) shows a very weak negative correlation between the measures (R=-.03), whereas the residual data show a weak positive correlation (R=.18).

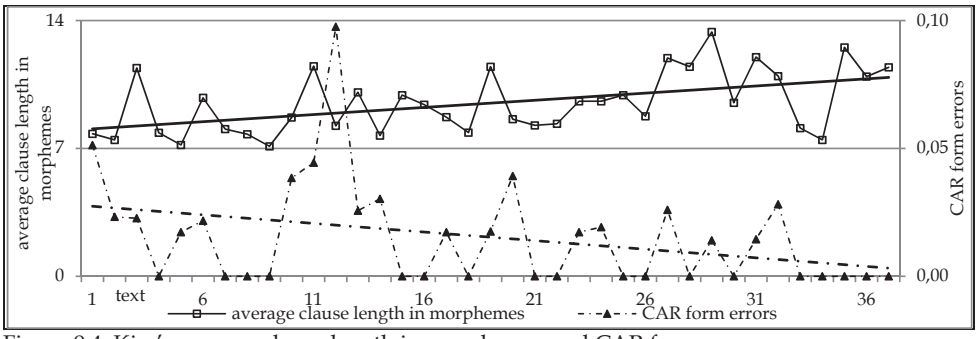

Figure 9.4: Kim's average clause length in morphemes and CAR form errors.

Andrea

Figure 9.5 shows Andrea's average clause length in morphemes and CAR form errors and shows decrease for the case form errors and growth for the average clause length. The data show large variability for both measures and the patterns show hardly any similarities.

Overall, the graph shows support and competition between the measures. Five times the CAR form errors remain at zero while the average clause length in morphemes changes. Andrea's correlation coefficient (raw data) shows a moderately strong negative correlation between the measures (R=-.24). Her residual data illustrate a weak negative correlation (R=-.12).

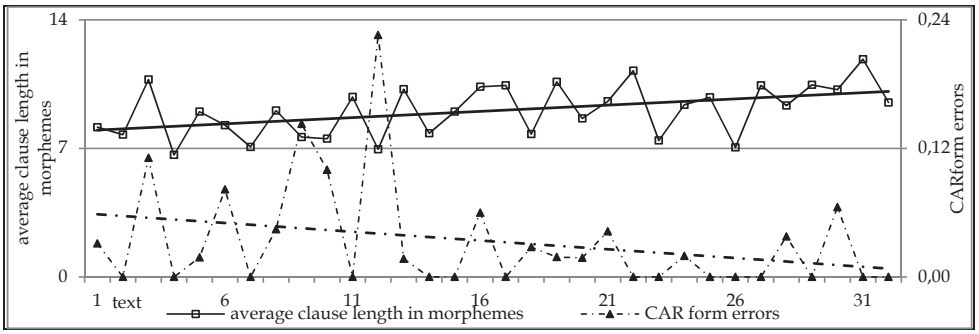

Figure 9.5: Andrea's average clause length in morphemes and CAR form errors.

Kim and Andrea compared

Figure 9.6 shows the interactions between both measures and illustrates that for both learners positive and negative correlations alternate. However, Kim and Andrea differ in their accuracy development. When Kim writes more complex clauses, this does not lead to more accuracy at the same time, while Andrea does become more accurate when she writes more complex clauses.

We had hypothesized that clause length in morphemes and CAR form errors are supportive growers; we expected more case form errors when the complexity of the clause length should grow (in morphemes) and fewer case

form errors when the complexity of the clause length should decline. For Kim's correlation coefficient this is true, for Andrea's not.

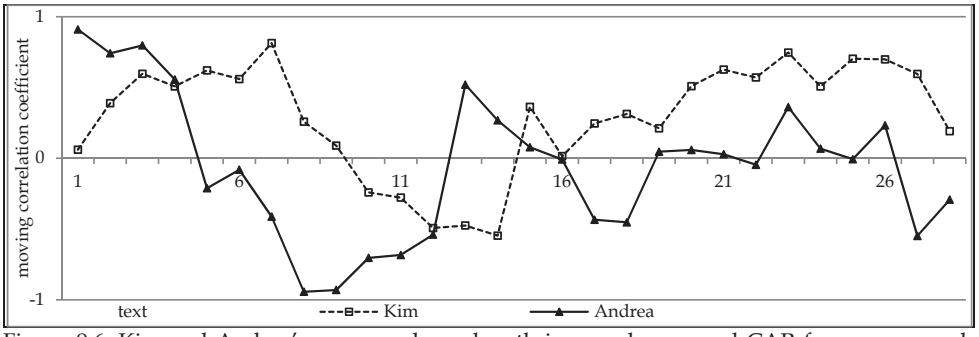

Figure 9.6: Kim and Andrea's average clause length in morphemes and CAR form errors and their interaction.

9.3 Interaction between average word length in morphemes and CAR use errors

In this section we look at the average word length in morphemes and CAR use errors. First we look at Kim and Andrea separately and then we compare their development. The section starts with example sentences.

The examples show the average word length in morphemes together with CAR use errors. Two case use errors are presented; example 5 shows a case use error, made in a one morpheme word that is part of a sentence consisting of several words with a small number of morphemes. Example 6 shows a case use error in a more morpheme word.

(5) Suomi e/i ole todella[lex] helpp**o**[pr.nom]. *'Finnish is not very easy.'* (Andrea-t4s1/mp4-WlMrph5/6(1.20))
(6) viiko/n/lopu**/ssa**[infl] *'in the weekend'* (Kim-t36s7/mp34-WlMrph1/4(4.00)-cmpw)

Kim

Figure 9.7 presents Kim's average word length in morphemes and CAR use errors and shows a small decrease for the case use errors and an increase for the average word length in morphemes. The data show quite some similarities but several differences as well.

Overall, the graph shows support and competition between the measures. Once, the average words length in morphemes remains at the same level while the CAR use errors grow (mp8-9). Kim's correlation coefficients for

both the raw and residual data illustrated very weak positive correlations between the measures (R=.05; R=.07).

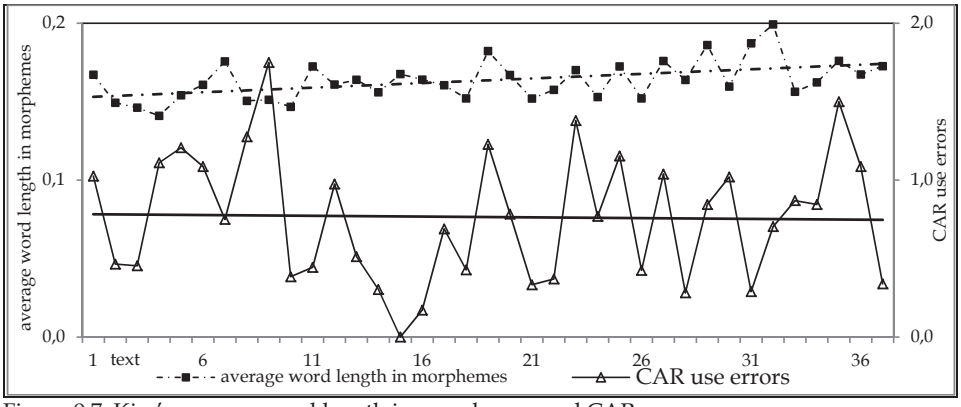

Figure 9.7: Kim's average word length in morphemes and CAR use errors.

Andrea

Figure 9.8 shows Andrea's average word length in morphemes and CAR use errors. The trend lines show considerable decrease for the case use errors and small increase for the word length in morphemes. Both measures illustrate quite some variability in the data and besides, the patterns show mainly differences.

Overall, the graph shows support and competition between the measures. Once, the average words length in morphemes remains at the same level while the CAR use errors decline (mp7-8) and once, the CAR use errors remain at the same level while the average words length in morphemes declines (mp23-24). Andrea's correlation coefficient (raw data) showed a weak negative correlation between the two measures (R=-.17); her residual data showed a very weak negative correlation, though (R=-.08).

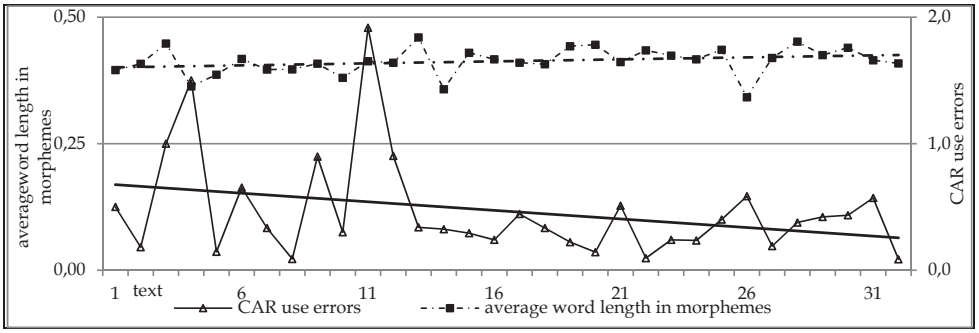

Figure 9.8: Andrea's average word length in morphemes and CAR use errors.

Kim and Andrea compared

Figure 9.9 illustrates the moving windows of correlations of the two measures and shows alternations of positive and negative correlations for both Kim and

Andrea. Again, their patterns show considerable differences, though. As was seen for the average clause length in morphemes in combination with the CAR use errors, also writing more complex words goes hand in hand with making more errors for Kim; she does not learn words and their case use simultaneously. Also as was seen for the average clause length in morphemes and the CAR use errors, Andrea writes more complex words while at the same time making fewer errors; she learns the words and their case use more holistically in context.

Beforehand we had hypothesized that word length in morphemes and case use errors support each other; when the average word length in morphemes would become more complex more case use errors were expected and less case use errors when the average word length in morphemes would be less complex. For Kim's correlation coefficient this is true but for Andrea's it is not.

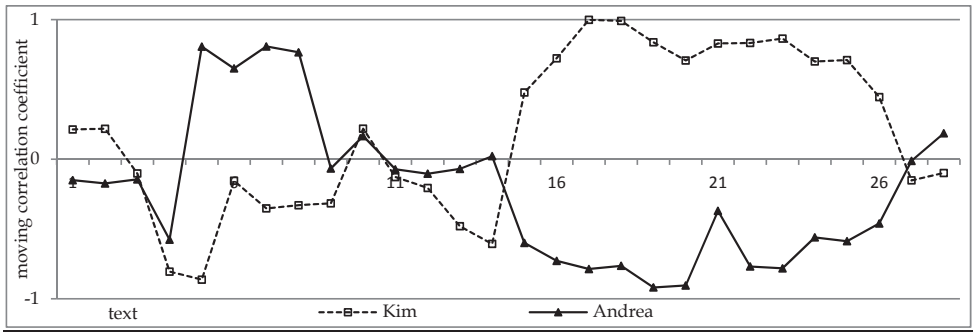

Figure 9.9: Kim and Andrea's average word length in morphemes and CAR use errors and their interaction.

9.4 Interaction between average word length in morphemes and CAR form errors

Now the average word length in morphemes is looked at in combination with CAR form errors. First we will look at Kim and Andrea separately after which we compare their development. The section starts with example sentences.

The examples show words in which we explore both measures. In the first three examples we show the two possibilities of incorrectness in the form of cases; examples 7 and 8 consist of a stem with one case marking and show incorrectly formed consonant gradations (-nt-→-nn-; -k-→-kk-) and example 9 contains two stems (compound word), a case marking and a possessive suffix and has a stem form error (-kk-→-k-) (see chapter 2). Finally, example 10 shows two correct formed cases in a (compound) word.

(7) Holla**nt**i/ssa[cons.gr] *'in Holland'* (Kim-t2s1/mp2-WlMrph1/2(2.00))
(8) ri**k**aa/ksi[cons.gr] *'rich'* (Andrea-t31s8/mp28-WlMrph1/2(2.00))
(9) poi**kk**a/ystävä/lle/ni[form] *'to my boyfriend'* (Andrea-t34s4/mp31-WlMrph1/4(4.00)-cmpw)
(10) kansa/n/musiikki/a *'folk music'* (Andrea-t25s1/mp24-WlMrph1/4(4.00)-cmpw)

Kim

Figure 9.10 illustrates Kim's average word length in morphemes and CAR form errors and shows decrease for the form errors in cases and increase for the word length in morphemes. Both measures show quite large variability. Besides, the patterns show mainly differences.

Overall, the graph shows support and also competition between the measures. Nine times the CAR form errors remain at zero while the average clause length in morphemes alters. Kim's correlations coefficient for the raw data showed a very weak positive correlation between the measures (R=.08), whereas her residual data showed a moderately strong positive correlation (R=.32).

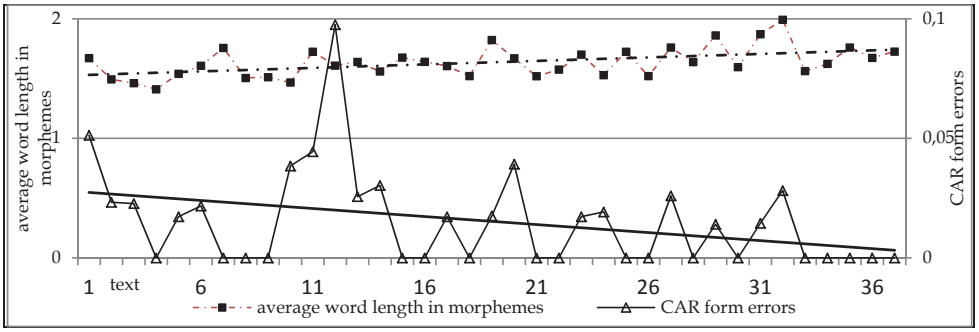

Figure 9.10: Kim's average word length in morphemes and CAR form errors.

Andrea

Figure 9.11 shows Andrea's case average word length in morphemes and CAR form errors and shows a decreasing trend line of the latter and a slightly growing trend line of the word length in morphemes. Both measures show variability and their patterns do not show any similarities.

Overall, the graph shows support and competition between the measures. Five times the CAR form errors remain at zero while the average clause length in morphemes alternates and once the average word length in morphemes remains the same while the CAR form errors grow (mp7-8). Andrea's correlations coefficient (raw data) showed a very weak (R=.07) correlation between the measures. Her residual data showed a weak positive correlation (R=.17).

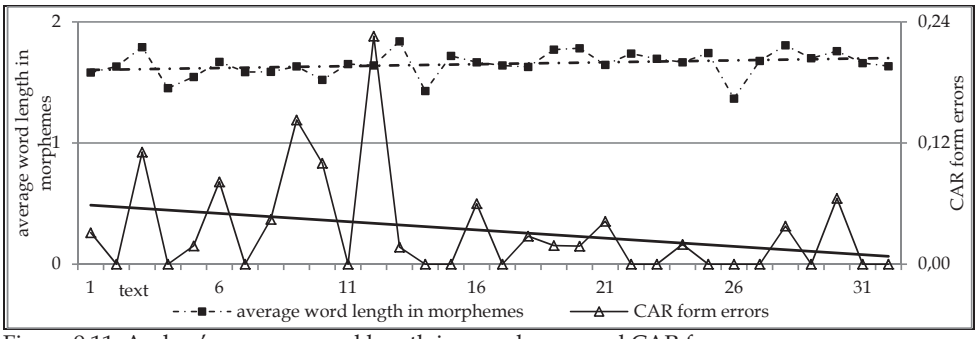

Figure 9.11: Andrea's average word length in morphemes and CAR form errors.

Kim and Andrea compared

Figure 9.12 presents the moving windows of correlations of the two measures. For both learners they show a changing relation, becoming more supportive over time and with a surprising similarity at the end of the period. It seems that longer words (including more cases) go hand in hand with more case form errors.

We had hypothesized that the average word length in morphemes and CAR form errors are supportive growers; growth of the complexity of the average word length in morphemes was expected to lead to more case form errors and less complexity of the average word length in morphemes to fewer case form errors. Both Kim's and Andrea's correlation coefficient show that this is true.

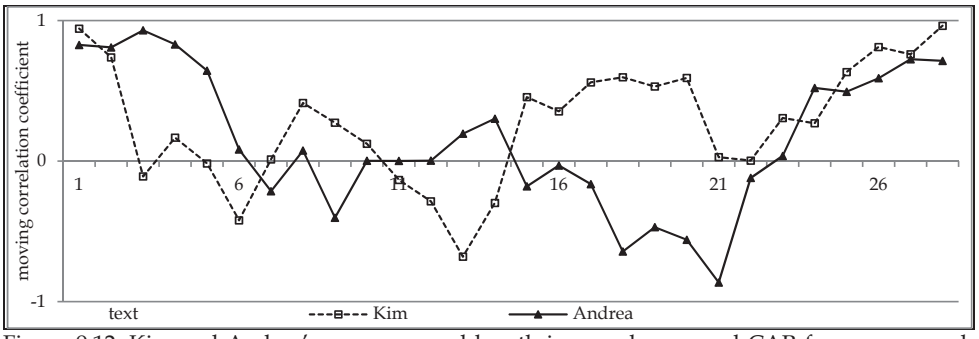

Figure 9.12: Kim and Andrea's average word length in morphemes and CAR form errors and their interaction.

9.5 Interaction between use of past and perfect tense and CAR use errors

In this section, the past and perfect tense use and CAR use errors are looked at. First Kim and Andrea are looked at separately and then their development is compared. The section starts with example sentences.

The examples show the complex tense use together with CAR use errors. Examples 11 and 12 show two different kinds of case use errors in past tense sentences. The fact that the sentences contain a logical subject in genitive may cause difficulties in finding the correct case for the remaining syntactic categories (see chapter 2). Example 13 shows the past tense together with a combination of words with incorrectly applied government (chapter 2). Example 14 shows a case use error in a perfect tense sentence.

(11) Meidän **täytyi** muuttaa muutam**a**[infl] kerta**a**[infl]. *We had to move several times.'* (Andrea-t14s2/mp13-past)
(12) Heidän **täytyi** suunnitella muistomerkki**a**[d.obj]. *'They had to design a monument.'* (Kim-t35s11/mp33-past)
(13) Me **asuimme** isossa asunnossa kesku**sta**[gov] lähellä. *'We lived in a big house near the city centre.'* (Kim-t14s6/mp13-past)
(14) …hän **on ollut** hyvin tärkeä suome**sta**[infl]. *'…he has been very important for Finland.'* (Kim-t26s5/mp24-perfect)

Kim

Figure 9.13 illustrates Kim's use of the past and perfect tense and CAR use errors and shows small decline for the latter measure and remarkable increase for the complex tense use. The variability of both measures is large and the patterns show mainly differences.

Overall, the graph shows support and also competition between the measures (mp1-9 are put aside, so: mp10-37). Once, the complex tense use remains at zero while the CAR use errors grow. Kim's correlation coefficients showed a weak negative correlation between the measures for the raw data (R=-.19) and a moderately strong negative correlation for the residual data (R=-.21).

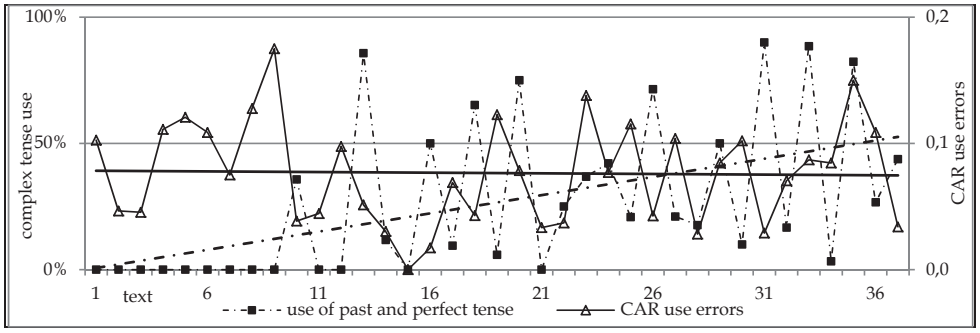

Figure 9.13: Kim's past and perfect tense use and CAR use errors.

Andrea

Figure 9.14 presents the use of the past and perfect tense and CAR use errors by Andrea and shows small increase for the use of the past and perfect tense and

remarkable decrease for the CAR use errors. The measures show quite some variability and the patterns show mainly differences.

Overall, the graph shows support and competition between the measures (mp1-3 are put aside, so: mp4-32). Once, the CAR use errors stay at the same level while the complex tense use alternates (mp23-24) and four times the complex tense use stays at zero while the CAR use errors alternate. Andrea's correlation coefficients of both the raw and residual data showed a weak negative correlation between the measures (R=-.17; R=-.15).

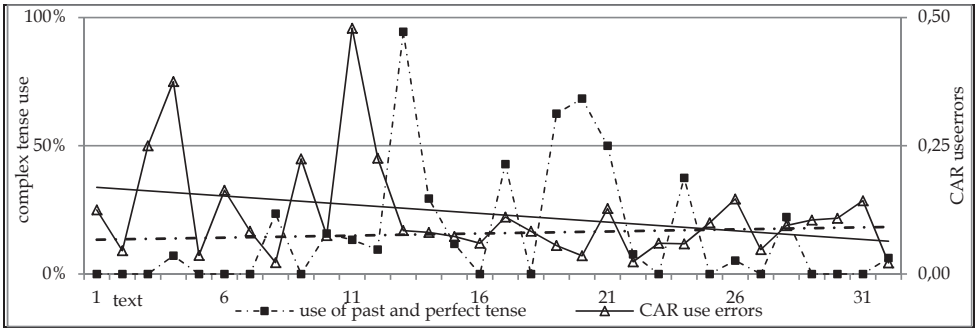

Figure 9.14: Andrea's past and perfect tense use and CAR use errors.

Kim and Andrea compared

Figure 9.15 illustrates the moving windows of correlations of both measures and shows alternations of positive and negative correlations for both Kim and Andrea. However, the patterns show considerable similarities.

We hypothesized that the complex tense use and the case use errors are supportive growers because the expectation was that more case use errors would be made when complex tenses were used and that less case use errors would be made when the tense use would be less complex (i.e. the present tense would (also) be used). For the correlation coefficients of both Kim and Andrea this is not true.

However, as was seen before for the complex tense use and the average word length in morphemes (chapter 8), the alternations between support and competition suggest that the measures do not actually influence each other over time. Both learners keep on making case use errors (fewer at the end) per word with and without using the past and perfect tense.

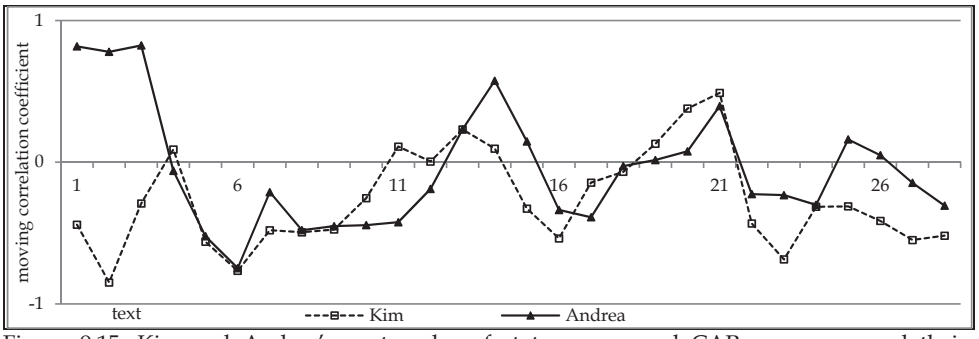

Figure 9.15: Kim and Andrea's past and perfect tense use and CAR use errors and their interaction.

9.6 Interaction between use of past and perfect tense and CAR form errors

In this section we look at the complex tense use in combination with CAR form errors. We will explore Kim and Andrea separately after which we will compare their development. The section starts with example sentences.

Examples 15 and 16 show complex tense use together with form errors in cases. Both examples illustrate a (part of a) sentence in a complex tense with different kinds of case form errors. Besides, Andrea's sentence was written in task 14 (about the childhood of the learners) in which both learners showed a peak in the use of the past and (plu)perfect but not in the incorrect form of cases.

(15) Norjassa me **olemme käyneet** pohjoisniem**issa**[cons.gr][infl]. *'In Norway we have visited the North Pole.'* (Kim-t22s8/mp20-perfect)
(16) …minä **tapasin** mun kaver**eja**[form][poss.suff] siellä ja minä **tykäsin**[cons.gr] oppia uusia asioita. *'I met my friends over there and I loved to learn new things.'* (Andrea-t14s5/mp13-past)

Kim

Figure 9.16 shows Kim's complex tenses and CAR form errors and shows a fast increase of the use of the past and perfect tense and a decrease of the CAR form errors and. The measures show large variability and their patterns show mainly differences.

Overall, the graph shows support and competition between the measures (mp1-9 are put aside, so: mp10-37). Five times the CAR form errors remain at zero while the complex tense use alters and the complex tense use remains at zero while the CAR form errors alters. Kim's correlation coefficient (raw data) showed a weak negative correlation between the measures (R=-.14), whereas the residual data showed a very weak positive correlation (R=.04).

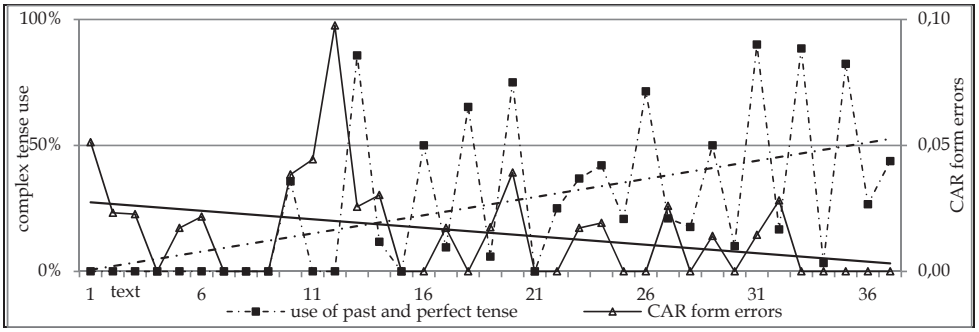

Figure 9.16: Kim's past and perfect tense use and CAR form errors.

Andrea

Figure 9.17 presents Andrea's complex tenses and CAR form errors and shows an increase of the use of the complex tenses and a decline of the trend line of the CAR form errors. The measures show a lot of variability and their patterns do hardly show any similarities.

Overall, the graph shows support and also competition between the measures (mp1-3 are put aside, so: mp4-32). Five times the CAR form errors remain at zero while the complex tense use alternates and four times the complex tense use remains at zero while the CAR form errors alternate. Andrea's correlation coefficients for both the raw and the residual data illustrated very weak negative correlations between the measures (R=-.09; R=-.08).

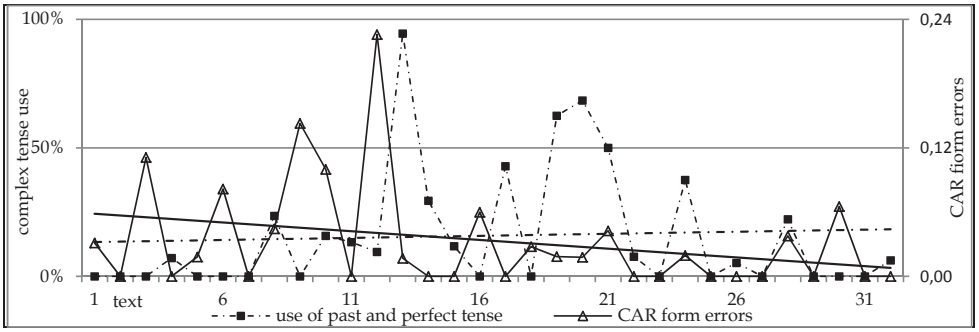

Figure 9.17: Andrea's past and perfect tense use and CAR form errors.

Kim and Andrea compared

Figure 9.18 illustrates the moving windows of correlations of the complex tense use and the CAR form errors for both learners. Both learners' patterns show remarkable similarities.

We had hypothesized that the morphological complexity measure use of the past and perfect tense and the morphological accuracy measure CAR form errors are supportive growers; our expectation was that complexity in tenses

would lead to more case form errors and that less complexity in the tenses (i.e. when the present tense would (also) be used) would lead to fewer case form errors. For Kim's correlation coefficient this is true though only marginally. For Andrea's correlation coefficient this is not true; her residual data show a very low level of competition. However, again the alternations between support and competition suggest no actual influence of the measures on each other over time. Kim and Andrea keep on making case form errors (though fewer at the end) per word, with and without writing in the past and perfect tense.

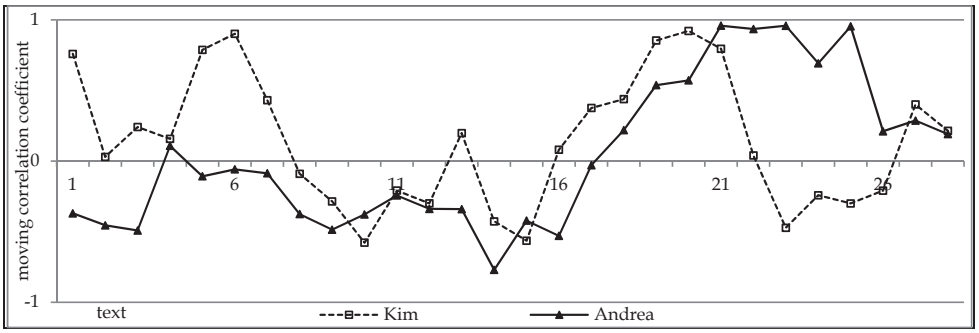

Figure 9.18: Kim and Andrea's past and perfect tense use and CAR form errors and their interaction.

9.7 Discussion

This chapter traced the development of five complexity and accuracy measures (average clause length and word length in morphemes, complex tense use and CAR use and CAR form errors) for the two learners. Different from the study of Spoelman and Verspoor (2010), who only found meaningful interactions between complexity measures over time, the current study also found a few interactions between accuracy and complexity measures.

With regard to the average clause length in morphemes versus the CAR use errors, the learners showed some differences. As pointed out in chapter 8, it appears that Kim did not learn words and their case use simultaneously and that for Andrea this process happened more holistically in context. For both Kim and Andrea the correlation coefficient was very weakly negative, indicating that the measures do affect each other negatively, though hardly.

With respect to the average clause length in morphemes versus the CAR form errors, Kim and Andrea's moving windows of correlation patterns looked quite similar. Kim and Andrea differ in their accuracy development, though. When Kim wrote more complex clauses, this did not lead to more accuracy at the same time, while for Andrea it did. For Kim the correlation coefficient was weakly positive and for Andrea it was weakly negative. As with the use errors, the difference between the patterns and the correlation coefficients might be due to conditions that affect different interactions between the subsystems.

With respect to the average word length in morphemes versus the CAR use errors, the moving windows of correlation patterns of the two learners showed similarities in the first part of the data but quite some differences in the second part. Again Kim did not learn words and their case use simultaneously while Andrea did so more holistically in context. The correlation coefficient was very weakly positive for Kim and very weakly negative for Andrea. The average word length in morphemes is hardly influenced by an increase of correct use of the cases.

With regard to the average word length in morphemes versus the CAR form errors, the moving windows of correlation patterns of the two learners showed a changing relation (more supportive over time). At the end of the period it showed surprising similarities. Longer words went together with more case form errors. The correlation coefficient was moderately strong positive for Kim and weakly positive for Andrea. In other words, when words became more complex, in proportion more errors were made in case forms. As would be expected from a dynamic perspective, development is a process of trial and error.

With respect to the complex tense use versus both the CAR use errors and the CAR form errors, the moving correlation patterns of Kim and Andrea showed considerable similarities and huge variability. However, as was seen before (chapter 7), the alternations between support and competition suggest that the measures do not affect each other much over time. Kim and Andrea both keep on making case use and form errors (though fewer at the end) per word, independent of the use of complex tenses.

10 CONCLUSION

This chapter first summarizes the main aims of the dissertation and the findings at both the group and the individual level. After linking the findings to the literature and interpreting them, the relevance of the current study to theories on second language development in general and to the dynamic approaches in second language development in specific will be given. Finally, we will make some suggestions for further research.

In the current study, we explored the differences in second language development between two groups: four learners of Finnish as a foreign (FL) and four learners of Finnish as a second (L2) language. The FL group took a Finnish proficiency course that focused heavily on explicit grammar teaching and had no immersion in that language in the Netherlands. The L2 group took a university language course with mainly a focus on meaning, and the learners were immersed in the language during their stay in Finland.

The development of Finnish as an FL or L2 is especially interesting because in the Finnish language there are several issues and constructions which are unfamiliar for learners who have an Indo-European language background; for them learning Finnish is learning a very different language.

The first question was if there were differences in outcomes in L2 development between the two groups at the end of one academic year. The second question was if there were differences in developmental patterns over time.

For the first question, a review of the literature (Ellis, 1994; Ellis, 2001) shows that explicitly instructed learners usually perform better in the end, but as Norris and Ortega (2000) point out, most of the studies have tests that may be biased in favor of instructed groups. Moreover, the interventions are usually very brief (a few weeks' time). To avoid the test bias, the current study uses free response data and to avoid the time limit, the current study spans one academic year. However, even this time span may actually have been too short. Rousse-Malpat and Verspoor (2012) showed different accuracy effects after one year of focus on form versus focus on meaning, but these had disappeared after two years.

For the two groups, we compared a number of complexity and accuracy measures longitudinally and discovered significant differences in their development; however, because of the different L1 backgrounds, three of which without case systems in their L1, we could not attribute the differences to context. As Spoelman (2013) showed, second language learners of Finnish cannot rely on their L1 when there is no cross linguistic similarity. Therefore, we compared the development of one particular learner from the L2 group whose L1 was German, with one learner from the FL group.

For the second question, we take a dynamic perspective. By tracing a great number of measures in two focal learners in detail, we focus on developmental trajectories. Do these learners develop in a similar manner or not? However, studies from a dynamic perspective (Larsen Freeman, 1997; Caspi, 2010; Van Geert, 1994; Verspoor et al., 2012; Bulté, 2013; Murakami, 2013) have already shown that there is considerable variability (intra individual variation) and variation (inter individual variation) in learners, especially at lower levels. Language development is a trial and error process and each learner will have to find his/her own learning path, so similar or different patterns may be due to normal variation. Still some studies (Verspoor et al., 2012; Murakami, 2013; Bulté, 2013) have also shown that learners will show more similar trends, especially in the broader measures (such as average sentence length, word length, total number of errors and so on). In the current study, several broad measures have been used; moreover, the two focal learners are first compared to their groups to see if they are representative for the groups.

We compared both the groups and focal learners statistically for differences in outcome and development. We also looked for significant peaks and dips in their data, which would indicate a developmental phase shift. Moreover, we explored interactions between several syntactic and morphological measures because we were interested in the question which subsystems would support each other or compete with each other.

We emphasize, however, that by presenting the results of these two small groups and two focal learners we do not have any intention to generalize.

10.1 Development over time from a traditional perspective

The research started with several comparisons between the FL and the L2 groups on syntactic complexity, morphological complexity and accuracy measures because we were interested in finding significant differences between the two groups.

The FL group consisted of three Dutch students and one Belgium student majoring in Finnish in the department of Finno-Ugrian Languages and Cultures at the University of Groningen in the Netherlands. The L2 group consisted of four students (German, Portuguese, Japanese and Indonesian as L1) who

learned the Finnish language through Finnish university courses in Finland. The data consisted of about 30 texts written per person over the course of 10 months that were assigned weekly. For the FL group they were part of the course; for the L2 group they were not. The university course in the Netherlands was grammar focused and the learners received detailed feedback on their errors in the writing with references to the checklist. The Dutch learners were asked to watch a short video every week with native speakers of Finnish as actors, but in class they communicated mainly in Dutch with the lecturer. In Finland the communication between the teachers and students was mainly in Finnish. The learners in Finland, who volunteered to take part in the project, were asked by the researcher to write weekly on the same topics as the FL learners had and received feedback on the errors in exchange for participating in the project. All errors were corrected but no explicit grammar rules were given, even though the focal learner (the German writer) early on would have preferred that.

The writings of each learner were corrected (for which a checklist was used) by a native speaker of Finnish and hand coded for 22 complexity and accuracy measures at the syntactic level and the morphological level: the types of sentences, the sentence and clause length in terms of morphemes and the use of cases (syntactic complexity), the word length in terms of morphemes and tenses (morphological complexity), the total of errors and the use errors in cases (syntactic accuracy), the form errors in cases (morphological accuracy) and the incorrect meaning and order of words.

The statistical analyses (paired samples tests and correlations) indicated that the FL and L2 groups were similar in many ways. Both groups showed more syntactic complexity (more complex sentence use) and morphological complexity (increase of the average sentence and clause length in morphemes) during the time span. Significant differences between the two groups were found in the use of cases and some related complexity and accuracy measures. As far as the use of cases is concerned, the FL group showed more development over time in the total use of the 15 Finnish cases, less use of the nominative case and correspondingly, relatively more use of the 12 cases other than the nominative, genitive and partitive. In other words, the FL learners used a wider range of cases. In addition and probably partially related to the wider range of cases used, the FL group used morphologically more complex words during the time span. Finally, the FL group was more accurate overall and specifically they applied the cases better and used more words correctly in terms of meaning and order in the context.

The causes for the differences could be condition (type of instruction, language learning environment) or the effect of L1. As mentioned above, the FL learners were given very explicit grammar lessons with rules and examples, especially in the use of cases. The L2 learners on the other hand, received hardly any explicit instruction, so they had to deduce the implicit rules mainly from the input. However, as was pointed out earlier, the group consisted of four learners, three of whom have an L1 without a case system. As Murakami (2013)

has shown, the L1 had a significant effect on the acquisition of morphemes in English and he stated that even advanced learners of L1's that do not have a particular morpheme in their L1, hardly ever attain an accuracy level over 90%. Also, as already mentioned, Spoelman (2013) showed that learners of Finnish can rely on their L1 when it contains the same features as there are in the Finnish language and cannot rely on their L1 when it does not. This was shown by Spoelman (2013) for the use of the partitive by respectively Estonian and Dutch learners.

Another cause for the overall differences could be the limited amount of time of the study. As already seen, Rousse-Malpat and Verspoor (2012) compared a focus on form with a focus on meaning group. They found differences in accuracy after one year in favor of the focus on form group, but these differences had disappeared after two years, so it may take more time for the focus on meaning learners to distill the patterns.

Finally the groups may have differed in motivation because the FL learners were all Finnish majors who would be graded on their writing texts. The L2 group consisted of motivated learners (as they took the course and wanted to interact in Finnish) but they were not language majors, nor did they receive grades for their work.

Having found that the groups showed differences in several complexity and accuracy measures, we became interested in the question whether individuals with a similar language background differ too. To examine this, one learner was selected from each group. Kim from the FL group was Dutch and Andrea from the L2 group was German (Dutch and German are Indo-European, Germanic languages), both with the highest sample mean of their group, both with a great interest in Finnish grammar and an eagerness to learn and understand the grammatical rules. Because German has a more detailed case system with more overt case markings than Dutch, Andrea may have been at a slight advantage to recognizing and using cases in Finnish.

The exploration started with the question whether there were differences between them at the beginning and the end of the study. Two texts for each learner written at the beginning of the study served as pre-test and two texts for each learner written at the end of the study served as post-test. These eight texts were first scored holistically by seven native speakers of Finnish and experts on Finnish as an L2. They were scored on four criteria: sentence complexity, morphological complexity, accuracy in general and authenticity (idiomaticity). Except for the second text, where Kim was found to score higher on sentence complexity, there were no significant differences found in the texts written by them; in other words, there were no differences between the two focal learners at the end of the academic study year when scored holistically.

We then compared the two texts per writer that had been used for the post test on all the quantitative data available and found a rather substantial difference in the use of the other 12 cases (used as indicator of overall complexity) and the total of errors (used as indicator of overall accuracy). As a statistical test is not possible on only two texts per person, we took the last eight

texts per learner. No statistical difference was found in the complexity measure but there was a difference in the total number of errors (normalized to 100 words). In other words, the focal learners were equally complex, but the learner with the focus on grammar was generally more accurate. These findings could be related to context, explicit instruction, or time span. As their holistic scores showed, the learners had progressed equally in many different ways. The statistical analyses showed only one real difference in overall accuracy.

The focal learner in the FL group, Kim, was similar to her group members in all the measures that were tested statistically. Sometimes the development over time of isolated measures showed different patterns, but that is to be expected from a dynamic perspective. For example, Kim's writing showed several peaks in several measures and the one in the other 12 cases use was almost significant (and significantly so for Sanne in the FL group), which suggests that these two FL learners have moments of overusing these cases. The fact that the four FL learners are so similar could be attributed to the combination of similar context, instruction, L1, motivation and time. This shows that language development, at least in the foreign language learning context, may not be as chaotic or unpredictable as a DST approach might suggest.

The focal learner in the L2 group, Andrea, was similar to her group members in terms of general complexity measures, but clearly different in the measures related to cases, both in using them and accuracy rates. Andrea kept using the nominative (default case) less than her L2 group members during the whole time and used more cases (total use of cases and the use of the 12 other cases) than the other members of the L2 group. Andrea was also different from her L2 group members in that early on she made more errors, but these later disappeared and she became more accurate than her group members. This was probably due to the fact that Andrea used more other cases than her group and thus had more chance of making case errors. The fact that the only clear differences between Andrea and her group members are to be found in the area of cases suggest that this is an L1 effect. With German as an L1, she was able to recognize and use them more easily.

The two focal learners, Kim and Andrea, demonstrated striking similarities in final outcomes but some interesting differences in the trajectories of different measures, which will be discussed in the next section. During the whole trajectory Kim used more cases (including the 12 other cases), and there was a significant difference for the absolute numbers per 100 words but not for the relative numbers.

Table 10.A summarizes the findings of the statistical tests on the group data and on the focal learners Kim for the FL group and Andrea for the L2 group. From the longitudinal data it appeared that Kim and Andrea also differed in the use of the past and perfect tense, which is not included in the table.

Table 10.A: All findings of the statistical tests on the group data and the focal learners Kim (FL) and Andrea (L2). CAR stands for case accuracy rate: the number of incorrect uses and or forms divided by the total use of cases.

	Similarities	Differences (in favor of first mentioned)	Possible causes
FL and L2 groups	Simple sent. Compound sent. Complex sent. Comp-complex sent. Ave SL in morphs Ave CL in morphs Genitive use Partitive use	Total use of cases Nominative use 12 other cases use Ave WL in morphs Case use errors (abs.) Total errors CAR use +form errors Incorrect meaning Incorrect word order	Differences in: Context Instruction L1 Motivation Time
FL and Kim	All measures including: Total use of cases Nominative use 12 other cases use Ave WL in morphs CAR use + form errors		Similarities in: Context Instruction L1 Motivation Time
L2 and Andrea	All measures including: Nominative use Ave WL in morphs	Total use of cases 12 other cases use CAR use + form errors	Differences in: L1 Motivation
Kim and Andrea	All measures including: Total use of cases Nominative use 12 other cases use Ave WL in morphs CAR use + form errors	Total errors	Differences in: Context Instruction Time Similarities in: L1 Motivation

To summarize, the FL and L2 learners behaved rather similarly, except in the use of cases, which we assumed was mainly an L1 effect. When we compared our two focal learners with similar L1's holistically, we found no differences at the end. When we compared their last eight writings statistically on the measures that were numerically different, we found one difference: total number of errors. Andrea was less accurate than Kim. It is quite remarkable that the two focal learners in such different contexts, with the one in the Netherlands with much less exposure to the Finnish language and the one in Finland with much less explicit instruction, were in the end so similar when regarded from a traditional perspective.

10.2 Development over time from a dynamic perspective

Even though the outcomes from a traditional perspective for our focal learners showed few differences, the question is whether they went about learning the language in a similar manner, which brings us to our second research question: Do these learners develop their language over time in a similar manner?

The purpose of an analysis from a dynamic perspective is to find general and individual developmental patterns, in our case mainly to see whether there are clearly different developmental patterns that could be due to instruction. Do the learners with the focus on grammar, who are encouraged early on to use certain forms, and the learners with the focus on meaning, who will have to discover the forms on their own, have clearly different patterns?

From a dynamic perspective, we want to find out how different subsystems of the language develop over time. Item-based constructions such as individual words may develop relatively smoothly as they have to be learned one by one, but other more schematic or more productive constructions such as the use of a rather regular morphological item may show a high degree of variability at one point and then a sudden shift once it has been mastered. The difference between a learner with the focus on grammar and a learner with the focus on meaning could be that the former recognizes the patterns earlier and will use them earlier with a clear shift, whereas the latter has to discover the forms in a more item-based manner, leading to fewer clear shifts.

Moreover, we want to find out how different subsystems of the language interact over time. For example, if words become more complex, do sentences become more complex at the same time or is there a competition between these two subsystems at first (the learner may focus on word complexity at the expense of sentence complexity)? It is also possible that later, when the learner has mastered both subsystems, these two subsystems become coordinated and develop at the same rate.

Following Van Geert (2008) and Van Geert and Van Dijk (2002), we looked at the trajectories of single measures by means of min-max graphs, and when needed, Monte Carlo analyses were performed to see if there were clear shifts over time. If the min-max graphs showed clear moments of widening or narrowing band widths, these moments were regarded as shifts in development. If the trajectories also showed clear peaks or dips (signs of overuse or underuse), they were tested by means of a Monte Carlo analysis and if found to be significant, they were taken to be developmental (rather than random) patterns.

Moreover, we looked at the interaction between subsystems such as support (grow at the same time), competition (compete), and changing relation over time (first competition, then support), which could be interpreted as a precursor relationship. In some of these cases, it may be argued that the subsystems are supportive not because they are separate subsystems that support each other, but because they are inherently connected in the language

(for example a longer word in letters is likely to have more morphemes). In other cases, it may be argued that some subsystems are not really connected growers in that they do not seem to influence each other in any consistent manner over time. For example we found that the use of more complex tenses seemed to have no connection with word complexity or accuracy measures.

We will summarize the main findings on the patterns between the explored subsystems and provide an overview of facts that might have influenced the results of these interactions.

The writings of the two focal students were traced longitudinally in detail and examined for general similarities or differences.

Kim and Andrea developed rather similarly in learning Finnish; they showed growth and decrease for the syntactic and morphological complexity and accuracy measures. Nevertheless, the developmental patterns of the two learners showed several differences. It seems that the learners with a focus on grammar showed steeper increases (peaks), some of which were (almost) significant after which they were tested for all participants of the group (use of the other cases: Kim (almost significant) and Sanne; case accuracy ratio (CAR) form errors: Kim). Overall Kim showed more complexity and accuracy than Andrea early on, but towards the end most of these differences disappeared and the learners seemed to converge, but not entirely yet.

10.2.1 Similar patterns for the two learners

The interaction over time for several measures was similar in the two learners, which may be cautiously regarded as general developmental patterns for learners for Finnish.

A supportive relation was found between the more complex sentence patterns and the average sentence length in morphemes, suggesting that when a sentence consists of more morphemes, it is more likely to be a complex sentence. At the beginning both focal learners wrote more simple sentences consisting of fewer morphemes, while at the end they wrote more complex sentences consisting of more morphemes. However, the data also showed a few competitive points, suggesting that these growers are not inherently connected.

Another supportive relation was found between the average clause and word length in morphemes, suggesting that the number of morphemes in a clause is higher when the number of morphemes in the words is higher. At the beginning, both learners wrote short clauses and short words, while at the end they wrote longer clauses and longer words in terms of morphemes. However, again, the data showed a few competitive points, suggesting that these growers are not inherently connected.

Similar changing relations, from supportive to competitive, for both learners were found between the more complex sentence patterns and the average clause length in morphemes. At the beginning, both learners used both fewer complex sentences and fewer morphemes in the clauses. However, later on there were competitive interactions between these measures, suggesting competition in attentional resources. Either the sentence became more complex

with dependent clauses, or the clause became more complex with more complex words in terms of morphemes.

Also the CAR form errors and the average word length in morphemes showed a changing relation that became more supportive over time, suggesting that early on, the longer words became (which of course included case endings) the more errors were made in form. At the end of the period both learners wrote more morphemes per word, applied more cases and had relatively more case form errors even though overall complexity and accuracy increased.

There were also some measures that seemed to have no clear effect on each other and are probably not competing for attention. The development of the past and perfect tense use seemed to be rather random and not connected with other complexity and accuracy measures. The three interactions (moving windows) between the complex tenses and the average word length in morphemes, the CAR use errors and the CAR form errors showed quite some alternations in support and competition, suggesting that interaction is rather random. During the time span, both learners wrote more morphemes per word and kept on making errors in case use and form (though fewer at the end) with and without using the past and perfect tense. However, it was noted that Kim and Andrea had relatively different patterns in the moving window of correlations between the complex tenses and the average clause length in morphemes.

10.2.2 Different patterns for the two learners

For several measures the interaction over time was different in the two learners, which may be very cautiously regarded as differences due to context and/or instruction for learners of Finnish.

One of the clearest differences between the two learners is that Kim's writing showed several peaks in several measures and the one in the other 12 cases use was almost significant (and significantly so for Sanne in the FL group), which suggests that these two FL learners have moments of overusing these cases. It is likely that the focused attention on these forms encourages these learners to use them, which may result in a degree of overuse early on. Andrea's measures did not show such strong peaks in these measures.

Different patterns are also found in the general trend lines and the interaction between the CAR use errors and the CAR form errors. The general trend lines show that Kim's case form errors decreased quite suddenly, but her case use errors did not. For Andrea both showed a sudden decline at the same time. This would suggest that there is a difference in the processing of these two items in the two learners. The learner with the focus on grammar can reason out form errors, which are rather schematic, so once the "rule" has been discovered it can be applied over a range of cases, leading to the sudden drop. However, she cannot reason out the use errors, which are more idiomatic and therefore item-based, so these form-meaning mappings are learned one by one. The learner with the focus on meaning has to learn both the forms and uses more holistically in an item-based manner through the input and only after enough

frequency of occurrence will the items with the form and use interconnected be used.

Another analysis that supports the different learning trajectories of case use is found in the moving window of correlation of the CAR use errors and two complexity measures: the average clause length and the average word length in morphemes. Kim showed a supportive interaction between the two complexity measures and accuracy measure in the second part, indicating that she wrote more complex clauses and words, but made more errors at the same time. This suggests that complexity and case uses are not learned simultaneously. Andrea's competitive interaction in that same time period indicates that she also wrote more complex clauses and words, but making fewer errors at the same time, suggesting that she learned the words and their use more holistically in context.

Also in the CAR form errors there was a difference between the two focal learners as indicated by the different moving windows of correlation patterns between the CAR form errors and the average clause length in morphemes. Again, Kim's supportive interaction in the second part of the period indicates that she started writing more complex at both the clause and word level, but did not become more accurate in case form at the same time. Andrea, on the other hand, wrote more complex at both the clause and word level, and she became more accurate in case form at the same time.

To summarize, Kim and Andrea's developmental patterns were quite similar for most measures examined. There was only one clear difference. For Kim, more complexity would go hand in hand with more case errors, whereas for Andrea more complexity did not mean more case errors. On the whole we may conclude that for these two learners the difference in focus on grammar and focus on meaning resulted in some differences in the developmental trajectories but not so much in the final outcomes.

10.3 Relevance to theories on second language development in general

Overall we found a few major differences between the FL and L2 groups, which we attributed to L1 differences (Murakami, 2013; Spoelman, 2013). The two focal learners with a similar L1 background in the end seemed to differ only on the total of errors, so the FL learner was on the whole more accurate.

We may assume that the effect of focus on grammar helped the learners to discover and notice the forms and enabled them to use them earlier, but the use may also have led to more case errors in those forms initially. The effect of focus on meaning was that it took longer for the L2 learner to notice the forms and use them, but once they had been noticed and used they were used relatively more correctly than in the FL condition. The total of errors (i.e. lexical, syntactic, morphological and word order errors) remained higher though for

the L2 learner during the time span of one academic year. This result is in line with other research, particularly which was done in the Cefling project where thousands of L2 writers have been studied[35]. The overall conclusion in this project is that frequent use of a grammatical form precedes accuracy by about one CEFR level (see e.g. Lesonen, 2013; Martin et al., 2010). According to these studies, a one-year span is only long enough to bring about the growth of accuracy for the most frequent grammatical structures.

Finally, the current study may contribute to further research in Finnish as an L2 or FL. We started with the exploration of several general measures in order to find interesting measures to examine. We found that for the Finnish agglutinative language the average number of morphemes per finite verb gives insight into complexity (Verspoor et al., 2008).

10.4 Relevance to dynamic approaches in second language development

As far as developmental patterns are concerned, all learners' language use became more complex over time. Furthermore, even though the FL group used more cases than the L2 group, both groups showed a similar order of case use during the time span (chapter 4). Moreover, the FL learners showed similar trends, though with some local variability and variation, but overall, they showed similar directions. The FL focal learner used more complex constructions early on, accompanied by seemingly developmental peaks in use of the 12 other cases, suggesting overuse which appeared to be a clear sign of discontinuity (by two FL learners), indicating a developmental phase shift (Van Dijk, 2004). The L2 focal learner was slower to use more complex constructions and her accuracy rates were lower early on, but as time went on she used the more complex constructions without more case errors. The learners also showed different patterns of development for CAR use and form errors. The FL learner had separate trajectories for them, whereas the L2 learner learned them more together. Apparently, the case forms were easy to learn for the FL learner; after a start with a pattern of many incorrectly formed cases, she learned how to form them, which she then applied in her writings (despite the fact that the communicative need was not present in her everyday life (Haley & Rentz, 2002)). We may conclude that neither the type of instruction nor the language learning context affects order of acquisition, but that they do affect the overall accuracy and form-meaning trajectories.

In line with dynamic thinking we emphasize the findings of Spoelman and Verspoor (2010), Verspoor et al. (2012) and Murakami (2013) that language development for beginner learners of a second language is a rather variable

[35] Cefling: https://www.jyu.fi/hum/laitokset/kielet/tutkimus/hankkeet/paattyneet-hankkeet/cefling/en (visited on 05-06-2014).

process with a large amount of non-linear development, variation and changing relationships among the measures. Moreover, and again in line with dynamic thinking, we found developmental trends of growth in all general and specific syntactic and morphological measures for the FL group and for the two individuals (FL and L2), in line with findings by Verspoor et al. (2004), Verspoor et al. (2008), Spoelman and Verspoor (2010), Verspoor et al. (2012) and Bulté (2013). Moreover, we found several meaningful interactions between complexity measures and between accuracy measures and also, unlike Spoelman and Verspoor (2010), between complexity and accuracy measures. Still, overall relatively few strongly connected subsystems were found. We found that the measures mainly interact in a similar way for the FL and the L2 learner, which may be due to their similar L1. Besides, we found only one precursor effect: a declension of case form errors before case use errors by the FL learner (Spoelman & Verspoor, 2010; Caspi, 2010).

From this, together with the findings that the L2 learner from Germany acted more in the way of the FL group and the FL learner Kim, we may conclude that the L1 is of enormous importance in acquiring an FL/L2. Also the fact that the FL learner (who learned in exactly the same conditions as the learner Spoelman and Verspoors' study (2010)) as well as the L2 learner (in totally different conditions) both showed quick decrease of the error rates, underwrites the importance of the L1. Still, we found that case use and form errors do not disappear in the learners' written texts during the time span, despite the fact that they paid attention to the correct forms in the writing process (Swain and Lapkin, 1995). In addition, for the FL learner holds that noticing the rules did not mean that no errors were made (Schmidt and Frota, 1986). However, unlike Coleman (1997) and Ortega (2003), we found that overall the FL learner showed more accuracy than the L2 learner. It seems that, at least for the first year of learning, the form of instruction (focus on grammar or on meaning) plays a role as well (Verspoor et al., 2004) and that just communicative practice seems to be not sufficient to become proficient and accurate in the second language (Lightbown, 2000). However, all learners seemed to have the urge to express themselves from the start which may have gone beyond their language proficiency limits, as was also found by Seilonen (2013).

However, if the difference in accuracy can be attributed to instruction, it is not necessarily the difference between explicit (focus on form) versus implicit (focus on meaning). Instruction may also have differed in terms of the kind, type and amount of input. A communicative class usually focuses on every day spoken language, whereas the focus on forms class also deals with written and spoken texts, which are usually discussed in more depth. Moreover, a communicative class usually encourages interaction among the learners, so the input may not always have been correct. And as Andrea was apparently the more successful learner, the course may not have given her enough challenge. Still, we cannot know the amount of influence by the types of instruction; also the language learning environment between the two groups is different.

Concerning the latter, Andrea might have been at the advantage because being immersed by Finnish helps to acquire idiomatic language acquisition (Seilonen, 2013).

Finally, the differences found could also have disappeared in time. As mentioned before, the focus on meaning learners in the Rousse-Malpat and Verspoor (2012) study were less accurate after one year, but equally accurate after two years. Kajander (2013) showed similarities in the sequence for the development of accuracy measures, though at different levels, as well: in this case for young learners and adults. It might take an implicit learner just longer to figure out the patterns. Besides, rule descriptions on morphological issues do not encompass all areas of morphology (Martin, 1995).

The question now is whether the L2 learners with a different L1 would have benefited relatively more from instruction. From the fact that the FL learner showed more accuracy in the Finnish cases in the short term than the L2 learner, we may conclude that explicit instruction does have an effect on use. However, we do not know whether it is preferable for extracting difficult constructions, especially for grammatical items that do not exist in the L1 (de Bot *et al*. (2005). However, the L2 learner with a similar L1 regained the accuracy in cases in the long term and it probably just took her some more time to learn to "see the world in a Finnish way" (Martin *et al*. 2010: 64) and to understand the opaque constructions of the Finnish language. This is also in line with Pavesi (1986) and Ellis (1989) who found that the acquisition of a second language by instructed and non-instructed learners was done in the same order but that the learners who had had instruction learned faster; however, in the current study, the learner with the focus on grammar did not proceed much further.

Finally, traditional statistical analyses could not have shown the differences in processes, as in such analyses the effect of instruction should have been measured over time but in our opinion, short-term and one-off interventions should be regarded with extreme caution.

10.5 Suggestions for further research

Now that the current study has been accomplished we take the opportunity to sum up questions that were not asked because of study limitations, but would be interesting to get the answers to in future research.

Because the current study only examined data of beginner learners of Finnish, it would be useful to examine data of more proficient learners of Finnish to find whether these learners would show less variability than beginners (Verspoor et al., 2012; Murakami, 2013).

Furthermore, it would be worthwhile to explore the development of the study's participants and the explored measures' interactions in the next couple of years; would we find morphological and syntactic difficulties after several

years of exposure to Finnish, like Harley and Swain (1984) and Lyster (1987) found for the French language and would we find the differences between the FL and the L2 groups to have disappeared after two years, like Rousse-Malpat and Verspoor (2012) found?

Unlike the availability of these data of the FL learners (data of three academic years is available), for the L2 learners there is no chance to get these data (most of them left Finland after the academic year) and for this group it might be helpful to use the technique of modeling; with this technique we could simulate the developmental process to test theoretical assumptions (Caspi, 2010; Verspoor et al., 2011).

Moreover, it would be interesting to take a deeper look into the data of the six not intensively discussed learners. This would provide answers to the questions whether the four FL learners with the same L1 improve in a similar manner and how the learners in the L2 context, with their totally different language backgrounds, develop: in similar trajectories or not?

Following on from the previous suggestion, it would be useful to find more learners with similar L1's that learned the Finnish language as L2 (in particular learners from the Finno-Ugric language group), so that they could be examined in combination with the four FL learners with their Dutch language background.

It would also be useful to examine the measures which were counted for the current study but not used, in order to find whether these measures would confirm the outcomes between the two learners. For this reason, several issues that exist in the Finnish language and that are not known as such in the Germanic languages (like the have-construction, the use of some specific cases or constructions with a logical subject) could be examined.

Finally, it would be worthwhile to explore the errors of the FL and L2 learners in order to find whether there were cross-linguistic influences, i.e. of L1 to FL/L2 and vice versa (Kellerman & Sharwood Smith, 1986) or language transfer influences, i.e. of particularly L1 on FL/L2 (Gass & Selinker, 1983) and vice versa. Besides, it would be interesting to find helpful similarities between the examined L1's and Finnish that could be used in acquiring the Finnish language.

SUMMARY

Dynamic Systems Theory (DST), the theoretical background of this study, is a rather new theory in applied linguistics. In a dynamic perspective on language and language development, patterns that emerge from language use are studied and individual differences and developmental processes are important. The assumption is that all factors involved (e.g. social, linguistic) continuously interact and that variability is needed to select those language forms that are needed for each learner's language development (Thelen & Smith, 1994). In a study from a dynamic perspective, variability is a sign of change and development and the learners and their context are studied together (Verspoor et al., 2011). In this way, useful information about changes in subsystems can be detected (Van Geert & Van Dijk, 2002).

The main aim of this dissertation is to explore the development of beginner learners in two different settings: learning Finnish as a foreign (FL) and as a second (L2) language. The four FL learners (three Dutch learners and one learner from Belgium) took a Finnish proficiency course in the Netherlands that focused heavily on explicit grammar teaching. These learners treated language primarily as an object of study and focused on several specific forms in order to learn and understand them. In class, they communicated mainly in Dutch with their lecturer. The four L2 learners (one German, one Brazilian, one Japanese and one Indonesian learner) took a university language course in Finland, with mainly a focus on meaning. These learners treated language as a way to communicate. In class, they communicated mainly in Finnish with their lecturer. The study uses free response data (about 30 texts written per person, assigned weekly) and spans one academic year (nine months). For the FL group, the writing tasks were part of the homework, while for the L2 group they were an extracurricular activity.

The study explores complexity and accuracy measures. For syntactic complexity we looked at types of sentences, sentence and clause length (in terms of morphemes) and the use of cases. For morphological complexity we looked at word length (in terms of morphemes) and the use of tenses. For syntactic accuracy we looked at all errors and specifically at errors in the use of cases and words and in word order. For morphological accuracy we looked at the incorrect case forms.

First we looked for differences in outcomes in L2 development of several complexity and accuracy measures between the two groups of learners (group study) at the end of one academic year. For this purpose we looked at the overall averages of the data per group. Second, we looked for differences in developmental patterns and interactions over time between several syntactic and morphological measures. For this purpose we explored the data of two focal learners Kim, a Dutch learner from the FL group, and Andrea, the German learner from the L2 group (focal learners study). These two learners have been chosen because they have similar L1's. Moreover, they wrote the largest

number of texts and used the largest number of words in their respective groups. Besides, Kim and Andrea are both quite interested in the Finnish grammar.

The findings of the group study show that the FL an L2 groups are similar in many ways: more syntactic complexity (more complex sentence use) and morphological complexity (increase of the average sentence and clause length in morphemes) during the time span. However, the groups show significant differences in the use of cases and some related complexity and accuracy measures. They are all in favor of the FL group: more development over time in the total use of the 15 Finnish cases, less use of the nominative case and correspondingly, relatively more use of the 12 cases other than the nominative, genitive and partitive. In other words, the FL learners use a wider range of cases. Besides, the FL group uses morphologically more complex words during the time span. Finally, the FL group is more accurate overall and specifically they apply the cases better and use more words correctly in terms of meaning and order in the context. Reasons for the differences between the two groups may be: the types of instruction, the language learning context, the effect of L1 (Murakami, 2013; Spoelman, 2013), the limited time of the study (Rousse-Malpat and Verspoor, 2012) or less motivation to write well.

The findings of the focal learners study are preceded by exploring whether the two focal learners, with a similar language background but in different settings, show the same differences as their groups: in other words, whether they are similar to their respective group members. Kim (FL) appears to be so in all the measures that we test statistically. The fact that the four FL learners are so similar may be attributed to the combination of similar language learning context, type of instruction, L1, the time of the study or motivation. Andrea (L2) appears to be similar to her group members in terms of general complexity measures. She clearly differs in the measures related to cases, both in using them and accuracy rates. The fact that the only clear differences between Andrea and her group members are to be found in the area of cases suggest that this is an L1 effect.

For the focal learners study we started with a pre-test (two texts for each learner written at the beginning of the study) and a post-test (two texts for each learner written at the end of the study) which were scored holistically by seven native speakers of Finnish and all experts on Finnish as an L2 on four criteria: sentence complexity, morphological complexity, accuracy in general and authenticity. The holistic scores show no differences between the two focal learners. Subsequently we analyzed the last eight texts of the two focal learners statistically. We found one difference between the two learners; the FL learner shows significant more overall (syntactic, morphological and lexical) accuracy. The findings may be related to the types of instruction, the language learning context or the time span. It is quite remarkable that the two focal learners in such different contexts are in the end so similar when regarded from a traditional perspective.

We continued the focal learners study with the exploration of the development over time from a dynamic perspective. The purpose of such an analysis is to find general and individual developmental patterns. We explored how subsystems of the language develop over time and subsequently whether the two learners develop their language over time in a similar manner. We did this by looking at the trajectories of single measures by means of min-max graphs and when needed, by performing Monte Carlo analyses to look for clear shifts over time (Van Geert, 2008; Van Geert & Van Dijk, 2002). First, we found that language development for beginner learners of a second language is a rather variable process with a large amount of non-linear development, variation and changing relationships among the measures (Spoelman and Verspoor, 2010; Verspoor et al., 2012; Murakami, 2013). Besides, we found developmental trends of growth in all general and specific syntactic and morphological measures for the learners (Verspoor et al., 2004; Verspoor et al., 2008; Spoelman and Verspoor, 2010; Verspoor et al., 2012; Bulté, 2013). Second, despite the fact that the focal learners develop rather similarly in learning Finnish, they show several differences as well. It seems that the FL learner shows steeper peaks, some of which are (almost) significant. After testing, it seems that the use of the 12 other cases is almost significant for Kim (and for Sanne, an FL group member, it is significant) and that the measure case form errors (case accuracy ratio) is significant for Kim. Overall, Kim shows more complexity and accuracy than Andrea early on; the difference in focus on grammar and focus on meaning results in some differences in the developmental trajectories. We may assume that the effect of focus on grammar helps the learners with such a focus to discover and notice the forms and enables them to use them earlier, but the use may also lead to more case errors in those forms initially. The effect of focus on meaning is that it takes longer for the L2 learner to notice the forms and use them, but once they have been noticed and used they are used relatively more correctly than in the FL condition. However, most of these differences disappear towards the end; the two focal learners seem to converge, though not entirely yet; the total of errors (i.e. syntactic, morphological, lexical and word order errors) remains higher for the L2 learner during the time span of one academic year.

Finally, we explored in the focal learners study how different subsystems of the language interact over time and explore whether they show support, competition or a changing relation over time. We found that the interaction over time is similar for several measures in the two learners: first, between the complex sentence use and the average sentence length in morphemes (supportive relation). This suggests that a sentence, consisting of more morphemes, is more likely to be a complex sentence. However, the data also show a few competitive points, suggesting that these growers are not inherently connected. The second similar interaction was found between the average clause and word length in morphemes (supportive relation). This suggests that the number of morphemes in a clause is higher when the number of morphemes in the words is higher. However, again, the data show a few

competitive points, suggesting that these growers are not inherently connected. The third similar interaction was found between the complex sentence use and the average clause length in morphemes (changing relations, from supportive to competitive). Both learners start with fewer complex sentences and fewer morphemes in the clauses and later the data show completion: the sentence becomes more complex with dependent clauses or the clause becomes more complex with more complex words in terms of morphemes. The last similar interaction was found between the case form errors and the average word length in morphemes (changing relation, more supportive over time). This suggests that the longer the words become, the more case form errors are made in proportion. We also found that some measures seem to have no clear effect on each other because they are probably not competing for attention: the past and perfect tense use versus the complexity and accuracy measures. These findings may be cautiously regarded as general developmental patterns for learners for Finnish.

Furthermore, interactions over time were found to be different for several measures in the two learners: first, between the case use errors (case accuracy ratio) and the case form errors. Kim's case form errors decrease quite suddenly but her case use errors do not while Andrea shows a sudden decline for both measures at the same time. This suggests a difference in the processing of these two items in the two learners. The learner with the focus on grammar Kim can reason out form errors, which are rather schematic, so once the "rule" has been discovered it can be applied over a range of cases, leading to the sudden drop. She cannot reason out the more idiomatic and therefore item-based use errors though. The learner with the focus on meaning Andrea has to learn both the case forms and uses more holistically in an item-based manner through the input. Only after enough frequency of occurrence will the items with the form and use interconnected be used. We found a second different interaction, and with this another analysis that supports the idea of different learning trajectories of case use, between the case use errors and the average clause length and the average word length in morphemes. Kim shows a supportive interaction between the measures in the second part; she writes more complex clauses and words, but makes more case use errors at the same time. This suggests that the learner with the focus on grammar does not learn complexity and case uses simultaneously. Andrea's competitive interaction in that same time period indicates that she also writes more complex clauses and words, but makes fewer case use errors at the same time. This suggests that the learner with the focus on meaning learns the words and their use more holistically in context. We found another different interaction between the case form errors and the average clause length in morphemes. Again, it shows that Kim starts writing more complex at the clause level, but does not become more accurate in case form at the same time while Andrea writes more complex at the clause level and becomes more accurate in case form at the same time. The differences in the interaction over time in the two learners may be very

cautiously regarded as differences due to the language learning context or the types of instruction for learners of Finnish.

With this exploration we hope to contribute to the field of applied linguistics; first because to date there are no studies in which differences in the language development of beginner learners of Finnish as foreign language and as second language in free writing are compared. Besides, no research has been done yet that surveys the exploration of development over time of beginner learners of Finnish as foreign and second language and there is no research on syntactic and morphological development over time of a foreign or second language yet, nor on the different roles and relations of subsystems in language.

TIIVISTELMÄ

Dynaamisten järjestelmien teoria (DST = Dynamic Systems Theory), jota olen käyttänyt tutkimukseni teoreettisena taustana, on melko uusi tulokas soveltavan kielitieteen alueella. Dynaaminen lähestymistapa kieleen ja sen kehittymiseen tarkoittaa, että kielen käytöstä esiin nousevia malleja tutkitaan kiinnittäen erityistä huomiota yksilöllisiin eroihin ja kehitysprosesseihin. Oletuksena on, että kaikki vaikuttavat tekijät (esim. sosiaaliset, kielitieteelliset) ovat jatkuvassa vuorovaikutuksessa keskenään. Vaihtelevuus on tarpeen kunkin oppijan kielen kehityksessä tarvittavien kielellisten muotojen valikoitumiseksi (Thelen & Smith, 1994). Tutkittaessa ilmiötä dynaamisesta näkökulmasta vaihtelevuus merkitsee muutosta ja kehitystä, ja oppijaa tutkitaan kontekstissaan (Verspoor et al., 2011). Tämä tuottaa hyödyllistä tietoa järjestelmän eri osissa tapahtuvista muutoksista (Van Geert & Van Dijk, 2002).

Väitöskirjani päätavoitteena on kartoittaa kahdessa erilaisessa ympäristössä kielen alkeita opiskelevien kehitystä: he opiskelevat suomea joko ”vieraana kielenä” Suomen ulkopuolella tai ”toisena kielenä” (S2) Suomessa. Neljä suomea vieraana kielenä opiskelevaa (kolme hollantilaista ja yksi belgialainen) suorittivat Alankomaissa suomen kielen kurssin, jolla pääpaino oli kieliopin opetuksessa. Näille oppijoille kieli oli ensisijaisesti opiskelun kohde, ja he keskittyivät tiettyjen muotojen oppimiseen ja ymmärtämiseen. Opettajan kanssa puhuttiin pääasiassa hollantia. Neljä S2-opiskelijaa (saksalainen, brasilialainen, japanilainen ja indonesialainen) osallistuivat Suomessa yliopiston järjestämälle kielikurssille, jolla pääpaino oli merkityksissä. Näille oppijoille kieli oli viestintäväline, ja opettajan kanssa puhuttiin pääasiassa suomea. Aineistona tässä yhden lukuvuoden (yhdeksän kuukautta) kestäneessä tutkimuksessa käytettiin opiskelijoiden vapaamuotoisia vastauksia (kultakin opiskelijalta noin 30 tekstiä, jotka heille annettiin tehtäväksi viikoittain). Suomea vieraana kielenä opiskellut ryhmä tuotti tekstit kurssin kotitehtävinä, kun taas S2-ryhmä kirjoitti ne kurssin ulkopuolella tätä tutkimusta varten.

Tutkimuksessa arvioidaan kielen kompleksisuutta (monimutkaisuutta) ja virheettömyyttä. Syntaktinen kompleksisuus saatiin selville kartoittamalla erilaiset virketyypit, virkkeiden ja lauseiden pituudet morfeemeina ja sijamuotojen käytön. Morfologinen kompleksisuus perustui sanojen pituuteen morfeemeina ja aikamuotojen käyttöön. Syntaktista virheettömyyttä arvioitiin etsimällä nimenomaan sijamuotojen ja sanojen käyttöön sekä sanajärjestykseen liittyvät virheet. Morfologista virheettömyyttä selvitettiin virheellisten sijamuotojen avulla.

Ensin etsittiin vieraan kielen omaksumiseen liittyviä eroja kahden opiskelijaryhmän tuloksista (”ryhmätutkimus”) vertaamalla tuotettujen tekstien kompleksisuutta ja virheettömyyttä lukuvuoden lopussa. Tätä varten selvitettiin ryhmien tulosten keskiarvot. Seuraavaksi keskityttiin syntaktisten ja morfologisten elementtien kehityksessä ja keskinäisessä vuorovaikutuksessa tutkimuksen aikana esiintyneisiin eroihin. Niiden määrittämiseksi ryhmistä valittiin

kaksi ”avainoppijaa”: hollantilainen Kim suomea vieraana kielenä opiskelevien ryhmästä ja saksalainen Andrea S2-ryhmästä (”avainoppijatutkimus”). Heidät valittiin, koska heidän ensikielensä ovat samanlaisia. Lisäksi he kirjoittivat omissa ryhmissään eniten tekstejä ja käyttivät eniten sanoja. Sekä Kim että Andrea ovat myös hyvin kiinnostuneita suomen kieliopista.

Ryhmätutkimuksen tulosten perusteella suomea vieraana ja toisena kielenä opiskelevat ryhmät olivat monessa suhteessa samanlaisia: sekä syntaktinen kompleksisuus (monimutkaisempien virkkeiden käyttö) että morfologinen kompleksisuus (keskimäärin pidemmät virkkeet ja lauseet morfeemeina mitattuna) lisääntyivät tutkimuksen aikana. Ryhmien sijamuotojen käytössä ja niihin liittyvissä kompleksisuus- ja virheettömyystasoissa oli kuitenkin huomattavia eroja. Suomea Alankomaissa vieraana kielenä opiskelevan ryhmän tulokset olivat poikkeuksetta parempia: he kehittyivät enemmän suomen 15 sijan käytössä ja alkoivat käyttää vähemmän nominatiivia ja vastaavasti muita 12 sijaa suhteessa enemmän kuin nominatiivia, genetiiviä ja partitiivia. Toisin sanoen suomea vieraana kielenä opiskelevat käyttivät laajempaa sijamuotovalikoimaa. Kyseinen ryhmä myös käytti morfologisesti monimutkaisempia sanoja. Kaiken kaikkiaan ryhmän jäsenet käyttivät suomea virheettömämmin kuin S2-ryhmä. Erityisesti sijoja he käyttivät taitavammin, mutta myös enemmän sanoja kontekstiin sopivassa merkityksessä ja järjestyksessä. Syynä ryhmien välisiin eroihin voi olla opetuksen toteutustapa, kielenoppimisen konteksti, ensikielen vaikutus (Murakami, 2013; Spoelman, 2013), tutkimuksen lyhyt kesto (Rousse-Malpat & Verspoor, 2012) tai toisen ryhmän heikompi motivaatio kirjoittaa hyvin.

Ennen avainoppijatutkimusta selvitettiin, oliko kahden tutkittavan oppijan (joiden kielitausta oli samanlainen mutta ympäristö erilainen) välillä samoja eroja kuin heidän ryhmiensä välillä, eli olivatko he samanlaisia kuin muut ryhmänsä jäsenet. Suomea vieraana kielenä opiskelleen Kimin kohdalla kaikki tilastollisesti mitatut muuttujat näyttivät vastaavan ryhmän tasoa. Hänen ryhmänsä neljän oppijan samanlaiset tulokset voivat olla samanlaisen kielenoppimiskontekstin, opetuksen toteutustavan, ensikielen, tutkimusajan ja motivaatiotekijöiden yhteisvaikutusta. S2-ryhmän Andrea vaikutti samanlaiselta kuin muut ryhmänsä jäsenet kielen yleisen kompleksisuuden suhteen. Hän kuitenkin erosi selvästi muista sijojen käytön ja niiden virheettömyystason suhteen. Se, että ainoat selvät erot Andrean ja hänen ryhmänsä välillä liittyivät sijamuotoihin, viittaa ensikielen vaikutukseen.

Avainoppijatutkimukseen kuului ennakkotestaus (kumpikin oppija kirjoitti kaksi tekstiä tutkimuksen alussa) ja lopputestaus (kumpikin kirjoitti kaksi tekstiä tutkimuksen lopussa). Seitsemän suomea äidinkielenään puhuvaa S2-asiantuntijaa antoi testeille kokonaisarvosanat neljän kriteerin pohjalta: virkkeiden kompleksisuus, morfologinen kompleksisuus, yleinen virheettömyys ja autenttisuus. Kyseisten kahden oppijan kokonaisarvosanoissa ei ole lainkaan eroja. Seuraavaksi analysoitiin heidän viimeiset kahdeksan tekstiään tilastollisesti. Kävi ilmi, että suomea vieraana kielenä opiskelevan Kimin tekstit olivat yleisesti (syntaksiltaan, morfologialtaan, fonologialtaan ja sanastoltaan) huo-

mattavasti virheettömämpiä. Tulos voi olla yhteydessä opetusmuotoon, kielenoppimisen kontekstiin tai tutkimuksen kestoon. On varsin huomionarvoista, että nämä kaksi niin erilaisissa konteksteissa toimivaa oppijaa ovat loppujen lopuksi perinteisestä näkökulmasta tarkasteltuina niin samanlaisia.

Avainoppijatutkimusta jatkettiin kartoittamalla pitkittäiskehitystä dynaamisesta näkökulmasta tarkoituksena löytää yleisen ja yksilötason kehitysmalleja. Tutkin, kuinka kielellisen järjestelmän osat kehittyvät ajan kuluessa sekä sitä, kehittyykö kahden tutkitun oppijan kielitaito samalla tavalla. Tämän tehtiin tarkastelemalla yksittäisten tasojen kehityskaaria minimi- ja maksimikaavioiden avulla ja suorittamalla tarvittaessa Monte Carlo -analyysejä selkeiden aikaan kytkettyjen muutosten havaitsemiseksi (Van Geert, 2008; Van Geert & Van Dijk, 2002). Ensin havaittiin, että toisen kielen alkeisopiskelijan kielellinen kehittyminen on melko monimuotoinen prosessi, johon sisältyy paljon epälineaarista kehitystä, vaihtelua ja mittaustulosten välisten suhteiden muutoksia (Spoelman & Verspoor, 2010; Verspoor et al., 2012; Murakami, 2013). Lisäksi löydettiin kehitystrendejä, jotka liittyvät oppijien yleisten sekä spesifisten syntaktisten ja morfologisten tasojen kasvuun (Verspoor et al., 2004; Verspoor et al., 2008; Spoelman & Verspoor, 2010; Verspoor et al., 2012; Bulté, 2013). Toinen havainto oli, että huolimatta oppijien suhteellisen samanlaisesta kehityksestä, heidän välillään oli myös useita eroavuuksia. Suomea vieraana kielenä opiskelevan kehityksessä näyttää olevan jyrkempiä piikkejä, joista jotkut ovat (lähes) merkitseviä. Testauksen jälkeen näyttää siltä, että 12 muun sijan käyttö on lähes merkitsevä piirre Kimille (ja se on merkitsevä Sannelle, toiselle suomi vieraana kielenä -ryhmän jäsenelle), ja että sijamuotovirheiden määrä (sijamuotojen virheettömyyssuhde) on Kimille merkitsevä piirre. Ylipäätään Kimin tekstit ovat alusta alkaen kompleksisempia ja virheettömämpiä kuin Andrean; opetuksen erilainen painotus (kielioppi tai merkitys) johtaa kehityskaarten eroihin. Voidaan olettaa, että kielioppipainotteisuus auttaa oppijaa tunnistamaan muodot ja käyttämään niitä aikaisemmin, mutta käyttö saattaa myös johtaa suurempaan määrään sijavirheitä näissä varhaisessa vaiheessa käytetyissä muodoissa. Merkityspainotteisuus taas vaikuttaa niin, että S2-oppija oppii hitaammin tunnistamaan ja käyttämään muotoja, mutta kun niin tapahtuu, hän käyttää niitä suhteellisesti oikeammin kuin toisen ryhmän edustaja. Suurin osa näistä eroista kuitenkin hävisi tutkimuksen loppua kohti – kaksi avainoppijaa näyttivät lähentyvän toisiaan, vaikkakaan eivät vielä täysin: S2-oppijan virheiden (syntaktiset, morfologiset, leksikaaliset ja sanajärjestysvirheet) kokonaismäärä pysyi suurempana yhden lukuvuoden kestäneen tutkimuksen aikana.

Lopuksi keskityttiin avainoppijatutkimuksessa siihen, kuinka kielellisen järjestelmän osat ovat vuorovaikutuksessa keskenään ajan kuluessa – tukevatko ne toisiaan, onko niiden välillä kilpailua vai muuttuuko niiden välinen suhde. Havaittiin, että näillä kahdella oppijalla oli useiden tutkittujen elementtien vuorovaikutussuhde samanlainen. Ensinnä suhde oli samanlainen kompleksisen virkkeiden käytön ja morfeemeina mitatun keskimääräisen virkkeen pituuden välillä (toisiaan tukeva suhde). Tämä viittaa siihen, että useammasta morfeemista koostuva virke on todennäköisemmin kompleksinen virke. Aineistoon

sisältyi kuitenkin myös joitakin keskenään kilpailevia elementtejä, jotka osoittavat, että kyseiset kasvavat muuttujat eivät ole luonnostaan yhteydessä toisiinsa. Toinen samanlainen vuorovaikutussuhde havaittiin morfeemeina mitatun keskimääräisen lause- ja sanapituuden välillä (toisiaan tukeva suhde). Näin ollen lauseen morfeemien määrä on suurempi, kun sanojen morfeemien määrä on suurempi. Myös tässä aineistossa oli mukana muutama keskenään kilpaileva elementti, joten nämäkään kasvavat elementit eivät ole luonnostaan yhteydessä toisiinsa. Kolmas samanlainen vuorovaikutussuhde löytyi kompleksisen virkkeiden käytön ja keskimääräisen lausepituuden välillä, edelleen morfeemeina mitattuna (tukevasta kilpailevaan muuttuva suhde). Molemmat oppijat käyttivät alussa vähemmän kompleksisia virkkeitä ja vähemmän morfeemeja lauseissaan, mutta myöhemmin aineistossa oli havaittavissa täydennystä: virkkeistä tuli monimutkaisempia, kun niihin lisättiin alisteisia lauseita, tai lauseet monimutkaistuivat morfeemeiltaan kompleksisemmista sanoista. Viimeinen samanlainen vuorovaikutussuhde löytyi sijamuotovirheiden ja keskimääräisen sanapituuden välillä (vähitellen vahvemmin tukevaksi muuttuva suhde): mitä pidempiä sanoja käytetään, sitä suurempi on sijamuotovirheiden suhteellinen määrä. Havaittiin myös, että jotkin tasot eivät näytä vaikuttavan lainkaan toisiinsa, koska ne todennäköisesti eivät kilpaile keskenään huomiosta – esimerkiksi imperfektin ja perfektin käyttö vastaan kompleksisuus- ja virheettömyystaso. Näitä tuloksia voidaan varauksella pitää yleisinä suomenkielen oppijia koskevina kehitysmalleina.

Lisäksi kahden tutkitun oppijan välillä havaittiin tutkimuksen kuluessa eroja useiden muuttujien vuorovaikutuksessa. Ensimmäinen ero löytyi sijojen käyttövirheiden (sijojen virheettömyyssuhde) ja sijamuotovirheiden välillä. Kimin sijamuotovirheet vähenivät melko äkillisesti, mutta hänen sijojen käyttövirheensä eivät. Andrealla nämä molemmat virhetasot taas laskivat yhtä aikaa. Tämä osoittaa heidän prosessoivan kyseisiä elementtejä eri tavalla. Kielioppipainotteisesti opiskellut Kim pystyy päättelemällä välttämään kaavamaisia muotovirheitä. Näin ollen hän ”säännön” kerran opittuaan osaa soveltaa sitä useisiin eri yhteyksiin, mikä johtaa nopeaan virheiden vähenemiseen. Päättelemällä hän ei kuitenkaan pysty välttämään idiomaattisiin ilmauksiin liittyviä, tapauskohtaisia käyttövirheitä. Vähemmän kielioppipainotteisesti opiskelevan Andrean on opeteltava sekä sijamuodot että niiden käyttö kokonaisvaltaisemmin ja tapauskohtaisemmin. Muoto ja käyttö yhdistyvät vasta, kun niitä on harjoiteltu tarpeeksi. Toinen havaittu oppijoiden välinen ero, joka tukee ajatusta sijojen käytön erilaisista oppimiskäyristä, liittyy sijojen käyttövirheiden ja keskimääräisen lause- ja sanapituuden väliseen vuorovaikutukseen. Kimillä suhde näiden tasojen välillä oli toisiaan tukeva tutkimusmateriaalin keruun jälkimmäisessä osassa: hän kirjoitti monimutkaisempia lauseita ja sanoja mutta teki samalla enemmän sijojen käyttövirheitä. Tämä osoittaa, että kielioppipainotteisesti opiskeleva ei opi kompleksisuutta ja sijojen käyttöä samanaikaisesti. Andrealla vuorovaikutussuhde samana ajanjaksona oli kilpaileva: hänkin kirjoitti monimutkaisempia lauseita ja sanoja, mutta teki samalla vähemmän sijojen käyttövirheitä. Tästä voidaan päätellä, että vähemmän kielioppipainotteisesti

opiskeleva oppii sanat ja niiden käytön kokonaisvaltaisemmin niiden oikeissa konteksteissa. Sijamuotovirheiden ja keskimääräisen lausepituuden väliltä löytyi vielä yksi vuorovaikutusero. Kim alkoi kirjoittaa monimutkaisempia lauseita, mutta ei hallinnut sijamuotoja aiempaa virheettömämmin. Andrea taas kirjoitti monimutkaisemmin lausetasolla ja hallitsi samalla sijamuodot virheettömämmin. Näiden kahden oppijan välisten, aikaan kytkettyjen vuorovaikutuserojen voidaan hyvin varovaisesti katsoa johtuvan kielenoppimiskontekstista tai opetuksen toteutustavasta.

Toivon tutkimuksen tuovan uutta tietoa ja uusia näkökulmia soveltavan kielitieteen tutkimukseen, koska suomenkielen alkeita vieraana ja toisena kielenä opiskelevien välisiä kielen kehityksen eroja ei aiemmin ole vertailtu vapaamuotoisia tekstejä analysoimalla. Näiden kohderyhmien kehitystä ei ole myöskään tutkittu tietyn ajanjakson kuluessa, samoin kuin ei toisen ja vieraan kielen syntaktista ja morfologista kehitystä eikä kielellisen järjestelmän eri osien erilaisia rooleja ja suhteita.

SAMENVATTING

Dynamische systeemtheorie (Dynamic Systems Theory, DST), de theoretische achtergrond van deze studie, is een vrij nieuwe theorie in toegepaste taalwetenschappen. In een dynamisch perspectief op taal en taalontwikkeling worden patronen bestudeerd die bovenkomen in het gebruik van taal en zijn individuele verschillen en ontwikkelingsprocessen belangrijk. De aanname is dat alle factoren (zoals sociale, taalkundige) continue interacteren en dat variabiliteit nodig is om die taalvormen te selecteren die nodig zijn voor de taalontwikkeling van elke leerder (Thelen & Smith, 1994). In een studie vanuit een dynamisch perspectief is variabiliteit een teken van verandering en ontwikkeling en de leerders en hun context worden gezamenlijk bestudeerd (Verspoor et al., 2011). Op deze manier kan waardevolle informatie over veranderingen in subsystemen worden achterhaald (Van Geert & Van Dijk, 2002).

Het hoofddoel van dit proefschrift is om de ontwikkeling van beginnende leerders in twee verschillende situaties te verkennen: Fins leren als vreemde taal (VT) en als tweede taal (T2). De vier VT leerders (drie uit Nederland en één uit België) studeerden Fins aan een universiteit in Nederland en bij deze studie lag de nadruk op het leren van de Finse grammatica. Voor deze leerders was taal voornamelijk een studieobject waarbij zij zich toelegden op een aantal specifieke vormen met als doel om deze te leren en begrijpen. Tijdens de lessen communiceerden zij voornamelijk in het Nederlands met de docent. De vier T2 leerders (één uit Duitsland, één uit Portugal, één uit Japan en één uit Indonesië) volgden een cursus Fins aan een universiteit in Finland en bij deze cursus lag de nadruk op het kunnen communiceren. Voor deze leerders was taal voornamelijk een manier om je te kunnen uiten. Tijdens de lessen communiceerden zij voornamelijk in het Fins met de docent.

De studie maakt gebruik van geschreven data (ongeveer 30 teksten per leerder, wekelijks geschreven) en beslaat een studiejaar (9 maanden). Voor de VT groep waren de schrijfopdrachten onderdeel van het huiswerk, terwijl ze voor de T2 groep een activiteit buiten de cursus om waren.

De studie onderzoekt variabelen die complexiteit en correctheid uitdrukken. Voor syntactisch complexiteit keken we naar zinstypen, zins- en zinsdeellengte (aantal morfemen) en het gebruik van naamvallen. Voor morfologische complexiteit keken we naar woordlengte (aantal morfemen) en naar tijden. Voor syntactische accuraatheid bestudeerden we alle fouten en specifiek het incorrecte gebruik van naamvallen en woorden en incorrecte woordvolgorde. Voor morfologische accuraatheid keken we naar naamvalvormfouten.

Eerst bestudeerden we verschillen in resultaten in tweede taal ontwikkeling tussen de twee groepen leerders aan het einde van een academisch jaar (groepstudie). We keken hierbij per groep naar de gemiddelden van variabelen die complexiteit en correctheid uitdrukken. Ook

keken we naar verschillen in ontwikkelingspatronen en interacties tussen verscheidene syntactische en morfologische variabelen. Hiervoor verkenden we de data van de twee hoofdpersonen Kim, een Nederlandse uit de VT groep en Andrea, de Duitse uit de T2 groep (individuele studie). Deze twee leerders zijn uitgekozen omdat ze een vergelijkbare T1 hebben. Bovendien schreven ze het grootst aantal teksten en gebruikten ze het grootst aantal woorden van hun groep. Daarnaast zijn Kim and Andrea allebei enorm geïnteresseerd in de Finse grammatica.

De bevindingen van de groepstudie tonen dat de VT en de T2 groep op vele manieren vergelijkbaar zijn: meer syntactische complexiteit (toename van complexe zinnen) en morfologische complexiteit (toename van de gemiddelde zin- en zinsdeellengte in morfemen) door de tijd heen. Echter, de groepen laten significante verschillen zien in het gebruik van naamvallen en van enkele daaraan gerelateerde variabelen die complexiteit en correctheid uitdrukken. De verschillen zijn alle ten gunste van de VT groep: meer ontwikkeling door de tijd heen in het gebruik van de 15 Finse naamvallen, minder gebruik van de nominatief en in overeenstemming daarmee relatief meer gebruik van de 12 naamvallen anders dan de nominatief, genitief en partitief. Met andere woorden, de VT leerders gebruiken meer verschillende naamvallen. Bovendien gebruikt de VT groep morfologisch complexere woorden door de tijd heen. Ook is de VT groep over het algemeen genomen accurater. Met name past de VT groep de naamvallen beter toe en gebruikt ze meer woorden correct (in termen van betekenis en plaats in de context). Oorzaken voor de verschillen tussen de twee groepen kunnen het type instructie zijn, de context waarin de taal wordt geleerd, het T1 effect (Murakami, 2013; Spoelman, 2013), de beperkte tijd waarin het onderzoek heeft plaatsgevonden (Rousse-Malpat and Verspoor, 2012) of minder motivatie om correct te schrijven.

De bevindingen van de individuele studie worden voorafgegaan door het verkennen van de vraag of de twee hoofdpersonen, met een vergelijkbare taalachtergrond maar in verschillende situaties, dezelfde verschillen laten zien als hun groepen; met andere woorden, of ze vergelijkbaar zijn met hun groepsgenoten. Kim (VT) blijkt dat inderdaad te zijn voor alle variabelen die we statistisch onderzoeken. Het feit dat de vier VT leerders zo vergelijkbaar zijn kan worden toegeschreven aan de combinatie van een vergelijkbare context waarin de taal wordt geleerd, type instructie. T1, tijdspanne van het onderzoek heeft plaatsgevonden of motivatie. Andrea (T2) blijkt vergelijkbaar te zijn met haar groepsgenoten waar het gaat over algemene variabelen die complexiteit uitdrukken. Ze verschilt duidelijk in de variabelen die gerelateerd zijn aan naamvallen, in zowel gebruik als in accuraatheid. Het feit dat de enige duidelijke verschillen tussen Andrea en haar groepsgenoten zich bevinden in het gebied van de naamvallen suggereert dat dit een T1 effect is.

Voor de individuele studie begonnen we met de holistische beoordeling van twee teksten voor elke leerder, geschreven in de beginperiode van de studie en twee teksten voor elke leerder, geschreven aan het einde van de studie, door zeven Finse deskundigen op het gebied van tweede

taalverwerving. Zij beoordeelden de teksten op vier criteria: zinscomplexiteit, morfologische complexiteit, algemene accuraatheid en authenticiteit. De holistische beoordelingen tonen geen verschillen tussen de twee hoofdpersonen aan het einde van het studiejaar. Vervolgens voerden we enkele statistische analyses uit op de laatste acht teksten van de twee hoofdpersonen. We vonden één verschil tussen de twee leerders; de VT leerder is significant accurater (in syntaxis, morfologie en woordenschat). De bevindingen kunnen in verband staan met het type instructie, de context waarin de taal wordt geleerd of de beperkte tijd waarin het onderzoek heeft plaatsgevonden. Het is vrij opvallend dat de twee hoofdpersonen in zulke verschillende omstandigheden zo gelijk zijn aan het eind, bekeken vanuit een traditioneel perspectief.

We vervolgden de individuele studie met het verkennen van de ontwikkeling door de tijd heen vanuit een dynamisch perspectief. Het doel van een dergelijke analyse is het vinden van algemene en individuele ontwikkelingspatronen. We onderzochten hoe verschillende subsystemen van de taal zich ontwikkelen door de tijd heen en vervolgens of the twee hoofdpersonen hun taal door de tijd heen op vergelijkbare wijze ontwikkelen. We deden dit door in min-max grafieken te kijken naar alle bewegingen van variabelen. Als het nodig was, voerden we Monte Carlo analyses uit om duidelijke kenteringen te kunnen constateren (Van Geert, 2008; Van Geert & Van Dijk, 2002). Allereerst vonden we dat taalontwikkeling voor beginnende tweede taalleerders een tamelijk variabel proces is met een grote hoeveelheid niet-lineaire ontwikkeling, variatie en veranderende relaties tussen de variabelen (Spoelman and Verspoor, 2010; Verspoor et al., 2012; Murakami, 2013). Bovendien vonden we ontwikkelingstendensen van groei in alle algemene en specifieke syntactische en morfologische variabelen (Verspoor et al., 2004; Verspoor et al., 2008; Spoelman and Verspoor, 2010; Verspoor et al., 2012; Bulté, 2013). Ten tweede, ondanks het feit dat de hoofdpersonen zich nogal vergelijkbaar ontwikkelen in het leren van de Finse taal, tonen ze ook een aantal verschillen. De data van de VT leerder vertoont steilere pieken waarvan enkele bijna significant zijn. Na testen blijkt dat het gebruik van de 12 overige naamvallen bijna significant is voor Kim (en voor Sanne, een VT groepsgenoot, significant) en de naamvalvormfouten (verhoudingsgetal) significant. Kim schrijft complexer en accurater dan Andrea in het begin; het verschil in nadruk op grammatica en nadruk op communicatie resulteert in enkele verschillen in de ontwikkelingsbanen. We mogen aannemen dat het effect van de nadruk op grammatica de leerders helpt om de vormen te ontdekken en op te merken en hen in staat te stellen ze eerder te gebruiken, maar het gebruik kan in eerste instantie ook leiden tot meer naamvalvormfouten. Het effect van de nadruk op communicatie is dat het langer duurt dat de T2 leerder de naamvalvormen opmerkt en toepast, maar als dit eenmaal het geval is dan worden ze in verhouding correcter toegepast dan in de VT situatie. Echter, het grootste deel van de verschillen verdwijnt richting het einde van de periode; de twee hoofdpersonen lijken samen te komen, hoewel nog niet helemaal; het totaal

aantal fouten (in woordenschat, syntaxis, morfologie en woordvolgorde) blijft hoger voor de T2 leerder tijdens het eerste studiejaar.

Tot slot onderzochten we in de individuele studie hoe verscheidene subsystemen in de taal interacteren door de tijd heen en of ze elkaar ondersteunen, elkaar beconcurreren of dat hun relatie verandert met de tijd. Voor een aantal variabelen blijkt de interactie door de tijd heen vergelijkbaar te zijn voor de twee leerders: ten eerste, tussen het gebruik van complexe zinnen en de gemiddelde zinslengte in morfemen (ondersteuning). Dit suggereert dat een zin die bestaat uit meerdere morfemen waarschijnlijk een complexe zin is. De data laat echter een aantal concurrerende punten zien, wat suggereert dat deze variabelen elkaar niet zondermeer ondersteunen. We vonden een tweede vergelijkbare interactie tussen de gemiddelde zinsdeellengte en woordlengte in morfemen (ondersteuning). Dit suggereert dat het aantal morfemen in een zinsdeel groter is wanneer het aantal morfemen in de woorden groter is. Echter, opnieuw laat de data een aantal concurrerende punten zien, wat suggereert dat deze variabelen elkaar niet zondermeer ondersteunen. We vonden een derde vergelijkbare interactie tussen het gebruik van complexe zinnen en de gemiddelde zinsdeellengte in morfemen (veranderende relatie, van ondersteuning tot concurrentie). Beide leerders beginnen met minder complexe zinnen en minder morfemen per zinsdeel waarna de data verderop in de tijd laat zien dat de zinnen complexer worden met bijzinnen en dat de zinsdelen complexer worden met meer complexe woorden door het gebruik van meer morfemen. De laatste vergelijkbare interactie vonden we tussen de naamvalvormfouten en de gemiddelde woordlengte in morfemen (veranderende relatie, meer ondersteuning met de tijd). Dit suggereert dat hoe langer de woorden worden, des te meer naamvalfouten er in verhouding in worden gemaakt. We vonden ook dat enkele variabelen geen duidelijk effect op elkaar hebben omdat zij waarschijnlijk niet strijden om aandacht: het gebruik van de verleden tijd versus variabelen die complexiteit en correctheid uitdrukken. Deze bevindingen kunnen voorzichtig worden beschouwd als algemene ontwikkelingspatronen voor leerders van de Finse taal.

We vonden ook dat de interactie door de tijd heen verschillend is voor verscheidene variabelen voor de twee leerders: ten eerste, tussen de naamvalgebruiksfouten (verhoudingsgetal) en de naamvalvormfouten. Kim's naamvalvormfouten nemen vrij plotseling af maar haar naamvalgebruiksfouten niet, terwijl Andrea een plotseling afname voor beide variabelen op hetzelfde moment laat zien. Dit suggereert een verschil in verwerking van deze twee subsystemen in de twee leerders. De VT leerder Kim kan naamvalvormfouten, die nogal schematisch zijn, onder de knie krijgen, dus wanneer eenmaal de "regel" is ontdekt kan het worden toegepast op de naamvallen, wat leidt tot een plotselinge stop. Ze kan de meer idiomatische naamvalgebruiksfouten niet op dezelfde wijze onder de knie krijgen. De T2 leerder Andrea moet zowel het gebruik als de vorm van de naamvallen door wat ze aan informatie binnenkrijgt op holistische wijze leren. Pas wanneer ze vaak genoeg iets in de Finse taal tot zich heeft gekregen zal ze het gebruik en de vorm van naamvallen onderling

met elkaar verbinden en de naamvallen op die manier toepassen. We vonden een tweede verschil in interactie, en hiermee weer een ondersteuning van het idee dat er verschil is in leertrajecten met betrekking tot het naamvalgebruik, tussen de naamvalgebruiksfouten en de gemiddelde zinsdeellengte en de gemiddelde woordlengte in morfemen. Kim laat een ondersteunende interactie zien tussen de variabelen in het tweede deel; ze schrijft complexere zinsdelen en woorden maar maakt ook meer naamval gebruiksfouten. Dit suggereert dat de VT leerder complexiteit en het gebruik van naamvallen niet tegelijkertijd aanleert. Andrea haar concurrerende interactie in diezelfde periode geeft aan dat zij ook complexere zinsdelen en woorden schrijft, maar tegelijkertijd minder naamval gebruiksfouten maakt. Dit suggereert dat de T2 leerder de woorden en hun gebruik holistisch in de context leert. We vonden een derde verschillende interactie tussen de naamvalvormfouten en de gemiddelde zinsdeellengte in morfemen. Weer schrijft Kim complexere zinsdelen maar wordt tegelijkertijd niet accurater in de naamvalvormen terwijl dat laatste bij Andrea wel het geval is. De verschillen in de interacties door de tijd heen in de twee leerders kunnen zeer voorzichtig worden beschouwd als verschillen die veroorzaakt worden door het type instructie of de context waarin de Finse taal door de leerders wordt geleerd.

Met deze verkenning hopen we bij te dragen aan het veld van toegepaste taalwetenschappen; ten eerste omdat er tot op de dag van vandaag nog geen studies zijn waarin verschillen in de taalontwikkeling van beginnende leerders van de Finse taal als vreemde en als tweede taal in het vrije schrijven worden vergeleken. Bovendien is er nog geen onderzoek gedaan naar de ontwikkeling op de lange termijn van beginnende leerders van de Finse taal, noch naar de syntactische en morfologische ontwikkeling van een vreemde of tweede taal, noch naar de verschillende rollen en relaties van subsystemen in taal.

REFERENCES

Aalto, E., Tukia, K. & Männikkö, H. (2003). *Matkalla jossakin suomessa*. Jyväskylä: Jyväskylän yliopiston kielikeskus.

Ahonen, L. M. T. (2006). *Kuulostaa hyvältä*. Groningen.

Anhava, J. (1998). *Maailman kielet ja kielikunnat*. Gaudeamus.

Bates, E. (1997). Origins of language disorders: A comparative approach. *Developmental Neuropsychology*, *13*(3), 447-476.

Behrens, H. (2009). Usage-based and emergentist approaches to language acquisition. *Linguistics*, *47*(2), 383-411.

Peat, D. F., & Briggs, J. (1989). The Turbulent Mirror: An Illustrated Guide to Chaos Theory & the Science of Wholeness.

Bulté, B. (2013). *The development of complexity in second language acquisition, a Dynamic Systems approach.* Vrije Universiteit Brussel.

Bulté, B. & Housen, A. (2012). Defining and operationalising L2 complexity. In A. Housen, F. Kuiken & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: complexity, accuracy and fluency in SLA,* 21-46. Amsterdam: John Benjamins.

Bybee, J. (2008). Usage-based grammar and second language acquisition. In P. Robinson & N.C. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition,* 216-236. New York: Routledge.

Caspi, T. (2010). *A dynamic perspective on second language development*. Rijksuniversiteit Groningen.

Coleman, J.A. (1997). Residence abroad within language study. *Language teaching, 30*(1), 1-20.

Collentine, J. (2004). The effects of learning contexts on morphosyntactic and lexical development. *Studies in Second Language Acquisition*, *26*(02), 227-248.

Dahl, Ö. (2008). Kuinka eksoottinen kieli suomi on. *Raukko, Jarno (suom.)*, 545-559.

d'Anglejan, A., & Renaud, C. (1985). Learner characteristics and second language acquisition: A multivariate study of adult immigrants and some thoughts on methodology. *Language Learning*, *35*(1), 1-19.

De Bot, K., Lowie, W. & Verspoor, M. (2005). *Second language acquisition: An advanced resource book*. London: Routledge.

De Bot, K., Lowie, W. & Verspoor, M. (2007). A Dynamic Systems Theory approach to second language acquisition. *Bilingualism: Language and Cognition 10*(1), 7-21.

DeKeyser, R. M. (2005). What Makes Learning Second-Language Grammar Difficult? A Review of Issues. *Language Learning*, *55*(S1), 1-25.

Dijkstra-Pruim, P. 2003. Speaking, writing, and explicit rule knowledge: toward an understanding of how they interrelate. *Foreign language annals, 36*(1), 66-76.

Ellis, N. C. (2007). Dynamic systems and SLA: The wood and the trees. *Bilingualism: Language and Cognition*, *10*(01), 23-25.

Ellis, R. (1989). Are classroom and naturalistic acquisition the same? *Studies in Second Language Acquisition*, *11*(03), 305-328.

Ellis, R. (1994). *The study of second language acquisition*. Oxford University Press.

Ellis, R. (2001). Introduction: Investigating Form-Focused instruction. *Language learning*, *51*(s1), 1-46.

Green, P. S., & Hecht, K. (1992). Implicit and explicit grammar: An empirical study. *Applied Linguistics*, *13*(2), 168-184.

Grünthal, W. (1941). *Itämerensuomalaisten kielten yksikön nominatiivi objektin edustajana aktiivin yhteydessä. Lauseopillinen tutkimuskoe*. Suomalaisen Kirjallisuuden Seuran Toimituksia 218. Helsinki.

Hahmo, S.L. & Liebe, Chr. (1998). *Finse grammatica*. Groningen.

Haley, M. H., Rentz, P., & Education, F. C. P. S. A. (2002). Applying SLA research and theory to practice: What can a teacher do. *TESL-EJ*, *5*(4), 1-9.

Harley, B. & Swain, M. (1984). The Interlanguage of immersion and its implications for second language teaching. In A. Davies, C. Criper & A.P.R. Howatt (Eds.), *Interlanguage,* 291-311. Edinburgh, Scotland: Edinburgh University Press.

Hasselblatt, C. Th. (1998). Zum Problem des finnougrischen Objekts. – UAJb NF 15, 23-42.

Herdina, P., & Jessner, U. (2002). *A dynamic model of multilingualism: Perspectives of change in psycholinguistics* (Vol. 121). Multilingual Matters.

Holmberg, A., & Nikanne, U. (Eds.). (1993). *Case and other functional categories in Finnish syntax* (Vol. 39). Walter de Gruyter.

Honko, M. (2013). *Alakouluikäisten leksikaalinen tieto ja taito: toisen sukupolven suomi ja S1-verrokit.* Tampereen yliopisto.

Housen, A. (2002). A corpus-based study of the L2-acquisition of the English verb system. *Computer learner corpora, second language acquisition and foreign language teaching*, *6*, 2002-77.

Kaivapalu, A. (2005). *Lähdekieli kielenoppimisen apuna*. Jyväskylän yliopisto.

Kajander, M. (2013). *Suomen eksistentiaalilause toisen kielen oppimisen polulla.* Jyväskylän yliopisto.

Karlsson, F. (1983). *Suomen kielen: äänne-ja muotorakenne*. Helsinki: Werner Söderström.

Karlsson, F. (1999). *Finnish: An essential grammar*. Routledge.

Kecskes, I., & nde Papp, T. (2000). *Foreign language and mother tongue*. Psychology Press.

Kiparsky, P. (2005). Absolutely a matter of degree: The semantics of structural case in Finnish. *handout of talk presented at CLS*, *41*.

Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon Press.

Lagus, K., Creutz, M. & Virpioja, S. (2005). Latent linguistic codes for morphemes using independent component analysis. In A. Cangelosi, G. Bugmann, & R. Borisyuk (Eds.), *Modeling language, cognition and action: Proceedings of the Ninth Neural Computation and Psychology Workshop (NCPW9)*. Plymouth, England.

Langacker, R. W. (1987). *Foundations of cognitive grammar: Theoretical prerequisites* (Vol. 1). Stanford university press.

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

Larsen-Freeman, D. (2002). Language acquisition and language use form a chaos/complexity theory perspective. In C. Kramsch (Ed.), *Language acquisition and language socialization: Ecological perspectives*, 33–46. Continuum.

Larsen-Freeman, D. & Long, M. H. (1991). *An introduction to second language acquisition research.* London: Longman.

Lesonen, S. (2013). Suomen kielen lingvististen piirteiden sijoittuminen Eurooppalaisen viitekehyksen taitotasoille.

Leow, R. P. (2001). Attention, awareness, and foreign language behavior. *Language Learning*, *51*(s1), 113-155.

Lightbown, P. M. (2000). Anniversary article. Classroom SLA research and second language teaching. *Applied linguistics*, *21*(4), 431-462.

Long, M. H. (1990). The least a second language acquisition theory needs to explain. *Tesol Quarterly*, *24*(4), 649-666.

Long, M. H. (1991). Focus on form: A design feature in language teaching methodology. In K. De Bot, R. Ginsberg & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspective*, 39-52. Amsterdam: John Benjamins.

Lyster, R. (1987). Speaking immersion. *The Canadian Modern Language Review 43*(4), 701-717.

MacWhinney, B. (1976). Hungarian research on the acquisition of morphology and syntax. *Journal of Child Language*, *3*(03), 397-410.

MacWhinney, B. (1997). Second language acquisition and the competition model. In A. De Groot & J. Kroll (Eds.), *Tutorials in Bilingualism: Pycholinguistic Perspectives*, 113-142. Mawwah, NJ: Lawrence Erlbaum.

MacWhinney, B. (2008). Enriching CHILDES for morphosyntactic analysis. In H. Behrens (Ed.), *Corpora in Language Acquisition Research: History, Methods and Perspectives,* 165-198. Benjamins, Amsterdam.

Marchman, V. & Thal, D. (2004). Words and grammar. In D. I. Slobin & M. Tomasello (Eds.). *Beyond nature-nurture: Essays in honor of Elizabeth Bates,* 141-164. Mahwah, NJ: Lawrence Erlbaum.

Martin, M. (1995). *The map and the rope: Finnish nominal inflection as a learning target*. Jyväskylän yliopisto.

Martin, M., Mustonen, S., Reiman, N., & Seilonen, M. (2010). On becoming an independent user. *Communicative Profiency and Linguistic Development: Intersections between SLA and Language Testing Research. Eurosla Monographs*, *1*, 57-79.

McKirahan, R. (2003). Presocratic Philosophy. In Chr. Shields (Ed.), *The Blackwell Guide To Ancient Philosophy,* 5-26. Blackwell Publishing Ltd.

Murakami, A. (2013). *Individual variation and the role of L1 in the L2 development of English grammatical morphemes: Insights from learner corpora.* University of Birmingham.

Nassaji, H., & Fotos, S. (2004). 6. Current developments in research on the teaching of grammar. *Annual Review of Applied Linguistics*, *24*, 126-145.

Newson, M., & Maunola, V. (2006). Word order in Finnish: Whose side is the focus on. *The Even Yearbook*, *7*, 1-32.

Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language learning*, *50*(3), 417-528.

Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics*, *30*(4), 555-578.

O'Grady, W. (2008). The emergentist program. *Lingua*, *118*(4), 447-464.

Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied linguistics*, *24*(4), 492-518.

Pavesi, M. (1986). Markedness, discoursal modes, and relative clause formation in a formal and an informal context. *Studies in Second Language Acquisition*, *8*(01), 38-55.

Pavlenko, A. & Lantolf, J. P. (2000). Second language learning as participation and the (re)construction of selves. In J.P. Lantolf (Ed.), *Sociocultural theory and second language learning*, 155-177. Oxford University Press.

Ringbom, H. (1987). *The role of first language in foreign language acquisition*. Clevedon: Multilingual Matters.

Rousse-Malpat, A., & Verspoor, M. (2012). Measuring effectiveness in Focus on Form versus Focus on Meaning. *Dutch Journal of Applied Linguistics*, *1*(2).

Schmidt, R. W. (1990). The Role of Consciousness in Second Language Learning. *Applied linguistics*, *11*(2), 129-158.

Schmidt, R.W. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. *AILA Review, 11*, 11-26.

Schmidt, R. W. & Frota, S. N. (1986). Developing basic conversational ability in a second language: a case study of an adult learner of Portugese. In R.R. Day (Ed.), *Talking to Learn: Conversation in second language acquisition,* 237-326. Rowley, MA: Newbury House.

Schoonen, R., Gelderen, A. V., Glopper, K. D., Hulstijn, J., Simis, A., Snellings, P., & Stevenson, M. (2003). First language and second language writing: The role of linguistic knowledge, speed of processing, and metacognitive knowledge. *Language learning*, *53*(1), 165-202.

Schot-Saikku, P. (1990). *Der Partitiv und die Kasusalternation: Zum Fall Partitiv in der Finnischen Syntax*. Hamburg: Buske.

Seilonen, M. (2013). *Epäsuora henkilöön viittaaminen oppijansuomessa.* Jyväskylän Yliopisto.

Spoelman, M. & Verspoor, M. (2010). Dynamic Patterns in Development of Accuracy and Complexity: A Longitudinal Case Study in the Acquisition of Finnish. *Applied Linguistics 31*(4), 532-553.

Spoelman, M. (2013). *Prior linguistic knowledge matters. The use of the partitive case in Finnish learner language.* Oulun yliopisto.

Thelen, E. & Smith, L.B. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT Press.

Thelen, E. & Smith, L.B. (1998). Dynamic Systems Theories. In W. Damon & R.M. Lerner (Eds.), *Handbook of Child Psychology*, (5th ed.), 563-634. New York: Wiley.

Van Dijk, M. (2004). *Child language cuts carpers: Variability and ambiguity in early child development*. Rijksuniversiteit Groningen.

Van Dijk, M. & Van Geert, P. (2007). Wobbles, humps and sudden jumps: A case study of continuity, discontinuity and variability in early language development. *Infant and Child Development 16*(1), 7-33.

Van Dijk, M., Verspoor M. & Lowie, W. (2011). Variability and DST. In M.Verspoor, K. De Bot & W. Lowie (Eds.), *A dynamic approach to second language development: methods and techniques,* 55–84.

Van Geert, P. (1991). A dynamic systems model of cognitive and language growth. *Psychological Review 98*(1), 3.

Van Geert, P. (1994). *Dynamic systems of development: Change between complexity and chaos*. New York: Harvester.

Van Geert, P. (2008). The dynamic systems approach in the study of L1 and L2 acquisition: An introduction. *The Modern Language Journal*, *92*(2), 179-199.

Van Geert, P. & Steenbeek, H. (2005). Explaining after by before: Basic aspects of a dynamic systems approach to the study of development. *Developmental Review 25*(3), 408-442.

Van Geert, P., & Van Dijk, M. (2002). Focus on variability: New tools to study intra-individual variability in developmental data. *Infant Behavior and Development*, *25*(4), 340-374.

Verspoor, M. & Sauter, K. (2000). *English sentence analysis: An introductory course*. Amsterdam/Philadelphia: John Benjamins.

Verspoor, M., De Bot, K. & Lowie, W. (2004). Dynamic Systems Theory and Variation: a Case study in L2 writing. In H. Aertsen, M. Hannay & R. Lyall (Eds.), *Words in their places. A Festschrift for J. Lachlan Mackenzie,* 407-421. Amsterdam: Free University Press.

Verspoor, M., De Bot, K. & Lowie, W. (2011). *A dynamic approach to second language development: methods and techniques*. Amsterdam: John Benjamins.

Verspoor, M., Lowie, W., & Van Dijk, M. (2008). Variability in second language development from a dynamic systems perspective. *The Modern Language Journal*, *92*(2), 214-231.

Verspoor, M., Schmid, M.S. & Xu, X. (2012). A dynamic usage based perspective on L2 writing. *Journal of Second Language Writing 21*(3), 239-263.

Vihonen, S. (1978). *Suomen kielen oppikirja 1600-luvulla*. Studio Philologica Jyväskyläensia 11. Jyväskylä.

Vilkuna, M. (1989). *Free word order in Finnish: Its syntax and discourse functions* (Vol. 500). Helsinki: Suomalaisen kirjallisuuden seura.

Volodin, A.P. (1997). Finnisch-ugrische (uralische) Sprachen in Kontakt. Zur allgemeinen Fragestellung. S.L. Hahmo et alii (Hgg). *Finnisch-ugrische sprachen in Kontaky. Vorträge des Symposiums aus Anlaß des 30-jährigen Bestehens der finnougristik an der Rijksuniversiteit Groningen 21. -23. November 1996*. Maastricht, 35-43.

Von Farkas, J. (1956). Der Genitiv und der Akkusativ in der uralischen Grundsprache. *UAJb, 28,* 5-17.

Waldrop, M. (1992). *Complexity: The Emerging Science at the Edge of Order and Chaos*. New York, Simon and Schuster.

Wiik, K. (1988). Kolme ja puoli vuosisataa suomalaista ja virolaista kielopinkirjoitusta: nominien taivutus. In P. Nemvalts & P. Rintala (Eds.), *Lähivertaluja* 3, 11-26. Publications of the Department of Finnish and General Linguistics of the University of Turku 14.

Williams, J. N. (2003). Implicit learning of form-meaning connections in vocabulary and grammar. *Eurosla Yearbook 13* (in press).

APPENDIX A: List of abbreviations

Cases

abess. = abessive
ablat. = ablative
acc. = accusative
adess. = adessive
allat. = allative
com. = comitative
elat. = elative
ess. = essive
gen. = genitive
illat. = illative
iness. = inessive
instr. = instructive
nom. = nominative
part. = partitive
transl. = translative

Example sentences

adj.	= adjective
attr.	= attributive
cmpw	= compound word
comp	= compound sentence
compl	= complex sentence
comp.compl	= compound-complex sentence
gramm.subj.	= grammatical subject
log.subj.	= logical subject
pl.	= plural
poss.th.	= possessed thing
pron.	= pronoun
que.cl.	= question clitic
sing.	= singular
subj.	= subject
subst.	= substantive
vb.	= verb
2.pl.	= second person plural
-V-	= vowel, the same as the nearest preceding vowel
/	= boundary (between stem and endings and between endings)
// . ; : ? !	= boundaries between two clauses
19mrph.	= 19 morphemes
(9:4)	= the item consists of nine words of which four involve the main subject

(av)ClMrph = average clause length in morphemes
(av)WlMrph = average word length in morphemes
ClMrph1/13 = 1 clause, consisting of 13 morphemes
ClMrph2/13(4-9:6.5) = two clauses, together consisting of 13 morphemes, four in the first clause, nine in the second clause, with a mean of 6.5 morphemes per clause
SlMrph2/25:12.5 = two sentences, together consisting of 25 morphemes, with a mean of 12.5 morphemes per sentence
WlMrph8/10(1.25) = eight words, together consisting of 10 morphemes, with an average word length of 1.25
[past4mrph] = past tense, consisting of 4 morphemes
[past2/2mrph] = two words (together the past tense), both consisting of two morphemes
Andrea-t4s6/mp17 = written by Andrea, text 4, sentence 6, measure point 17

Case use errors
[gov] = incorrect application of government
[d.obj] = incorrect case for the direct object
[subj] = incorrect case for the grammatical subject or incorrect application of a sentence with no subject
[pr.nom] = incorrect case for the predicate nominal
[infl] = incorrect inflection
[congr] = incorrect application of congruence

Case form errors
[cons.gr] = incorrect consonant gradation
[form] = incorrect stem form

Remaining errors
[vow.h] = incorrect application of vowel harmony
[w.miss] = word missing
[poss.suff] = incorrect application of the possessive suffix
[w.cl] = incorrect use of word class
[lex] = lexical error (contextual)
[w.ord] = word order error (contextual)
[tns] = incorrect used tense (contextual)
[noncns] = non consisting word
[no subj] = no subject should be in the sentence

APPENDIX B: Checklist (consolidated version)

Case use errors

[01](synt.) incorrect application of government (olen ylpeä *sinulle*)
[02](synt.) incorrect case for the direct object (minä ostan *kirja*)
[03](synt.) incorrect case for the grammatical subject (keittiössä on *kahvi;* lempiruokani on *kalaa*) or incorrect application of a sentence with no subject (täällä on *tylsä*)
[04](synt.) incorrect case for the predicate nominal (kahvi on *hyvä;* vaatteet eivät ole *tärkeää*)
[05](synt.) incorrect inflection (hän soittaa *palomiehet*)
[06](synt.) incorrect application of congruence (*kallis* kirjat)

Case form errors

[01](morph.) incorrect consonant gradation (Helsin*k*issä)
[02](morph.) incorrect stem form (asi*a*ita)

APPENDIX C: Writing tasks 1 to 39

Text	Topic	FL Questions	L2 Questions	Kim mp	An-drea mp
1	Myself	Kuka sinä olet? Mitä sinä opiskelet? Missä sinä opiskelet?	idem	1	1
2	Place of residence	Missä sinä asut? Missä maassa? Missä kaupungissa tai kylässä? Millä kadulla tai tiellä?	idem	2	2
3	Going to the grocery	Mitä sinä ostat, kun sinä käyt kaupassa?	idem	3	3
4	The Finnish language	Onko suomi vaikea kieli? Mikä on vaikeaa, mikä on helppoa? Miksi sinä opiskelet suomea?	idem	4	4
5	My house	Asutko sinä keskustassa vai kaukana keskustasta? Onko se kiva asunto? Onko sen lähellä kauppa? Asutko yksin vai yhdessä jonkun kanssa?	idem	5	5
6	Weekend activities	Mitä sinä teet ensi viikonloppuna (lauantaina ja sunnuntaina)?	idem	6	6
7	Travelling	Asutko sinä kaukana yliopistosta? Voitko sinne kävellä vai kuljetko bussilla tai polkupyörällä?	idem	7	7
8	Cooking	Laitatko itse ruokaa vai syötkö usein ravintolassa? Onko sinulla jokin lempiruoka? Mikä se on?	idem	8	x
9	Bying clothes	Tarvitsetko sinä nyt uuden takin ja uudet kengät? Kenen kanssa sinusta on hauska ostaa vaatteita?	idem	x	8
10	Fashion	Onko sinulla hyvä maku? Onko sinusta tärkeää tietää, mikä on muotia? Miksi?	idem	9	9
11	Spare time/ hobbies	Mitä sinä teet, jos olet väsynyt tai vähän masentunut? Mitä sinä teet mielelläsi vapaa-aikana?	idem	10	10
12	Bying presents	Onko sinusta hauska vai vaikea ostaa lahjoja? Miksi?	idem	11	11
13	Poetry	Luetko sinä runoja? Jos sinä luet, niin kirjoita siitä, miksi pidät runoista. Jos et, niin kirjoita siitä, miksi et pidä runoista.	idem	12	12
14	Childhood	Kerro omasta perheestäsi. Missä sinä asuit lapsena? Millainen olit pienenä?	idem	13	13
15	Moods	Oletko sinä joskus erilainen kuin tavallisesti? Kerro kerrasta, kun sinulla oli siihen jokin hyvä syy.	idem	x	14
16	What if…	Mitä sinä tekisit kun viikon-loppuna sinä heräisit jo kello seitsemän kamalaan meteliin?	idem	14	15
17	Literature	Onko kirjallisuus sinusta mielenkiintoista? Minkälaisesta	idem	15	16

		kirjallisuudesta sinä pidät?			
18	Sauna	Oletko sinä ollut suomalaisessa saunassa? Millainen sauna se oli?	idem	16	17
19	Finland	Mitä sinä pidät Suomesta ja suomalaisista? Mikä sinusta Suomessa on erilaista kuin kotimaassasi?	idem	17	18
20	FL: Indepen-dency of Finland L2: Independent day	Milloin Suomi tuli itsenäiseksi? Millainen Suomen tilanne oli sitä ennen?	Mitä sinä tiedät siitä, miten Suomessa vietetään itsenäisyys-päivää?	18	19
21	FL: Independent day L2: Christmas	Mitä sinä tiedät siitä, miten Suomessa vietetään itsenäisyyspäivää?	Mitä sinä tiedät suomalaisesta joulusta? Vietetäänkö sitä eri tavalla vai samalla tavalla kuin joulua vietetään sinun koti-maassasi?	19	20
22	Lapland	Oletko sinä käynyt joskus Lapissa?	idem	20	21
23	New Year's eve	Miten sinun kotimaassasi vietetään uutta vuotta?	idem	21	22
24	Midsum-mer night	Vietetäänkö sinun kotimaassasi juhannusta? Mitä sinä tiedät juhannuksenvietosta Suomessa?	idem	22	23
25	Folk music	Oletko sinä kuullut suomalaista kansanmusiikkia? Mitä, missä, milloin?	idem	23	24
26	Jean Sibelius	Tiedätkö sinä kuka Jean Sibelius oli? Mikä, sinusta, hänen merkityksensä oli Suomelle?	idem	24	x
27	Tango	Tiedätkö sinä jotakin suomalaisesta tangosta ja Suomen tanssikulttuurista yleensä?	idem	25	25
28	Sickness	Milloin sinä olit sairas vai etkö sinä ole koskaan sairas?	idem	26	26
29	Saami people	Mitä sinä tiedät saamelaisista ja saamenkielestä?	idem	27	x
30	Name day	Suunnittele itsellesi nimipäiväjuhlat. Missä ne vietettäisiin? Keitä sinä kutsuisit sinne?	idem	28	27
31	Arts	Katso Suomen taiteen kultakauden maalauksia ja valitse taiteilija, jonka tauluista sinä pidät eniten. Kerro, miksi sinä pidät hänen maalauksistaan.	idem	29	28
32	Internship	Haluaisitko sinä mennä	idem	30	29

		harjoittelijaksi Suomeen? Millaisen harjoittelupaikan haluaisit? Miksi?			
33	FL: Lönnrot L2: Picknick	Mitä tiedät Elias Lönnrotin lapsuudesta ja nuoruudesta?	Kuvittele, että sinä olet järjestä-mässä vappupik-nikkiä ystävillesi. Säätiedotuk-sen mukaan vapun-päivänä voi sataa. Mitä teet? (=FL 34)	31	30
34	FL: Picknick L2: Summer residence	Kuvittele, että sinä olet järjestämässä vappupiknikkiä ystävillesi. Säätiedotuksen mukaan vapunpäivänä voi sataa. Mitä teet?(=L2 33)	Kuvittele, että sinä saat valita itsellesi kesäpaikan Suomessa. Missä se olisi, millainen se olisi?(=FL 38)	32	31
35	FL: Saarinen L2: Future plans	Kuka oli Eliel Saarinen?	Kirjoita siitä, miten sinä kuvittelet sinun elämäsi jatkuvan tästä eteenpäin.	33	32
36	Finnish novel	Kuvittele, mitä novellissa tämän jälkeen tapahtuu.	-	34	x
37	Engel	Ota selvää siitä, kuka Carl Ludvig Engel oli?	-	35	x
38	Summer residence	Kuvittele, että saat valita itsellesi kesäpaikan Suomessa. Missä se olisi, millainen se olisi? (=L2 34)	-	36	x
39	Future main characters (course book)	Kirjoita siitä, miten kuvittelet Jutan ja Anssin tarinan jatkuvan tästä eteenpäin.	-	37	x

APPENDIX D: The actual written words per participant

text	FL	Kim	Sanne	Cleo	Annet	L2	Andrea	Chiya	Bowo	Clara
1		70	17	x[36]	56		67	16	14	11
2		75	30	x	44		38	9	9	9
3		78	40	x	40		72	16	x	24
4		78	48	x	x		64	21	33	53
5		98	50	x	x		99	32	32	22
6		79	49	x	x		94	32	23	19
7		78	39	71	x		57	25	23	17
8		93	45	78	52		x	22	x	16
9		x	60	72	x		97	32	53	13
10		80	53	66	43		98	55	55	76
11		77	47	47	76		94	52	63	52
12		80	44	66	39		89	50	55	63
13		87	57	63	26		89	53	54	71
14		86	62	78	78		100	72	70	57
15		x	64	65	36		93	51	53	53
16		84	54	60	x		89	54	51	51
17		71	55	58	x		87	55	60	51
18		103	41	43	x		89	55	59	54
19		114	102	x	x		86	62	64	60
20		119	90	94	35		96	65	55	34
21		107	79	92	51		92	81	66	50
22		103	82	x	100		93	70	x	49
23		125	60	95	54		84	54	56	52
24		106	82	x	71		92	52	58	x
25		107	105	74	66		90	52	78	43
26		119	81	99	40		x	50	53	96
27		138	69	94	36		101	76	55	44
28		121	95	96	x		98	66	65	52
29		129	67	98	x		x	53	60	x
30		119	104	100	27		87	50	63	x
31		115	71	x	38		93	55	55	48
32		119	77	x	x		80	69	51	37
33		109	85	x	53		87	57	59	40
34		132	98	x	x		100	55	59	44
35		135	100	99	x		93	62	62	48
36		138	X	x	x					
37		121	103	x	x					
38		98	86	x	x					
39		106	104	x	x					

[36] x = the participant did not do this writing task.

APPENDIX E: Correlation coefficients

	Kim (FL) / Andrea (L2)
Syntactic complexity – Syntactic complexity	
Total of compound, complex and comp-complex sentences - Average sentence length in morphemes	0.67, d[37] 0.65 / 0.71, d 0.64
Total of compound, complex and comp-complex sentences - Average clause length in morphemes	0.05, d *-0.10*[38] / 0.02, d *-0.19*
Syntactic complexity- Morphological complexity	
Average clause length in morphemes - Average word length in morphemes	0.66, d 0.56 / 0.61, d 0.57
Syntactic complexity- Morphological complexity	
Average clause length in morphemes - Use of past and perfect tense	0.22, d *-0.03* / 0.16, d 0.15
Morphological complexity - Morphological complexity	
Average word length in morphemes - Use of past and perfect tense	0.17, d *-0.08* / 0.32, d 0.31
Syntactic accuracy - Morphological accuracy	
CAR use errors – CAR form errors	0.01, d 0.00 / 0.25, d 0.17
Syntactic accuracy – Syntactic complexity	
CAR use errors - Average clause length in morphemes	*-0.08*, d*-0.08* / *-0.16*, d *-0.02*
Morphological accuracy – Syntactic complexity	
CAR form errors - Average clause length in morphemes	*-0.03*, d 0.18 / *-0.24*, d *-0.12*
Syntactic accuracy - Morphological complexity	
CAR use errors - Average word length in morphemes	0.05, d 0.07 / *-0.17*, d *-0.08*
Morphological accuracy - Morphological complexity	
CAR form errors - Average word length in morphemes	0.08, d 0.32 / 0.07, d 0.17
Syntactic accuracy - Morphological complexity	
CAR use errors - Use of past and perfect tense	*-0.19*, d *-0.21* / *0.17*, d *-0.15*
Morphological accuracy - Morphological complexity	
CAR form errors - Use of past and perfect tense	*-0.14*, d 0.04 / *-0.09*, d *-0.08*

[37] d=detrended (residual) data.
[38] Roman: value is positive, Italic: value is negative.