# Difficulty in an online language learning course and its impact on learner tactics and strategies

**by**

**Dionysios P. Arkadianos**

MA (Intercultural Education), University of Patras, 2013
BSc (Computational Linguistics), University of Manchester, 2003
BA (Education), University of Patras, 2000

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

in the
Educational Technology and Learning Design Program
Faculty of Education


SIMON FRASER UNIVERSITY
Spring 2023

## Declaration of Committee

| | |
|---|---|
| **Name:** | **Dionysios Arkadianos** |
| **Degree:** | **Doctor of Philisophy** |
| **Title:** | **Difficulty in an online language learning course and its impact on learner tactics and strategies** |
| **Committee:** | **Chair: David Chang**<br>Lecturer, Education |
| | **John Nesbit**<br>Supervisor<br>Professor, Education |
| | **Phil Winne**<br>Committee Member<br>Distinguished Professor, Education |
| | **Panayiotis Pappas**<br>Committee Member<br>Professor, Linguistics |
| | **Tenzin Doleck**<br>Examiner<br>Assistant Professor, Education |
| | **Marina Tzakosta**<br>External Examiner<br>Associate Professor, Pre-school Education<br>University of Crete |

# Ethics Statement

The author, whose name appears on the title page of this work, has obtained, for the research described in this work, either:

a. human research ethics approval from the Simon Fraser University Office of Research Ethics

or

b. advance approval of the animal care protocol from the University Animal Care Committee of Simon Fraser University

or has conducted the research

c. as a co-investigator, collaborator, or research assistant in a research project approved in advance.

A copy of the approval letter has been filed with the Theses Office of the University Library at the time of submission of this thesis or project.

The original application for approval and letter of approval are filed with the relevant offices. Inquiries may be directed to those authorities.

## Abstract

The notion of difficulty in Second Language Acquisition research is a highly discussed topic, and it has been extensively explored in the literature. However, research studies investigating difficulty in the context of Computer Assisted Language Learning instructional design are scarce. This thesis explores the concept of difficulty in an online Modern Greek language course and how it can be implemented in an adaptive dashboard for a digital learning environment. The thesis examines different metrics of difficulty for a language learning unit. It introduces a new metric for difficulty, the Linguistic Complexity Index, consisting of three indicator indices connected respectively to lexical, morphological, and syntactic complexity. The results indicated that only morphological and syntactic complexity have a statistically detectable correlation to the effort required by the learner as operationalized by *time of completion* for each unit of the course. Additionally, the study revealed changes in learners' behavior depending on the variability of each of the three indicator indices. Increased lexical complexity relates to an increase in the use of deductive learning tactics, whereas increased syntactic complexity relates to an increase in the use of inductive learning tactics. Variability of morphological complexity showed no statistically detectable connection to the use of deductive versus inductive learning strategies. These findings have interesting implications for applying the three complexity indicator indices in designing adaptive presentation of language learning resources.



**Keywords**: difficulty; linguistic complexity; CALL instructional design; learning tactics; learning strategies; induction - deduction

## Dedication

I dedicate this work to my daughter Ariadne, my wife Maria, and my parents Panagiotis and Pege.

## Acknowledgements

First and foremost, I express my sincerest gratitude to my senior supervisor, Dr. John Nesbit. Dr. Nesbit has been a great guide and mentor throughout this long journey, always ready to provide the support I needed, especially during the challenging times of the program. John, you have been a great mentor and educator, and it has been a privilege for me to be one of your PhD students.

Next, I would like to sincerely thank Dr. Phil Winne and Dr. Panagiotis Pappas, who also participated in my supervisory committee. Their comments and feedback have been crucial for me to successfully complete the program.

Additionally, I want to thank my colleagues and friends, Costa Dedegikas and Polychronis Kambylis, for supporting my research with the development of the online language learning platform, as well as some scripted tools for extracting the data from the student logs.

Finally, I want to acknowledge Dr. Katerina Lagos, for trusting me with the secondary data from the course she was teaching, her participation to the Rebooting the Greek Language project and her overall guidance and support.

# Table of Contents

## List of Tables

## List of Figures

# Chapter 1.

# Introduction

In the last three decades, the field of Computer-Assisted Language Learning (CALL) has experienced rapid development, mainly due to the expanded use of computers in every aspect of human activity including language learning (Deutschmann & Vu, 2015). The advance of digital devices and educational technologies nurtured the development of this interdisciplinary field, showcasing its increasing research potential and diverse range of topics (Chen et al., 2021) The increasing computational power of computer devices, combined with their increased portability and affordability, as well as the expansion of internet usage, have been decisive factors affecting the growth of the field (Blake, 2013). Today, a wide variety of software and technology resources is available to language educators and learners, as well as many different theoretical approaches and methodologies on how to implement them (Garrett, 2009).

During the evolution of the field, CALL underwent a significant conceptual shift. The traditional view of CALL as a discipline focusing on the creation and curation of digital resources that can be used by language instructors, instead of focusing on the actual process of language learning using technology and the emerging theoretical and methodological issues, has been widely challenged. More modern approaches emphasize the fact that contemporary language learning digital environments are effectively transforming language learning and they are changing language teaching practice in new and compelling ways. This conceptual shift is reflected in the definition of the field of CALL. Earlier definitions referred in general to any activity or process which involved the use of a computer and resulted in language learning (i.e., improvement in the learner's language ability). Modern approaches define the field as the intersection of language education and technology, focusing on investigating the development, selection, use and evaluation of language learning activities mediated by technology (Chapelle, 2010).

The aforementioned conceptual shift had also a great impact on the field at both theoretical and epistemological levels. Deutschmann and Vu (2015) noted that the different phases of CALL, as they were determined by Warschauer (2000) and Bax

(2003), correspond to a simultaneous theoretical shift from behaviorist approaches advocating CALL activities such as repeated practice of fundamental skills to cognitive and sociocultural approaches promoting socially situated and collaborative learning experiences through CALL. Additionally, Garrett (2009) argued that earlier epistemological views of the field of CALL advocated the supremacy of pedagogy over technology. However, modern approaches consider all three major components of CALL (theory, pedagogy, and technology) as equal, and introduce a fourth component, infrastructure, which refers to the contexts of CALL that have a major influence on how the other three components work.

The research space of CALL was also influenced by these changes. Blake (2013) distinguishes three major directions in CALL research. The first one is the comparative approach, which attempts to evaluate the learning results of CALL-based methods and resources in comparison to more traditional, face-to-face language instruction. This approach has been heavily criticized to the point that it has now been almost abandoned. Chapelle (2010) notes that such critiques focus on the limited generalizability of the results, as they are difficult to replicate, are not connected to learning theory, and there is uncertainty about the causes of observed differences. It has been noted that, in most cases, CALL applications and research aim at expanding the learning goals of traditional approaches rather than improving performance on the same goals. Hence, the two different approaches should not be directly compared, as their instructional goals are different.

The second direction of CALL research refers to best practices and instructional designs, with a focus on teacher training and redefining the role of the language instructor, as the language learning paradigm is shifting towards hybrid and blended learning modes (Blake, 2013). This direction puts more emphasis on methodological issues of CALL implementation in the classroom, rather than investigating issues of instructional design for CALL applications. This research approach comes as a consequence of a conceptual shift of the field, where the focus moved from CALL software and resources towards a wider perspective, considering CALL language processes and their contexts as a whole, instead of examining specific instances of CALL implementation. A side effect of this approach is that the features and attributes of language learning platforms are rarely the object of research investigation. As Gillespie

(2020) noted, research studies on the characteristics and design decisions regarding Virtual Learning Environments are scarcely published in CALL literature.

The third direction of CALL research according to Blake (2013) is referred to as *process oriented*. Its focus is on the examination of learning procedures and the different ways language learners use CALL applications and resources. The conceptual shift in the field of CALL, which suggested a new language learning paradigm as the result of the introduction of computational methods in language instruction, encourages research on the learning processes in digital language learning environments in an effort to explain how this paradigm change occurred. However, even though several research studies investigate learning processes in CALL applications with a focus on specific language skills, there is generally limited discussion on the research methodologies used, or how the empirical findings of these studies inform CALL pedagogy (Gillespie, 2020).

## 1.1. Problem Statement

From elementary schools to universities, the COVID-19 pandemic has been the cause of an abrupt paradigm shift in language instruction as face-to-face and even blended instructional modes were substituted by fully online instruction (Payne, 2020). The challenges raised by this development have been great, not only because many language instructors were unprepared for transitioning to fully online delivery of their courses, but most importantly due to the difficulty of building class community and harmonious relationships between teachers and students, or between peers, without any face-to-face contact (Maldonado–Mahauad et al., 2017).

During the pandemic, most language teachers configured their fully online teaching by combining synchronous online teaching sessions with self-studying learning tasks and assignments for learners which used asynchronous language learning software (Payne, 2020). The asynchronous components of these systems implement an instructional setup where the teacher is not immediately present to provide support and appropriate feedback. Previous research suggested that an asynchronous language learning experience, with limited teacher presence, poses great challenge to learners, and especially to their motivation (Hromalik & Koszalka, 2018). One of the solutions addressing these challenges has been the implementation of adaptive language learning systems (Slavuj et al., 2017).

The major feature of adaptive language learning systems is that, instead of providing a static and identical learning experience to all learners, they tailor the content and the way it is presented according to the learner's characteristics, providing an individualized and specialized learning experience to each learner (Wauters et al., 2010). A learning system achieves adaptivity in various ways. Slavuj et al. (2017) emphasize the potential of adjusting how the content is presented, by adaptively manipulating the sequence of learning items or tasks, as well as providing high quality automated feedback, usually by identifying errors and determining their cause as gaps in learner's knowledge. Ma et al. (2014) highlight more features of intelligent tutoring systems that benefit from adaptivity, like offering guidance to students in the form of hints or prompts, or answering questions posed by learners.

Learner motivation is an important factor to consider in designing adaptive sequencing systems. Self-determination theory (Deci & Ryan, 2012) emphasizes the importance of learner autonomy as a key element for success in learning, and for psychological growth in general. A study by Schneider et al. (2018) showed that providing simple choice options as a design feature is an efficient motivation enhancing strategy, which also results in better retention and transfer performance for the learners. It is crucial for an instructional design to provide the necessary tools and information, to enable learners to feel, and actually be, in control of their learning activity and their learning goals. This means that adaptive systems should not remove control from the learner, but instead should be designed to afford learner choice.

A student dashboard is an interface that provides an overview of learning progress and often affords the students a modicum of control (Levy & Stockwell, 2013). Student dashboards have the potential to scaffold learners' self-monitoring and decision-making, and, especially from the perspective of this thesis, their decisions about the sequencing of instructional modules or unit. Unfortunately, research on language learning dashboard design is scarce (Gillespie, 2020) and focuses on the visualization of student activity in specific features of the digital learning environment, such as their views of text pages and instructional videos (Youngs et al, 2018). Usually, student dashboards present information on the content of the course in the form of an overview of the instructional goals of a unit, or as a list of the grammatical phenomena the material refers to (Gelan et al., 2018). These overviews usually contain linguistic terminology learners may not be familiar with. This is especially true for students with limited prior knowledge on linguistics.  Additionally, the overviews are not overly informative about the difficulty of a particular unit and the amount of effort it requires. The estimation of the difficulty of a particular unit or linguistic phenomenon is a highly discussed topic in Second Language Acquisition Research. Housen and Simoens (2016) argue that the concept of difficulty in Second Language Acquisition is addressed in multiple and often ambiguous ways. For example, some approaches consider difficulty solely as a reflection of the complexity of the linguistic phenomenon addressed by the research, whereas others relate difficulty directly to learner individual differences and how cognitively demanding a language learning unit is to a given language learner. This conceptual ambiguity is mainly the cause for vague and unclarified definitions and operationalizations of difficulty. Additionally, only a small number of operationalizations have been suggested by experts as estimates of linguistic complexity, with most of them using language-specific processes for their calculations. Ideally, we should have complexity metrics for course units or linguistic instructional goals, which are easy to calculate, universal to all languages, and predictive of the amount of time and effort required from the learner. Finally, there is the issue of how to represent numerically the construct of difficulty. Several researchers (Wauters et al., 2010; Wainer & Mislevy, 2000; Pandarova et al. 2019) adopt a single metric approach to difficulty. Housen and Simoens (2016) raise the issue of whether such an approach is adequate or oversimplifies an otherwise complex and multifaceted construct.

In terms of adaptive sequencing of learning items and tasks, many methods have been suggested on how to develop a model that determines the order in which the system provides the learners with information to be learned or tasks to be completed. Wauters at al. (2010) argued that an adaptive sequencing system based on Item Response Theory applied in a learning environment may result in higher learner motivation and more efficient learning. In the suggested learning environment, the calibration of the items or tasks in the system bank (meaning how their difficulty is estimated) is usually performed with a non-adaptive test taken by a number of students. The difficulty is determined by the number of wrong answers or longest time taken by the students to complete them (Wainer & Mislevy, 2000). This method poses some challenges, like the case of missing data. More specifically, the item pool may contain a high number of items and, because of the randomised nature of the non-adaptive testing, a number of items may not get answered by the students. Additionally, the ability of the learners to skip items may also lead to a high number of missing difficulty values for the items or tasks.

Another feature of an adaptive learning system is adjusting the parameters of the presentation of the content, like the order learning resources are presented, selection of the type and difficulty of resources, providing hints and guidance to the learners and sequencing the questions of the exercises to achieve efficient and effective learning (Wauters et al., 2010). These adjustments on the representation of the learning material and the form of the learning environment should be founded on empirical evidence from studies in the field of CALL that investigate learners' interactions with digital language learning environments. The rise of learning analytics methods and techniques has been a significant catalyst to CALL experts' efforts to explain and theorize language learners' behavior in digital language learning platforms and systems (Gelan et al., 2018). Investigation of learner tactics and strategies, as they engage in language learning tasks in a digital learning environment, can provide empirical data to inform CALL pedagogy, which is an under-researched topic in the field (Gillespie, 2020), and can guide the design of adaptivity features for the presentation of the content in a language course.

## 1.2. Purpose of the Thesis

The purpose of the present thesis is twofold: (a) to propose a metric that will serve as an estimate of the difficulty of a language learning unit to be implemented in an adaptive language learning system, and (b) to demonstrate the utility of the metric for investigating students' interactions and choices in a CALL application, especially in the form of learning tactics and strategies.

These goals will be pursued in the context of an online language learning platform which was designed to offer a Beginner's Modern Greek language course to university-level students. This learning management system serves both as the context in which language learning takes place and as the research instrument, which captures aspects of the learners' learning behavior, to be implemented in the statistical analyses used to address the research questions. In view of the dependence of the research on the course and its digital learning environment, a comprehensive review of their features will be provided in subsequent chapters. Even though this language learning management system is not adaptive per se, the research questions will be formulated in the context of adding adaptivity features to a future iteration of this system.

For the first goal of the thesis, the intent is to investigate candidate difficulty metrics to be used for adaptive sequencing of language learning units. These metrics should be language independent, i.e., the procedures or measures used to estimate the difficulty of learning items in a course should be applicable to any language, and they must refer to the internal linguistic properties of the language learning units instead of student-dependent properties. The online language learning platform under investigation already includes information about the units of the Modern Greek language course which can be used to estimate the difficulty of these units. This study examines each one of these metrics, evaluates their usefulness as estimates of difficulty, before suggesting a different process of estimating the difficulty of a unit based on its linguistic properties.

The second goal of the thesis involves investigating how student behaviors, resulting from learning tactics and strategies adopted by the students enrolled in the Modern Greek language course, correspond to the varying complexity of the units in the course as estimated by the suggested metric of the previous goal. These learning tactics and strategies are described as interactions learners have with the features of the digital learning environment. The study identifies these tactics, as traced in the student logs created by the learning management system, and relates them with learning strategies defined and operationalized in prior research. The implications of these findings for designing adaptive features in learning systems such as the Beginner’s Modern Greek language course are considered.

## 1.3. Research Questions

The first research question refers to the problem of estimating the difficulty of a language learning unit. As has been mentioned, the online language learning platform under investigation already includes information that can be used to estimate the difficulty of each language learning unit, i.e., the type of the unit (vocabulary or grammar), the number of prerequisite units, and the number of levels of linguistic description involved in the instructional goals (morphology, syntax, semantics, and pragmatics). Additionally, this thesis will propose another estimate of the difficulty of a language learning unit, namely the *Linguistic Complexity Index*, which is based on the inherent characteristics of the language structures or features involved, and investigate whether such a metric should be a single, composite measurement or a set of multiple subindices. As an additional point of investigation, this thesis examines how the metrics for estimating difficulty correlate to a measure of how cognitively demanding a unit is to the learners, so that the metrics can be used in a student-faced dashboard to scaffold learners in their decision of which language learning unit they should tackle next. Housen and Simoens (2016) use the term *cognitive complexity* to refer to the cognitive demand of a language unit on the learners and suggest the *time of completion of a unit* as a method to operationalize it. Considering all of the above, the first research question is formulated as:

RQ. 1: Which metrics based on the features of a language learning unit in the Modern Greek language course are appropriate estimates of the difficulty of the unit and also good predictors of the cognitive complexity of that unit?

The second research question examines the relationship between language learning units of varying difficulty and learners' interactive behaviors while they are logged in the system and are engaged with the content and activities of the Beginner's Modern Greek language course. The examination of learning behaviors will be conducted in two levels as described in Matcha et al. (2019) – a *tactical* one, where the learners' actions and the use of the system features will be considered, and a *strategic* one, which will regard the learner's behavior at a higher level, delineating the bigger picture of these interactions, with reference to language learning strategies identified in prior research in the field. Thus, the second research question is formulated as:

RQ 2 – How do the *learning tactics* and *learning strategies* adopted by students of the Beginner's Modern Greek online language course relate to the difficulty of language learning units?

## 1.4. Significance of the Thesis

Even though research in the field of CALL, especially in the last decades, focuses on illuminating the learning processes taking place during the students' interactions with the various features of the digital language learning environments, the empirical evidence presented in those studies was used primarily to identify different learner types or determine which learner attributes facilitate learning in CALL applications (Blake, 2013). Their findings provide insight on how learning occurs and which factors may have an impact in the learning process, but they usually stop short of discussing application to CALL pedagogy or to instructional design, which are under-researched areas of the field (Gillespie, 2020). This thesis maintains a pedagogical and learning design perspective, either by considering instructional design decisions and learning environment features as the starting point for formulating the research questions or by discussing the implications of the findings to instructional design and pedagogical considerations in the field.

Reviews of the CALL research space also reveal a certain disengagement of these studies from the field of linguistics, even though CALL is regarded as a field of Applied Linguistics (Gillespie, 2020). While addressing its research questions, this thesis attempts to maintain a linguistics-based approach by utilizing elements of Linguistic Theory and theoretical approaches to the concept of linguistic complexity to arrive at a numerical estimate for the construct. Additionally, this study aims to connect learning behavior to the linguistic attributes of the content to be studied and investigates how learners regulate and adjust their learning tactics and strategies according to the linguistic nature of the content they are working on.

Finally, this thesis attempts to bring awareness to aspects of CALL research which are generally overlooked. Issues of content development and how learners in CALL applications should be scaffolded, or which type of learning resources are required for the acquisition of particular language structures, are seldom investigated or discussed in the literature. Additionally, certain elements of instructional design, such as the structure and function of student dashboards, usually have a miniscule footprint in CALL research, typically being mentioned briefly if at all. This study seeks to provide the foundation of further study in the design of such features of language learning environments, as well as providing guidelines for developing content for unsupervised online language learning courses.

## 1.5. Outline of the Thesis

The present thesis consists of seven chapters. Chapter 1 contextualizes the research topic within the current field of Computer-Assisted Language Learning, outlines the problem which will be investigated in this study, and explains the purpose and the significance of the research. Chapter 2 reviews definitions and operationalizations of the various types of linguistic complexity and discusses the issues and challenges in both defining and operationalizing complexity. Chapter 3 reviews CALL literature on navigational patterns of learners, as well as learning tactics and strategies adopted by them, and it examines definitions and operationalizations of inductive and deductive language learning strategies. Chapter 4 provides details about the educational context of the Beginner's Modern Greek language course under investigation. Chapter 5 presents the methodology for the research and defines the metrics utilized in the study, as well as data screening and analysis processes, and revisits the research questions, with reference to the concepts discussed in Chapters 2 and 3. Chapter 6 presents the results of the quantitative analyses that address the research questions. Chapter 7 provides an interpretation of the study findings, discusses the implications of the empirical evidence, and suggests future directions and opportunities for further research on the topic.

# Chapter 2.

# Defining Difficulty as Linguistic Complexity

The notion of difficulty in linguistics has received great attention in second language acquisition (SLA) studies and has been hotly debated among many linguists and researchers (Ehret & Szmrecsanyi 2016, Palloti 2015, Bulte & Housen 2012). Housen and Simoens (2016) note that one of the most common themes in the conceptualization of difficulty is its association with the construct of linguistic complexity. In research, complexity has been employed either as an independent variable, characterizing a linguistic task, or a dependent variable, to describe the quality of a student's language production. In applied SLA, this construct has been used to assess task or content sequences and provide teachers with some preliminary information about which parts of a language course might be more challenging for the students, in order to prepare their instructional approach accordingly. In intelligent language tutoring systems, complexity has been used in algorithms to predict the difficulty of exercise questions for dynamic difficulty adaptation. However, applying the construct of complexity has not been without several issues and challenges, both theoretical and methodological (Palotti, 2015). Since linguistic complexity, as the conceptualization of difficulty, constitutes one of the key concepts in the present study, it is important to review how it is defined in the literature, how it has been operationalized in various research studies in SLA, and major challenges and issues when implementing the construct.

## 2.1. Definition of Linguistic Complexity

At the core of debate about linguistic complexity lie different approaches on how to clearly define the concept (Palotti 2015). Complexity, being a term used widely even in non-academic contexts, may represent different constructs across different areas of research and theoretical approaches. Even terminology varies, e.g., "complexity" and "difficulty" (Housen & Simoens, 2016), with little clarification about the concepts these terms represent. Therefore, it is important to review the most prominent definitions and terms for linguistic complexity to avoid confusion and to situate the operationalization used in the study. A classification system will be used to categorize the various definitions and facilitate the review.

A first distinction between approaches to define complexity in linguistics is the one between relative and absolute approaches (Dahl 2004). The relative approach defines complexity in relation to language users. Higher degrees of complexity imply more resources the learners need to invest in order to process and internalize a target linguistic structure (Bulte & Housen, 2012). A more specific definition adopting the relative approach was suggested by Hukstijn and DeGraaf (1994), who argue complexity is related to the mental effort of language learners to learn, process or verbalize the target linguistic items. This type of complexity is referred to as relative complexity, cognitive complexity or difficulty (Bulte & Housen 2012). Relative complexity considers individual differences of learners, especially differences in their cognitive abilities, such as language aptitude, working memory, implicit and procedural learning ability, prior linguistic and language knowledge, socioaffective and personality factors and others (Housen & Simoens, 2016). As discussed in the next section, relative complexity tends to be operationalized either subjectively, using ratings made by experts, or objectively, using variables like time spent on task.

On the other hand, the absolute approach takes a more language-focused stance when defining complexity. Linguistic complexity according to this approach, which is referred to as absolute complexity, structural complexity, intrinsic complexity or simply complexity (Bulte & Housen 2012), is an inherent attribute of the language system or language structure to be learned by the learners. Structural complexity is defined in quantitative terms related to the number of distinct components a language unit or structure consists of and the number of connections between these components.

Especially in the case of structural complexity, different definitions have been suggested based on the level of linguistic description that is the focus of the approach. According to this categorization, two different types of structural complexity can be distinguished. Formal complexity considers lower levels of linguistic analysis, mainly morphology and syntax. According to this definitional approach, complexity is related to “the number of operations to be applied on a base structure to arrive at the target structure” (Bulte & Housen, 2012, p. 25). However, as DeKeyser (2016) argues, linguistic structures and their functions or meanings do not exist in isolation but as a whole, and learners need to consider them as such. Therefore, definitions and operationalizations need to include both aspects of a linguistic unit, form and meaning, considering also how these two interact. Bulte and Housen (2012) refer to that approach as functional complexity, “the number of meanings and functions of a linguistic structure and, to the degree of transparency, or multiplicity, of the mapping between the form and meanings / functions of a linguistic feature” (p 25). Very interesting notions in this definition are those of transparency or multiplicity, which relate to the mapping between form and meaning(s). In particular, lower transparency is connected to language features that have more than one linguistic interpretation. For example, the suffix -s may denote plural form of a noun or a third person singular form of a verb in simple present tense. Researchers relate the transparency of this mapping with higher levels of functional complexity. Housen and Simoens (2016) note there are two cases of less transparent mapping between form and meaning that result in higher functional complexity. The first is related to irregular forms mapped to specific meanings or language features, which are connected with higher complexity, and the second refers to multiple mappings of the same form.

Additionally, complexity definitions have been proposed that refer to higher levels of linguistic description, specifically semantics and pragmatics, for the target linguistic feature to be learned. Bulte and Housen (2012) distinguish two such definitions referred to as propositional complexity and discourse or interactional complexity. Propositional complexity refers to how complicated is the meaning the learner is trying to convey in the utterance. Discourse or interactional complexity relates to how complex is the communicational context of the target linguistic item.

Another set of definitions suggested by Bulte and Housen (2012) refers to the different levels of examination of a particular linguistic structure or feature to be processed or acquired by the learner. First, we have definitions that focus on an analysis at a theoretical level. These definitions are more abstract and relate closely to a specific theory of language. Palotti (2015), for example, distinguishes *system complexity* which refers to the Saussurean notion of langue – the language system as a whole, with all its components and features and the relationships between them and which is independent of the individual language user – and *text complexity* which relates to the Saussurrean notion of parole – the concrete instances of language, highly dependent on the language user who articulates them. Palotti elaborates this distinction further, connecting those two terms with objectivity. Grammatical complexity reflects how complex are the linguistic rules which apply to form a specific linguistic construct or to manifest a specific linguistic feature, constituting a more objective approach. On the other hand, stylistic complexity reflects the language user's choice of form enabling expression of a proposition. Second, we have definitions which approach the construct at an observational level. These definitions are more concrete and focus on the actual language use and authentic linguistic utterances. Under this analytical lens, complexity is related to how different forms serve different communication purposes (specialized vs common vocabulary, academic language and grammatical convention, etc.). Third, we have the most concrete types of definitions which adopt a more quantitative approach to define complexity. These definitions provide the foundation for the operationalizations of the complexity construct, which will be reviewed in the next section. These definitions use quantitative measures like frequency, ratios, indices and other statistical metrics to determine the degree of complexity. They allow for analysis and comparison of different linguistic features or items in an objective way, especially when applied to linguistic samples derived from different languages.

A different type of definition for complexity incorporates terms and concepts from research, and particularly research in the area of second language acquisition. Housen and Kuiken (2009) distinguish two strands of empirical research which consider the construct of complexity in an entirely different way. In the first set of studies, complexity is considered an independent variable influencing learner performance in a foreign language. These studies investigate the impact of complexity on the teachability of the target structure or on the effectiveness of the different types of instruction. The second strand of studies regards complexity as a dependent variable, usually implemented along with fluency and accuracy as an aspect of second language performance or an indicator of second language proficiency.

Finally, Revesz et al. (2017) relate their definition of linguistic complexity to the specific linguistic task performed by foreign language learners. They refer to that particular approach as task complexity and they consider two different aspects of the construct. The first one involves the inherent complexity of the language structure or feature that is involved in the task (just like in the case of structural complexity, which has been discussed previously in this section). The second aspect involves attributes of the task itself that contribute to the overall complexity. Such attributes of the task include elements like planning time (how much time is given to the learners to plan their actions for the task), revising conditions (whether certain revisions of the involved linguistic features are allowed prior to the task), provision of support (various levels of scaffolding for the task, spelling and grammar checker, thesaurus, and other resources to support the task), storyline complexity, etc.

Considering all the previous definitions of linguistic complexity, and for the purpose of addressing the research questions formulated in the previous chapter of the thesis, we will be adopting the absolute approach to defining linguistic complexity, i.e., considering it as an internal attribute of the linguistic task or structure involved in a particular unit of a language course, as it is the definition that directly relates to the difficulty of a language learning item or task that we attempt to estimate.

## 2.2. Operationalization of Linguistic Complexity

A next step after defining a concept or construct in empirical research is to determine a method to operationalize it. Operationalization is a crucial part of an empirical study, as it involves defining the measurement of a phenomenon or concept that cannot be measured directly, though its existence is inferred by other concepts or variables. Palotti (2015) emphasizes that any operationalization of a concept or construct needs to be founded on a conceptual definition. This section reviews the most prominent attempts to operationalize linguistic complexity, connecting those operationalizations to their related conceptual definitions.

Some operational definitions use mathematical formulas and formal rules to determine the complexity of a linguistic item. Dahl (2004) mentions one such operational definition, which is referred to as Kolmogorov complexity, the length of the shortest description of a string of symbols. For example, the strings abcd and abababab both have a Kolmogorov complexity of 4 because the latter has 4 × ab as a minimum length, a 4-symbol descriptor. Kolmogorov complexity has been used not only to determine the complexity of lexical items represented as strings of letters, but also of syntactic constructions and patterns represented as strings of constituents. Other researchers suggest operationalizations based on statistical methods applied to empirical data, like regressing item difficulties on item features (Pandarova et al., 2019).

Ehret and Szmrecsanyj (2016) suggest determining complexity by using methods adopted from information theory. These present striking similarities to linguistic operational definitions like Kolmogorov complexity. These researchers utilize features of file compression software to operationalize complexity. They argue this type of software implements algorithms that use a type of adaptive entropy estimation, which approximates the operationalization of Kolmogorov complexity. The algorithm describes new strings based on information extracted from previous strings, thus measuring both new information included and redundant information for a given string.

The most fundamental conceptual divide between differing definitions of linguistic complexity is that between structural and cognitive complexity, which arise from language-related and learner-related factors and attributes respectively. Rodriguez-Silva and Roehr-Brackin (2016) suggest cognitive complexity can be empirically measured using subjective, holistic ratings made by researchers and foreign language acquisition experts, language teachers and language learners themselves, as well as objective measurements. The latter may refer specifically to the linguistic task per se (time spent on task, use of resources offered to the learner, order of actions, reaction times, etc.), or they may be general psychophysiological measures, such as eye movement, brain activity via MRI, etc. On the other hand, Pandarova et al. (2019) argue cognitive complexity should not be considered in isolation, but instead should be conjoined with structural complexity. For that reason, they suggest psychometric models that incorporate both aspects of linguistic complexity. One such model adopts the one-parameter Rasch model, which connects the construct of complexity to the probability of a correct answer to an item. This probability is calculated as “a logistic function of the difference between the person’s ability parameter ($\theta_p$) and the item difficulty parameter ($\beta_i$)” (p. 345).

Other operationalizations focus only on structural complexity and its various types depending on the level of linguistic description that is the focus of the definition: lexical, morphological, or syntactic complexity. Most researchers consider those aspects separately due to the different nature of their attributes and features. Hence, they define complexity as a construct of multiple components, each of them reflecting a different analytical level.

Lexical complexity has been operationalized using various metrics that reflect attributes of the lexical items involved in a specific task or linguistic feature. Palotti (2015) distinguishes two different foci in measuring lexical complexity. The first one considers the number of lexical items involved in the target linguistic unit (either the language system in general or a specific linguistic feature or structure), whereas the other considers the complexity of individual items such as compounds and lexical items containing derivational morphemes. Pandarova et al. (2019) suggest an estimation of lexical complexity using the frequency of the involved lexical items. Specifically, less frequent words are connected to higher levels of linguistic complexity. Finally, Revesz et al. (2017) relate lexical complexity to the lexical variability of the target linguistic structures to be processed by the learners. There are various ways to capture lexical variability of a language structure, examples being the D-Value (Malvern & Richards, 1997), Nation's vocabulary range (Nation & Kyongho, 1995), and Coh-Metrix indices of frequency and concreteness (McNamara, Louwerse, Cai & Graesser, 2004). Bulte and Housen (2012) argue that the number of lexical items alone cannot reflect the lexical complexity of a linguistic item.

Morphological complexity in most cases is calculated focusing mainly on inflectional rather than derivational morphology (Housen & Simoens, 2016). The reason for such an approach is that inflectional morphology is related to various grammatical attributes and features whereas derivational morphology is connected to the formation of new lexemes, bringing it closer to lexical complexity instead. Haspelmath and Sims (2010) propose a functional approach to the operationalization of morphological complexity. Their operational definition considers the number of form-function relationships, which are collected under the term morphological patterns. Those morphological patterns relate features reflecting syntactic or semantic properties of a morphological form with the number of different morphemes conveying those features.

Syntactic complexity involves the order in which different constituents combine to form different clauses and/or sentences. Palotti (2015) suggests combining two different metrics to estimate syntactic complexity of a linguistic structure. The first one considers the number of constituents in a syntactic structure and the number of combinations these constituents can take. To determine these numbers, the verb is considered the core of the structure, so only the immediate constituents to the verb are used in the estimation, meaning no embedded clauses are considered. The second metric refers to the number of clauses per structure which, according to Palotti, is one of the most used measures of syntactic complexity in second language acquisition research.

Finally, some researchers prefer to consider morphological and syntactic complexity as a composite construct, given that many linguistic item features are connected to both levels of linguistic description. The close relation of both levels is reflected in language acquisition and linguistics research, as both morphological and syntactic attributes are referred to under the term grammatical features. Pandarova et al. (2019) suggests several grammatical features to be considered when estimating complexity: voice, adverb placement, subject–verb agreement, etc. One of the indices these researchers propose to be used when operationalizing morphosyntactic complexity is the morphosyntactic edit distance (MSED). The MSED refers to the number of syntactic and morphological transformations, which need to be performed in order to create the target linguistic structure.

## 2.3. Issues and Challenges in Defining and Operationalizing Complexity

As was mentioned, complexity is a highly debated construct, that has led to controversy and disagreement among linguists. Bulte and Housen (2012) emphasize that "language complexity is a multifaceted, multidimensional and multilayered construct, a fact that is still insufficiently acknowledged in L2 research" (p 41). The researchers identify three aspects of the construct that promote debate about its nature. Firstly, complexity has both cognitive and linguistic dimensions, being related to the learner and to the target linguistic structure to be learned. Secondly, it can be considered in either a performance or a developmental context. Finally, it may be considered in all levels of linguistic analysis, from morphology to pragmatics and discourse linguistics. Additionally, Housen and Simoens (2016) note another characteristic which makes complexity a difficult concept to define is that it is not monolithic or static, but a dynamic construct affected by a wide variety of factors.

In Section 2.1, which reviewed the different terms for complexity that appear in the literature, it became apparent there is a wide variety of definitions for the construct, which may refer to entirely different concepts. As Palotti (2015) suggests, the polysemy of the construct is one reason for the multiple theoretical and methodological problems emerging when trying to define or operationalize linguistic complexity. According to Palotti, the existence of so many definitions is evidence that they refer to entirely different constructs. Whether these constructs are related or not is an empirical issue.

Even if all those constructs are found to correlate strongly with each other, this is not sufficient evidence that they are different aspects of the same construct, as they present asymmetries in the relationship. For example, literature suggests that the construct of cognitive complexity reflects rather than creates complexity.

Another characteristic of complexity that raises issues when defining this construct is that the definitions are influenced by the theoretical context they assume. In structural complexity, for example, definitions may depend on the linguistic theory that is used as a theoretical framework. Different theories (generative, typological, cognitive linguistics theory, etc.) may result in different degrees of complexity for the same linguistic structure or feature.

Additionally, some definitions of linguistic complexity are not clear or explicit enough, resulting in ambiguities. This is especially true in the case of functional complexity. Hendriks and Watorek (2011) argue not all meanings connected to a specific linguistic form are well defined and unambiguous. Some of them are less prototypical or more interconnected or multilayered, so they appear as more complex than others.

Finally, the nature of instruction may impact the determination of complexity of a target linguistic unit. An example of such a case is provided by Housen and Simoens (2016), who note that implicit instruction – instruction based on the presentation and processing of actual language articulations without any reference to rules – relates primarily to the implicit complexity of the instructional goal, i.e., the complexity of the linguistic construct itself. On the other hand, explicit instruction – instruction based on the explicit description of linguistic phenomena and presentation of all the involved rules to the learner – may depend on the clarity of the rules and the overall pedagogical conditions of the current educational context.

Attempts to operationalize complexity have also been plagued by various issues and challenges, as has been emphasized by various experts and researchers. Palotti (2015) pinpoints the lack of theoretical foundation for the proposed operationalizations, arguing that these operationalizations have been proposed without consideration of their theoretical underpinnings or of the issues of construct validity that have been raised. Deutcher (2009) also argues that both definitions and operationalizations of linguistic complexity involve the concept of a system. The concept is very challenging to define in an exact manner because such an attempt needs to consider what a system's boundaries are and how to assess the complexity of the system, specifically, which components to include and how to evaluate the interactions between them.

Additionally, the polysemy of the concept of linguistic complexity raises another challenge to properly operationalize it. There are several factors that need to be considered in order to evaluate all the different aspects of complexity, whether these factors are feature-related, learner-related or context-related. Considering each of the factors in isolation results in a partial investigation of the construct, as a more complete examination should also consider the potential interaction between the factors (Housen & Simoens, 2016).

Furthermore, some experts argue that not all aspects of complexity have been covered in the same degree by the various attempts to operationalize the construct. Bulte and Housen (2012) note some aspects of complexity have been operationalised by a wide range of measures (especially syntactic and morphological complexity), while others by very few measures (like lexical complexity) and some others are covered by none at all (collocational lexical complexity, derivational morphological complexity, etc.).

Finally, another issue related to the various attempts to operationalize the construct of complexity is how to arrive at a universal complexity index which will reflect all the different aspects of the construct. Palotti (2015) notes that this is not as simple an issue as just "adding" the various complexity scores, as different weights should reflect the different degree of contribution of each complexity metric to the overall index. Since this is not only a theoretical, but also a methodological issue, we will be revisiting it in the Methodology chapter, where several suggested solutions to the problem of calculating a composite variable from several indicator variables will be reviewed and discussed.

# Chapter 3.

# Investigating Learning Tactics and Strategies in CALL

The last two decades have seen a rising deployment of clickstream data analysis and learning analytics in Computer Assisted Language Learning (CALL) research. As Debski (2003) reported in his review that analyzed data from 91 CALL research articles, the use of computerized tracking methods in CALL was already rising steadily at that early stage. Many researchers pinpointed the benefits and potential of big data analysis in language learning on digital learning platforms. Hwu (2013) emphasized the importance of online clickstream data to provide more accurate quantitative analysis, since state-of-the-art tracking technologies allow for a large body of data to be collected and processed, in real and not experimental conditions. Additionally, even when informed about data collection through a research consent protocol, learners tend to ignore that data capturing takes place during their learning sessions, so more spontaneous and authentic learning behaviors can be observed.

A compelling observation about learning analytic research studies in CALL made by various researchers is that research in Second Language Acquisition (SLA) and CALL focuses mainly on the products of foreign language acquisition and not on the learning process per se. Youngs et al. (2018) note that most research studies based on the analysis of clickstream data investigate mainly the end products of language learning (assessment quizzes and essays, linguistic artifacts etc.), while the language learning process remains opaque and invisible to researchers and practitioners alike. Chun (2013) argues that traced learner actions can be used to detect patterns in learning behavior, underlying learning strategies and navigation paths employed by learners and, also, determining whether these patterns of behavior are organized or chaotic.

Tracing learner behavior patterns and learning strategies and tactics is a very important aspect of learning analytics research in CALL as it provides insight into the learning processes taking place during language acquisition. According to Matcha et al. (2019), a *learning tactic* is “a learning technique or cognitive operation that is used by a student to perform a particular task” (p. 462). On the other hand, a *learning strategy* involves the regularity in the application of particular learning tactics, depending on specific conditions during the learning process. Considering the definitions above, it is evident that learning tactics relate to short-term learning actions, whereas learning strategies refer to long-term learning orientations. Whether learners select and implement appropriate learning tactics and / or strategies, according to any emerging needs or conditions from the learning context depends on their Self Regulated Learning (SRL) skills (Gasevic et al. 2017). Selection of a learning tactic or strategy may depend on either internal (cognitive) conditions, or external (task-related) conditions (Winne & Hadwin, 1998).

There are several ways learning analytics can be used to research specific learning behavior patterns and learning strategies. The most significant distinction is that between hypothesis-driven methods and data-driven methods (Shirvani-Boroujeni & Dillenbourg, 2018). *Hypothesis-driven* methods are theoretically grounded and detect predetermined learning strategies from interaction sequences. The challenge in implementing hypothesis-driven methods is that learning behaviors can be so complex, that is often not feasible to accurately define a priori a specific learning pattern. *Data-driven* methods entail the discovery of learning behavior patterns that emerge from the interaction data. The challenge in implementing data-driven methods is that, without a theoretical framework, sometimes it is unclear whether a detected pattern is an actual behavioral pattern, meaning that it reflects the learner's internal processes while trying to accomplish a learning goal and it is not just a random action. Hwu and Tzseng (2013) make a valid remark on which of the two approaches should be adopted in a CALL research study, noting that when a digital language learning environment includes activity types, interactions, features and, in general, affordances which are uncommon in CALL practice, no research hypothesis can be formulated *a priori*. Hence, the most usual approach is to identify patterns from the collected behavioral data which can inform *a posteriori* research hypotheses to be tested in subsequent studies.

## 3.1. Theoretical considerations and issues

Several CALL researchers have noted a number of considerations and challenges that need to be addressed in a study examining learning strategies in a CALL educational context. Heift and Chapelle (2013) cite some of the issues that researchers should consider when conducting their studies or interpreting the results. For example, even though most CALL environments offer supporting resources to learners to help them in their comprehension and production of the target language, learners have also opportunities to seek help on the internet or other digital language learning resources (word processors, digital dictionaries, thesauri, etc.). These actions usually are not captured by most language learning LMS and, therefore, remain hidden from the researchers. Second, when investigating the utilization of computer-generated feedback by the students, an important fact to consider is that automated feedback is usually available for explicit language activities, which focus on specific vocabulary and grammar goals. However, in more free-form activities, such as free conversation or writing essays, such feedback is difficult to obtain, especially for less common languages such as Greek. Therefore, for those activities, data which indicate little or no consultation of feedback by the student could be misinterpreted, in the sense that absence of consultation doesn't necessarily mean a conscious decision by the student but rather an inevitable trajectory due to the absence or low quality of the available feedback. Another issue, reported by Hwu and Tzseng (2013) involves time-related clickstream data, i.e., time spent on a specific section or activity of the digital language learning environment. Inclusion of this type of data might provide misleading information, if, for example, a learner seemingly spends much time on an activity, when in fact the learner merely left open an interface window without performing any action.

Additionally, some researchers observed a methodological shift in investigating learning strategies. Hwu (2013) noted that early research studies on learning strategies implemented in CALL investigated (a) whether the students used the language learning system as intended by the instructional designer or (b) which learning strategies wielded the best learning outcomes, with the purpose of suggesting such strategies to learners as the optimal learning process leading to success in the language course. The recent trend for such research studies is to examine learning strategies in relation to individual differences and learner preferences, as empirical evidence suggests that different learners may implement different learning strategies with comparable results. Therefore, the research purpose is to create CALL designs that accommodate the variety of approaches adopted by different groups of language learners. Heift and Chapelle (2013) also emphasize the importance of investigating the effect of various individual difference factors on the choice of learning strategies.

Zhou and Wei (2018) also offer an important methodological suggestion, as they argue that learning behavior and learning strategies in language learning should not only be considered in a holistic way for the entirety of the learning process, but also separately, for each of the four different language competencies:

- comprehension of oral speech – listening
- comprehension of written speech – reading
- oral speech production – speaking, and
- written speech production – writing,

Each competence has different characteristics and, as such, may require a different approach or learning strategy. Additionally, studies investigating language learning strategies should distinguish learner behavior depending on the different language subsystems involved in a learning session (vocabulary or grammar). Finally, the researchers offer another classification of learning strategies in language learning, distinguishing between *cognitive strategies* (applied directly during the interaction with the language learning resources and content, i.e., taking notes, reading aloud, highlighting important points in the text) and *metacognitive strategies* (applied by learners upon reflection on the learning process, i.e., advanced planning, comprehension monitoring, reflecting on encountered issues and challenges).

In terms of interpreting behavioral sequences and patterns, in order to identify the underlying learning strategies adopted by learners, language learning experts follow the methodological approach of connecting action sequences and behavioral patterns to theoretical constructs adopted from various learning theories. Martin-Monje et al. (2018) linked navigational patterns of students participating in a language Massive Open Online Course (MOOC) to the students' engagement patterns, adopting a learner type classification proposed by Anderson et al. (2014). Lin et al. (2017) mapped learning strategies adopted from self-regulated learning theory (i.e., goal setting, task strategies, help-seeking, self-evaluation) to specific action sequences observed during learner interactions with the system. Desmarais et al. (1998) focused on learner individual differences as the theoretical framework to be used in the interpretation of the learning strategies adopted by students in a digital language learning course, especially prior knowledge and language proficiency, distinguishing between novice and advanced learners. Hwu (2007) linked learning strategies to individual personality differences using the Jung-Myers-Briggs typology. Finally, Payne (2020) associated learners' activity sequences with their cognitive load. The researcher investigated speech production activity sequences, to determine which of these may result in lower cognitive load. Again, this research wasn't aiming to examine the learning process and shed light on the thought processes of learners following a specific sequence. The purpose was to evaluate methods of optimally sequencing language learning activities for students taking an online course.

Another trend emerging in research studies on learning strategies in CALL digital environments is associating specific learning behavior patterns, and subsequently strategies, with success in a course. This is seen as especially important due to the large drop-out rates typically observed for language MOOCs and MOOCs in general. One of the most comprehensive studies for online language courses at the post-secondary educational level, in terms of the aspects of learning behavior investigated, is the one by Gelan et al. (2018). According to the results of that study, the key aspect of a learner's behavior for success in an online language course is that of "timeliness". In particular, students who prepared for an assessment in a timely manner achieved higher scores. Successful learners had significantly more online sessions, showing higher regularity in interacting with the digital learning environment than unsuccessful learners. Successful learners also attempted and completed more activities in the course, spent more time logged-in, and revisited both theory and exercises. Li et al. (2018) reported that successful students are very competent in managing their time, carefully studying and extensively reviewing the course content, completing assignments in a timely manner and self-evaluating their learning. Similarly, Keskin et al. (2016) observed that students who passed an online language course logged in more frequently and had more content interactions than failed students. In a more specific account of successful learners' learning behavior, the researchers noted that they spent more time on themes of content as well as messaging and seeking help from the instructor. On the other hand, failed students invested more time to reading the discussion fora, attempting assessments, and reading the feedback provided by the system. Veletsianos et al. (2021) reported that learners who failed the course spent less time studying resources and content, and more in assessment activities such as quizzes. They also found "the non-completers engage more sporadically, with distinct and dramatic spikes and dips in their activity" (p. 25). Finally, Martin-Monje et al. (2018) also noted that successful learners are more active in their online interactions. However, they observed that not all the interactions explain student success. For example, forum interaction and the submission of peer feedback do not contribute significantly to explain learner success in language MOOCs.

Finally, in terms of the operationalization of learning behavior in a digital language learning environment, most metrics used in relevant studies relied on unobtrusively captured student log data and, either focus on time spent on task, or on how many times students access specific activity types or course sections in the online language course. The study by Gelan et al. (2018) provides a multilevel capture of the learning behavior of students in a university level language learning course. At a learning session level, the metrics include number of sessions per student, learning session frequency, total and average time duration of learning sessions. At an activity level, the LMS captures data such as times accessing a particular type of activity, total time spend on a specific type of activity and activity type sequences. Student navigational patterns were presented as visualizations derived from process-mining data analysis. These visualizations were of the form of transition graphs, which expressed the order of activity types in a learning session. Additionally, the various learning behavior metrics implemented in the study were analyzed in contrast to particular event milestones of the language course (review exams, mid-term exams, final exam), which provided an additional set of time-related data referring to the timeliness of the learning sessions in relation to these dates.

The aforementioned study aimed towards addressing many different aspects of students' learning behavior. However, there are studies with a narrower focus, which implement more in-depth metrics derived from learners' clickstream data. Veletsianos et al. (2021) implemented five different metrics to analyze the learners' time management behavior: *login frequency* (the raw number of logins for each learner), *time chunking* (average amount of time in each learning session), *activity speed* (number of activities per minute logged in the digital platform), *login consistency* (the percentage of learner logins that occurred during the most common hour / day), and *early emphasis* (the percentage of a learner's activity which occurred within the first three weeks of their activity in the course).

Other CALL research studies on learning patterns and strategies use operationalizations of learning behavior that are very task specific. Youngs et al. (2018) investigated the learning behavior of language learners in relation to how they interacted with video study materials. The researchers investigated the total number of students watching the videos, their actions while watching them, whether there were a single or multiple viewings, whether they watched the video wholly or partially, if they engaged with the exercises after watching the video or in parallel, etc. Another study by Lan (2013) focused on the specific affordances of the Learning Management System (LMS) under investigation. The study gathered data from an application of the LMS in which students could choose among learning strategies to achieve the instructional goals. They could look up the strategies used by their peers or select one of the available strategies and get appropriate scaffolding on their use by the system. It is evident that in studies like these, the evidence obtained is greatly dependent on the educational context, and the insights of the learning procedures offered by the results are not easily transferable to other contexts, especially due to the absence of an explicit theoretical framework.

Finally, certain research studies on learning strategies in digital language learning courses implement operationalizations of learning patterns that combine student clickstream data captured by the LMS with self-report instruments integrated in the learning environment. Lin et al. (2017) implemented this dual combination of a self-report instrument (the Online Self-Regulated Learning Questionnaire) with student log data, to map observed learner action sequences to specific Self-Regulated Learning strategies: goal setting, task completion strategies, help-seeking, and self-evaluation. In that respect, data obtained from self-report were used to provide insight regarding which underlying learning strategies determined particular action sequences during a learning session. Desmarais et al. (1998) used self-report data paired with student log data to gain further insight on the internal learning processes of students demonstrating linear or chaotic learning patterns. As the researchers note, even though chaotic patterns are often the result of non-effective learning techniques or of difficulties with the content of the course, they might also indicate a learner’s intention to find more appropriate or effective activities in order to fulfill an instructional goal. Therefore, self-report data is used to clarify the circumstances under which a chaotic learning behavior pattern is observed.

In conclusion, most research studies investigating student learning strategies in Computer Assisted Language Learning environments focus on clickstream data related to time aspects of learning behavior (time spent on task, frequency of learning sessions, duration of learning sessions, time spent on specific sections or activity types) or navigational aspects (most accessed activity types, order of accessing different activity types or different sections of the learning environment). Two major trends can be identified when reviewing such studies: connecting learning patterns and strategies to course efficiency (completion time) and effectiveness (academic performance), or relating learning strategies to specific types of learners, depending on their individual differences. In the vast majority of these studies, the log data are aggregated for the entire language course, whereas only a few of them involve a longitudinal examination of potential shifts in students' learning behavior patterns. It is important to note that learning behavior, and consequently learning strategies, are not static constructs, but rather may change over time. Shirvani-Boroujeni and Dillenbourg (2018) mention that different longitudinal learning profiles have emerged in various studies. Some learners temporarily change their study behavior and then revert to their original one, while others permanently switch to a new learning behavior. This dynamic nature of learning strategies and behavioral patterns is an aspect that should always be taken into consideration in research studies investigating emerging learning behaviors and implemented learning tactics and strategies.

## 3.2. Inductive and deductive language learning strategies

The contrast between inductive and deductive strategies in language learning has been a widely discussed topic among experts in the domain of Second Language Acquisition (SLA), as well as in Computer Assisted Language Learning. At the core of this debate lie the two different approaches to second and foreign language teaching, the traditional teacher-centered and the modern learner-centered approach.

Traditionalist second language acquisition experts and practitioners follow a deductive language learning paradigm, where the various vocabulary elements and grammar rules are explicitly presented to the learners, followed by authentic linguistic examples which showcase the target linguistic phenomenon or element to be learned. Gollin (1998) refers to this approach as the "grammar-translation" method of learning a language, indicating one of the most prominent activities of this mode of language instruction, which involved the transformation of a linguistic utterance in the learner's native language into a grammatically sound sentence to the target language. The deductive approach to language learning offers a more structured learning experience to students, providing them with all the necessary resources and rules beforehand.

The inductive approach to second / foreign language learning is more learner-centered and involves a more experiential and exploratory learning experience. Rules and meanings are not presented explicitly to learners, but inferred through exposure to various authentic language examples. Gollin (1998) uses the term “audiolingualism” for this type of instructional approach to language, where the language systems and mechanics are not explicitly explained, but induced by carefully graded exposure to the target language.

Comparisons between those two approaches have been numerous, highlighting strong points and weaknesses of these two different language learning strategies. Brown (2000) argues that inductive language learning has the advantage of being more learner-centered in nature and encourages active learner participation. The increased learner engagement results in better understanding of the language system and mechanics and enhances learner autonomy and motivation. On the other hand, inductive language learning can be more time-consuming and cognitively demanding for language learners. Also, there is a high probability that the learners arrive at an incorrect or incomplete inference. Additionally, the exploratory nature of inductive learning strategies may also be the cause for frustration, which could be avoided in a more structured and rigid learning experience, where goals and expectations are explicitly defined and determined.

In the last decade, there has been a shift to how research in CALL approaches the two different learning strategies (Tsai, 2019). Up until then, most studies focused on comparing inductive and deductive language learning in terms of effectiveness and efficiency in specific language learning contexts. The purpose of these studies has been to suggest optimal instructional strategies and activity sequencing guidelines to instructors, in order to ensure learner success and fulfilment of the instructional goals. This approach has been slowly abandoned, as more second / foreign language acquisition experts and researchers noted various issues with research studies such as these. Sik (2015) argues that there is no consensus about the effectiveness of one approach over the other. Additionally, the researcher notes that there have been studies overlapping both approaches, claiming that both learning strategies may be implemented by a learner, switching to one or to the other according to various factors or conditions in the language learning educational setting. Mallia (2014) investigated the efficiency of the two approaches in adult learners. The researcher reported that there was no statistically detectable difference in performance for both groups. Tsai (2019) also noted that empirical data continue to show mixed results regarding which of the two approaches is more effective. The researcher attributes these controversial results to the different ways in which the deductive and inductive approach are operationalized. Shaffer (1998) notes that, even though many different research studies have been conducted that compared the two different language learning strategies, in some cases there has been conceptual inconsistency with the definition of inductive approach, treating it as habit formation by exposure to the language rather than as inference of rules and meaning from language examples. Lee and Lin (2019) investigated differences in performance between the two approaches among university-level language learners. The results showed that both the inductive and deductive groups showed vocabulary acquisition and retention, with no statistically detectable differences. The researchers attributed the similar performance of the two approaches to the fact that both learner groups (inductive and deductive) achieved deep level processing of the provided learning resources (for example, concordance lines), the only difference between them being that each group reached that stage in a different order in the learning process.

Recent research studies have emphasized the fact that these two learning strategies are not necessarily antagonistic but may coexist, and, in combination, may provide the necessary scaffolding to language learners. Gollin (1998) advocates for such a combination, referring to it as “guided discovery”. According to this approach, explicit focus on grammar rules and vocabulary items is combined with inference from examples. Thus, learners are actively engaged in the learning process, increasing motivation and performance. Lee and Lin (2019) also suggest that the combination of the two learning strategies might give much better results than relying solely on either one of them. Sik (2015) also notes that there have been studies approaching induction and deduction as complementing strategies, claiming that both learning strategies may be implemented by a learner, switching to one or to the other according to various factors or conditions in the language learning educational setting.

This conceptual shift in CALL research, regarding both deductive and inductive language learning in a more complementary and less antagonistic prism, directed study investigations towards the area of identifying factors that require implementation of one strategy over the other. The factor most discussed in the literature is the individual differences between language learners. Sik (2015) mentions that the effectiveness of each approach depends on the individual characteristics of each learner. Language learners with prior knowledge of the target language or linguistics in general (linguistic terminology, understanding of grammatical principles and concepts) will benefit more from an inductive learning strategy, as they possess the necessary information and skills to make inferences by working on authentic linguistic examples. Lee and Lin (2019) emphasize the importance of learner preference when a specific learning strategy is implemented. Evaluating an inductive approach in vocabulary acquisition, the researchers note that many learners feel uncomfortable with not knowing the exact meaning of a keyword, as they try to infer it from various sentence examples, especially when the word has an abstract sense. Additionally, learner preference may also depend on previous learning experiences, with students taught in traditional language teaching methods feeling more comfortable when implementing deductive language learning strategies. Sik (2015) also pinpoints that in her research with university students, even though the pre and post-tests do not show any statistically detectable difference between the average scores of the two groups (inductive and deductive), the perception of both teachers and learners was that the deductive approach is more effective. Mallia (2014) also argues that learners vary in their preferences and possibly differ in the benefit they obtain from each approach. The researcher notes that teacher perceptions also favour one approach over the other and advocates against *a priori* selection of a particular approach, since either of them may be appropriate, depending on the educational context.

Another set of factors favouring one language learning strategy over the other is related to the nature and complexity of the instructional goal and the target language or linguistic feature to be studied. Haight et al. (2007) argue that inductive learning is an option for languages with salient features, consistency and simplicity of use and form. Shaffer (1998) suggests that inductive language learning strategies may facilitate learning when focusing on difficult to describe grammatical concepts, which can be presented easier as working examples of language use to learners. Tsai (2019), investigating the two different learning approaches in a vocabulary acquisition context, suggests that the inductive approach is more appropriate for learning collocations and expressions, while the deductive approach facilitates definitional knowledge building. In terms of grammar learning, Abuseileek (2009) argues that for complex structures and sentences the deductive approach is more effective and leads to higher performance scores.

Regarding the operationalization of the two different learning strategies, the approaches reviewed in all the different research studies differ, which, as Tsai (2019) notes, leads to mixed and controversial results when evaluating them against learning outcomes. The most common approach is related to the order students access various tools and learning resources provided by the online language courses. Lee and Lin (2019) differentiated between the different approaches by operationalizing the inductive approach as accessing the concordance lines of vocabulary items first and then visiting the vocabulary page, which includes all the definitions of the target lexical items to be studied, where the opposite direction described a deductive language learning approach. Tsai (2019) also used student log data to categorize students to groups favoring one approach over the other. Both groups used the same set of tools to achieve the vocabulary learning goals. However, the inductive group started their learning sessions by querying the educational corpus before consulting the online dictionary provided by the digital learning environment, while the deductive group used the same tools in reverse order. Gelan et al. (2018) operationalized learning strategy by observing which of the features of the online course were first accessed in each learning session. When learners accessed exercise pages first, this was an indication of an inductive learning strategy, while, when students accessed the theory pages first, the researchers considered this action an indication of a deductive learning strategy. In research on vocabulary acquisition, Poole (2012) related the use of specific tools to each language learning strategy. The participants were given online texts containing the lexical items to be studied. The lexical items were highlighted in the text and, when clicked, provided two different types of gloss, a concordance gloss, which was connected to inductive learning, and a dictionary gloss, which was connected to deductive learning. Other studies, aiming at assessing the efficiency of each approach, artificially restricted learner access to specific tools and features of the learning environment, which were connected to each of the two different learning strategies. AbuSeileek (2009) created two different versions of a language course, one inductive and one deductive. In the inductive version of the course there were no rules section or theory pages included. In the deductive version of the course, the session begun with the video of a teacher explaining the grammar rules and explaining examples, before allowing access to the exercise section.

In conclusion, the research paradigm followed by the majority of earlier studies (Shaffer, 1998; AbuSeileek, 2009) was the creation of two groups of participants, each one interacting with the digital learning environment in predetermined way, which operationalized either the inductive or the deductive learning, and comparing the performance of these two groups. This particular research design essentially reduces student autonomy, randomly assigning a student to one of the two groups without considering their individual differences and restricting their learning actions and interactions to match the intended learning strategy. Most recent research studies on the implementation of learning strategies in digital language learning environments (Tsai, 2019; Lee & Lin, 2019; Gelan, 2018) adopt a more open and observational research paradigm, where the participants have the autonomy of implementing a variety of action sequences without any limitation or restriction beyond the affordances of the system. This approach also comes with certain issues and weaknesses. The most important limitation is that these studies do not provide much insight into causal relations between the investigated variables. In the example of language learning strategies, it is not certain whereas a specific strategy is successful for the achievement of a particular instructional goal, or whether the students with the higher ability usually choose that particular method. Another issue lies in the operationalization of the deductive and inductive language learning strategies adopted in these studies. Clickstream data captured during these interactions are analyzed, and the emerging patterns are *a posteriori* associated to either the inductive or the deductive learning strategy. The operationalizations used to perform those associations are based on the order of the actions performed or the order of the sections of the digital learning environment accessed by the learners (Tsai, 2019; Gelan et al., 2018). Even though the order in which a learner may use the various available learning resources may be a strong indication of the implemented learning strategy, it might also be misleading. For example, language learners may briefly access the exercise section of a course, in order to get an idea about what the instructional goal entails and direct their focus on the appropriate sections of the theory section. Even though, according to the order of actions, learners followed an inductive approach, this is actually an example of a deductive learning strategy. Hence, a different method of operationalization of the two different strategies may lead to more secure conclusions and interpretations about the inductive or deductive approach being implemented by the students. My thesis research adopts operationalization of these language learning strategies at a finer-grained level of analysis (activity level as opposed to overall course level), where combinations of induction and deduction may be possible. More details on this research method are presented in Chapter 5.

# Chapter 4.

# Educational and Research Setting

This chapter provides detailed information on the educational setting where the research was conducted, including the Learning Management System, the participants and the university course which implemented the Modern Greek language learning platform as part of the curriculum. The first section will focus on describing the design, features, and limitations of the online learning platform, elaborating on both structure and content. The following section refers to the participants' characteristics with the caveat that, since the researcher had access only to pre-existing data, the available information about the participants was limited. The final section of the chapter describes the university course from which the data were generated, further clarifying the way the learning management system was integrated in the learning process.

## 4.1. The "Rebooting the Greek Language" Learning Platform

*Rebooting the Greek language* is an online language learning platform with Learning Management System capabilities, which was designed and developed by the New Media Lab of the Stavros Niarchos Foundation Center for Hellenic Studies at Simon Fraser University. It is a web-based application created with the purpose of providing a digital learning environment to support learners of the Modern Greek language. It is intended to be used by students in (a) charter and day schools where Greek is part of the curriculum; (b) afternoon and Saturday Greek schools which are operated by Greek communities in the diaspora, where Greek is learned as a heritage language; and (c) post-secondary educational institutions, where Modern Greek courses are offered.

The purpose of the learning platform is to provide a complementary learning resource for the acquisition of basic knowledge and proficiency in Modern Greek language. In particular, the learning content that has been developed targets the A1 language proficiency level, as determined by the Common European Framework of Reference for Languages (Council of Europe, 2001). This framework provides a list of instructional goals for each level, which can be classified in several categories: phonetic and phonological, communicational, vocabulary, grammar, language use and cultural knowledge.

The learning material for the platform is organized in large thematic modules. Each module is based on a specific theme, approximately equivalent to the ones determined by the Common European Framework: introductions and greetings, describing a person, presenting your family, talking about your daily routine, etc. These themes are slightly modified depending on the target audience for which the online course is intended. For example, modules on school life have been included in the version that addresses younger learners whereas for adult learners there are modules on the working environment. The content of the modules has also been adjusted, reflecting the different needs and ability of each learner age group. Versions which are intended for younger audiences have a more rigid structure to reflect the greater need for scaffolding, whereas those addressing adults have less structure and a wider range of thematic ties for the content (as advocated in Garrett, 2009). A certain differentiation is evident in the instructional goals as well, with literacy-related goals and wider goal variety (i.e., communication goals in "adult" settings, like job interviews, or advanced thematic topics, like social and economic issues) appearing in the upper range of the learner age span. Finally, there is a different approach in the definitions and explicit language explanations depending on the learner age group, with the versions for younger learners implementing minimal linguistic metalanguage (linguistic and grammatical terminology) while focusing less on the description of the language system and more on examples of language use.

Finally, for the structure and implementation of the learning content and resources, I adopted the task-based language learning approach as described by Doughty and Long (2003). According to this approach, the basic level of analysis is the *task*, which corresponds to a complex instructional goal about communication in the Modern Greek language (for example, how to formally introduce yourself). Thus, the learning material and the various resources available to learners are organized in a way that prepares them to be able to perform the addressed task. That specific approach implements the whole-task approach to complex learning, advocated by several researchers (e.g., van Merrienboer & Kirschner, 2017). Hence, each module in the learning platform corresponds to a complex target learning task that has thematic ties to its topic. For example, the module about person description has at its center the task of describing to someone else the physical appearance and personality traits of a person. All the vocabulary, expressions, grammatical rules, and pragmatic/cultural information necessary to perform this task are made accessible to the learners as part of the content of the module. In this way, all the important linguistic information and the language resources connected to the task are presented to the learners in a highly contextual manner.

### 4.1.1. Module Structure

As already mentioned, the module is the core element of the learning system and serves as the basis for describing the structure and organization of the learning content in the online platform. The module is a distinct thematic and learning entity. Each one has hierarchical relationships with the other modules in the platform (some modules include linguistic knowledge which is required by other modules further up in the module hierarchy).

In order to decide on a specific organizational schema for the internal structure of the modules which would aptly fit the characteristics and special conditions of the learning context, the structure suggested by Heller et al. (2006) was adopted. Each module has a specific internal structure, which consists of smaller learning entities called *nodes*. The nodes are organized in a tree-like formation, *the skill tree*. A visual representation of a skill tree of a module is shown in Figure 4.1. For each module, there is a ranging number of *initial nodes*, but always one terminal node, which will be referred to as the *end node* of the module. In order to better explain this structural schema and how it connects to the learning concepts and the instructional goals of the system, it is important to discuss first how these skill trees derive from the learning content of the course.

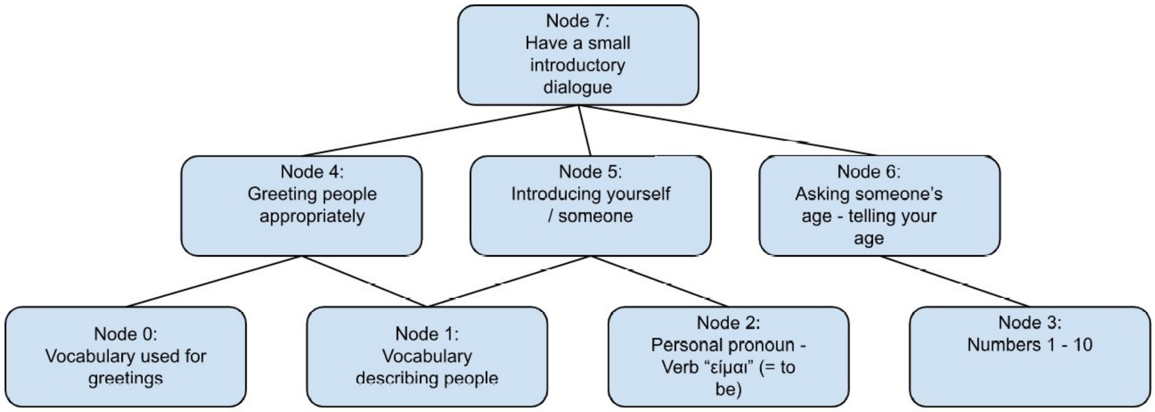



**Figure 4.1. A tree representation of how the nodes of a module in the Modern Greek course are organized**

At the centre of each module is a complex target learning task which is related to the thematic identity of that module. By means of a task analysis, this complex task is further divided into simpler (in terms of complexity) subtasks. As suggested by van Merrienboer et al. (2003), those subtasks derive naturally from the target task, and they are not arbitrary and subjective chunks of knowledge, which do not exist independently in the real world. Each node in the skill tree corresponds to a subtask, and the relations between the subtasks eventually determine the internal structure of the module. The hierarchical relationship between subtasks is determined by the prerequisite knowledge for each of these subtasks. Subtasks with more prerequisites appear in the upper levels of the hierarchy, while those with fewer prerequisites appear in the lower levels of the hierarchy.

However, the above structural schema appears somewhat incomplete, as it fails to acknowledge an important aspect of second language acquisition. Nikolov and Djigunovic (2006) argue that the processes involved in acquiring a second language can be distinguished in two broad categories relating to the procedural/declarative dimension. This distinction is based on the co-existence of two different cognitive systems which are at work during second language acquisition. The first is a rule-based system which contains powerful generative rules responsible for constructing grammatically sound sentences. The second one is a formulaic, example-based declarative system, which stores language examples, with the function of some rules operating on chunks of knowledge. This system is responsible for storing vocabulary and specific phrases, such as expressions and collocations. Of relevance to the present research is the fact that adults tend to rely more on the rule-based system, whereas children rely more on the example-based system. Nevertheless, all learners use both systems when learning a foreign language.

The existence of these two cognitive systems is represented in the structural schema that is used in the learning platform through the distinction of two different classes of nodes, namely declarative nodes and procedural nodes. The first class is related to the declarative cognitive system that is used to store vocabulary and idiomatic sentences and phrases and corresponds to *vocabulary-based* instructional goals. Vocabulary is presented in a contextualized way, as both a part of a larger thematic entity (module) and as a prerequisite for communicative tasks corresponding to nodes further up in the hierarchy. This contextual perspective for vocabulary presentation is advocated by various researchers (Levy, 2009; Groot, 2000). In the module example of Figure 4.1, the theme is about having a small introductory dialogue. Node 0 includes vocabulary on greetings, which are part of the theme (you start the dialogue by greeting someone). Node 0 also serves as a prerequisite for Node 4, which has as its instructional goal “how to greet people appropriately” (formally or informally). Even though the most prominent type of instructional goals for this class of nodes is related to vocabulary, there are also declarative nodes involving grammatical word elements, such as pronouns, prepositions or inflection of verbs, nouns and adjectives. In Figure 4.1, Node 2 is such a node, addressing the personal pronouns and the verb *είμαι* (to be).

The second class of nodes is related to the rule-based cognitive system which uses stored rules to generate grammatically sound sentences to be used in various communicative activities (i.e., activities in a defined communicative context, for example introducing yourself). These nodes correspond to *procedure-based* communicational goals which relate to the subtasks that have been mentioned above. In this way, the various grammar-based goals are not presented out of context, but with relevance to a specific task or communicational intention (Garrett, 2009). This functional approach to grammar, which allows for the integration of grammar-based goals in the procedural nodes of the hierarchy, deals with the problem of learner demotivation when studying grammar in a disconnected and decontextualized manner (Nikolov & Djigunovic, 2006). The distinction between those two different classes of nodes (declarative and procedural) is also equivalent to their structural complexity as defined by Housen and Simoens (2016). In particular, they define *structural complexity* (also *linguistic* or *absolute complexity*) as the complexity directly related to the inherent linguistic characteristics of a language feature or subsystem. Hence, since procedural nodes involve instructional goals on the sentence or contextual level of the Greek language, addressing multiple levels of linguistic description (morphology, syntax, sentence semantics and pragmatics), they can be considered as nodes of higher structural complexity. In contrast, declarative nodes, which involve instructional goals on the word level of the Greek language, usually involving just a single level of linguistic description (semantics or morphology), may be considered as nodes of lower structural complexity.

Hence, the structural schema that was adopted has the following major characteristics:

- Each module has a single goal node, which corresponds to the complex target learning task that is the object of that specific module, and it is related to its thematic identity. It appears at the top of the hierarchy. In Figure 4.1, Node 7 is the goal node of the hierarchy of the module.

- The goal node has only prerequisite nodes and itself does not constitute a prerequisite for another node in the hierarchy of the module.

- The initial nodes (which can be one or many) are nodes in the hierarchy that do not have any prerequisite nodes and they appear on the bottom of the hierarchy. In Figure 4.1, Nodes 0 to 3 are the initial nodes of the module.

- A top-down approach was followed when creating a module hierarchy. The initial point of analysis is the complex task, which constitutes the instructional goal for the whole module. All the other subtasks or subgoals naturally derive from the end task (goal)

- A bottom-up approach is followed by the learners, as they navigate the module. The learners begin with the initial nodes at the bottom of the hierarchy and, as they complete prerequisite nodes, they gain access to nodes further up in the tree.

- When learners first access a module hierarchy, only the initial nodes are unlocked (and therefore accessible to them). The nodes unlock as the learners complete their prerequisite nodes. The algorithmic rule that determines how the learners progress in the hierarchy is that a node becomes unlocked when all its prerequisite nodes are completed. In the example of Figure 4.1, Node 0 and Node 1 need to be completed to unlock Node 4.

- There are two different classes of nodes, related to two different types of instructional goals: declarative nodes and procedural nodes. The main difference between those two types of nodes is related to their structural complexity, with procedural nodes having higher structural complexity than declarative nodes. In terms of the structure of the node hierarchy inside a module, declarative nodes always precede procedural nodes, as they are considered prerequisite linguistic knowledge. In the example of Figure 4.1, Node 3 is a declarative node, while Node 4 is a procedural node.

- When an end node is completed, the whole module is considered completed and the learner may proceed to other modules. The algorithmic rule that determines how the learners progress in the hierarchy of modules is identical to the one governing progression in the hierarchy of nodes.

This structural schema presents several advantages. First, it is an accurate visualization of the knowledge space for each module in the learning platform, as it represents all its constituents as well as the relationships between them. Second, it corresponds to the specific instructional goals that comprise the Modern Greek language curriculum for the A1 proficiency level. Each of the goals may be mapped to one of the nodes inside the hierarchy of the module. Therefore, teachers are able to identify goals, as well as the learning material connected to them, and integrate the use of the learning platform seamlessly in their teaching practice. Third, the tree-form representation of the content structure can accommodate multiple purposes by serving as three different tools: (a) a navigational tool through which the learners access the content; (b) a representation of the learner's knowledge state showing completed, in-progress, and non-accessed goals; and (c) a goal setting tool used by learners to decide future steps in their learning.

### 4.1.2. Node Structure

The internal structure of a node in the Modern Greek language course consists of three different types of learning elements, all of which are connected to that instructional goal. The objects inside a node can be classified into the following distinct categories: study materials, learning activities and assessments. This internal structure accommodates the alignment of instructional goals, learning strategies and assessment, which is a fundamental feature for the design of an instructional intervention (Larson & Lockee, 2013; Garrett, 2009).

*Study materials* are learning resources that provide the learners the necessary information to prepare for the task at hand. Using the taxonomy of design elements for language learning contexts, suggested by Rienties et al. (2018), these activities are considered *assimilative*, i.e., the learners read or listen to linguistic information found in the material and resources provided. These resources are of various types, following a multi-modal approach to the language resources that are presented to the learners (Stepp-Greany, 2002). The study materials may include authentic texts, audio files of authentic dialogues, video files, vocabulary lists, vocabulary flashcards, diagrams, inflection tables, grammar rules, lists of linguistic examples and cultural texts. The learners are not required to access all these resources or study them extensively. These study materials can also be referenced after the learners have accessed the various learning activities included in the node.

*Learning activities* are learning objects inside the nodes that require some kind of active interaction on behalf of the learner. These activities can be characterized as Tutorial Computer Assisted Language Learning (CALL) (Garrett, 2009), and they are closed-response exercises (multiple choice, fill in the blanks, reading words or sentences, drag and drop, etc.), which are highly structured and do not allow for a high degree of independent learner interaction. These activities are automatically corrected by the system which also provides timely feedback. Their main purpose is for language learners to practice on. There is always a place for Tutorial CALL activities in second language acquisition as they offer learners targeted practice on the different linguistic forms and phenomena in the new language (Garrett, 2009). In the Rienties et al. (2018) taxonomy, this type of activity is referred to as *interactive/adaptive*, where learners are required to apply their knowledge, skills and understanding of the language system in order to complete communication tasks in simulated settings and get immediate feedback on their performance.

*Assessments* are learning activities that function as evaluation tools indicating whether the instructional goal has been accomplished. Assessments include closed type activities and correspond to different language skills (reading, listening, writing, and speaking), that are related to the general goal of the node. Assessments correspond to the *assessment* learning type of activities in the taxonomy by Rienties et al. (2018), where learners are assessed on their learning based on specific instructional goals. At this point it is important to note that learners do not get any feedback on quizzes, other than their final score, which includes the number of wrong answers and their score percentage. The final score indicates whether they succeeded or failed the assessment. The score threshold for success is 70%.

All the different learning objects included in each node are of specific number and type. In particular, each node contains two *study materials*. One explicitly provides the necessary vocabulary or language system descriptions that are required by the equivalent instructional goal. The other provides all the previous linguistic information implicitly through sentence examples, dialogs or authentic texts. Each node contains four *learning activities*, one for each of the four different language competences: listening, speaking, writing, and reading. Each activity includes 10 different questions, to achieve greater uniformity between different activities/nodes/modules. Finally, each node includes three different *assessments*, one for reading comprehension, one for listening comprehension and one for writing. Speaking assessments were not included in the content due to the technical limits of the platform. Voice capture and recognition technology were unavailable. Figure 4.2 shows the learning environment layout of a node, with the three different tabs (Study – Learn – Assess) which categorize the three different types of activities (study materials, learning activities and assessments).


**Figure 4.2. Screenshot of the learning environment interface when working on a node.**

## 4.2. Participants

Archival data were obtained from two cohorts of undergraduate students at a U.S. university in California, one cohort in the Fall 2020 semester (*n* = 37) and the other in the Spring 2021 semester (*n* = 35). All these students were enrolled in an undergraduate course on Greek language and history. Since the research study involved the analysis of data extracted from a Learning Management System (LMS), permission to access the data was sought from the instructor and coordinator of the course. The professor also provided some limited demographic information about the participants since this information could not be obtained from the LMS student data. From the total of 72 students of the sample, 30 were male and 42 were female students. During the first three weeks of the term, 12 students quit the course (9 from the first cohort and 3 from the second), and they were not considered for the research, reducing the sample size (*n* = 60). Additionally, two of the remaining students didn't complete all the nodes of the course, bringing the sample size down to *n* = 58, since these two are considered cases with missing data. The students' ages ranged from 18 to 30 years old, and they come from various ethnic backgrounds. It is important to note that this U.S. university is certified as a Hispanic-serving institution. None of the students had any prior knowledge of Modern Greek or any previous instruction in the language.

## 4.3. Educational Context

The online Modern Greek language course was used as part of an undergraduate course on Greek History and Language. As a dual topic combined course, it includes two separate elements, one focusing on the history and culture of Greece in the 19$^{th}$ and early 20$^{th}$ centuries. The other is a language learning component introducing the basic concepts of the Modern Greek language. This duality on the thematic nature of the course is also reflected in the duality of the mode of instruction. The course implemented a hybrid model that incorporated independent asynchronous instruction for the Greek language element together with live instruction for the history and culture element.

Since the present research study focuses on the online language learning platform supporting this course, more details on the language component of the course will be given. During the term, the students were required to complete all 8 modules in the Modern Greek language course. The instruction on Greek language was completely asynchronous and it was based on the content and resources provided in the online course. In terms of content, all the modules of the course abide by some universal standards, except for the number of nodes contained in each one of them. Hence, the number of nodes for each module varies, depending on the analysis of the specific linguistic task the module addresses. An overview of all the different features for all the nodes in the Modern Greek language course is provided in Table 4.1. The *type* column refers to the type of a node and it can either be declarative or procedural. The *prerequisite* column refers to the number of prerequisite nodes each node has. The *levels* column shows how many levels of linguistic description are involved in the instructional goal of each node (morphology, syntax, semantics, or pragmatics).

**Table 4.1. Features of all the nodes included in the Modern Greek course**

| node | type | Prerequisite nodes | Levels of linguistic description |
|---|---|---|---|
| **m1n0** | declarative | 0 | 2 |
| **m1n1** | declarative | 0 | 1 |
| **m1n2** | procedural | 0 | 1 |
| **m1n3** | procedural | 0 | 2 |
| **m1n4** | procedural | 2 | 3 |
| **m1n5** | procedural | 3 | 3 |
| **m2n0** | declarative | 0 | 1 |
| **m2n1** | procedural | 0 | 2 |
| **m2n2** | declarative | 0 | 1 |
| **m2n3** | declarative | 0 | 1 |
| **m2n4** | procedural | 1 | 3 |
| **m2n5** | procedural | 1 | 3 |
| **m2n6** | procedural | 2 | 2 |
| **m2n7** | procedural | 3 | 4 |
| **m3n0** | declarative | 0 | 2 |
| **m3n1** | declarative | 0 | 3 |
| **m3n2** | declarative | 0 | 2 |
| **m3n3** | procedural | 2 | 3 |
| **m3n4** | procedural | 1 | 3 |
| **m3n5** | procedural | 2 | 3 |
| **m4n0** | declarative | 0 | 2 |
| **m4n1** | declarative | 0 | 3 |
| **m4n2** | declarative | 0 | 1 |
| **m4n3** | declarative | 0 | 2 |
| **m4n4** | procedural | 3 | 4 |

| node | type | Prerequisite nodes | Levels of linguistic description |
|---|---|---|---|
| **m4n5** | procedural | 3 | 3 |
| **m4n6** | procedural | 2 | 3 |
| **m5n0** | declarative | 0 | 3 |
| **m5n1** | declarative | 1 | 2 |
| **m5n2** | declarative | 0 | 1 |
| **m5n3** | procedural | 1 | 3 |
| **m5n4** | procedural | 2 | 3 |
| **m5n5** | procedural | 2 | 3 |
| **m6n0** | declarative | 0 | 2 |
| **m6n1** | declarative | 0 | 1 |
| **m6n2** | declarative | 0 | 2 |
| **m6n3** | declarative | 0 | 2 |
| **m6n4** | procedural | 2 | 3 |
| **m6n5** | procedural | 2 | 3 |
| **m6n6** | procedural | 2 | 3 |
| **m7n0** | declarative | 0 | 1 |
| **m7n1** | procedural | 0 | 1 |
| **m7n2** | declarative | 0 | 2 |
| **m7n3** | declarative | 0 | 1 |
| **m7n4** | declarative | 0 | 1 |
| **m7n5** | procedural | 4 | 3 |
| **m7n6** | procedural | 3 | 3 |
| **m7n7** | Procedural | 2 | 3 |
| **m8n0** | declarative | 0 | 1 |
| **m8n1** | declarative | 0 | 1 |
| **m8n2** | declarative | 0 | 1 |
| **m8n3** | declarative | 0 | 1 |
| **m8n4** | declarative | 0 | 2 |
| **m8n5** | procedural | 2 | 2 |
| **m8n6** | procedural | 3 | 3 |
| **m8n7** | procedural | 2 | 3 |
| | | | |

## 4.4. Limitations Imposed on the Study

The nature and characteristics of the instructional design for the Modern Greek online language learning platform, as well as the technical limitations of the implementation of the design, research restrictions, and various conditions of the educational setting impose certain limitations on the research study. It is important to review these limitations and discuss how they may affect the research.

As it has been mentioned, the Modern Greek language course is a beginner's course in the language. Hence, there are certain *content limitations*. First, the vocabulary to be acquired is basic, with most words having concrete rather than abstract senses. Hence, cases of semantic ambiguity are practically non-existent. Additionally, due to the structure and the task-based language learning approach to the instructional design, the vocabulary elements for each node are closely connected to each other, sharing common properties or attributes, instead of being totally unrelated to each other. The elementary level of the Modern Greek language course has also implications for the sentence complexity encountered by the learners. Most sentences have simple syntactic structure, consisting mainly of a single clause with a small number of constituents. Sentences become more complex in the latter modules of the course, without reaching extremely complicated structures. This fact also limits the cases of syntactic ambiguity to be found in the course.

Another form of content limitation derives from the technical limitations of the online language platform, as well as the intentional purpose of the educational software. As mentioned previously, the platform implements closed-response, tutorial CALL learning activities and assessments, while more open-ended language production tasks do not appear in the course. This decision was made due to the lack of a Greek language parser in the system, which would be able to provide real time automated feedback on the grammatical structure and meaning of students' answers. Additionally, the intended use of the platform was as a supplementary resource rather than a standalone language learning tool to be implemented in a flipped classroom foreign language learning model. In such a blended language learning setting, learners will use the platform to gain insight and basic knowledge of a particular task or instructional goal before extending their knowledge and understanding of the linguistic concepts and structures involved by participating in more open and collaborative activities in the classroom (Evseeva & Solozhenko, 2015).

Additionally, certain *research limitations* were due to access restrictions to data. In particular, the researcher was granted permission to access only archived data of students in the course, i.e., the student logs that capture learner activity in the Modern

Greek online language learning platform. Apart from this data, the instructor shared some basic demographic information on the student cohorts under investigation in the present research. However, no further information and data on the students were obtainable, such as their specific GPA or prior language learning experiences. It was not possible to conduct further investigation using other research instruments such as questionnaires.

Finally, certain *research limitations* were imposed by restrictions on the educational setting. In particular, the language element of the undergraduate course implemented the Modern Greek online language learning platform as the standalone and only learning tool. Learners used the language resources provided by the learning environment to acquire the necessary knowledge, practice their skills, and assess their proficiency in Modern Greek. The final grade derived directly from the student's performance in the assessment tests of the platform. Therefore, no other metrics of student performance on the course, such as a cumulative final exam, were available to the researcher.

# Chapter 5.

# Methodology

The present chapter presents the research design and methods I adopted to address the research questions introduced in Chapter 1. The first section presents the philosophical assumptions that shaped the method of inquiry, and describes the research design that was implemented for the investigation. The second section presents the concept of Linguistic Complexity Index, which aims to measure the structural complexity of language course components (in this research, the nodes of the Modern Greek Language course). It also describes the process of calculating the index for each individual node. The third section introduces the data collection instrument and discusses the format of the captured log data, the process of data screening and the technical limitations imposed during the data collection process. The final section revisits the research questions in light of the aforementioned methodological features, and it reformulates them using the specified metrics and variables introduced in Chapters 2 and 3.

## 5.1. Epistemological considerations and method of inquiry

The philosophical framework that guided the purpose and research questions of this thesis is pragmatism as developed by philosophers such as Dewey (1908) and Peirce (1905/1998) and applied to research methodology by Parvais et al. (2016) and others. The pragmatic orientation is mostly evident in the focus of this study on investigating an educational phenomenon under the lens of a particular problem which, in this case, is informing and improving the instructional design of a computer assisted language learning application. Hence, the goal is not to uncover an "objective truth" about processes involved in second language learning, but, rather, to discover "what works" and provide solutions to established problems and challenges in the field (Parvaiz et al., 2016).

The pragmatic focus of this research is reflected in the specified research questions which guided the research design, and the data collection instruments. The research strategy was selected to accommodate the purpose of this study and the context bound (i.e., the Modern Greek online language course) nature of the data, and it is oriented towards specific, real-world goals, which are improving the learning platform and the cognitive scaffolds provided to learners. The intention is to understand learner interactions and behavior within the system in order to evaluate and further improve the learning experience.

Considering the aforementioned theoretical and epistemological considerations, as well as the aims of the research study, a quantitative observational research approach, based on data captured during the learners' interactions with the Modern Greek online language learning platform, is the most appropriate for addressing the research questions. Hence, this research relies heavily on the analysis of unobtrusively collected clickstream data.

The emergence and subsequent extensive use of digital learning environments in various educational settings allowed for easier and more efficient collection of clickstream data. Many researchers have emphasized the advantages of such data for research on the learning process, while pinpointing certain issues and challenges presented by the implementation of such methods. Gibson (2018) claims that learner attributes should be mapped to automated data collection instead of test measurements and assessments, since the learner's behavior is not burdened by the awareness of being evaluated. Blikstein et al. (2014) also argue that unobtrusive, automated, real-time collection allows for the use of instrumentation that provide fine-grained data, which has the potential of advancing research in learning by revealing detailed trajectories during a learning activity. However, they point out that openness of the learning context, which provides learners with the freedom to generate and pursue different solutions to a problem, is an important requirement for the general usefulness of these data. Matcha et al. (2019) pinpoint the importance of unobtrusive data collection methods in the detection of learning strategies and tactics, as they do not increase the learners cognitive load, as well as their advantages over self-report instruments (surveys or think aloud protocols), since learners are not always accurate in reporting how they learn. The researchers also argue that self reports may fail to capture how strategies are developed over time.

Winne (2017) notes that clickstream data are the surface manifestation (and for this reason observable and tangible) of the internal cognitive processes of learning, whose characteristics can only be inferred by the researchers. For example, when learners highlight a particular sentence or paragraph, the analysis may assume that they have identified certain elements in these bodies of texts that met specific criteria which correspond to the requirements (standards) learners hold for completing a task. In comparison, self-report data may be more representative of the thinking procedures and reasoning of the learner, but these data are also influenced by the learner's efforts at impression management, biases, as well as memory restrictions, especially in the case of greater lag between the learning activity and the self-reporting activity. Another limitation mentioned by Winne (2017) is related to a frequent inability of an LMS to determine more in-depth information about the data collected. For example, it is very difficult to get any information about whether the learner actually studied the material he or she accessed, or the reasoning process (or absence thereof) that resulted in the implementation of a particular strategy during the interaction materials provided within the learning environment.

Duval et al. (2012) also raise the issue of determining the relevance of a learner action to a certain learning process. For example, a certain action (a click) may be accidental or due to confusion about the interface, and therefore not meaningfully related to an ongoing learning activity. Gelan et al. (2018) advise against using clickstream data in isolation because, even though they may provide a detailed picture of the surface behavior of the learners, they usually offer minimum insight on the reasons and planning behind the learners' actions. Rienties et al. (2018) also comment on the same issue, suggesting that clickstream data analysis should be accompanied by additional data capable of providing insight on the affordances of the digital environment and the ways students learn when interacting with it. Information on the learning context in which learning takes place is crucial for the adequate analysis and interpretation of the clickstream data collected during the research.

## 5.2. Linguistic Complexity Index

As it was mentioned in the introductory chapter, one of the purposes of the current thesis is to propose metrics that reflect the difficulty of a unit in the Modern Greek language course and to evaluate their utility for both learners and instructors in predicting the *cognitive complexity* of the respective content unit, i.e., the cost to a learner's cognitive assets to complete the unit. Currently in the digital learning environment dashboard of the online Modern Greek language course, some estimates of difficulty for a node are included. However, these estimates present two major issues:

- Some estimates are not directly related to the nature and characteristics of the linguistic content of each course unit. For example, the *number of prerequisite nodes* is related to the hierarchical relations between the nodes in a module, but not to the inherent complexity of the node per se
- Some other estimates refer to the inherent complexity of a node, but they do so in a rather superficial way. For example, the *number of levels of linguistic description* indicates only how many of these levels are involved in the instructional goal of a node, without actually naming them.

Therefore, an additional aim of the present study is to propose a metric which will reflect difficulty as the *structural complexity* of a node (as it was discussed in Chapter 2), based on the nature and the various characteristics of its linguistic content.

In that respect, the operationalization of the Linguistic Complexity Index (LCI) derives from the relevant literature and research on defining and measuring the structural complexity of a linguistic object on multiple levels, from a simple linguistic construct like the plural form of a noun in a language to a whole language system. In the case of the Modern Greek language course, the LCI will be used to assess the difficulty of an instructional goal for each node. The calculation of the index for a given node depends on the type of that node. For declarative nodes, the instructional goal translates to a set of lexical items (words) to be acquired by the learner. For procedural nodes, the linguistic goal is translated into a target language utterance (phrase or sentence). It is important to note that only linguistic knowledge new to the learner is considered in these calculations. Prerequisite linguistic knowledge is not to be considered in the calculation of the LCI for a node. The reason behind this decision lies in the fact that the LCI measures the inherent complexity of the node, while prior linguistic knowledge is related to learners' individual differences, i.e., their ability to maintain knowledge already acquired and not forgetting it. However, the relationship between these two variables may be the focus for a future study.

The fundamental characteristic of this index is that it represents a construct, and it consists of three different aspects, thus reflecting the multi-aspect operationalization of linguistic complexity suggested by various researchers (Palotti, 2015; Pandarova et al., 2019; Housen & Simoens, 2016). These different aspects, as described in the relevant literature, are lexical complexity, morphological complexity and syntactic complexity, and they need to be considered separately, since each of them involves different attributes and features of the language unit.

Lexical complexity is directly related to the lexical items involved in the instructional goal of a specific node. In the calculation of that particular subindex, the number of newly introduced lexical items (i.e., the vocabulary) of a particular node is considered (Palotti, 2015; Housen & Simoens, 2016). To account for the fact that lexical complexity is also dependent on the semantic nature of a given word, with words with a concrete sense corresponding to less lexical complexity than the ones with an abstract sense (Palotti 2015), the number of lexical items is multiplied by 1 for words of the first case and by 2 for words of the latter. Thus, for a node introducing a new vocabulary of 12 words, from which 7 have a concrete sense and 5 have an abstract sense, the lexical complexity is calculated to be

$$(7 \times 1) + (5 \times 2) = 17.$$

There are a couple of points that need to be further discussed, considering some of the observations made by researchers on the aspect of lexical complexity. Bulte and Housen (2012) propose a different categorization of lexical items, as *lexical* words and *function* words, to account for differentiated complexity, with function words having mainly grammatical attributes and not a specific semantic sense (for example, personal pronouns or articles). In the present approach, function words are considered as abstract lexical items, while their grammatical attributes (morphology and syntax features) will be considered when estimating the other two aspects of the complexity construct (morphological and syntactic complexity). Additionally, some researchers (Bulte & Housen 2012; Pandarova et. al. 2019; Housen & Simoens, 2016; Revesz et al. 2017) argued that the frequency of a lexical item should be considered when estimating lexical complexity, as less frequent words present more difficulty in their acquisition in comparison to more frequent ones. However, there is not a widely accepted metric to determine this frequency, and when such metrics are suggested (such as in the case of Coh-Metrix index, McNamara et al., 2005), the metric is highly dependent on discourse. For example, certain terms might be more frequent in a scientific text, while others might have higher frequency in a fiction novel. Since the present approach is implemented in a beginner's language course in which the vocabulary introduced is a basic one, as determined by the Common European Framework of Reference for Languages (Council of Europe, 2001), all the lexical items included are of relatively high frequency and frequency was not considered in the estimation of the lexical complexity index. Nevertheless, this is an issue that needs to be revisited in the future when different proficiency levels or domain specific language courses might be investigated. Another issue to be considered is related to determining the weights for concrete and abstract words when calculating their lexical complexity. Unfortunately, literature doesn't provide a validated operationalization for this calculation. Instead of adopting arbitrary weights for each word sense type, I have used the ratio of average time of completion of nodes with arbitrary vocabulary items to nodes with concrete vocabulary items, which is approximately 2:1. Finally, in the case of different morphological variations for the same lexeme (for example different gender types, like ψηλός – ψηλή (masculine tall – feminine tall), those variations have been considered as the same word for the estimation of lexical complexity, since the different allomorphs are to be considered when estimating morphological complexity. The only exception is in derivational morphemes, which formulate different lexemes and are considered different words (for example βιβλίο (book) – βιβλιοθήκη (bookcase)), as has been suggested by Housen and Simoens (2016). Finally, lexical compounds (multi-word phrases) and expressions are also considered as single lexical units, as suggested by Palotti (2015), since their meaning does not derive as the sum of all the different senses of the included words.

Morphological complexity is related to the different morphemes attaching to a word which give it certain grammatical properties. As has already been mentioned, the focus will be solely on inflectional morphology, since derivational morphology is considered in the estimation of lexical complexity. The morphological complexity index is estimated by multiplying the number of different morphemes introduced to the learners in a particular node to a composite number calculated by multiplying the number of different values for each of the attributes involved in the instructional goal. This is a mathematical interpretation of the *morphological patterns* suggested by Haspelmath and Sims (2010), which refers to the form-function relationship for each group of allomorphs. As an example of a morphological complexity calculation, for a linguistic goal that involves the inflection of the verb μένω in the present tense, there are six different suffixes (allomorphs) which manifest two different attributes of a verb, the number (singular or plural) and the person (first, second or third). Hence, the morphological complexity for that instructional goal is calculated as

$$6 \times (2 \times 3) = 36.$$

Finally, syntactic complexity refers to the proper order of lexical items when forming phrases or sentences. Thus, this type of complexity is directly related to the complexity of the target sentence. In the syntactic level, there are two different dimensions when estimating the complexity of a particular sentence. The first is the number of clauses in the sentence and the second is the maximum number of constituents for each of these clauses. These metrics reflect not only the length of the sentence (which is the metric for syntactic complexity suggested by various researchers (Bulte & Housen 2012; Palotti 2015), but also the number of arguments attached to the verb of each clause. Calculation of the syntactic complexity index is performed by multiplying those two numbers. For example, the target sentence Ο Νίκος μένει στην Αθήνα και ο Γιάννης μένει στην Πάτρα (Nikos lives in Athens, and Giannis lives in Patra) has two clauses and each clause has two constituents attached to the verb. Hence, the syntactic complexity is calculated as

$$2 \times 3 = 6$$

It is important to note that syntactic complexity is considered entirely as a within-sentence linguistic phenomenon. Reference words connecting lexical units from different sentences and similar discourse syntactic phenomena were beyond the scope of this study. The reason behind this decision lies in the elementary nature of the Modern Greek course, which does not include any instructional goals referring to discourse linguistics. Table 5.1 shows all the nodes in the Modern Greek course, with their three calculated subindices corresponding to the three types of complexity.

Another point among the limitations of the procedure to estimate linguistic complexity that requires further discussion concerns the issue of ambiguity in all the different levels of linguistic description (morphological, syntactic, semantic). Ambiguous linguistic structures add to the complexity of the language unit, since the learner needs to determine which of the different alternatives applies in a specific case of language use. Since the Modern Greek language course investigated in the present study is essentially a beginner's course, it contains basic vocabulary and simple morphological forms and syntactic structures which do not present any type of linguistic ambiguity. Hence, dealing with this issue is beyond the scope of the present thesis. However, this is an issue that needs to be addressed in the future, as the linguistic complexity index will be implemented in more advanced courses or even in different languages.

As a final point of discussion in the present section, the issue of how linguistic complexity as a construct is translated by the LCI is revisited. The relevant literature suggests that the construct of complexity can only be considered as a set of different dimensions, each one of them focusing on a specific level of linguistic description, from morphology to semantics. The challenging part is to clearly define the mapping of those individual dimensions to a single composite variable. Song et al. (2013) distinguish different approaches for creating a composite variable from several "indicator" variables. In simple averaging, the composite variable derives from the sum of the *z* scores of the original variables. One important thing to keep in mind for that approach is that the contribution of each indicator variable to the composite is considered equal. Therefore, the approach is not appropriate for indicator variables with dissimilar relationships to any outside variables involved in the study (Song et al., 2013). The weighted averaging approach applies a principal component analysis to the standardized values of the original variables. It is important to note that, since each of the suggested composite variables (different components created by the analysis) are orthogonal to each other, choosing one of these components as the composite variable has an impact to the predictive power of the composite variable. Each one of these approaches has its advantages and drawbacks, and it needs to be considered against the specific conditions and variable relationships of this research study. Therefore, one of the goals of the present research is to determine which of these approaches to the formation of the LCI composite variable is the most appropriate under the specific circumstances presented by the Modern Greek Language course.

**Table 5.1. Values of individual complexity indices for all the nodes in the Modern Greek Language course**

| NODE | LEXICAL COMPLEXITY | SYNTACTIC COMPLEXITY | MORPHOLOGICAL COMPLEXITY |
|---|---|---|---|
| M1N0 | 24 | 0 | 4 |
| M1N1 | 30 | 0 | 0 |
| M1N2 | 2 | 2 | 24 |
| M1N3 | 0 | 2 | 8 |
| M1N4 | 0 | 3 | 4 |
| M1N5 | 4 | 4 | 28 |
| M2N0 | 14 | 0 | 0 |
| M2N1 | 9 | 2 | 11 |
| M2N2 | 21 | 0 | 0 |
| M2N3 | 10 | 0 | 0 |
| M2N4 | 2 | 5 | 12 |
| M2N5 | 7 | 3 | 6 |
| M2N6 | 27 | 4 | 21 |
| M2N7 | 0 | 4 | 21 |
| M3N0 | 0 | 0 | 10 |
| M3N1 | 0 | 0 | 7 |
| M3N2 | 32 | 0 | 0 |
| M3N3 | 0 | 4 | 3 |
| M3N4 | 5 | 4 | 12 |
| M3N5 | 0 | 5 | 29 |
| M4N0 | 9 | 0 | 26 |
| M4N1 | 10 | 0 | 5 |
| M4N2 | 7 | 0 | 0 |
| M4N3 | 24 | 0 | 5 |
| M4N4 | 36 | 4 | 21 |
| M4N5 | 0 | 4 | 23 |
| M4N6 | 0 | 4 | 28 |
| M5N0 | 15 | 0 | 6 |
| M5N1 | 15 | 0 | 3 |
| M5N2 | 15 | 0 | 0 |
| M5N3 | 0 | 4 | 12 |
| M5N4 | 0 | 4 | 21 |
| M5N5 | 0 | 4 | 21 |
| M6N0 | 8 | 0 | 0 |
| M6N1 | 12 | 0 | 0 |
| M6N2 | 15 | 0 | 5 |
| M6N3 | 24 | 0 | 7 |
| M6N4 | 0 | 5 | 13 |
| M6N5 | 0 | 5 | 13 |
| M6N6 | 0 | 6 | 13 |

| NODE | LEXICAL COMPLEXITY | SYNTACTIC COMPLEXITY | MORPHOLOGICAL COMPLEXITY |
|---|---|---|---|
| **M7N0** | 8 | 0 | 0 |
| **M7N1** | 12 | 4 | 0 |
| **M7N2** | 15 | 0 | 0 |
| **M7N3** | 12 | 0 | 0 |
| **M7N4** | 12 | 0 | 0 |
| **M7N5** | 0 | 8 | 12 |
| **M7N6** | 3 | 8 | 24 |
| **M7N7** | 0 | 10 | 30 |
| **M8N0** | 19 | 0 | 0 |
| **M8N1** | 11 | 0 | 0 |
| **M8N2** | 11 | 0 | 0 |
| **M8N3** | 10 | 0 | 0 |
| **M8N4** | 10 | 0 | 6 |
| **M8N5** | 0 | 8 | 0 |
| **M8N6** | 0 | 5 | 12 |
| **M8N7** | 0 | 8 | 12 |

## 5.3. Data Collection and Formatting

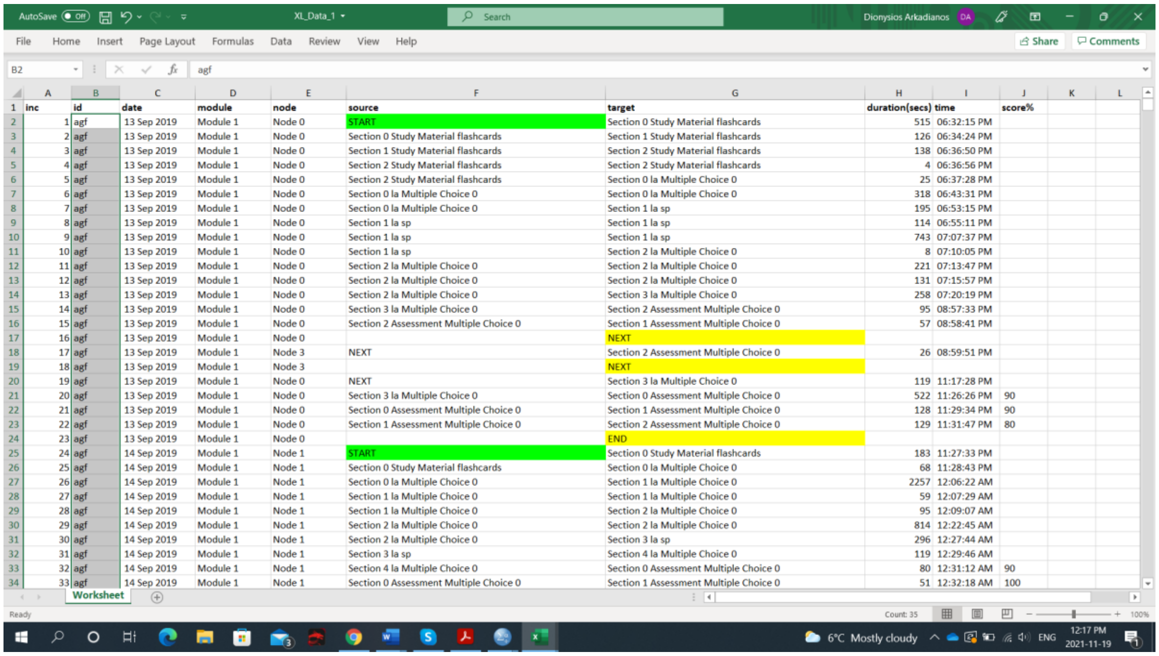
The digital learning platform which hosts the Modern Greek language course for this study implements a mechanism that tracks learner behavior in the digital learning environment via data logs. The data logs are extracted for each student participating in the course as text files with the format shown in Figure 5.1.

```
PATH: Level 3 Module 3 Node 0 Section 0 Study Material flashcards TIME: 27 seconds 08:57:29 AM 17 Sep 2020
PATH: Level 3 Module 3 Node 0 Section 0 la Multiple Choice 0 TIME: 148 seconds 09:00:04 AM 17 Sep 2020
PATH: Level 3 Module 3 Node 0 Section 0 Assessment Multiple Choice Assessment 0 TIME: 93 seconds 09:01:47 AM 17 Sep 2020
PATH: Level 3 Module 3 Node 0 Section 0 Assessment Multiple Choice Assessment 0 TIME: 89 seconds 09:03:18 AM 17 Sep 2020
PATH: Level 3 Module 3 Node 0 Section 1 Assessment Multiple Choice Assessment 0 TIME: 57 seconds 09:04:23 AM 17 Sep 2020
PATH: Level 3 Module 3 Node 0 Section 0 Study Material flashcards TIME: 181 seconds 09:07:33 AM 17 Sep 2020
PATH: Level 3 Module 3 Node 0 Section 0 la Multiple Choice 0 TIME: 4 seconds 09:07:41 AM 17 Sep 2020
PATH: Level 3 Module 3 Node 0 Section 1 la Multiple Choice 0 TIME: 7 seconds 09:07:52 AM 17 Sep 2020
PATH: Level 3 Module 3 Node 0 Section 2 la sp TIME: 4 seconds 09:07:59 AM 17 Sep 2020
PATH: Level 3 Module 3 Node 0 Section 0 Assessment Multiple Choice Assessment 0 TIME: 15 seconds 09:08:17 AM 17 Sep 2020
PATH: Level 3 Module 3 Node 0 Section 1 la Multiple Choice 0 TIME: 6 seconds 09:08:26 AM 17 Sep 2020
PATH: Level 3 Module 3 Node 0 Section 0 Study Material flashcards TIME: 2 seconds 09:08:29 AM 17 Sep 2020
PATH: Level 3 Module 3 Node 0 Section 3 la Multiple Choice 0 TIME: 9 seconds 09:08:43 AM 17 Sep 2020
PATH: Level 3 Module 3 Node 0 Section 0 Assessment Multiple Choice Assessment 0 TIME: 92 seconds 09:10:16 AM 17 Sep 2020
PATH: Level 3 Module 3 Node 0 Section 0 la Multiple Choice 0 TIME: 17 seconds 09:10:43 AM 17 Sep 2020
PATH: Level 3 Module 3 Node 0 Section 1 la Multiple Choice 0 TIME: 114 seconds 09:12:39 AM 17 Sep 2020
PATH: Level 3 Module 3 Node 0 Section 0 Assessment Multiple Choice Assessment 0 TIME: 101 seconds 09:14:23 AM 17 Sep 2020 SCORE: 70%
PATH: Level 3 Module 3 Node 0 Section 1 Assessment Multiple Choice Assessment 0 TIME: 815 seconds 09:27:42 AM 17 Sep 2020
PATH: Level 3 Module 3 Node 0 Section 3 la Multiple Choice 0 TIME: 99 seconds 09:29:31 AM 17 Sep 2020
PATH: Level 3 Module 3 Node 0 Section 1 la Multiple Choice 0 TIME: 5 seconds 09:29:41 AM 17 Sep 2020
PATH: Level 3 Module 3 Node 0 Section 0 Study Material flashcards TIME: 777 seconds 09:42:40 AM 17 Sep 2020
PATH: Level 3 Module 3 Node 0 Section 0 Study Material flashcards TIME: 33 seconds 09:43:14 AM 17 Sep 2020
```

**Figure 5.1. Example of a student log.**

In the data logs, each action performed by the learner is recorded as a single line. The first part of the record is the path to the activity that the learner attempts. The path indicates the module number, node number, and activity name. The record also includes the time the learner spent in the activity, measured in seconds, a time stamp indicating when the learner accessed the activity, and the score the learner achieved when successfully completing an assessment activity.

The raw text data logs had to be formatted into a form that would allow further data analysis which provided metrics aligned with the research questions. Using simple scripts, the text log files were formatted into Excel files of a specific structure. Figure 5.2 shows an example of such an Excel file. The module and node columns indicate the specific module and node the learner was working on when the particular action took place. The source and target columns have been created as such, to provide data for the transitions the learners made during their interaction with the system. Source is the activity the learner was interacting with, and target is the activity the learner accessed next. The duration column indicates the time in seconds the learner spent in the target activity. The date and time columns comprise the time stamp indicating when the learner accessed the target activity. Finally, the score column shows the score the learners achieved when successfully completed an assessment activity. It is important to emphasize that an assessment is considered completed by the system when the learner achieves a score of at least 70%, hence, the scores presented in this column are 70% or higher



| inc | id | date | module | node | source | target | duration(secs) | time | score% |
|---|---|---|---|---|---|---|---|---|---|
| 1 | agf | 13 Sep 2019 | Module 1 | Node 0 | START | Section 0 Study Material flashcards | 515 | 06:32:15 PM | |
| 2 | agf | 13 Sep 2019 | Module 1 | Node 0 | Section 0 Study Material flashcards | Section 1 Study Material flashcards | 126 | 06:34:24 PM | |
| 3 | agf | 13 Sep 2019 | Module 1 | Node 0 | Section 1 Study Material flashcards | Section 2 Study Material flashcards | 138 | 06:36:50 PM | |
| 4 | agf | 13 Sep 2019 | Module 1 | Node 0 | Section 2 Study Material flashcards | Section 2 Study Material flashcards | 4 | 06:36:56 PM | |
| 5 | agf | 13 Sep 2019 | Module 1 | Node 0 | Section 2 Study Material flashcards | Section 0 la Multiple Choice 0 | 25 | 06:37:28 PM | |
| 6 | agf | 13 Sep 2019 | Module 1 | Node 0 | Section 0 la Multiple Choice 0 | Section 0 la Multiple Choice 0 | 318 | 06:43:31 PM | |
| 7 | agf | 13 Sep 2019 | Module 1 | Node 0 | Section 0 la Multiple Choice 0 | Section 1 la sp | 195 | 06:53:15 PM | |
| 8 | agf | 13 Sep 2019 | Module 1 | Node 0 | Section 1 la sp | Section 1 la sp | 114 | 06:55:11 PM | |
| 9 | agf | 13 Sep 2019 | Module 1 | Node 0 | Section 1 la sp | Section 1 la sp | 743 | 07:07:37 PM | |
| 10 | agf | 13 Sep 2019 | Module 1 | Node 0 | Section 1 la sp | Section 2 la Multiple Choice 0 | 8 | 07:10:05 PM | |
| 11 | agf | 13 Sep 2019 | Module 1 | Node 0 | Section 2 la Multiple Choice 0 | Section 2 la Multiple Choice 0 | 221 | 07:13:47 PM | |
| 12 | agf | 13 Sep 2019 | Module 1 | Node 0 | Section 2 la Multiple Choice 0 | Section 2 la Multiple Choice 0 | 131 | 07:15:57 PM | |
| 13 | agf | 13 Sep 2019 | Module 1 | Node 0 | Section 2 la Multiple Choice 0 | Section 3 la Multiple Choice 0 | 258 | 07:20:19 PM | |
| 14 | agf | 13 Sep 2019 | Module 1 | Node 0 | Section 3 la Multiple Choice 0 | Section 2 Assessment Multiple Choice 0 | 95 | 08:57:33 PM | |
| 15 | agf | 13 Sep 2019 | Module 1 | Node 0 | Section 2 Assessment Multiple Choice 0 | Section 1 Assessment Multiple Choice 0 | 57 | 08:58:41 PM | |
| 16 | agf | 13 Sep 2019 | Module 1 | Node 0 | | NEXT | | | |
| 17 | agf | 13 Sep 2019 | Module 1 | Node 3 | NEXT | Section 2 Assessment Multiple Choice 0 | 26 | 08:59:51 PM | |
| 18 | agf | 13 Sep 2019 | Module 1 | Node 3 | | NEXT | | | |
| 19 | agf | 13 Sep 2019 | Module 1 | Node 0 | NEXT | Section 3 la Multiple Choice 0 | 119 | 11:17:28 PM | |
| 20 | agf | 13 Sep 2019 | Module 1 | Node 0 | Section 3 la Multiple Choice 0 | Section 0 Assessment Multiple Choice 0 | 522 | 11:26:26 PM | 90 |
| 21 | agf | 13 Sep 2019 | Module 1 | Node 0 | Section 0 Assessment Multiple Choice 0 | Section 1 Assessment Multiple Choice 0 | 128 | 11:29:34 PM | 90 |
| 22 | agf | 13 Sep 2019 | Module 1 | Node 0 | Section 1 Assessment Multiple Choice 0 | Section 2 Assessment Multiple Choice 0 | 129 | 11:31:47 PM | 80 |
| 23 | agf | 13 Sep 2019 | Module 1 | Node 0 | | END | | | |
| 24 | agf | 14 Sep 2019 | Module 1 | Node 1 | START | Section 0 Study Material flashcards | 183 | 11:27:33 PM | |
| 25 | agf | 14 Sep 2019 | Module 1 | Node 1 | Section 0 Study Material flashcards | Section 0 la Multiple Choice 0 | 68 | 11:28:43 PM | |
| 26 | agf | 14 Sep 2019 | Module 1 | Node 1 | Section 0 la Multiple Choice 0 | Section 1 la Multiple Choice 0 | 2257 | 12:06:22 AM | |
| 27 | agf | 14 Sep 2019 | Module 1 | Node 1 | Section 1 la Multiple Choice 0 | Section 1 la Multiple Choice 0 | 59 | 12:07:29 AM | |
| 28 | agf | 14 Sep 2019 | Module 1 | Node 1 | Section 1 la Multiple Choice 0 | Section 2 la Multiple Choice 0 | 95 | 12:09:07 AM | |
| 29 | agf | 14 Sep 2019 | Module 1 | Node 1 | Section 2 la Multiple Choice 0 | Section 2 la Multiple Choice 0 | 814 | 12:22:45 AM | |
| 30 | agf | 14 Sep 2019 | Module 1 | Node 1 | Section 2 la Multiple Choice 0 | Section 3 la sp | 296 | 12:27:44 AM | |
| 31 | agf | 14 Sep 2019 | Module 1 | Node 1 | Section 3 la sp | Section 4 la Multiple Choice 0 | 119 | 12:29:46 AM | |
| 32 | agf | 14 Sep 2019 | Module 1 | Node 1 | Section 4 la Multiple Choice 0 | Section 0 Assessment Multiple Choice 0 | 80 | 12:31:12 AM | 90 |
| 33 | agf | 14 Sep 2019 | Module 1 | Node 1 | Section 0 Assessment Multiple Choice 0 | Section 1 Assessment Multiple Choice 0 | 51 | 12:32:18 AM | 100 |

**Figure 5.2. Example of an Excel file extracted from the raw log data (in txt format).**

At this point, it is important to discuss some technical limitations of the data collection algorithm used in the study, which also delimited the data analysis procedures implemented. First, the time stamps created in the student logs refer to a student accessing the three different types of activities: study, learning and assessment activities. However, the algorithm doesn't capture any other learner interactions with the system. For example, accessing the different navigation views (module or node view in the dashboard) to review the description and instructional goals of a particular module or node is an important element of learner behavior which unfortunately could not be recorded in the student logs. Accessing the performance view is also a learner interaction which was not captured by the Learning Management System. The data collection instrument also lacked the capability to capture the performance scores of failed assessments. Thus, only scores equal to or greater than 70% (the performance threshold that determines the successful completion of an assessment activity) are present in the student logs. Finally, aside from activity type, the data capture algorithm failed to capture other information on the activities a student accessed. Therefore, there is no information about whether the learner reviewed study material with explicit or implicit linguistic information, or which language skill the activity was focused on.

Another limitation of the data capturing instrument is related to learner activity outside the system. The language learning platform captures learner activity only about activities provided by the learning environment. This fact creates two major implications to consider in the analysis of the collected data. Firstly, there is no way to determine whether the learners are engaged in the accessed activity, or they are doing something totally irrelevant to the learning task (Kovanovic et al., 2015). Secondly, the instrument does not have the potential to capture learner activity outside the system which may relate to the learning process, i.e., creating notes when reviewing some study materials or consulting notes when working on learning or assessment activities.

No learner self-reports were embedded in the architecture of the digital learning environment. Coupled with the secondary nature of the data used in the research study, the absence of these features prevents any insight to learner beliefs, reports, and interpretations of their learning activity. This leaves the student log data as the only fingerprint of the students' learning behavior in the Modern Greek online language learning course.

## 5.4. Data screening and preparation

This section of the chapter presents the data screening and preparation procedures which were implemented prior to the statistical analysis of the data. These procedures deal with certain issues that have been addressed in prior similar research.

One of the student log metrics that was heavily implemented in the statistical analyses conducted for the present study was the *duration* of a particular task measured in seconds. This metric was operationalized as the difference between the time stamp when the learner exited an activity minus the time stamp when the learner accessed that activity. Kovanovic et al. (2015) argue that this approach to time-on task-based measures is one widely adopted in many research studies. However, they also note that not many of these studies provide more details on such operationalizations, especially when dealing with the issue of upper-limit outliers of time-on-task values. In particular, the researchers point that, while the learners have an activity page of the digital learning environment open, they may engage in some alternative activity outside the system or in some off-task activity. This will result in upper-limit time-on-task outliers which may distort the statistical analyses results. Several time-oriented heuristics have been implemented in prior research to address this issue. However, there is not a method which has a clear advantage over the others, even though it has been demonstrated that the selection of such a strategy can alter the outcome of the statistical analyses (Kovanovic et al., 2015).

One of the metrics needed to address the research questions that were introduced in the first chapter of the thesis refers to the time of completion for a particular node, meaning the time in seconds a learner needed to complete all the assessment activities for that particular node and achieve a minimum score of 70% in each of them. Of particular importance are two descriptive statistics based on time of completion, namely the mean and the standard deviation of the time of completion for a node. To compute these statistics, a specific procedure has been followed when extracting and screening data from the student logs.

For each student in the sample, I extracted the data in the duration *(*in seconds*)* column of the student log Excel file. Subsequently, the data were screened for extreme outliers, using the *z*-scores method (as outlined in Meyers et al., 2016, p. 49). The *z*-scores were calculated for each of the 58 students who successfully completed the course, and the cases with scores $z \geq \pm 3.29$ were considered outliers (Rotelli & Monreale, 2022). To deal with outliers, the method of Winsorizing was implemented (Dixon, 1960). A duration with $z > 3.29$ was substituted with the value of the next lower duration for the same student, a method suggested in different references (Kovanovic et al., 2015, Rotelli & Monreale, 2022). Subsequently, for each of the 58 students, the different duration times for all the activities in the course were organized per node and added together to calculate the *total time on node* for each of the students. Thus, distributions of the *total time on node* for all 58 students of the cohort were created, one for each of the 56 nodes of the course. For each of these distributions, the *mean* and *standard deviation* were calculated.

The second research question refers to learning tactics and strategies learners implement to address the challenges and succeed in the instructional goals posed by the different nodes in the Modern Greek language course. The usual procedure adopted in similar research studies investigating learning strategies and tactics involves extraction of action sequences from the student log and implementation of different methods of analysis, such as network analysis, sequence analysis, or process mining (Matcha et al.

2019) to identify such tactics and strategies. For the extraction process, most studies used a unit of analysis and segmentation for the student logs which is most frequently referenced as a *learning session.* Operationalization of a learning session differs from study to study. Some research studies consider a learning session as learner activity between two consecutive learner logins in the digital learning system (Gelan et al., 2018) while other researchers operationalize learning sessions as learner activity between two consecutive periods of inactivity, using different heuristics to define the duration of these inactive periods (Siadaty et al., 2016).

In the present study, the focus was on investigating learner tactics and strategies in relation to the different characteristics of the nodes in the Modern Greek online language learning course. For that reason, a different unit of analysis/segmentation of learner activity is implemented, the *learning episode.* Learning episodes are operationalized as student activity in a specific node. Therefore, Excel scripts were used to extract action sequences corresponding to each one of these *learning episodes* per student. For each action sequence, the user id and the node id were recorded.

As a next step in data preparation, the action sequences with the highest frequencies in the distribution of all the student logs were extracted. As an arbitrary rule for determining these sequences, which are referred as *learning tactics*, the frequency threshold of 100 was implemented, meaning only learning tactics with frequencies equal to or greater than 100 were considered. The seven different learning tactics derived from this process and their frequencies are shown in Table 5.2. The learning tactics were given the same labels used for the three different tabs of the node interface (STUDY, LEARN, TEST), as shown in Figure 4.2. It is important to note that the eighth highest learning tactic had a frequency of 47, far below the 100-frequency threshold. In terms of learning strategies, we considered the operationalizations reviewed in Chapter 3 to determine whether one of these learning tactics may refer to an inductive or a deductive strategic approach to language learning. The strategy type of each of the tactics is also presented in Table 5.2.

**Table 5.2. Learning tactics with frequencies greater than 100.**

| tactic symbol | tactic | strategy | frequency |
|---|---|---|---|
| A | TEST | INDUCTION | 645 |
| B | TEST – STUDY - TEST | INDUCTION | 215 |
| C | LEARN – TEST | INDUCTION | 120 |
| D | LEARN | INDUCTION | 192 |
| E | STUDY – TEST | DEDUCTION | 536 |
| F | STUDY – LEARN – TEST | DEDUCTION | 1407 |
| G | STUDY | DEDUCTION | 664 |

In Chapter 3, it was mentioned that operationalizing deductive or inductive learning strategies based on the order of learner actions might be misleading. Therefore, an additional, temporal aspect of learner behavior was considered in the study in an effort to triangulate the underlying learning strategy implemented by the students. Time-related learning behavior metrics implemented in prior research (Gelan et al., 2018; Veletsianos et al., 2021) have been considered, to estimate the learners' focus on the different types of activities. Hence, a variation of the *activity speed* variable (in Veletsianos et al., 2012) was adopted, operationalized as the *mean time in seconds* per specific activity type (study material, learning activity or assessment). The intent was to examine how the differences in the mean time on a specific activity type per node correspond to the linguistic complexity of that node, suggesting change in the adopted learning strategy. The theoretical reasoning is that an increase of the *mean time on study materials* is suggestive of a deductive learning strategy, while an increase of the *mean time on learning activities* indicates an implemented inductive learning strategy (Tsai, 2019).

## 5.5. Research Questions

As it has been mentioned, the present study aims at proposing different approaches to measure and represent learning content difficulty and learner behavioral patterns in an online Modern Greek language learning course. In that respect, two major research questions are involved, each referring to language learning content and to student learning patterns respectively. Those two overarching research questions are subsequently analyzed to sets of sub-questions, to better account for all related issues to be investigated by this research study.

The first research question entails a two-part procedure. In the first part, all the different estimates of difficulty for a node that are already featured in the dashboard of the digital learning environment (type of node: declarative or procedural, number of prerequisite nodes, number of levels of linguistic description involved in the instructional goal) are evaluated in terms of how well they predict the cognitive complexity of that node. The cognitive complexity of the node is operationalized as the time the learners need to complete the node by successfully completing all the assessment activities for that node, an operationalization suggested in Housen and Simoens (2016). The second part is essentially an investigation of the process to calculate the proposed Linguistic Complexity Index (LCI) and evaluate its prediction of the cognitive complexity of a node. An additional point of investigation is how much overlap exists between the structural complexity metrics related to the explained variance of the dependent variable.

The second research question focuses on the students' learning behavior and the strategies they implement to successfully complete the Modern Greek language course. The student logs created by the Learning Management System were used to extract the data that outline those behavioral patterns exhibited by the learners during their interactions with the online platform. The purpose of this research question is to investigate emerging behavioral patterns, both on the tactical and strategic levels, and correlate them to the three subindices of Linguistic Complexity (lexical, morphological or syntactic), to discover interesting connections that may inform the instructional design. This approach aims to investigate how to present learning resources depending on the complexity characteristics of a particular node. For investigating the tactical level of learning behavior exhibited by the students in the Modern Greek online language course, the list of the most frequent learning tactics presented in Table 5.3 will be used. For the strategic level of learning behavior, the tactics in Table 5.3 were classified as deductive or inductive, based on operationalizations adopted in prior research. Additionally, the *activity speed* for each activity type (mean time on activity type) was considered as an additional indication of the underlying learning strategy. Regression analyses were conducted to evaluate the association of the exhibited learning behavior with the different aspects of complexity of each of the nodes in the course.

The final form of the research questions of the thesis, as well as the sub-questions deriving from each one of them, are presented in Chapter 6, along with the results of the research.

# Chapter 6.

# Results

This chapter presents the results of the statistical analyses implemented to address the two research questions of the present study. The chapter is organized in two sections, each dedicated to a research question. The first section presents the results of the analysis concerning the different structural complexity metrics for each of the nodes in the course, as well as the Linguistic Complexity Index introduced in the study. The second section addresses the learning tactics and strategies research question, with the statistical analyses attempting to associate learning behavior patterns to the different Linguistic Complexity subindices of the different nodes in the Modern Greek online language platform. The statistical analyses were conducted using the IBM SPSS 27.0 statistics software.

## 6.1. Results for Research Question 1

The first research question focuses on investigating the appropriateness of various node features, existing or suggested, as estimates of the effort that is required from the learner to successfully complete that node.

**RQ1. Which features of a node are good predictors of the cognitive complexity of that node?**

### 6.1.1. Are type of node (declarative or procedural), number of prerequisite nodes and number of levels of linguistic description good predictors of time of completion for a particular node?

In order to predict the mean and standard deviation of time of completion for a node in the Modern Greek language course, two multiple regression analyses were performed using type of node, number of prerequisite nodes and number of levels of linguistic description as predictor variables. The sample for each analysis consisted of $N$ = 56 nodes. The mean and standard deviation of time of completion for each node was calculated from a sample of $n$ = 58 students.

First, the assumptions for making valid inferences from the regression analyses were tested. The normality of the residuals of the regression was determined by inspecting the normal P – P plots of the regression standardized residuals. As can be seen in Figures 6.1 and 6.2, the residuals approximately conform to the diagonal normality line. The assumption of homoscedasticity was checked by examining the scatterplots of the residuals. Figures 6.3 and 6.4 show that the residuals are randomly distributed, suggesting homoscedastic distributions. Finally, for multicollinearity, the VIF values for the predictors are well bellow 10, indicating that there is no multicollinearity in both regressions. The VIF values are shown in Tables 6.3 and 6.4.

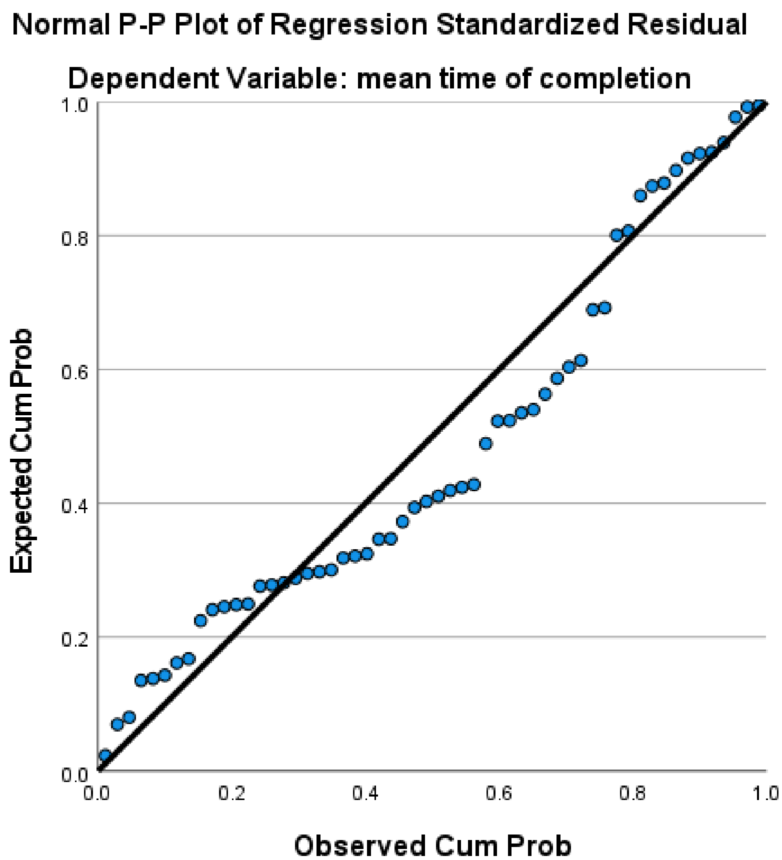



**Figure 6.1. Normal P – P plots of the regression standardized residuals for *the mean time of completion* for a node.**

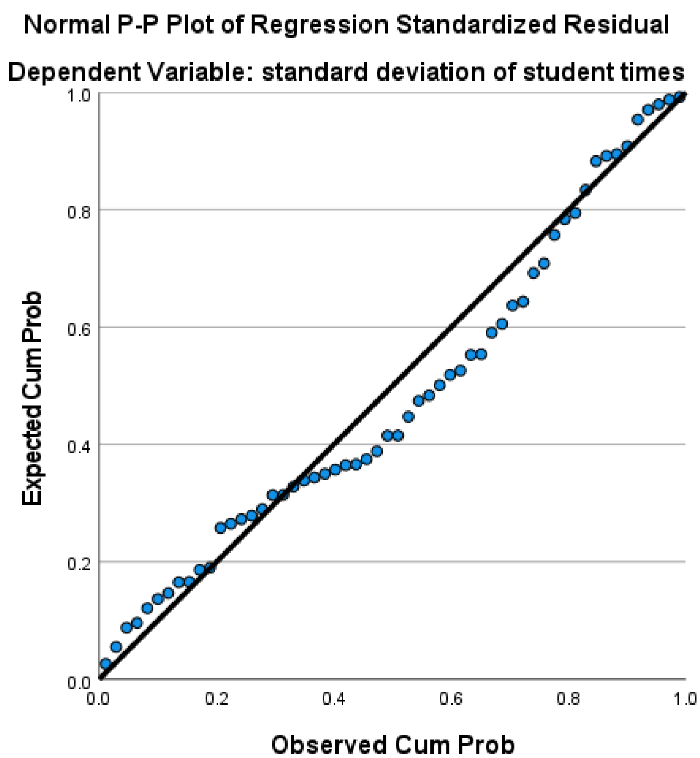



**Figure 6.2.** **Normal P – P plots of the regression standardized residuals for the *standard deviation of time of completion* for a node.**

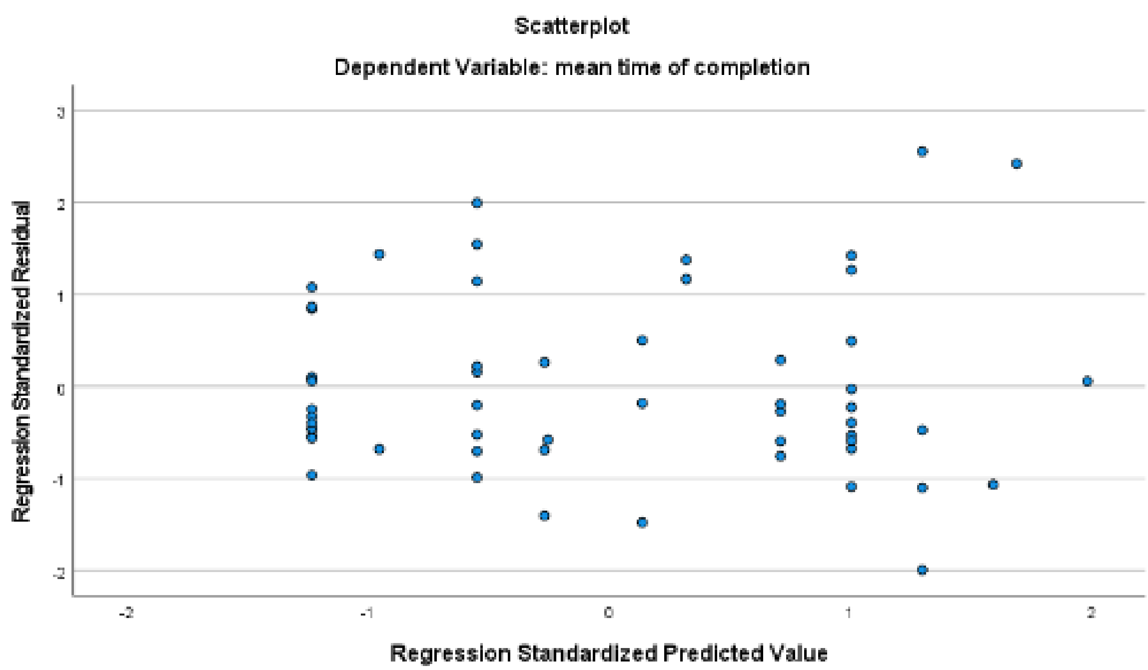



**Figure 6.3.** **Scatterplot of the residuals for the *mean time of completion* for a node.**

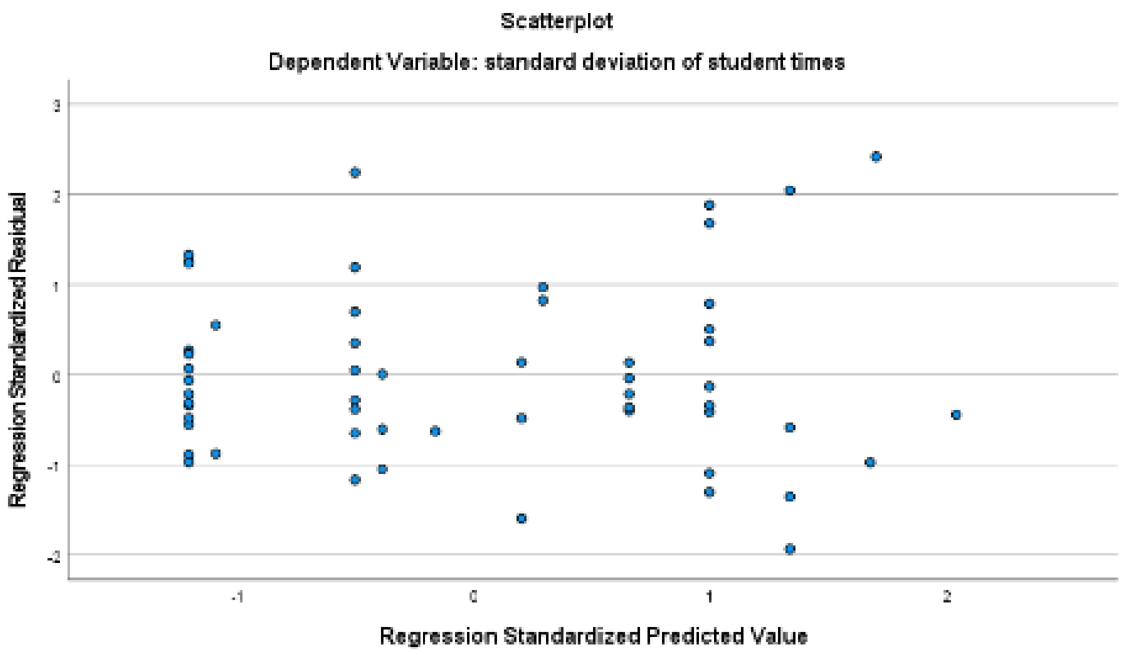



**Figure 6.4. Scatterplot of the residuals for the *standard deviation of time of completion* for a node.**

The correlations of the variables are presented in Tables 6.1 and 6.2. All the correlations in the table were statistically detectable ($p < .001$).

**Table 6.1. Correlations among predictors of mean time to completion.**

| Variable | 2 | 3 | 4 |
|---|---|---|---|
| 1. mean time of completion | .536 | .643 | .599 |
| 2. node type | - | .651 | .741 |
| 3. levels of linguistic description | | - | .719 |
| 4. prerequisites | | | - |

Note: Sample size is N = 56. All correlations were statistically detectable ($p < .001$).

**Table 6.2. Correlations among predictors of standard deviation of time to completion.**

| Variable | 2 | 3 | 4 |
|---|---|---|---|
| 1. standard deviation of time of completion | .470 | .590 | .550 |
| 2. node type | - | .651 | .741 |
| 3. levels of linguistic description | | - | .719 |
| 4. prerequisites | | | - |

Note: Sample size is $N = 56$. All correlations were statistically detectable ($p < .001$).

In the regression analysis predicting mean time to completion, the model was statistically detectable, $F(3, 52) = 14.542$, $p < .001$, and accounted for 42.5% of the variance of the predicted variable ($R^2 = .456$ and *adjusted* $R^2 = .425$). Higher number of levels of linguistic description per instructional goal primarily predicted higher mean time to completion for a given node. Table 6.3 presents the raw and standardized regression coefficients of the predictors, along with their correlations with the mean time to completion, the squared semipartial correlations and the structure coefficients.

**Table 6.3. Regression predicting mean time to completion.**

| Model | *b* | *SE b* | β | Pearson *r* | $sr^2$ | Structure Coefficient | VIF |
|---|---|---|---|---|---|---|---|
| **Constant** | 954.384 | 285.771 | | | | | |
| **Type of node** | 177.932 | 294.283 | .095 | .536 | .004 | .794 | 2.368 |
| **Levels of linguistic description*** | 436.115 | 158.758 | .418 | .643 | .079 | .953 | 2.211 |
| **Prerequisites** | 187.156 | 140.975 | .228 | .599 | .018 | .887 | 2.820 |

Note: Sample size is $N = 56$. $sr^2$ is the squared semi-partial correlation. *$p < .05$

Further inspection of the individual predictors revealed that only the number of levels of linguistic description involved in an instructional goal was a statistically detectable predictor within the model ($t = 2.747$, $p = .008$) and it was given a substantial weight in the model ($\beta = .418$). Additionally, the correlations between the different predictor variables are quite high, as it can be seen in Table 6.1, hence the unique variance explained by each of the variables indexed by the squared semipartial correlations was quite low. The strongest predictive power is connected to the levels of linguistic description variable, which uniquely accounts for approximately 8% of the variance of the dependent variable.

In the regression analysis predicting standard deviation, the model was statistically detectable, $F(3, 52) = 10.683$, $p < .001$, and accounted for approximately one third of the variance of the predicted variable ($R^2 = .381$ and *adjusted* $R^2 = .346$). Higher number of levels of linguistic description per instructional goal primarily predicted higher standard deviation of time to completion for a given node. Table 6.4 presents the raw and standardized regression coefficients of the predictors, along with their correlations with the standard deviation of time to completion, the squared semipartial correlations and the structure coefficients.

**Table 6.4.** **Regression predicting standard deviation of time to completion.**

| Model | b | SE-b | β | Pearson r | $sr^2$ | Structure Coefficient | VIF |
|---|---|---|---|---|---|---|---|
| **Constant** | 656.495 | 215.852 | | | | | |
| **Type of node** | 47.446 | 222.281 | .036 | .470 | .001 | .761 | 2.368 |
| **Levels of linguistic description*** | 290.567 | 119.914 | .393 | .590 | .070 | .955 | 2.211 |
| **Prerequisites** | 140.246 | 106.482 | .241 | .550 | .021 | .890 | 2.820 |

Note: Sample size is $N$ = 56. $sr^2$ is the squared semi-partial correlation. *$p$ < .05

The individual predictors were examined further, revealing that only the number of levels of linguistic description involved in an instructional goal was a statistically detectable predictor within the model ($t$ = 2.423, $p$ = .019) and it was given a substantial weight ($\beta$ = .393). Also, the unique variance of the standard deviation explained by the predictor variables is very low (the highest percent being connected to the levels of linguistic description variable, which uniquely accounts for 7% of the variance of the dependent variable).

### 6.1.2. For the *Linguistic Complexity Index,* is a single composite index or a set of three subindices for *lexical complexity, morphological complexity* and *structural complexity* respectively more appropriate for estimating the *structural complexity* of a node?

To answer this research sub-question and determine the most appropriate methodological approach to the formation of the Linguistic Complexity Index (LCI), the three indicator variables (*lexical complexity, morphological complexity* and *syntactic complexity*) were calculated for each one of the 56 nodes of the Modern Greek language course, as described in Chapter 3. After this step, a number of statistical analyses were conducted, in order to determine which of the approaches – simple or weighted averaging– will be implemented to determine the LCI for each of the nodes.

First, correlations between the indicator variables and the predicted variables (mean time to completion and standard deviation of time to completion) were calculated, to determine the magnitude of their relationship. The results are presented in Table 6.5. There is a statistically detectable positive correlation of syntactic complexity with mean time to completion and standard deviation of time to completion ($r$(54) = .576, $p$ < .001 and $r$(54) = .531, $p$ < .001 respectively), as well as a statistically detectable positive correlation of morphological complexity with mean time to completion and standard deviation of time to completion ($r(54) = .549$, $p < .001$ and $r(54) = .518$, $p < .001$ respectively). Additionally, there are statistically detectable correlations among the indicator variables, a negative correlation of lexical complexity with syntactic complexity and lexical complexity with morphological complexity, and a positive correlation of syntactic complexity with morphological complexity. Since the nature and the magnitude of the relationships of the three indicator variables with the predicted variables greatly differ, the simple averaging approach to the formation of the composite variable was judged not appropriate (Song et. al., 2013).

**Table 6.5. Correlations of the indicator complexity variables (*lexical complexity, syntactic complexity and morphological complexity*) with the predicted variables (*mean time to completion* and *standard deviation of time to completion*) and with each other.**

| Variable | 3 | 4 | 5 |
|---|---|---|---|
| 1. mean time of completion | -.101 | .576** | .549** |
| 2. standard deviation of time of completion | -.082 | .531** | .518** |
| 3. lexical complexity | - | -.532** | -.379* |
| 4. syntactic complexity | | - | .606** |
| 5. morphological complexity | | | - |

Note: Sample is $N = 56$. *$p < .05$ **$p < .001$

In order to proceed with the weighted averaging approach, the standardized $z$ scores of the three indicator variables need to be obtained. This standardization is necessary when calculating the weights of the composite variable, so that the association between the composite variable and the predicted variables (mean time to completion and standard deviation of time to completion) will not be overly affected by an indicator variable with large variance (Song et al., 2013).

As a next step, a principal component analysis of the three standardized indicator variables was performed on the data of 56 nodes. Because of the relatively large sample size, the variables to cases ratio was deemed adequate. The Kaiser-Meyer-Olkin measure of sampling adequacy was .642, which is close to the threshold of .70 which was suggested by Meyers et. al. (2016) for data suitable for a principal component analysis. Also, sufficient correlation between the variables was indicated by the statistically detectable result of Bartlett's test of sphericity ($p < .001$), in order to proceed with the analysis.

Only one component had an eigenvalue greater than 1.00 (Eigenvalue = 2.017), accounting for 67.23% of the total variance. Since only one component was extracted by the analyses, there was no need for rotation. The variable loadings that will be used as weights for the calculation of the composite LCI variable are shown in Table 6.6***.***

**Table 6.6. Structure coefficients: one factor unrotated solution.**

| Indicator Variable | Factor 1 |
|---|---|
| 1. Lexical Complexity | -.764 |
| 2. Syntactic Complexity | .882 |
| 3. Morphological Complexity | .810 |

The weights for each of the indicator variables presented in Table 6.6 were used to calculate the overall Linguistic Complexity Index (LCI), by multiplying the value of each variable with its equivalent weight and adding all the results together. Thus, each node has a unique LCI value, which will be used when addressing the second sub-question for research question 1, in the next subsection of this chapter. Additionally, the LCI will be also considered as a set of three different values, each one corresponding to one of the different aspects of linguistic complexity (lexical, morphological and syntactic). However, for each of these three indicator variables, their standardized *z* scores will be considered instead of their actual values, to allow for a more direct comparison of these aspects on the same scale.

Two pairs of multiple regression analyses were performed to investigate the adequacy of the *LCI* as a composite variable as well as a set of the three different subindices (lexical complexity, morphological complexity and syntactic complexity) as predictors of the two metrics of time to completion, mean and standard deviation.

First, the assumptions for making valid inferences from the regression analyses were tested. The normality of the residuals of the regression was determined by inspecting the normal P – P plots of the regression standardized residuals. As can be seen in Figures 6.5 through 6.8, the residuals approximately conform to the diagonal normality line. The assumption of homoscedasticity was checked by examining the scatterplots of the residuals. Figures 6.9 through 6.12 show that the residuals are randomly distributed, suggesting homoscedastic distributions. Finally, for multicollinearity, the VIF values of the three indicator variables are well bellow 10, indicating that there is no multicollinearity in both regressions. The VFI values are shown in Tables 6.7 and 6.8.

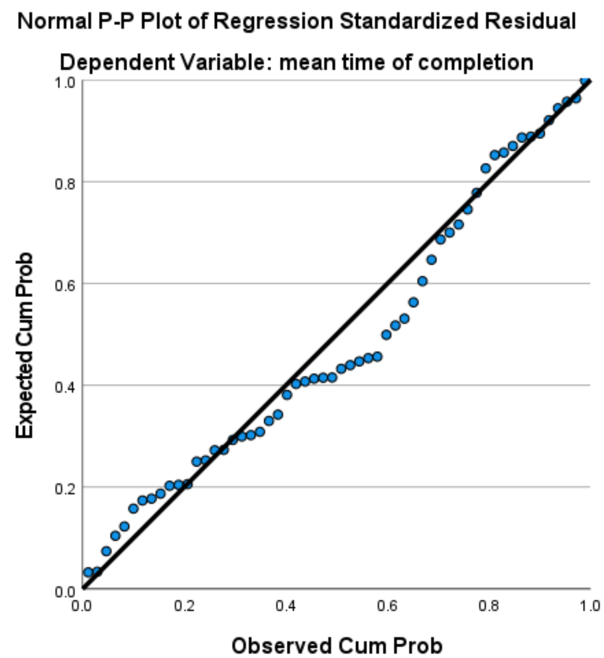



**Figure 6.5.** **Normal P – P plot of the regression standardized residuals for the mean time to completion using the three linguistic complexity subindices as predictors.**

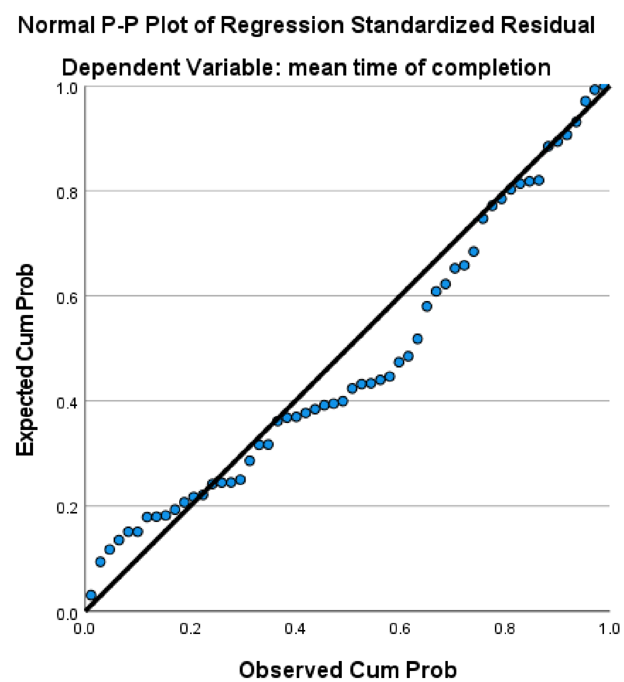



**Figure 6.6.** **Normal P – P plot of the regression standardized residuals for the mean time to completion using the single LCI as predictor.**

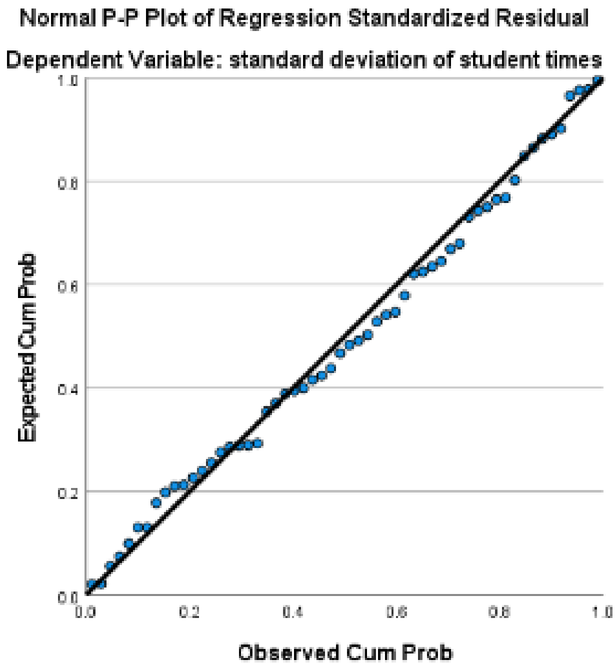



**Figure 6.7.** **Normal P – P plot of the regression standardized residuals for the standard deviation of time to completion using the three linguistic complexity subindices as predictors.**

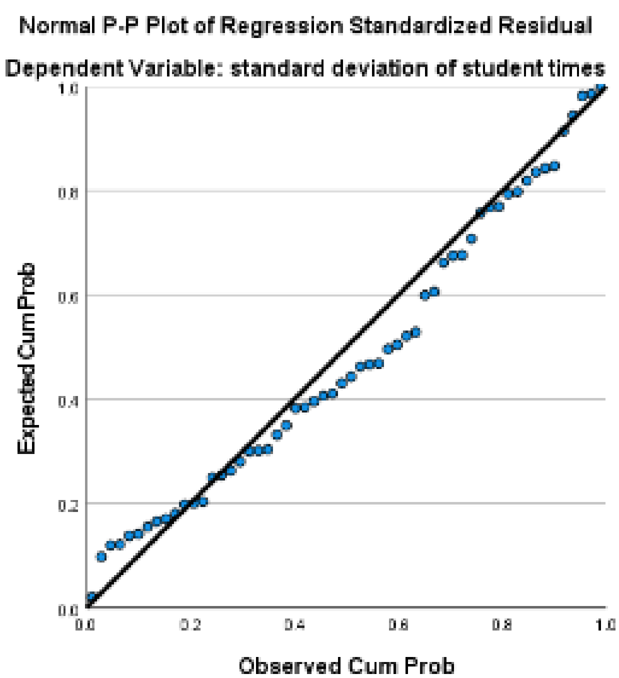



**Figure 6.8.** **Normal P – P plot of the regression standardized residuals for the standard deviation of time to completion using the single LCI as predictor.**

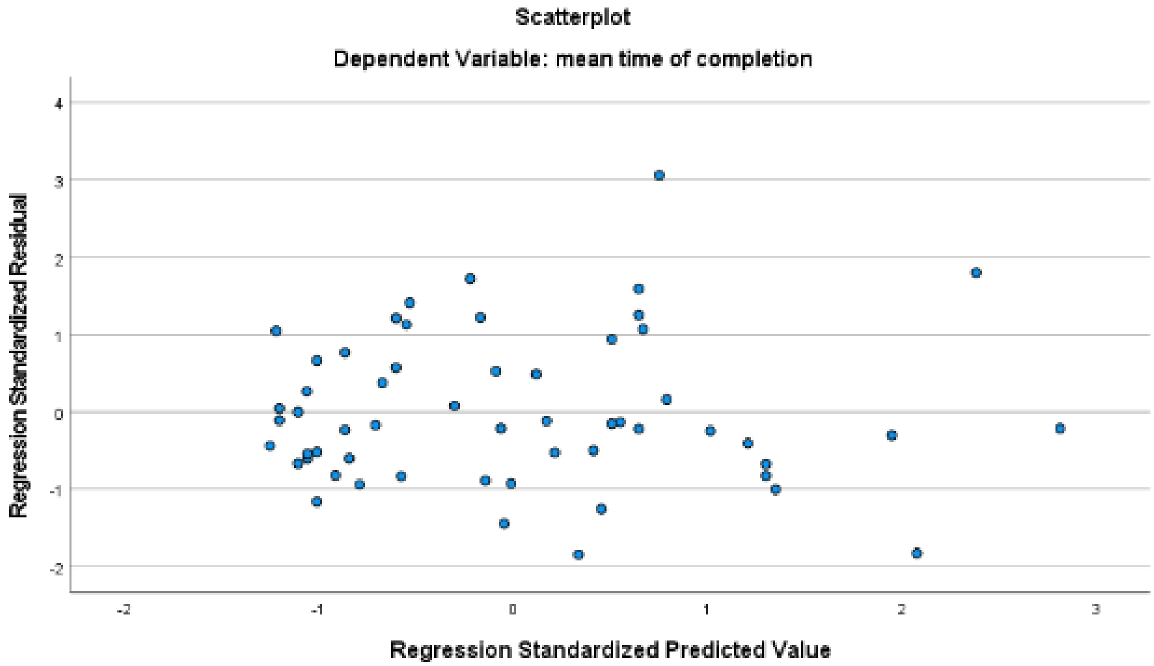



**Figure 6.9.** **Scatterplot of the residuals for the mean time to completion using the three linguistic complexity subindices as predictors.**

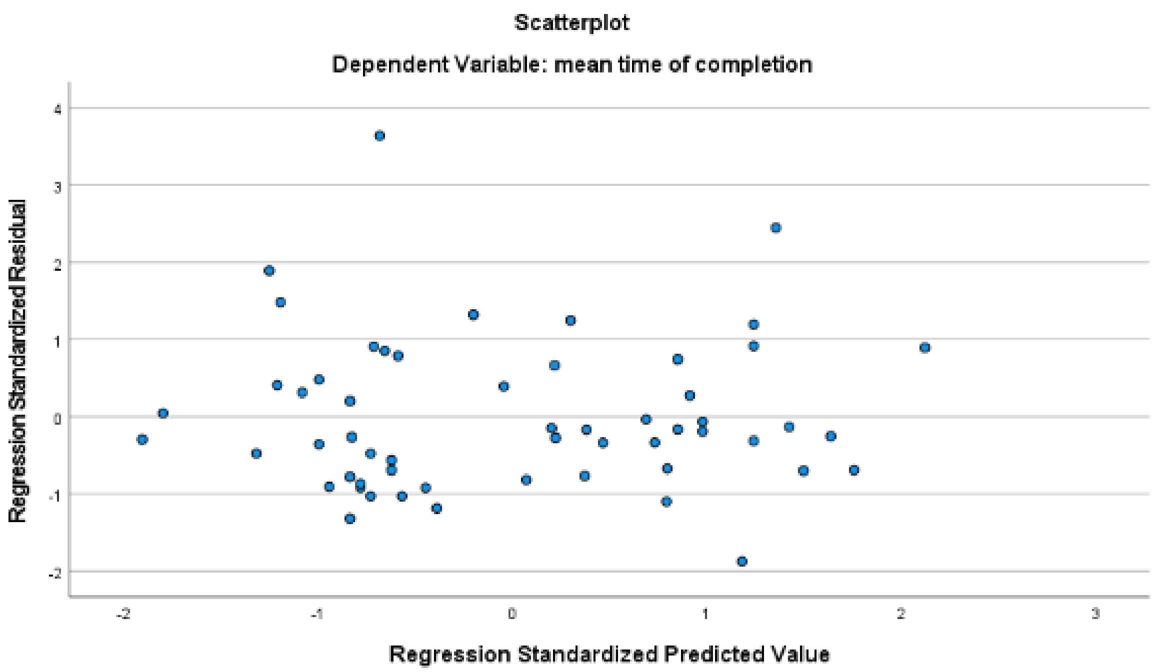



**Figure 6.10.** **Scatterplot of the residuals for the mean time to completion using the single LCI as predictor.**

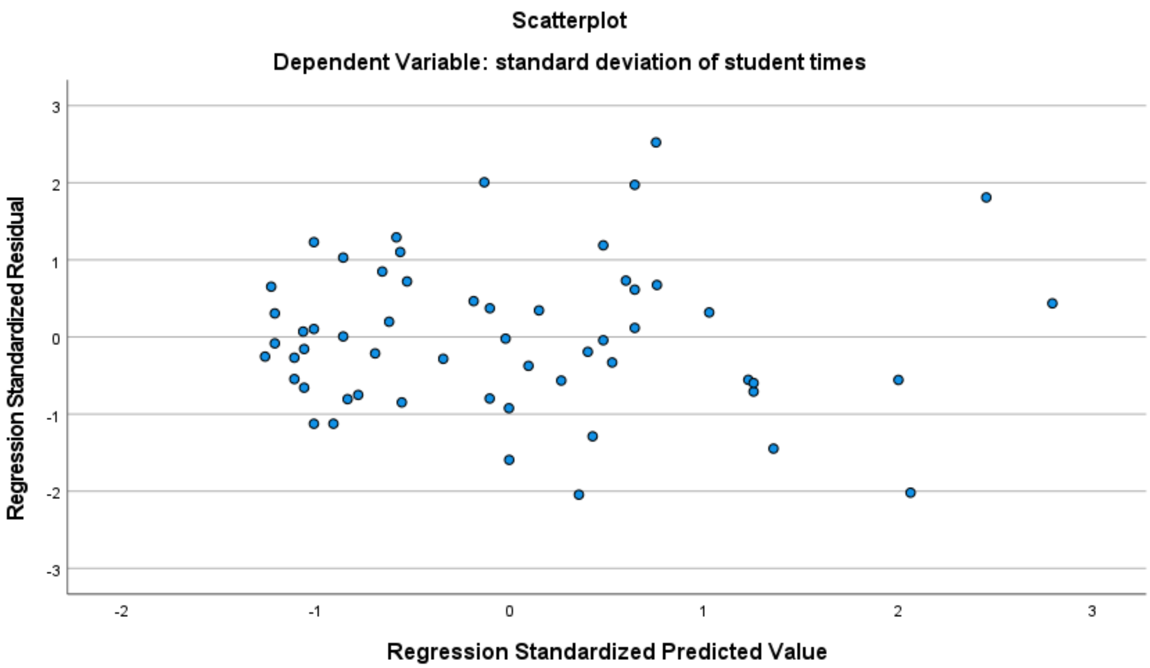



**Figure 6.11.** **Scatterplot of the residuals for the standard deviation of time to completion using the three linguistic complexity subindices as predictors.**

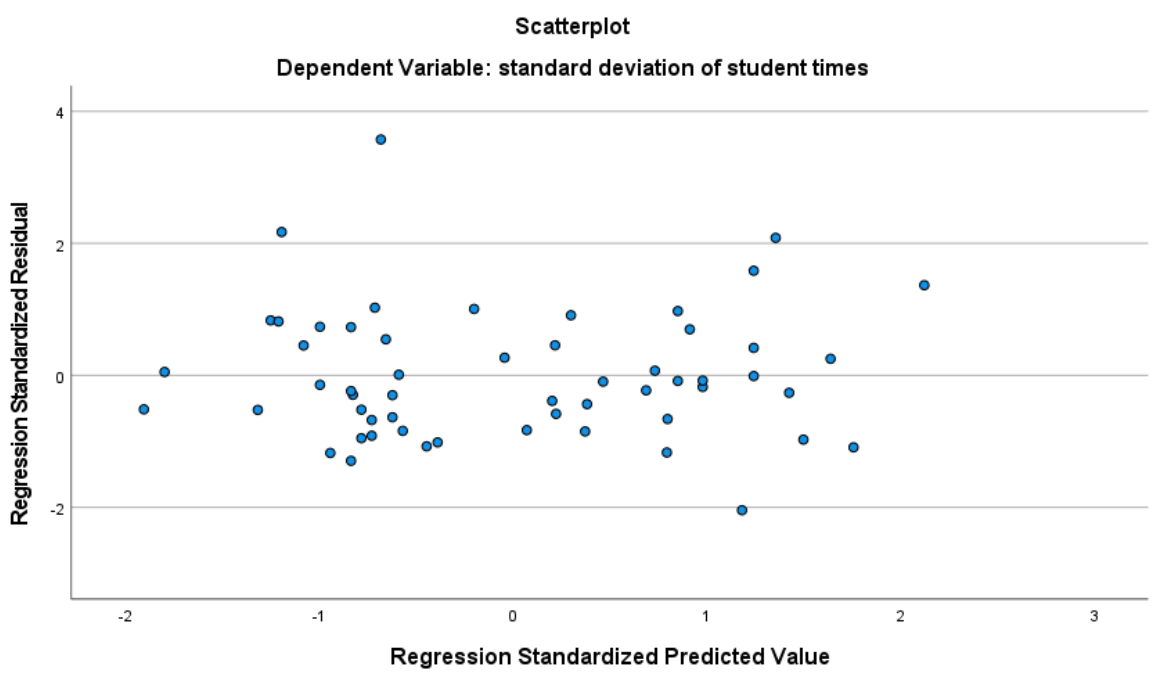



**Figure 6.12.** **Scatterplot of the residuals for the standard deviation of time to completion using the single LCI as predictor.**

In the first regression analysis using the three Lexical Complexity Subindices as predictors (with *mean* as the predicted variable), the prediction model was statistically detectable, $F(3, 52) = 15.074$, $p < .001$, and accounted for 43.4% of the variance of the predicted variable ($R^2 = .465$ and adjusted $R^2 = .434$). Higher syntactic and morphological complexity of a node predicted higher mean time to completion for a given node. Table 6.7 presents the raw and standardized regression coefficients of the predictors, along with their correlations with the mean time to completion, the squared semipartial correlations, and the structure coefficients.

**Table 6.7. Regression predicting mean time to completion using the three linguistic complexity subindices as predictors.**

| Model | b | SE-b | β | Pearson r | $sr^2$ | Structure Coefficient | VIF |
|---|---|---|---|---|---|---|---|
| **Constant** | 1153.627 | 211.897 | | | | | |
| **Syntactic complexity**** | 181.492 | 47.533 | .534 | .576 | .150 | .845 | 1.901 |
| **Morphological complexity*** | 33.878 | 12.564 | .345 | .549 | .074 | .805 | 1.590 |
| **Lexical complexity** | 31.005 | 11.890 | -.114 | -.101 | .070 | -.148 | 1.406 |

Note: Sample size is $N = 56$. $sr^2$ is the squared semi-partial correlation. *$p < .05$ ** $p < .001$

Further examination of the individual predictors revealed two of the three indicator complexity variables were statistically detectable predictors for the model ($t = 3.818$, $p < .001$ for syntactic complexity, $t = 2.696$, $p = .009$ for morphological complexity). In terms of weights in the model, syntactic complexity received a quite substantial one ($\beta = .534$), while morphological complexity received a weight of $\beta = .345$. Additionally, the unique variance explained by syntactic complexity indexed by the squared semipartial correlation was quite substantial (approximately 15%), almost double than the unique variance explained by morphological complexity (around 7%).

In the second regression analysis using the three Lexical Complexity Subindices as predictors (with standard deviation as the predicted variable), the prediction model was statistically detectable, $F(3, 52) = 12.036, p < .001$, and accounted for approximately 40% of the variance of the predicted variable ($R^2 = .410$ and *adjusted* $R^2 = .376$). Higher syntactic and morphological complexity of a node predicted higher standard deviation of time to completion for a given node. Table 6.8 presents the raw and standardized regression coefficients of the predictors, along with their correlations with the standard deviation of time to completion, the squared semipartial correlations and the structure coefficients.

**Table 6.8. Regression predicting standard deviation of time to completion, using the three linguistic complexity subindices as predictors.**

| Model | b | SE-b | β | Pearson r | $sr^2$ | Structure Coefficient | VIF |
|---|---|---|---|---|---|---|---|
| **Constant** | 757.828 | 157.622 | | | | | |
| **Syntactic complexity*** | 118.115 | 35.358 | .491 | .531 | .127 | .830 | 1.901 |
| **Morphological complexity*** | 23.426 | 9.346 | .337 | .518 | .071 | .809 | 1.590 |
| **Lexical complexity** | 21.501 | 8.845 | -.107 | -.082 | .067 | -.128 | 1.406 |

Note: Sample size is $N$ = 56. $sr^2$ is the squared semi-partial correlation. *$p < .05$

Further investigation of the individual predictors revealed that only two of them were statistically detectable predictors for the model ($t$ = 3.341, $p$ = .002 for syntactic complexity, $t$ = 2.507, $p$ = .015 for morphological) and both receive substantial weights in the regression model (for syntactic complexity $\beta$ = .491, for morphological complexity $\beta$ = .337). Finally, syntactic complexity explains uniquely approximately 13% of the variance of the dependent variable, while morphological complexity explains uniquely variance in the vicinity of 7%.

In the first regression analysis using the single LCI as a predictor (with mean as the predicted variable), the prediction model was statistically detectable, $F(1, 54)$ = 13.634*,* $p < .001$, and accounted for approximately 20% of the variance of the predicted variable ($R^2$ = .202 and *adjusted* $R^2$ = .187). Higher *LCI* value of a node predicted higher mean time to completion for a given node. Table 6.9 presents the raw and standardized regression coefficients of the predictor. The *LCI* of a node is a statistically detectable predictor of the mean time to completion for that node ($t$ = 3.692, $p < .001$) and it has a substantial weight in the model ($\beta$ = .449).

**Table 6.9. Regression predicting *mean time to completion*, using the single LCI as predictor.**

| Model | b | SE-b | β |
|---|---|---|---|
| **Constant** | 2092.981 | 115.757 | |
| **LCI**** | 29.607 | 8.018 | .449 |

Note: Sample size is $N$ = 56. $sr^2$ is the squared semi-partial correlation. ** $p < .001$

In the second regression analysis using the single LCI as a predictor (with standard deviation as the predicted variable), the prediction model was statistically detectable, $F(1, 54)$ = 11.172*,* $p$ = .002, and accounted for approximately 15% of the variance of the predicted variable ($R^2$ = .171 and *adjusted* $R^2$ = .156). Higher LCI value of a node predicted higher standard deviation of time to completion for a given node. Table 6.10 presents the raw and standardized regression coefficients of the predictor. The LCI of a node is a statistically detectable predictor of the standard deviation of time to completion for that node ($t$ = 3.343, $p$ = .002) and it has a substantial weight in the model ($\beta$ = .414).

**Table 6.10. Regression predicting standard deviation of time to completion as the dependent variable, using the single LCI as predictor.**

| Model | b | SE-b | Beta |
|---|---|---|---|
| **Constant** | 1393.534 | 83.506 | |
| **LCI*** | 19.334 | 5.784 | .414 |

Note: Sample size is $N$ = 56. $sr^2$ is the squared semi-partial correlation. * $p$ < .05

## 6.2. Results for Research Question 2

The second research question focuses on investigating changes in the learners' behavior, as they work on nodes of varying lexical, morphological and syntactic difficulty.

**RQ 2 – How do the *learning tactics* and *learning strategies* adopted by students of the Modern Greek online language course relate to the suggested estimates of *difficulty* for a particular node?**

### 6.2.1. How did the learning tactics adopted by the learners when studying a node relate to the lexical, morphological and syntactic complexity of the node?

In order to investigate which learning tactics are usually implemented in a node, depending on the different Linguistic Complexity Subindices for that node, three multiple regression analyses were performed, using the frequencies of the most frequent learning tactics extracted by the student logs (presented in Table 5.3) as the predictor variables, and each of the indices for the three indices of linguistic complexity as the dependent variable.

Before proceeding to the statistical analyses, the assumptions for making valid inferences from the regression analyses were tested. Inspection of the normal P-P plots of the regression standardized residuals was conducted to test the normality of residuals assumption. In Figures 6.13, 6.14 and 6.15 the residuals approximately follow the diagonal normality line, validating the assumption of the normality of residuals. For the assumption of multicollinearity, the VIF values for the predictors were checked and, since they were far below 10, no violation for this assumption was detected (the VIF values are presented in Tables 6.11, 6.12 and 6.13. However, inspection of the scatterplots of the residuals in Figures 6.16, 6.17 and 6.18 reveals a funnel-pattern in the distribution of the residuals, suggesting heteroscedastic distributions. Heteroscedasticity is known to distort the estimators of the standard errors of the regression coefficients, thereby invalidating the *t*-tests and *F*-tests of the regression (Astivia & Zumbo, 2019). To address this issue, the weighted least squares regression solution suggested by Astivia and Zumbo was adopted.

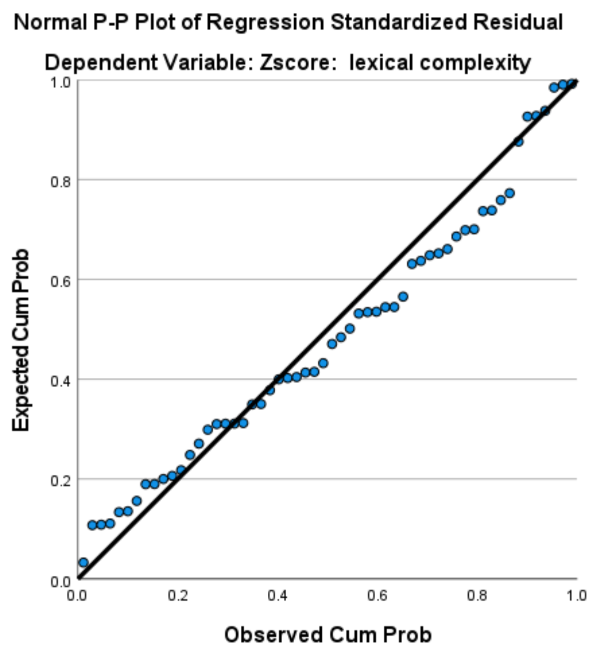



**Figure 6.13. Normal P – P plot of the regression standardized residuals for the standardized lexical complexity subindex.**

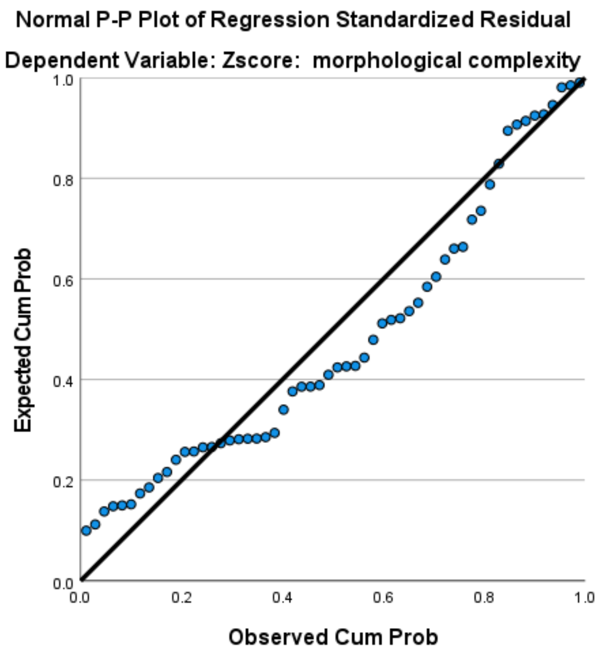



**Figure 6.14. Normal P – P plots of the regression standardized residuals for the standardized morphological complexity subindex.**

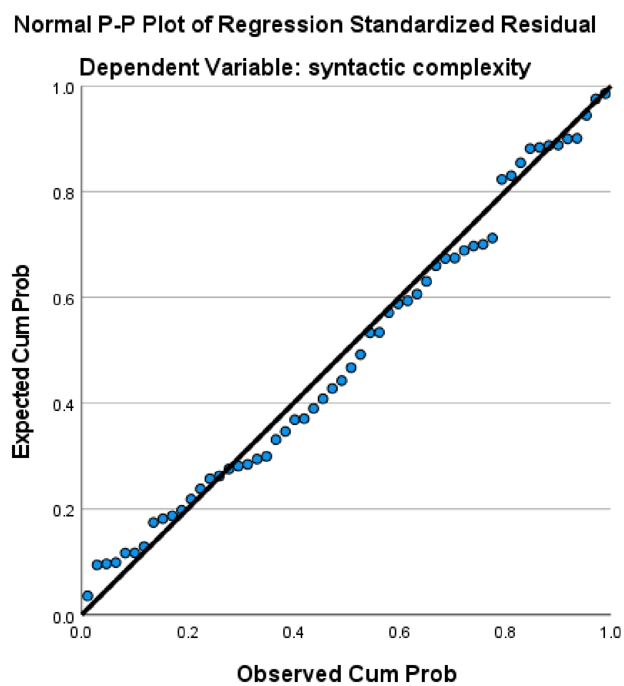



**Figure 6.15. Normal P – P plot of the regression standardized residuals for the standardized syntactic complexity subindex.**

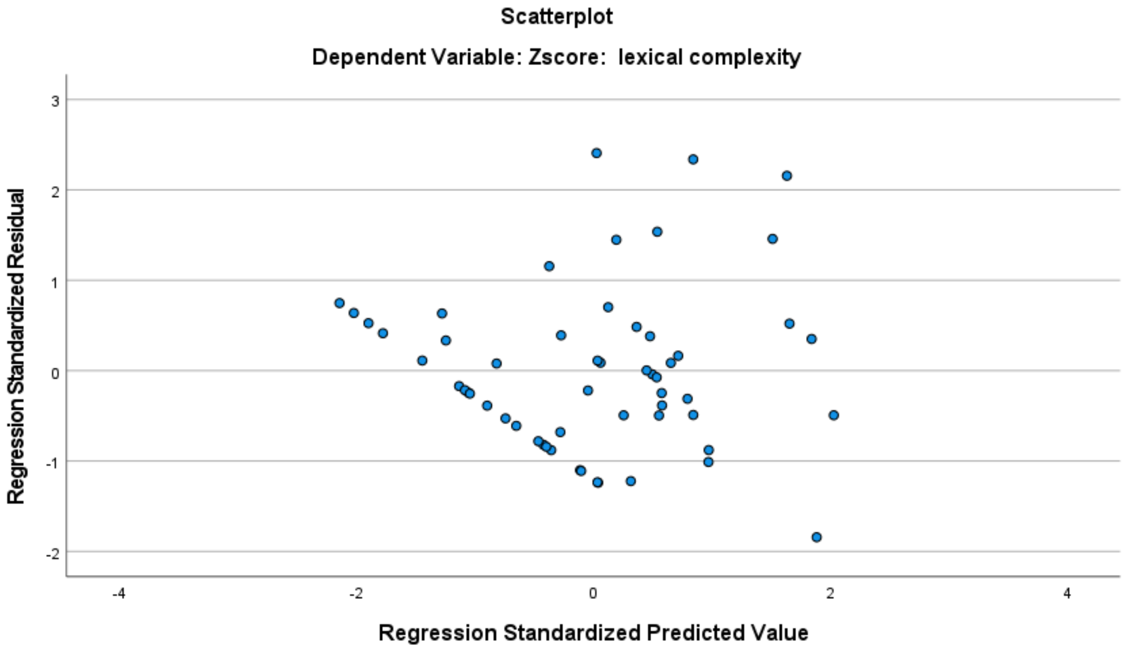



**Figure 6.16. Scatterplot of the residuals for the standardized lexical complexity subindex.**

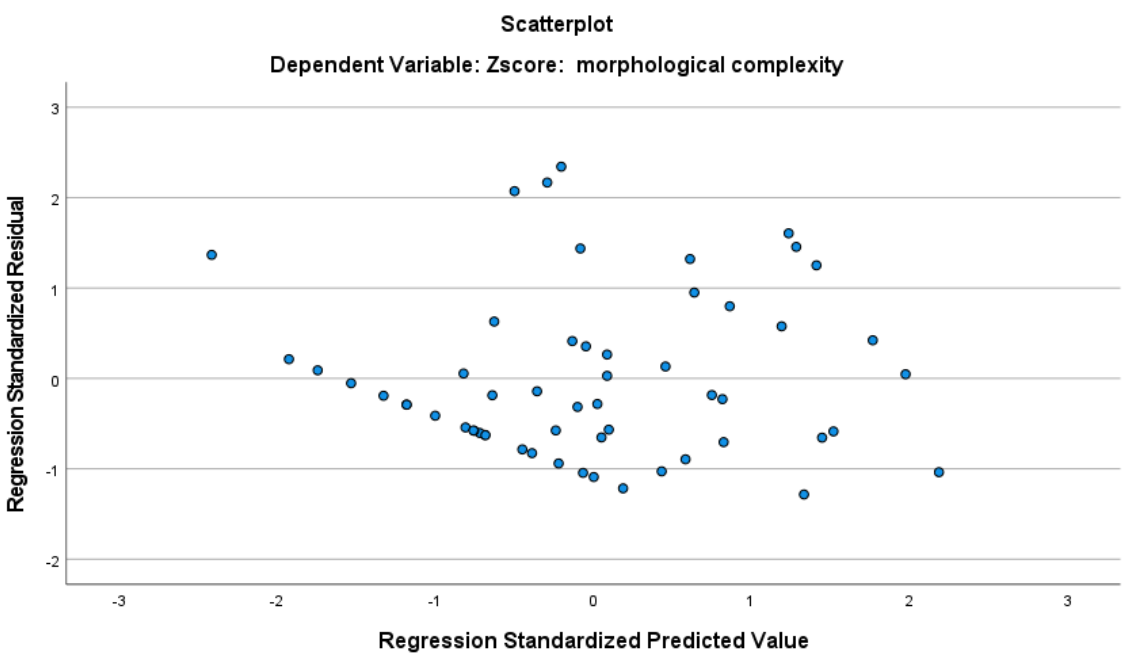



**Figure 6.17. Scatterplot of the residuals for the standardized morphological complexity subindex.**

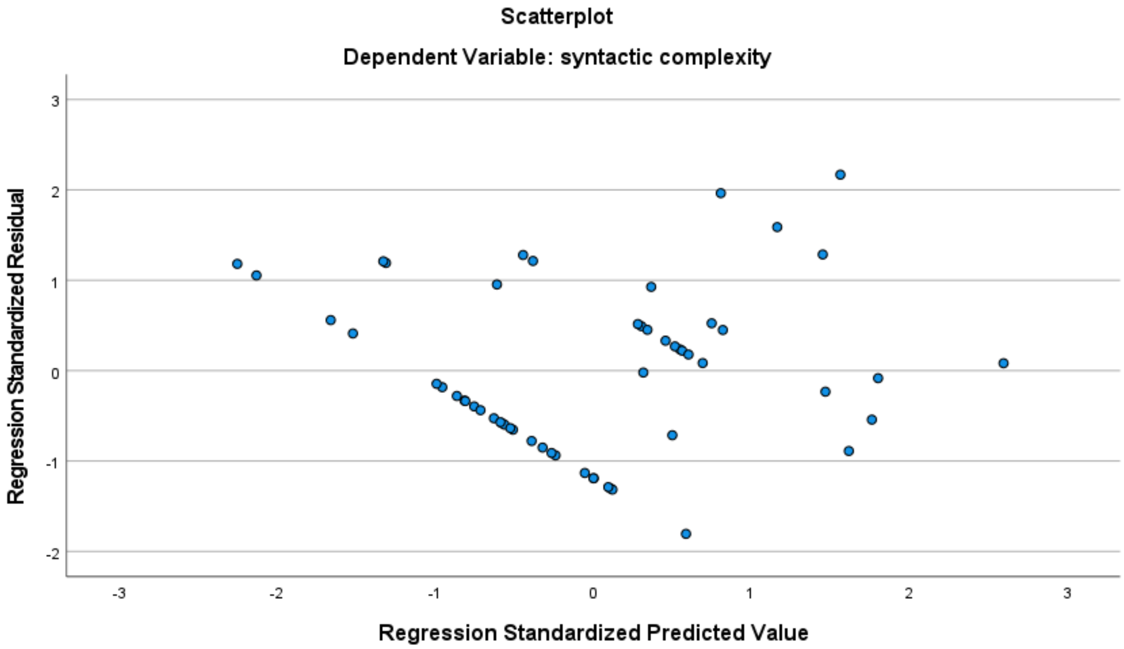



**Figure 6.18. Scatterplot of the residuals for the standardized syntactic complexity subindex.**

The first regression analysis was conducted to investigate how adopted learning tactics correspond to the standardized lexical complexity subindex of a node. The prediction model was statistically detectable, $F(7, 48) = 5.346$, $p < .001$, and accounted for approximately 35% of the variance of the predicted variable ($R^2 = .438$ and *adjusted* $R^2 = .356$). Table 6.11 presents the raw and standardized regression coefficients of the predictors, along with their correlations with the lexical complexity subindex, the squared semipartial correlations and the structure coefficients.

**Table 6.11. Regression predicting standardized lexical complexity subindex.**

| Model | b | SE-b | β | Pearson r | $sr^2$ | Structure Coefficient | VIF |
|---|---|---|---|---|---|---|---|
| **Constant** | -1.692 | .787 | | | | | |
| **Tactic A** | -.018 | .008 | **-.292*** | -.131 | .063 | -.198 | 1.364 |
| **Tactic B** | .023 | .034 | .123 | -.086 | .005 | .130 | 2.857 |
| **Tactic C** | -.045 | .036 | -.200 | -.100 | .018 | -.151 | 2.185 |
| **Tactic D** | .071 | .078 | .162 | .058 | .010 | .088 | 2.672 |
| **Tactic E** | -.020 | .036 | -.084 | -.065 | .003 | -.098 | 2.062 |
| **Tactic F** | .041 | .031 | .181 | .161 | .020 | .243 | 1.629 |
| **Tactic G** | .119 | .023 | **.587**** | .554 | .326 | .837 | 1.058 |

Note: Sample size is $N = 56$. $sr^2$ is the squared semi-partial correlation. $^*p < .05$ $^{**}p < .001$

Further inspection of the individual predictors revealed that only tactics A (accessing assessments only, $t$ = -2.311, $p$ = .025) and G (accessing study materials only, $t$ = 5.274, $p$ < .001) were statistically detectable predictors and they have been given substantial weights in the model ($\beta$ = -.292 and $\beta$ = .587 respectively). When students work on nodes with higher values of lexical complexity, they tend to have fewer learning episodes accessing only assessments and more learning episodes accessing only study materials. The adaptation of tactic G (only accessing the study materials) has the strongest predictive power, as it uniquely accounts for approximately 33% of the variance of the dependent variable), followed by the adaptation of tactic A (only accessing the assessments), which uniquely explains approximately 6% of the variance of the dependent variable.

In the second regression analysis (with the standardized morphological complexity subindex as the predicted variable), the prediction model was statistically detectable, $F(7, 48)$ = 3.431, $p$ = .005, and accounted for approximately one fourth of the variance of the predicted variable ($R^2$ = .333 and *adjusted* $R^2$ = .236). Table 6.12 presents the raw and standardized regression coefficients of the predictors, along with their correlations with the standardized morphological complexity subindex, the squared semipartial correlations and the structure coefficients.

**Table 6.12. Regression predicting standardized morphological complexity subindex.**

| Model | b | SE-b | β | Pearson r | $sr^2$ | Structure Coefficient | VIF |
|---|---|---|---|---|---|---|---|
| **Constant** | .021 | .857 | | | | | |
| **Tactic A** | .011 | .008 | .186 | .088 | .025 | .153 | 1.364 |
| **Tactic B** | -.028 | .037 | -.152 | .063 | .008 | .109 | 2.857 |
| **Tactic C** | -.004 | .039 | -.018 | .090 | .000 | .156 | 2.185 |
| **Tactic D** | -8.368E-5 | .085 | .000 | -.019 | .000 | -.034 | 2.672 |
| **Tactic E** | .095 | .039 | **.406*** | .286 | .080 | .496 | 2.062 |
| **Tactic F** | -.004 | .034 | -.019 | -.060 | .000 | -.104 | 1.629 |
| **Tactic G** | -.097 | .025 | **-.481**** | .438 | .218 | -.759 | 1.058 |

Note: Sample size is $N$ = 56. $sr^2$ is the squared semi-partial correlation. *$p$ < .05 ** $p$ < .001

The individual predictors were examined further, revealing that only tactics E (accessing study materials and then assessments, $t$ = 2.401, $p$ = .020) and G (accessing only study materials, $t$ = -3.966, $p$ < .001) are statistically detectable predictors of the morphological complexity subindex of a node, and they have substantial weights in the model ($\beta$ = .406 and $\beta$ = -.481 respectively). When students study nodes with high morphological complexity, they tend to have more learning episodes accessing study materials followed by assessments and fewer learning episodes accessing only study materials. The unique variance of the morphological complexity subindex explained by these two predictor variables is approximately 20% for tactic G and 8% for tactic E.

The third regression analysis was conducted to investigate how adopted learning tactics correspond to the standardized syntactic complexity subindex of a node. The prediction model was statistically detectable, $F(7, 48) = 8.903$, $p < .001$, and accounted for approximately half the variance of the predicted variable ($R^2$ = .565 and *adjusted* $R^2$ = .501). Table 6.13 presents the raw and standardized regression coefficients of the predictors, along with their correlations with the syntactic complexity subindex, the squared semipartial correlations and the structure coefficients.

**Table 6.13. Regression predicting standardized syntactic complexity subindex.**

| Model | b | SE-b | β | Pearson r | $sr^2$ | Structure Coefficient | VIF |
|---|---|---|---|---|---|---|---|
| **Constant** | 2.898 | .693 | | | | | |
| **Tactic A** | .016 | .007 | **.256*** | .064 | .048 | .085 | 1.364 |
| **Tactic B** | .017 | .030 | .092 | -.022 | .003 | -.029 | 2.857 |
| **Tactic C** | .040 | .032 | .176 | .080 | .014 | .106 | 2.185 |
| **Tactic D** | -.077 | .068 | -.176 | -.158 | .012 | -.210 | 2.672 |
| **Tactic E** | .015 | .032 | .062 | .185 | .002 | .246 | 2.062 |
| **Tactic F** | -.104 | .027 | **-.462**** | -.378 | .131 | -.503 | 1.629 |
| **Tactic G** | -.121 | .020 | **-.598**** | -.539 | .339 | -.718 | 1.058 |

Note: Sample size is $N$ = 56. $sr^2$ is the squared semi-partial correlation. *$p < .05$ ** $p < .001$

Further inspection of the individual predictors revealed that tactics A (accessing assessments only, $t$ = 2.299, $p$ = .026), F (accessing study materials, followed by learning activities and finally assessments, $t$ = -3.800, $p$ < .001) and G (accessing study materials only, $t$ = -6.110, $p$ < .001) were statistically detectable predictors for the model and they have been given substantial weights in the model ($\beta$ = .256, $\beta$ = -.462 and $\beta$ = -.598 respectively). When students work on nodes with higher syntactic complexity, they tend to have more learning episodes accessing only the assessments, fewer learning episodes accessing only the study materials, and fewer learning episodes accessing all three-sections in the order: study materials – learning activities – assessments. The adoption of tactic G has the strongest predictive power, as it uniquely accounts for approximately 33% of the variance of the predicted variable), followed by the adoption of tactic F, which uniquely explains approximately 13% of the variance of the predicted variable and finally the adoption of tactic A, which uniquely accounts for approximately 5% of the variance of the predicted variable.

Overall, the learners' behavior related to tactic G (spending the whole learning episode working solely on study materials) was a consistently strong predictor of the node differences in all three Linguistic Complexity Subindices. Increased adoption of the tactic was observed in nodes with high values of the lexical complexity subindex, while decreased usage of the same tactic was observed in nodes with high values of morphological and syntactic complexity. Additionally, the weight calculated by the regression model with syntactic complexity subindex as the predicted variable was substantially higher than the weight in the regression model with morphological complexity subindex as the predicted variable.

### 6.2.2. How did the lexical, morphological, and syntactic complexity of nodes relate to learners' adoption of deductive or inductive learning strategies?

Investigation of the relationship of deductive or inductive language learning strategies adopted by learners when studying a node with the three Linguistic Complexity Subindices of that node was performed by conducting two sets of regression analysis. The first set included three regression analyses, one for each of the Linguistic Complexity Subindices of a node, with the number of deductive learning tactics and the number of *inductive* learning tactics (as they were classified in Table 5.3) as the predictor variables. The second set included three regression analyses, one for each of the Linguistic Complexity Subindices of a node, with mean time spent on study materials and mean time spent on learning activities as predictor variables.

First, the assumptions for making valid inferences from the regression analyses were tested. The normality of the residuals of the regression was determined by inspecting the normal P – P plots of the regression standardized residuals. As it can be seen in Figures 6.19 through 6.24, the residuals approximate the diagonal normality line. The assumption of homoscedasticity was checked by examining the scatterplots of the residuals. Figures 6.25 and 6.26 present scatterplots with a funnel-like effect in the distribution of the residuals, suggesting heteroscedastic data for the regression analyses with lexical complexity subindex as the predicted variable. Therefore, we conducted a weighted least squares regression for these analyses, as suggested by Astivia and Zumbo (2019). Figures 6.27 through 6.30 show that the residuals are randomly distributed, suggesting homoscedastic distributions. Finally, for multicollinearity, the VIF values of the three indicator variables are well bellow 10, indicating that there is no multicollinearity in both regressions. The VFI values are shown in Tables 6.14 through 6.19.

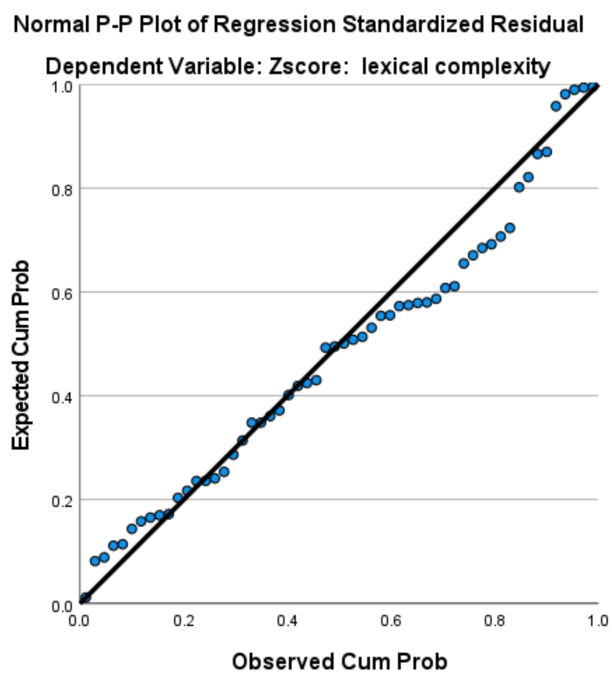



**Figure 6.19. Normal P – P plot of the regression standardized residuals for the standardized lexical complexity subindex, with the number of deductive and the number of inductive learning tactics as predictors.**

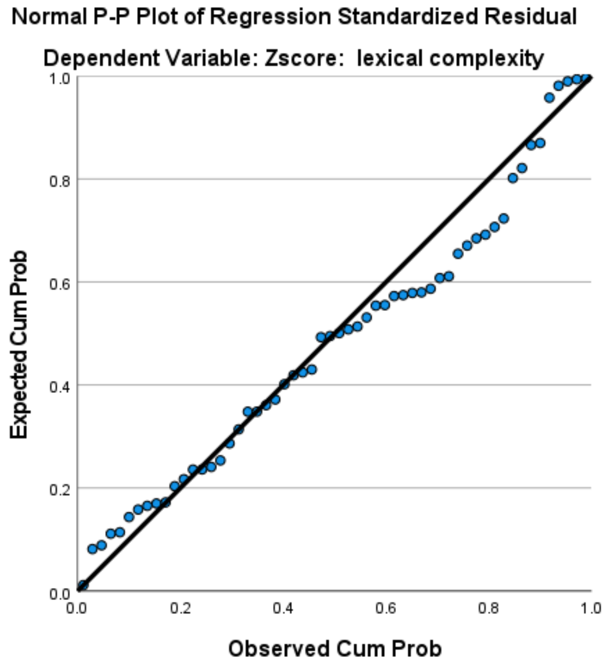



**Figure 6.20. Normal P – P plot of the regression standardized residuals for the standardized lexical complexity subindex, with the mean time spent on study materials and mean time spent on learning activities as predictors.**

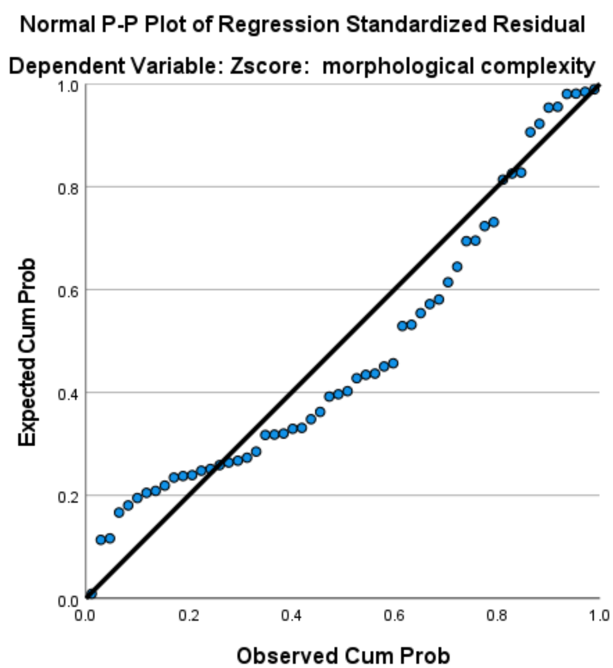



**Figure 6.21. Normal P – P plot of the regression standardized residuals for the standardized morphological complexity subindex, with the number of deductive and the number of inductive learning tactics as predictors.**

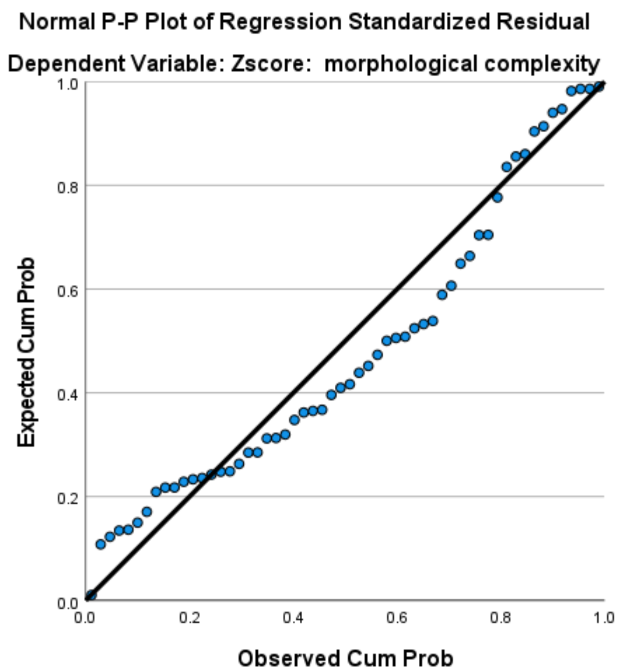



**Figure 6.22. Normal P – P plot of the regression standardized residuals for the standardized morphological complexity subindex, with the mean time spent on study materials and mean time spent on learning activities as predictors.**

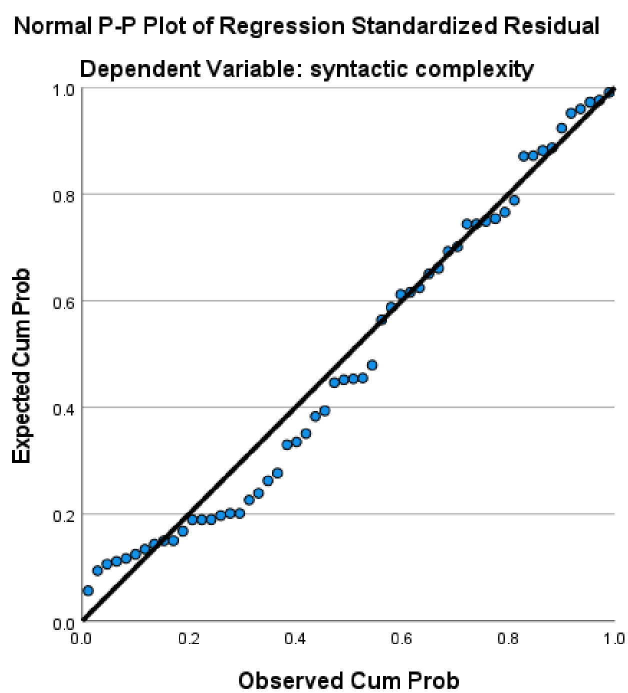



**Figure 6.23. Normal P – P plot of the regression standardized residuals for the standardized syntactic complexity subindex, with the number of deductive and the number of inductive learning tactics as predictors.**

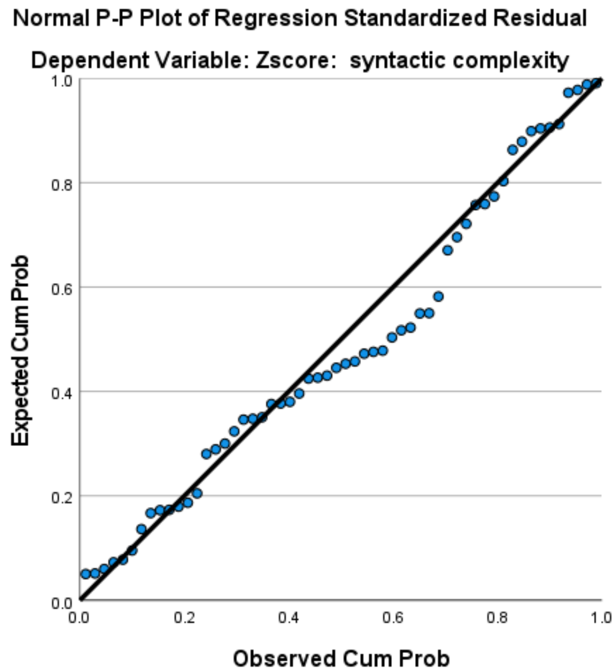



**Figure 6.24.** **Normal P – P plot of the regression standardized residuals for the standardized syntactic complexity subindex, with the mean time spent on study materials and mean time spent on learning activities as predictors.**

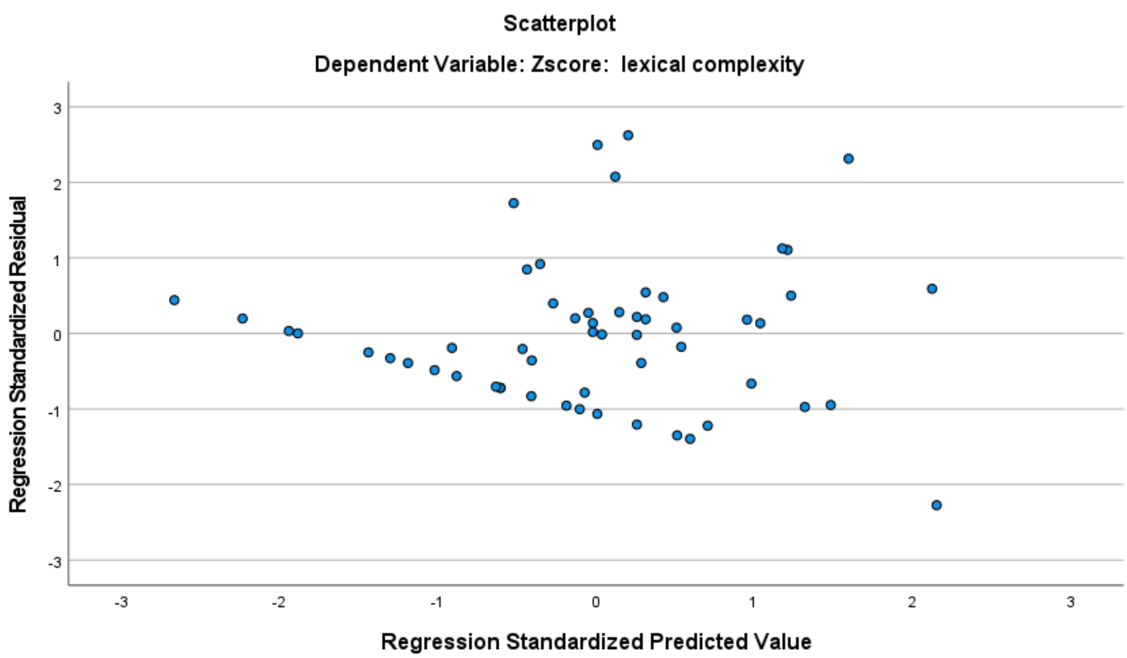



**Figure 6.25.** **Scatterplot of the residuals for the standardized lexical complexity subindex, with the number of deductive and the number of inductive learning tactics as predictors.**

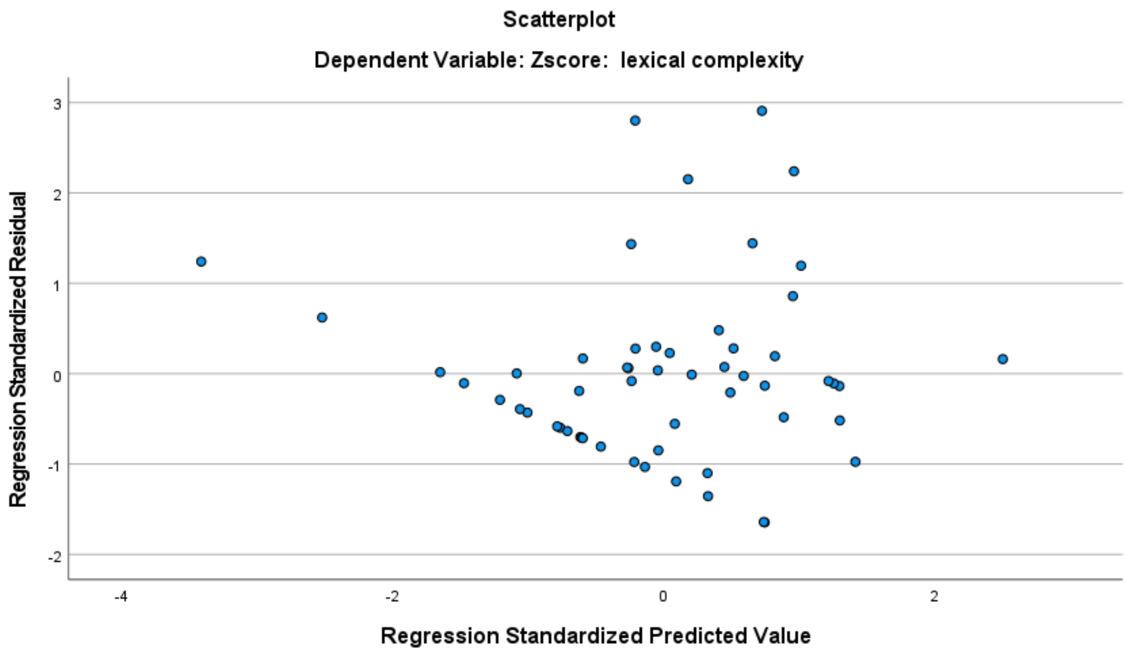



**Figure 6.26. Scatterplot of the residuals for the standardized lexical complexity subindex, with the mean time spent on study materials and mean time spent on learning activities as predictors.**

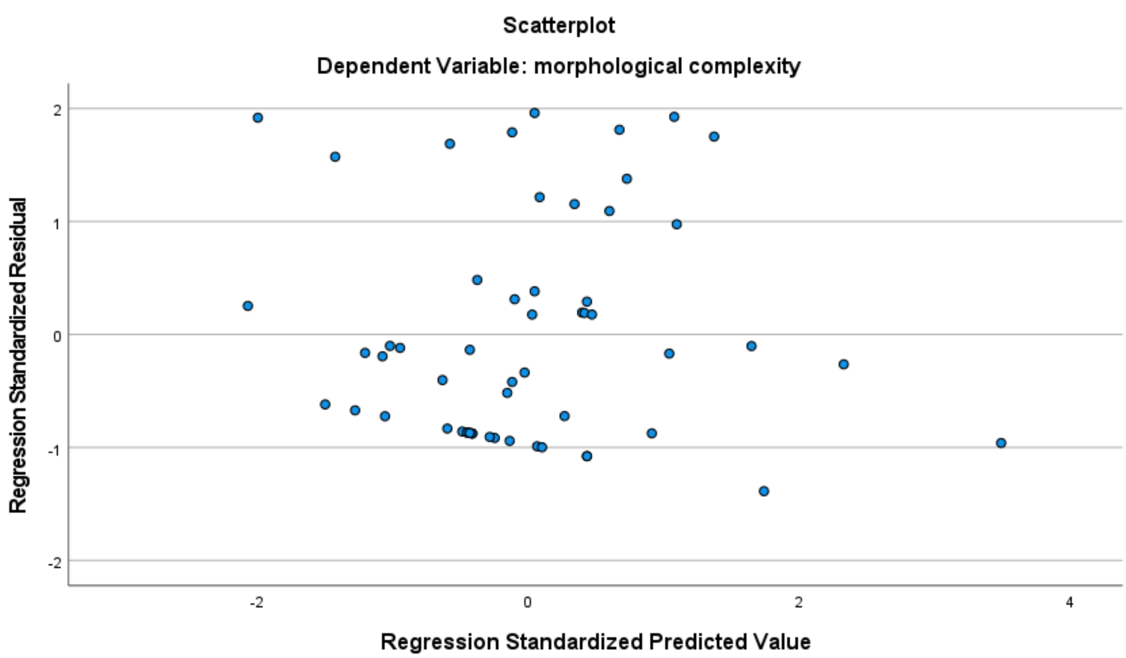



**Figure 6.27. Scatterplot of the residuals for the standardized morphological complexity subindex, with the number of deductive and the number of inductive learning tactics as predictors.**

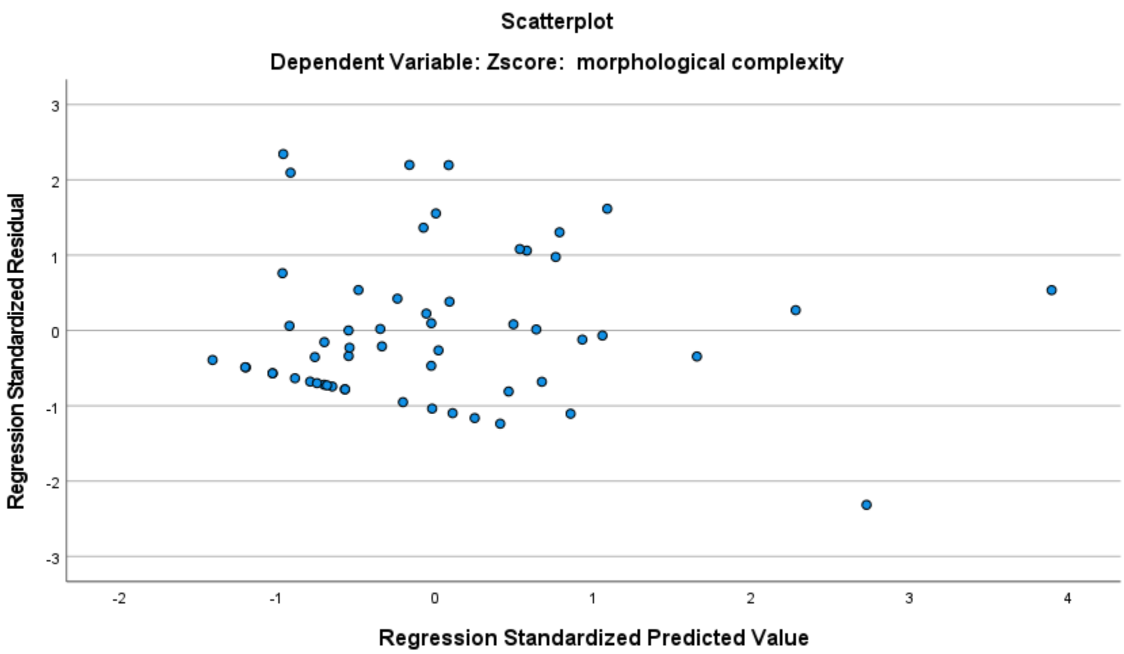



**Figure 6.28. Scatterplot of the residuals for the standardized morphological complexity subindex, with the mean time spent on study materials and mean time spent on learning activities as predictors.**

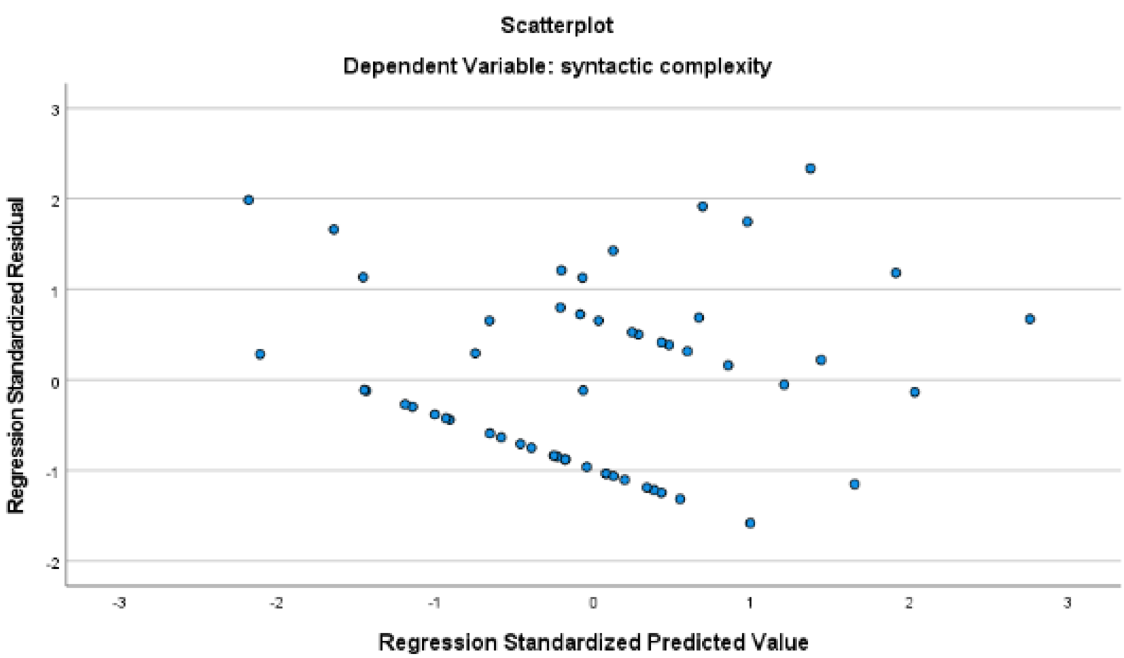



**Figure 6.29. Scatterplot of the residuals for the standardized syntactic complexity subindex, with the number of deductive and the number of inductive learning tactics as predictors.**

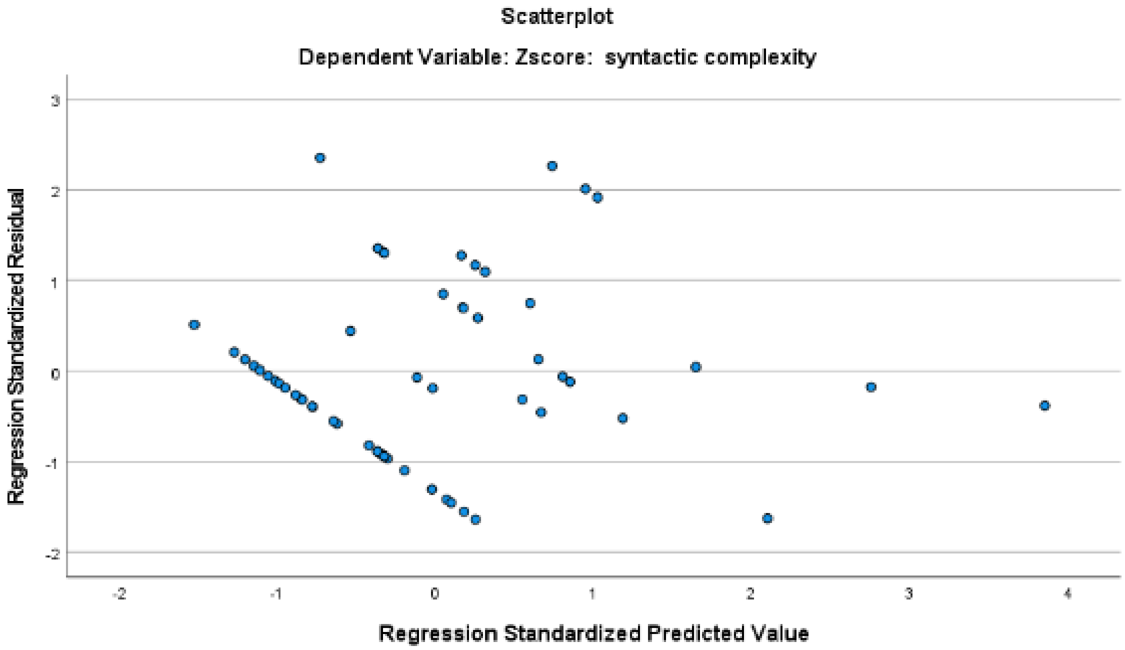



**Figure 6.30. Scatterplot of the residuals for the standardized syntactic complexity subindex, with the mean time spent on study materials and mean time spent on learning activities as predictors.**

In the first set of regression analyses, the predictor variables were the number of *inductive* and *deductive* learning tactics adopted by the learners when studying a node. Three regression analyses were conducted, one for each of the Linguistic Complexity Subindices of a node. The prediction model with the lexical complexity subindex as the predicted variable was statistically detectable, $F(2, 53) = 8.730$, $p < .001$, and accounted for approximately 20% of the variance of the predicted variable ($R^2 = .248$ and *adjusted* $R^2 = .219$). Higher linguistic complexity of a node was related to higher number of *deductive* learning tactics and lower number of *inductive* learning tactics adopted by students working on that particular node. Table 6.14 presents the raw and standardized regression coefficients of the predictors, along with their correlations with the lexical complexity subindex, the squared semipartial correlations and the structure coefficients. Both the number of deductive learning tactics ($t = 4.107$, $p < .001$) and number of inductive learning tactics ($t = -2.342$, $p = .023$) are statistically detectable predictors of the linguistic complexity subindex of a node and have substantial weights in the model ($\beta = .533$ and $\beta = -.304$ respectively).

**Table 6.14. Weighted least squares regression predicting standardized lexical complexity subindex, using number of deductive and number of inductive learning tactics as predictors.**

| Model | b | SE-b | β | Pearson r | $sr^2$ | Structure Coefficient | VIF |
|---|---|---|---|---|---|---|---|
| **Constant** | -2.505 | .654 | | | | | |
| **Number of inductive tactics** | -.014 | .006 | **-.304*** | -.092 | .078 | -.185 | 1.189 |
| **Number of deductive tactics**** | .069 | .017 | **.533**** | .412 | .239 | .827 | 1.189 |

Note: Sample size is $N$ = 56. $sr^2$ is the squared semi-partial correlation. *$p < .05$ ** $p < .001$

The prediction model with the morphological complexity subindex as the predicted variable was not statistically detectable, $F(2, 53) = 1.547$, $p = .222$. Also, the prediction model accounted for a very small portion of the variance of the predicted variable, approximately 2% ($R^2 = .055$ and *adjusted* $R^2 = .019$). Therefore, there was no statistically detectable difference in how often the students adopted deductive or inductive learning tactics when working on nodes of different morphological complexity. The raw and standardized regression coefficients of the predictors, along with their correlations with the morphological complexity subindex, the squared semipartial correlations and the structure coefficients are presented in Table 6.15.

**Table 6.15. Regression predicting standardized morphological complexity subindex, using number of deductive and number of inductive learning tactics as predictors.**

| Model | b | SE-b | β | Pearson r | $sr^2$ | Structure Coefficient | VIF |
|---|---|---|---|---|---|---|---|
| **Constant** | 19.190 | 7.038 | | | | | |
| **Number of inductive tactics** | .083 | .063 | .190 | .098 | .031 | .417 | 1.189 |
| **Number of deductive tactics** | -.290 | .181 | -.233 | -.157 | .046 | -.668 | 1.189 |

Note: Sample size is $N$ = 56. $sr^2$ is the squared semi-partial correlation.

The third prediction model, with the syntactic complexity subindex as the predicted variable was statistically detectable, $F(2, 53) = 9.977$, $p < .001$, accounted for almost one quarter of the variance of the predicted variable ($R^2 = .274$ and *adjusted* $R^2 = .246$). Higher values of syntactic complexity of a node were related to higher number of inductive learning tactics and lower number of deductive learning tactics adopted by students working on that particular node. Table 6.16 presents the raw and standardized regression coefficients of the predictors, along with their correlations with the syntactic complexity subindex*,* the squared semipartial correlations and the structure coefficients. Both the number of deductive learning tactics (*t* = -4.452, *p* < .001) and number of inductive learning tactics (*t* = 2.106, *p* = .040) are statistically detectable predictors of the syntactic complexity subindex of a node and have substantial weights in the model (*β* = -.568 and *β*= .269 respectively).

**Table 6.16. Regression predicting standardized syntactic complexity subindex, using number of deductive and number of inductive learning tactics as predictors.**

| Model | b | SE-b | β | Pearson r | $sr^2$ | Structure Coefficient | VIF |
|---|---|---|---|---|---|---|---|
| **Constant** | 9.932 | 1.783 | | | | | |
| **Number of inductive tactics** | .034 | .016 | **.269*** | .042 | .061 | .080 | 1.189 |
| **Number of deductive tactics** | -.205 | .046 | **-.568**** | -.461 | .271 | -.996 | 1.189 |

Note: Sample size is *N* = 56. $sr^2$ is the squared semi-partial correlation. **p* < .05 ** *p* < .001

In the second set of regression analyses, the predictor variables were the mean time spent on study materials and the mean time spent on learning activities when working on a node. Again, three regression analyses were conducted, one for each of the Linguistic Complexity Subindices of a node, a weighted least squares regression analysis for the lexical complexity subindex as the predicted variable (since there was a violation of the homoscedasticity assumption) and standard regression analyses for the remaining two subindices.

The regression analysis with the lexical complexity subindex as the predicted variable provided a statistically detectable prediction model $F(2, 53) = 9.165$*,* $p < .001$*,* accounting for approximately 30% of the variance of the predicted variable ($R^2 = .333$ and *adjusted* $R^2 = .308$). Higher mean times on study materials and lower mean times on learning activities spent while students studied a node, predicted higher lexical complexity of that node. The raw and standardized regression coefficients of the predictors, along with their correlations with the lexical complexity subindex*,* the squared semipartial correlations and the structure coefficients are presented in Table 6.17***.*** Both predictors were statistically detectable (with *t* = 3.616, *p* < .001 for mean time spent on study materials, and *t* = -3.784, *p* < .001 for mean time spent on learning activities), while their weights in the prediction model are quite high ($\beta$ = .406 and $\beta$ = -.425 respectively).

**Table 6.17. Weighted least squares regression predicting standardized lexical complexity subindex, using mean time spent on study materials and mean time spent on learning activities when working on a node as predictors.**

| Model | b | SE-b | β | Pearson r | $sr^2$ | Structure Coefficient | VIF |
|---|---|---|---|---|---|---|---|
| **Constant** | .207 | .437 | | | | | |
| **Mean time on study materials** | .003 | .001 | **.406*** | .391 | .165 | .678 | 1.001 |
| **Mean time on learning activities** | -.002 | .001 | **-.425*** | -.411 | .180 | -.712 | 1.001 |

Note: Sample size is $N$ = 56. $sr^2$ is the squared semi-partial correlation. *$p$ < .001

The second prediction model, with the morphological complexity subindex as the predicted variable was also statistically detectable, $F(2, 53) = 5.945$, $p = .005$. Additionally, the prediction model accounted for approximately 15% of the variance of the predicted variable ($R^2$ = .183 and *adjusted* $R^2$ = .152). Lower mean times on study materials and higher mean times on learning activities spent while students studied a node, predicted higher morphological complexity of that node. The raw and standardized regression coefficients of the predictors, along with their correlations with the morphological complexity subindex, the squared semipartial correlations and the structure coefficients are presented in Table 6.18. Only one of the predictor variables was statistically detectable, the mean time spent on learning activities when working on a node ($t$ = 3.408, $p$ < .001), and it has been given a high weight value in the regression equation ($\beta$ = .423).

**Table 6.18. Regression predicting standardized morphological complexity index, using mean time spent on study materials and mean time spent on learning activities when working on a node as predictors.**

| Model | b | SE-b | β | Pearson r | $sr^2$ | Structure Coefficient | VIF |
|---|---|---|---|---|---|---|---|
| **Constant** | -1.023 | .484 | | | | | |
| **Mean time on study materials** | -.001 | .001 | -.080 | -.065 | .006 | -.152 | 1.001 |
| **Mean time on learning activities** | .002 | .001 | **.423*** | .421 | .179 | .988 | 1.001 |

Note: Sample size is $N$ = 56. $sr^2$ is the squared semi-partial correlation. *$p$ < .001

Finally, the third regression analysis, with the syntactic complexity subindex as the predicted variable provided a statistically detectable prediction model, $F(2, 53) = 40.310$, $p < .001$, accounting for approximately 60% of the variance of the predicted variable ($R^2 = .603$ and *adjusted* $R^2 = .588$). Lower mean times on study materials and higher mean times on learning activities spent while students studied a node, predicted higher syntactic complexity of that node. Table 6.19 presents the raw and standardized regression coefficients of the predictors, along with their correlations with the syntactic complexity subindex, the squared semipartial correlations and the structure coefficients. Both the predictors, mean time spent on study materials (with t = -3.778, p < .001) and mean time spent on learning activities (t = 8.269, p < .001) when working on a node, were statistically detectable and they were given substantial weights in the regression equation ($\beta = -.327$ and $\beta = .716$ respectively).

**Table 6.19. Regression predicting standardized syntactic complexity subindex, using mean time spent on study materials and mean time spent on learning activities when working on a node as predictors.**

| Model | b | SE-b | β | Pearson r | $sr^2$ | Structure Coefficient | VIF |
|---|---|---|---|---|---|---|---|
| **Constant** | -1.246 | .337 | | | | | |
| **Mean time on study materials** | -.003 | .001 | **-.327*** | -.303 | .107 | -.390 | 1.001 |
| **Mean time on learning activities** | .003 | .000 | **.716*** | .705 | .511 | .907 | 1.001 |

Note: Sample is $N = 56$. $sr^2$ is the squared semi-partial correlation. *$p < .001$

# Chapter 7.

# Discussion

## 7.1. Study Findings

This thesis evaluated existing and proposed estimates of difficulty for the content of a beginning Modern Greek language course. These estimates will be used to expand the adaptability features for the student-faced dashboard of the digital language learning environment in future iterations of the course. The goal is to support students in their time management and learning pathway decisions while providing useful hints and feedback so that they may overcome the challenges in the course. A secondary goal of the research, which should be regarded as a case study, was to examine how certain design decisions for the course platform affected students' behavior. The following sections will discuss the study findings and examine the implications of these findings in the field of Computer Assisted Language Learning research, using an organizational schema based on the thesis research questions.

### 7.1.1. Evaluation of existing difficulty estimates

In Chapter 2 I discussed how the notion of difficulty is treated in the second language acquisition literature. Housen and Simoens (2016) note that discussion on this topic has been challenged by definitional issues, conceptual confusion, and misunderstandings on the relation of difficulty with other related constructs, such as linguistic complexity. My research has treated difficulty as a concept corresponding to the structural complexity of a node in the Modern Greek language learning course. In that respect, the various difficulty estimates are examined as estimates of the structural complexity of a node.

As was mentioned, the student dashboard for the Modern Greek language learning platform already includes three difficulty estimates for the language content of each of the nodes in a particular module. These are designed to help learners in their decision making, particularly on the order of accessing the nodes of the course.

### *Declarative or procedural node type*

The first estimate of node difficulty is the type of the node, which can be declarative or procedural. The results of the study showed that the type of a node strongly predicts both the mean and the standard deviation of the time of completion for that node. Declarative nodes are associated with lower mean time of completion for the learners whereas procedural nodes are associated with higher completion times. This fact accords with the language learning model of Nikolov and Djigunovic (2006). The researchers suggested that two different systems are at work in language acquisition: a declarative system that relates to the different lexical units, phrases, and expressions; and a procedural system, related to the grammatical rules of the language and the formation of grammatically sound utterances. In a procedural node, both systems are operating, theoretically resulting in a greater cognitive load which translates to higher completion times for the node. Therefore, the type of node provides a strong indication of the cognitive complexity for that node, which is operationalized as the time of completion for the node.

With regard to the usefulness of the index, though, the type of node may not provide enough information to help students decide which nodes to study first. Firstly, the type of node is a binary categorization of nodes, therefore there is no further distinction among nodes of the same type. Additionally, declarative nodes appear earlier in the module tree, while procedural nodes occupy spots in subsequent levels of that tree. Hence, the only meaningful decision a learner could make is whether to start working on a procedural node the moment it becomes available, or to wait until all the declarative nodes of the module are completed before continuing to the procedural nodes for that module. Finally, the range of the times of completion of nodes of a particular type is high, as it can be deduced by the high standard deviation. Without information about variance of time to complete a node, learners might infer a declarative node can be completed relatively faster than a procedural node. Without additional information about times of completion, but they are unaware how different declarative nodes compare to each other in terms of completion times (and, consequently, the mental and time resources needed to complete them).

### *Number of prerequisites*

The second difficulty estimate incorporated in the dashboard of the Modern Greek language learning platform is the number of prerequisites for a particular node in the hierarchy, meaning the number of nodes in the module tree that need to be completed before that node becomes available to the learner. The results of the study showed that there is also a high positive correlation between the number of prerequisites and the mean time of completion per node. This is also an expected effect, as a higher number of prerequisites in a hierarchy result in higher memory load for that sequence (Nesbit & Hunka 1987, McEneaney 2016) and, consequently, in higher completion times for a particular node. Therefore, the number of prerequisites for a particular node helps the learners to make an informed decision about the order in which they will complete the nodes of a module, appropriately managing the time they have in their disposal.

However, there are several issues when considering the number of prerequisites of a node as an index of structural complexity. First, the number of prerequisites is not a complete metric in terms of how the reported number is operationalized in the design. Only the direct prerequisite nodes in the competence tree are considered, hence the metric does not include prerequisite nodes at greater depth within or across modules. Therefore, there are prerequisites hidden either in the lower levels of the hierarchy or in previously completed competence trees from past modules, leaving learners with a partial picture of the prior knowledge required for completing the instructional goal associated with a particular node.

Additionally, even though the high number of prerequisites provides an estimate of the structural complexity of a node, it does not reflect any of the inherent complexity of a particular node. A higher number of prerequisites may be indicative of more complex node, as each of the prerequisites refers to a different element or construct involved in the instructional goal of the node. However, this is not always the case. For example, node 4 of module 4 (describing daily routines) has only two prerequisites. Nevertheless, the internal complexity of that node is rather high, as the learner is required to form the present tense for various verbs describing daily routines. The suffixes required for the formation of the present tense (all three persons for each number, singular and plural) are different, depending on which conjugation each verb belongs to (first or second conjugation). Furthermore, second conjugation verbs are further classified into type 1 and type 2 verbs, which also use different suffixes to form the present tense. Thus, there are several different parameters a learner needs to consider when forming the inflectional paradigm for a particular person and number in the present tense in addition to other syntactic assumptions that need to be fulfilled for a grammatical sentence (article – noun agreement and subject – verb agreement). This inherent complexity of the node makes it very difficult to complete, a fact that is reflected in the mean time of completion for that particular node (4982 sec), which is the highest among all 56 nodes in the distribution.

Finally, this metric doesn't provide any other information about the prerequisites for a particular node. All the prerequisites are considered equal, when in fact they may differ in complexity, and this individual structural complexity of each of the prerequisites contributes to the overall structural complexity of the node under consideration. Additionally, Nesbit and Hunka (1987) emphasize that the order in which the prerequisites are fulfilled can also be important when calculating the memory load of a learning sequence, as it affects the ability of a student to recall prerequisite objectives. However, no information about the order of completing prerequisite nodes or the date of completion of the prerequisite nodes is provided to the students. All this missing information likely makes this metric quite weak in terms of providing sufficient scaffolding for learners' decision making.

### *Levels of linguistic description*

The final difficulty estimate included in the initial design of the digital learning environment for the Modern Greek language course is the number of levels of linguistic description involved in the instructional goal of a node. This metric only refers to the number of levels involved in a given node, without further information on which these levels are (morphology, syntax, semantics, or pragmatics). The results of the analysis show a strong positive correlation with the mean time of completion, an anticipated effect, since a higher number of levels of linguistic description involved in the instructional goal of a node relate to different elements or attributes of the linguistic utterances required by the learners to successfully complete the node. Hence, the number of levels of linguistic description is an adequate indicator of the time of completion of a node, which the learners should consider when making time management related decisions.

In terms of usability, the greatest issue with this metric is that it provides no information on which of the levels of linguistic description are involved in the specific instructional goal. Hence, even though the number of levels involved is an indication of the complexity of a particular node, there isn't enough additional information to evaluate the complexity of two nodes with the same value of that metric. For example, node 2 of module 3 and node 6 of module 2 have both instructional goals involving two levels of linguistic description. However, the former node involves morphology and semantics, whereas the latter morphology and syntax.

Another issue with the number of levels of linguistic description as an estimate of the difficulty of a node is the lack of information provided to the student in terms of the linguistic components / phenomena involved in a particular node. Even though the learner is aware that the instructional goal involves multiple levels of analysis, there is no information on how many different components / phenomena in a particular level of analysis are involved. For example, a node in the morphology level might focus on multiple features, for example inflection for both number and case of a particular noun. This level of detail for morphological complexity is not reflected in this metric.

### *Combining type of node, number of prerequisites, and levels of linguistic description*

In terms of the evaluation of a prediction model for mean time of completion for a node and the standard deviation of time of completion for a node, using these three indexes as predictors, the regression analysis showed that this prediction model is statistically detectable. Of the three variables used as predictors for the model, the number of levels of linguistic description is given more weight. This is somewhat anticipated, since the number of levels of linguistic description is the more indicative of the complexity of a node among the three already incorporated features in the digital learning environment, even though, as it was discussed previously, there are still some issues when using this metric to evaluate the complexity of a node.

However, the greatest issue when using these three variables to infer information about the nodes they describe is that they do not identify clearly the differences between the nodes of a module (and of the Modern Greek course in general). As an example, half the nodes in the course are declarative and the other half procedural. In terms of comparing nodes of the same type, though, the learner doesn't receive more information than that. There is no further indication of which of two declarative nodes is more difficult. Similarly, node prerequisites range from 0 to 4, but there is no way for the learner to further differentiate nodes of the same number of prerequisites. For the levels of linguistic description, again the different categories of nodes range from 0 to 4 different levels, which again doesn't allow for further distinction between nodes with the same number of levels of linguistic description. This is a limitation when taking into consideration that, at a particular time, when a learner needs to decide what the next node to be accessed will be, the available nodes usually have the same attributes in terms of these three variables. This may hinder the decision-making capabilities of the learners, as they are not provided with enough information to be sufficiently supported in that decision.

Another interesting result of the regression analysis is the high correlations among these three features, which suggest a degree of overlap among them. This is also apparent in the squared semipartial correlations associated with each of these features and which represent the percentage of the variance of the predicted variable uniquely explained by each of the predictor variables in the model. The values of the squared semipartial correlations are relatively low, suggesting small percentages of the variance of the predicted variables (mean time of completion for a node and of the standard deviation for time of completion of a node) that are uniquely explained by each of these metrics. This effect can be explained if the particular characteristics of each of the three variables, as they are translated in the given digital educational setting, were to be examined in parallel. In the instructional design implemented for the Modern Greek language course, the declarative nodes are in the lowest level of a learning sequence, which means that usually they don't have any prerequisites, or if they do, their number will be very low. Additionally, since declarative nodes are usually vocabulary nodes, introducing new words or phrases to the learner, they often involve only one level of linguistic description in their instructional goal, i.e., semantics. In a few rare cases there might be another level involved, morphology, when the instructional goal requires the knowledge of different morphological variations of the words in the vocabulary. Hence, both correlations between type of node and prerequisites and type of node and levels of linguistic description are explained when considering the above conditions. Finally, in many nodes, the instructional goal consists of several different elements, which are equivalent to both the prerequisites of that node (a prerequisite matching each of the elements) and, at the same time, each of these elements corresponds to a different level of linguistic restriction. This is especially true in nodes with low numbers of prerequisites and low levels of linguistic description. Considering that the majority of nodes actually belong to this category, this likely explains the high positive correlation between the number of prerequisites and the number of levels of linguistic description variables.

As a final point, it is crucial to emphasize that the evaluation of these three structural complexity metrics already implemented in the dashboard of the Modern Greek language learning course was conducted specifically in terms of their predictability of the time of completion for a particular node, and it does not reflect any utility of a variable other than completion time. The type of node, for example, is an important piece of information for language learners, as they implement different learning strategies when acquiring new vocabulary and when implementing a grammatical rule to form a linguistic construction in the target language (Nikolov & Djigunovic, 2006; Tsai, 2019). The knowledge of prerequisites is a scaffolding feature to guide learners when deciding which nodes to complete first, so they will acquire the necessary knowledge for achieving instructional goals higher up in the learning sequence. Knowing how many levels of linguistic description are involved in a specific instructional goal may greatly help language learners with directing their focus on multiple levels of a linguistic structure or examining a singular level. Therefore, my investigation of these metrics for node completion time doesn’t entail that these variables should be ignored or excluded from the dashboard due to their limited predictive power.

### 7.1.2. Estimation and evaluation of Linguistic Complexity Index (LCI)

One of the aims of the present research study was to propose an alternative metric for assessing the structural complexity of a content unit, in this case, a node. This alternative structural complexity metric should reflect more information about the parameters of the content affecting learning, compared to the existing ones (type of node, number of prerequisites and number of levels of linguistic descriptions), while maintaining high predictability of the time needed for a learner to complete that node. The proposed metric is the Linguistic Complexity Index, which was defined and operationalized in Chapter 5.

According to the literature on linguistic complexity or difficulty, this concept is more adequately described as a construct, as it presents multiple aspects corresponding to the different layers of analysis of a particular linguistic structure (Housen & Simoens, 2016). Each of these different levels may be a source of complexity, as it involves different elements incorporated into a word, a phrase, or a sentence. At the morphological level, for example, there are several affixes that may attach to the beginning (prefix), the end (suffix) or the middle (infix) of a word, manifesting a particular grammatical attribute, such as mood, voice, number, person, or tense in the case of a verb, or number, case or gender in the case of a noun. On top of that type of complexity, there is also complexity emerging due to the interaction between these different levels of linguistic description. For example, there are certain affixes that, when attached to a word, may alter its semantic properties. Thus, certain alterations at the morphological level have consequences at the semantic level for an utterance. An appropriate index of linguistic complexity needs to consider all these complications and reflect all these theoretical considerations in its operationalization procedure.

The proposed Linguistic Complexity Index focuses on three different aspects of complexity: lexical, morphological and syntactic complexity. These three aspects are the ones appearing more often in the literature (Bulte & Housen, 2012; Ehret and Szmrecsanyi, 2016; Housen & Simoens, 2016; Palloti, 2015) and they have been extensively analyzed and operationalized using different methods and approaches. However, there are levels of linguistic description that are not considered in the proposed metric, i.e., phonetics, phonology and pragmatics. The reason for this lies in the limitations of the Modern Greek language learning platform, as well as the interconnections between those different layers of analysis. This Modern Greek language course focuses on written speech instead of oral, mainly due to the fact that there are no reliable automated solutions for evaluating Modern Greek spoken utterances. Therefore, the instructional goals for each of the nodes in a module involve only three language competences, i.e., listening comprehension, reading comprehension and writing, which are the ones that can be properly assessed by the system. Hence, phonological complexity, which applies in the production of oral speech, is beyond the scope of this research. In terms of pragmatics, since this is a beginners’ Modern Greek language course, there are not that many linguistic concepts introduced in the modules that apply to that level of analysis. Moreover, most of the concepts and phenomena in pragmatics are usually manifested in other levels of linguistic description, such as syntax and morphology. For example, one of the linguistic concepts in this Modern Greek course that refers to pragmatics is formal and informal speech. Formal speech is realized in a Greek linguistic utterance by using the verb in the second person of the plural instead of the second person of the singular, when addressing a person. In terms of the formation of the utterance, this involves the addition of a specific suffix to the verb, adding to the morphological complexity of an instructional goal that involves formal speech in communication instances. Thus, introducing the aspect of pragmatic complexity does not accomplish a better description of the linguistic complexity of an utterance, as it is already captured by other complexity aspects, like morphological complexity.

Since the concept of linguistic complexity is a theoretical construct, operationalizing it in order to determine its value for a particular linguistic content unit poses both a statistical and a theoretical challenge. The statistical issue of how to measure a construct by calculating a specific numeric value has been adequately addressed by Song et al. (2013). Those researchers proposed different approaches when dealing with the problem of estimating a "construct value" using two or more "indicator" variables, each one of them with its own advantages and disadvantages. One approach is simple averaging, where the sum of the *z*-scores of the indicator variables is the actual score of the construct variable. This is an appropriate approach when the indicator variables have similar relationships with the dependent variable under analysis (Song et al. 2013). The analysis, however, showed that the relationship of each of the indicator variables with both predicted variables (mean time of completion for a node and standard deviation of time of completion for a node) differ, not only in magnitude, but also in direction. Lexical complexity has a negative correlation with the predicted variables, while morphological and syntactic complexity both have a positive correlation with them. Hence, simple averaging is not an appropriate approach to estimating the value of the construct.

Another approach to this challenging problem is the statistical method of weighted averaging (Sharma, 1996), which offers a mathematically plausible solution. The standardized scores of the three indicator variables corresponding to the three different aspects of linguistic complexity are used in a Principal Component Analysis, which provides the weights to be used when estimating the value of the construct.

Hence, each of the indicator variables is assigned a specific weight, which will be used in the process of aggregating the individual scores of the three aspects of complexity in order to calculate an overall Lexical Complexity Index.

However, there are two major issues with this approach to calculate a composite variable to successfully represent the construct of linguistic complexity. First, the fact that this solution is mathematically plausible doesn't necessarily entail that it is also theoretically plausible. Aggregating the scores of the three aspects of complexity results in loss of important data and information that can be used to further explore and understand learner behavior. This becomes apparent considering the semi-partial correlations of the variables in the prediction models for both mean and standard deviation of time to completion, with the unique variance explained ranging between 7 and 15 percent.

The second issue with the weighted averaging approach is related to that exact mathematical nature of the solution. The weights that attach to the scores of the different indicator variables derive from statistical formulae and analysis procedures (principal component analysis) that are highly dependent on the specific data in this specific research context. Therefore, transferring the process in a different context might result in completely different weights, in terms of both the value and the sign (positive or negative). This can be made clearer when considering the weights given in this case, in relation with the nature of the content of a node, as it was determined by the instructional design of the Modern Greek course. The weights for morphological complexity and syntactic complexity are very similar (.810 and .882 respectively). This is in accordance with theory, since morphology and syntax are the two most interconnected levels of linguistic description, and highly interactive with each other (Philippaki-Warburton, 1992). Several syntactic assumptions, like subject-verb agreement, are fulfilled by the appropriate use of specific affixes, determining the number and the gender of a particular noun, or the number and the person of a particular verb. The two levels are so closely related, that in many theoretical linguistic textbooks are referred together as "grammatical" level of analysis (Philippaki-Warburton, 1992). On the other hand, the weight for lexical complexity is similar in magnitude but opposite in direction (-.764). This value suggests that lexical complexity has an opposite contribution to the general complexity of a linguistic unit. However, as it has already been mentioned, this mathematical solution which was used for the determination of these weights is highly contextual. In the current LMS, students followed the default order of nodes instead of differing learning paths. The weights obtained from the current constrained and fixed order could differ from those obtained if the students were encouraged to choose diverging sequences. Additionally, in the present design, the distinction between declarative and procedural nodes entails that there are vocabulary heavy nodes, which consequently present higher values of lexical complexity, and nodes more grammatically heavy, with higher morphological and syntactic complexity. There are only a few nodes that combine lexical and morphological complexity (those nodes that introduce functional words as vocabulary, such as prepositions). In most cases, when a node has a high value for lexical complexity, morphological and syntactic complexity are absent, and vice-versa. This is another reason behind the difference in the weights of the indicator variables. It is apparent that further research is needed to explore how the three different components might be jointly used to create a composite index, before attempting to generalize conclusions. The composition rules may be more complex, with different weights for each indicator variable depending on its value. For example, an indicator variable may be required to have a value greater than some threshold to be included in the calculation of the composite index.

Another theoretical issue refers to the loss of important information when estimating a single numerical value for the composite variable of linguistic complexity. With the process of weighted averaging, the aim is to arrive at a single score that characterizes the overall complexity of a node. This quantification of complexity may downgrade the importance of the qualitative dimension of its different aspects, as they are collapsed under a single number. Even though this has the advantage of providing a single metric to the learner, allowing them to compare at a glance the complexity of two nodes, at the same time it hides certain characteristics of the nodes which are also important for this same comparison. There are advantages to maintaining these three complexity subindices as it provides a multidimensional framework to use when creating student models for the LMS dashboard, i.e., to indicate learner performance on each of these three complexity directions and take these measures into account when recommending next nodes for the learner to access.

Apart from the theoretical issues of the weighted averaging approach to estimating lexical complexity as a composite variable, there is also a statistical issue related to the predictability of that metric in terms of the time required by a learner to complete a particular node. The loss of predictability has been one of the drawbacks of the method, noted by Song et al. (2013). The regression analyses for both time variables of the study, mean time to completion and standard deviation of the time to completion, have shown substantial reduction in predictive power of the regression equation (reduced almost in half), when using the composite Linguistic Complexity Index rather than the set of the three subindices of linguistic complexity. Therefore, aggregating the three subindices into a single index, not only provides reduced information to the language learner, but it also lacks predictability for the time to completion for the node.

Hence, my research suggests the Linguistic Complexity Index be operationalized as a set of three different subindices, each one corresponding to the lexical, morphological, and syntactic complexity of a node. The scores of these indices are standardized, to provide a common scale that may allow for comparisons between them. Therefore, each node in the Modern Greek course will be described by this set of three complexity subindices. The value of 0 for a particular index corresponds to the absence of linguistic elements that contribute to the complexity of that level of linguistic description. For example, lexical complexity of 0 for a node is interpreted as no new vocabulary items that need to be learned in that node.

### 7.1.3. Navigational Patterns and Learning Strategies

It has already been discussed that the field of learner attitudes and behaviors in a digital language learning environment is an under-researched area in the CALL research space (Gillespie, 2020). The few research studies focusing on this topic have issues of limited scope, like small sample sizes (sometimes as low as 10 participants) and limited duration of the research. Additionally, the level of examination of the learner behavior in CALL applications adopted by these studies is also the focus of some criticism (Blake, 2013). Some studies implement a high-level investigation of learner behavior, with the purpose of identifying language learner types. They typically use metrics that provide little detail on the learning strategies incorporated by learners such as log-in frequency, accessing lectures vs accessing exercises, total completion of activities, and forum activity (Veletsianos et al., 2021; Martin-Monje et al., 2018; Li et al., 2018; Keskin et al., 2016). On the other hand, there are a few studies (Desmarais, 1998; Zhou & Wei, 2018; Youngs et al., 2018), which investigate learner behavior in detail, but use metrics on specific affordances of the system that are not transferable to other CALL contexts where such affordances are absent.

In the topic of emerging language learning patterns and behaviors in CALL, the findings are presented either in isolation, or in correspondence to various academic success metrics such as passing or failing a course or, in the case of language MOOCs, completing 100% of the course. A much more involved approach, which promises to facilitate the advancement of research in instructional design for digital language courses (and digital courses in general) is to contrast these findings with certain design decisions that help explain the emerging learning behavior patterns. Hence, it is quite informative to consider how instructional design decisions might affect learners' decisions and navigation patterns.

In the Methodology section, the instructional design of the online Modern Greek language learning course was discussed, especially the design intention to create an open structure, where student progression wouldn't be locked in because of linear pathways. Student learning is theoretically enhanced by providing multiple study and learning options, as well as different learning pathways to explore. This design intention was realized in the tree-form structure of modules, where different nodes were available to the learner to access, allowing for completion of the module following different paths of node completion. A remarkable finding of the present study was that 100% of the learners in the sample followed exactly the same pathway. Another interesting observation was that this order was followed at all times, even when the learners failed to complete assignments and, consequently, nodes. Even though there was the option of attempting a different node and revisiting the failed one at a later time, every student in the course "locked" their progression to a linear path, not proceeding to the next node before successfully completing the previous one. This learning behavior may be explained by the inclusion of numbers in the node titles, as it can be seen in Figure 4.1 (i.e., node 0, node 1, node 2, etc.). The numbering of nodes created the impression of a specific order of the nodes, which the learners had to follow to successfully complete the course.

This kind of linearity can also be observed in the students' navigational patterns in the different sections of a specific node (study materials, learning activities and assessments). Table 5.3 shows that the frequency of the learning Tactic G ($f$ = 1407), in which students access the different sections in the order Study – Learn – Test far outweighed any other observed learning tactic extracted from the student logs. The second most frequent tactic has a frequency of $f$ = 664, less than half the frequency of Tactic G. This observation makes sense for two reasons. First the user interface design of the node sections navigation screen presents the three different categories of activities in a linear fashion, creating the impression that the different sections should be accessed in that order. The second reason has to do with the student preferences. As Mallia (2014) noted, learner preferences for learning strategies to be adopted when working on a language course are heavily dependent on the previous learning experiences. Most students prefer deductive strategies and tactics to inductive ones because deductive strategies were more prominent in the learners' earlier language learning experiences. Also, Sik (2015) argued that students' and teachers' dominant perception is that deduction is more effective and leads to better performance results. Hence, learners tend to first study theoretical pages to acquire conceptual linguistic knowledge, and then proceed to the examples and exercises.

Since the observed learning tactics (and consequently learning strategies) were overwhelmingly deductive, the investigation of learning behavior corresponding to the differences between the Linguistic Complexity Subindices of the different nodes relied on statistical analyses identifying statistically detectable changes in the frequencies of these tactics. The first analysis regressed the seven most adopted learning tactics against the lexical complexity subindex of the nodes in the Modern Greek language course. The analysis showed that, when learners encounter nodes of higher lexical complexity, they tend to dedicate whole learning sessions to studying vocabulary lists. At the same time, the number of learning sessions dedicated to assessments is significantly decreased. Students prefer revisiting the study materials in case of failure in the assessment, rather than attempting the assessment again. The analyses investigating prominence of deductive or inductive learning behaviors also follow the same pattern. In cases of nodes with higher values in their lexical complexity subindices, learners tend to rely more on deductive learning strategies and focus more on reviewing study material rather than working on learning activities. The learner preference for external linguistic information sources to attain vocabulary acquisition goals is an expected behavior, as argued by Tsai (2019), who noted that definitional knowledge of word senses is more efficiently accommodated by the use of deductive learning strategies. Lee and Lin (2019) also argued that in vocabulary acquisition most learners don't feel comfortable inferring the meaning of words from examples, concordance lists or other intrinsic linguistic information sources.

In the case of syntactic complexity, the regression analyses showed an almost opposite behavioral pattern. The number of learning episodes wholly dedicated to study materials was decreased when working on nodes with high syntactic complexity, and the same holds true for the number of learning episodes where learners traversed linearly through study materials, learning activities and assessments. In contrast, the number of learning episodes dedicated to assessments was increased. Learners relied more on inferring the syntactic rules based on the feedback given by the system, rather than revisiting or focusing on extrinsic linguistic information provided by the study materials. The learners' overall learning behavior in nodes of higher syntactic complexity also follows a similar pattern, suggesting a tendency towards adopting an inductive learning strategy rather than a deductive one. Also, students focused more on working on learning activities and less on studying the theoretical pages of such nodes. Syntax in Greek follows specific and consistent rules, so the occurrence of such learning behaviors agrees with the comments by Haight et al. (2007), who suggested that induction is an option for studying linguistic phenomena with salient and consistent features. Sheffer (1998) also suggested that induction facilitates learning when students focus on difficult to describe concepts or linguistic phenomena. This is especially true for the Greek syntax, where an implicit example of a well-formed Greek sentence can be less complex to process than an extrinsic description of the syntactic rules, which also may require knowledge of certain linguistic jargon used in the text.

The case of the morphological complexity of an instructional goal of a node holds the greatest interest, as it seems to be in the middle ground in terms of the learning behavior exhibited by the learners. The only statistically detectable predictor of morphological complexity was the number of learning episodes dedicated only to working on study materials, which appears significantly decreased for nodes of high morphological complexity. When focusing on a complex morphological concept, students need to study the different affixes attaching to a word, along with the grammatical features each represents (number, gender, case etc.), just like what they are doing when studying the vocabulary. On the other hand, the selection of the appropriate affix to be used in a sentence relies on the linguistic context of the word, i.e., the other lexical items in the sentence. Therefore, it is also important to look at intrinsic language examples, to examine these interactions between the different words of the sentence. This mixed approach to studying nodes of high morphological complexity is also evident in the regression analyses investigating the overall learning strategies adopted by the learners in the course. Neither the number of inductive nor the number of deductive learning tactics adopted by the learners was a statistically detectable predictor of the morphological complexity subindex of a node, suggesting that learners didn't favour a particular learning strategy when tackling nodes with higher morphological complexity. Also, the analyses showed an increased learner focus on working on learning activities, but not a decreased learner focus on reviewing studying materials, as it was the case in nodes with higher syntactic complexity.

## 7.2. Implications of the findings

As it was mentioned in Chapter 1, the main purpose of this thesis is to inform the instructional design of the Modern Greek online language learning platform and to suggest updates and improvements that will facilitate learning and support learner autonomy. At the same time, since there are aspects of this thesis focusing on topics in Computer Assisted Language Learning which are under-researched, like research methodology and language learning behavior in Virtual Learning Environments (Gillespie, 2020), the empirical findings may shed some light or provide direction when further investigating these fields.

### 7.2.1. Implications to the instructional design and content development

The data and the results of the statistical analysis conducted in this research suggest directions for certain improvements and updates to the digital learning environment of the Modern Greek online language course. The design of the structure and navigation patterns for the Modern Greek language course was intended to challenge the traditional language learning software layout, which follows the rigid format of a language learning textbook, with a series of units of the same form (text – grammar – exercises) that need to be completed in a specific order. Even though the online language learning platform gave learners the potential to follow different learning paths, especially in the order of accessing the various nodes inside a module, students didn't take advantage of that feature. The main reason behind this observation may have been the way nodes were represented in the user interface, where they were numbered and presented to the learners as icons arranged in a linear formation. These design choices suggested a single linear learning path, accessing the nodes in the order they were numbered, which have been followed by the entirety of the student cohorts enrolled in the Modern Greek course. Hence, in a future iteration of this digital learning environment, the graphical presentation of the nodes should be revisited, avoiding the inclusion of any kind of numbering in the title.

When an instructional design gives learners the freedom to pursue different learning tactics and strategies to accomplish instructional goals, it is also important to provide them with the appropriate scaffolding, so they are not overwhelmed by the number of choices they must make. In the case of adaptive intelligent language tutors, which incorporate more open navigational schemata, it is essential to support learner decisions on the order they access the course content (Slavuj et al., 2017). This is the case with the online Modern Greek language course, which allows learners to choose which from a number of nodes to access next. An adaptive approach to the student-facing dashboard would be a representation of the nodes based on specific criteria, such as the amount of temporal and cognitive resources needed to successfully complete them. The adaptive sequencing of nodes needs to be performed in a way that guides the learners in their decision making, while at the same time doesn't eliminate learner choice and their ability to control the learning experience (Schneider et al., 2018). In several intelligent language tutoring systems (Wauters et al., 2010, Slavuj et al., 2017, Susnjak et al., 2022), content sequencing is based on difficulty estimates using learner-related metrics, such as mean time of completion or average performance scores. The major issue of implementing estimates like these is that they obscure the influence of individual differences when working with the language learning material, while being totally dependent on them. For example, some language learners may be quite competent at completing a node of high syntactic complexity, even though the majority of learners may struggle with it, raising the mean completion time for that node. This fact is also evident in the empirical evidence provided by this thesis, where the standard deviation for the time of completion of many nodes in the Modern Greek course is very high, suggesting a large variation in completion times among the different language learners. The three Lexical Complexity Subindices suggested in my research provide an estimate of node difficulty that is independent of the learner individual differences, as they are calculated from the actual features of the language learning content, instead of learner dependent metrics. Representing difficulty in this multidimensional manner is advantageous for an intelligent tutoring system (Ma et al., 2014), especially since each separate dimension presents substantial differences in two aspects:

- In how well it predicts the time of completion of a node. The empirical evidence provided by this research suggest that syntactic complexity has the greatest predictability for time of completion, followed by morphological complexity, whereas lexical complexity appear to have no statistically detectable predictability.
- In the learning tactics and strategies implemented by the language learners. This research showed that learners favored deductive strategy for nodes with higher lexical complexity and inductive strategy for nodes with higher syntactic complexity, while no favored strategy was statistically detected for nodes with higher morphological complexity.

Sotillare et al. (2013) identify the domain model and the student model as two of the major conceptual components of an Intelligent Tutoring System (ITS). The domain model represents the knowledge space of the course delivered by the ITS, while the student model represents the learner's knowledge state on the learning objectives of the course. The Lexical Complexity Subindices may be introduced to the domain model of the online Modern Greek language course, as difficulty estimates of the content, but they may also be adopted as three different dimensions for the student model of the LMS. In the section of the student dashboard where the learners may find the nodes of a module that can study next, a three-coloring system (green for low difficulty, yellow for intermediate and red for high difficulty) may be used to classify these nodes according to their difficulty. A simple coloring system is preferable to more complex ones, which have been found to increase cognitive load and make the dashboards challenging to be used by the learners (Bera, 2016). In the determination of the color for a particular node in the course, the type of complexity can be factored, as well as learner-based metrics, such as learner performance metrics on nodes with similar complexity patterns. In this way, the predictive model adopts a more dynamic approach, giving it the flexibility to address issues like concept drift (Lu et al., 2018), where the accuracy of predictions is degraded due to changes in the learners' performance and learning behaviors. Future research may provide more empirical evidence on how these learner-related factors should be seeded in the difficulty estimation model.

Susnjak et al. (2022) noted in their study that the majority of learning analytics dashboards implement descriptive data rather than predictive (estimating future outcomes) or prescriptive (suggest ways of improvement). The empirical evidence resulted in the statistical analyses conducted in the present research may cater to the prescriptive aspect of the dashboard, by suggesting strategies and tactics according to the complexity characteristics of a particular node. This is especially crucial in the event of learners failing an assessment of a node, as it may guide them on how to overcome the challenges they encountered. For example, the research showed that in the case of nodes with high lexical complexity, most learners adopted a deductive strategy, focusing more on the study materials. Thus, learners should be guided towards reviewing the vocabulary lists in detail, before attempting the practice exercises. In the case of nodes with high morphological complexity, the empirical evidence suggests that learners don’t favor a particular learning strategy (inductive or deductive), or focus on a specific learning object type. Hence, a suggested learning strategy should involve working on study materials in parallel with the learning activities, for the learners to be exposed to authentic language use examples alongside the morphological formation rules explained in the theory pages. However, these hypotheses need to be further tested in future research.

In addition to informing the instructional design of this specific online language learning platform, the results presented in this thesis also have an impact to the wider field of instructional design and content development in Computer Assisted Language Learning, especially through the introduction of the Linguistic Complexity Subindices (LCS). In the case of computer-assisted language learning software, especially tutorial CALL software, the LCS may be a very useful tool for scaling the difficulty of exercise items in automatically generated sequencies of exercise questions. Calculation of the LCS for each item will provide customization parameters for increasing the corresponding aspect of the complexity of an exercise (lexical, morphological or syntactic). For example, in the case of an item which corresponds in high difficulty in all three subindices of linguistic complexity, the difficulty may be customized by selecting more simple vocabulary (lowering lexical complexity), selecting morphological alternatives of the same feature (number, person, case etc., to lower morphological complexity), or reducing the sentence constituents (lowering syntactic complexity).

The evaluation of the complexity of a particular linguistic phenomenon or structure using the LCS may also be a useful asset for second language acquisition experts developing content for language learning courses. LCS provide a numerical estimation of the linguistic complexity of a course unit or an instructional goal that can help content developers to organize the learning material with scaling difficulty, essentially controlling the complexity curve of the language course, or creating roughly equivalent units in the course, in terms of complexity. Additionally, content developers could also use the LCS to provide appropriate learning resources to the learners, according to the nature of the complexity of a particular unit. For example, for units with increased syntactic complexity there should be more working examples on how to form the specific linguistic structure or more authentic sample sentences with implicit linguistic information, and there should be less focus on extensively describing the formation process or on the explicit presentation of the syntactic rules involved.

### 7.2.2. Implications for research

The major contribution of this thesis to Computer Assisted Language Learning research, and Second Language Acquisition research in general, is the introduction of the Linguistic Complexity Subindices as a metric that estimates the difficulty of a particular linguistic structure or a unit of a language learning course. This set of subindices, each corresponding to a different aspect of structural complexity (lexical, morphological and syntactic) is theoretically founded in the relative literature and uses for its operationalization procedure linguistic attributes which are readily available and universal to all the language systems. This feature may allow for research on the field of applied linguistics (especially Second Language Acquisition) to be more language independent when investigating how the difficulty of the language learning content relates to the design of learning materials; more specifically, how it affects difficulty of learning the material, optimal acquisition order, and the efficacy of instructional methods. Also, the analyses conducted in this study provide empirical evidence to support claims made in previous research that higher structural complexity (difficulty of content) doesn't necessarily entail higher cognitive complexity (the resources needed by the learner to acquire designated linguistic knowledge).

Additionally, this thesis offers a different perspective on the theory and research of deductive and inductive language learning strategies in Computer Assisted Language

Learning. Literature on this field focuses on determining the effectiveness of each of these two approaches to language learning in terms of learner performance, knowledge transfer, or retention. Other studies approach these different learning strategies as an aspect of the individual differences between learners, with groups of students favoring the one or the other. This study examines how adoption of these learning strategies may change when learners work on instructional goals of different difficulty (as operationalized by the three Linguistic Complexity subindices).

This thesis also suggests a different approach to define and operationalize deductive and inductive language learning strategies. Previous research would characterize an action sequence in a digital language learning environment as deductive or inductive based on the order learners accessed specific types of activities and learning resources. The present study suggested an additional time sensitive metric to triangulate the classification of an overall language learning strategy as deductive or inductive: the mean time on study materials and the mean time on learning activities (using the terminology adopted in the online Modern Greek language course), an increase of which is an additional indication of deductive or inductive learning strategy respectively.

Finally, this research provided empirical evidence of how the learners' learning behavior is modified in nodes with different values of their Linguistic Complexity Subindices. It provides empirical evidence that challenge the claims of prior research that deductive learning strategies are adopted by learners working on linguistic structures or sentences of higher complexity Abuseileek (2009). The results of the analyses showed that the type of structural complexity may also affect the choice of a particular language learning strategy. These results may inform design of computer assisted language learning applications, as well as direct future research to further investigation of how learners change their adopted language learning tactics and strategies when working on language learning units or instructional goals that differ in structural complexity.

## 7.3. Limitations and future research

One of the limitations of this study stems from the operationalization of the three Linguistic Complexity Subindices. The purpose of the suggested procedure which calculates the values for each index was to provide a simple enough algorithm for this calculation, which will reflect the major theoretical underpinnings and considerations on the referred aspect of linguistic complexity (lexical, morphological, or syntactic) and will provide an adequate estimation of the construct for the various nodes of the Modern Greek language course. The trade-off for the simplification and versatility of the suggested complexity metric was that certain factors which affect linguistic complexity are not considered in the calculation procedure. For example, in the case of the lexical complexity subindex, we do not consider how different or similar are the senses of the words to be learned by the students. As Palotti (2015) claims, the two extremes (words with very similar senses and words with completely non-relevant senses) add to the lexical complexity of a vocabulary list. Hence, a future iteration of the research study might also address such issues by revisiting the calculation process of these indices.

Another factor that delimited the research study in certain ways was the nature of the Modern Greek language course content. As it has been mentioned in Chapter 4, the course targets absolute beginners in Modern Greek and covers an elementary level of Greek language proficiency. As such, the content does not include advanced linguistic phenomena, vocabulary, and extended linguistic products (learners focus on single sentences rather than paragraphs or longer essays). An interesting future direction for this type of research would be to examine language courses addressing more advanced levels of linguistic proficiency, or even different target languages, taking advantage of the universality and versatility of the Linguistic Complexity Subindices.

The online language learning platform incorporated for the Modern Greek language learning course was also lacking capabilities for more open-ended activities, as well as activities involving oral language production. Inclusion of these types of activities in a future research iteration will also provide a more complete picture of the learning tactics and strategies adopted by the learners. Additionally, the Modern Greek language course included a limited number of nodes (56 in total). A future iteration of the research might also involve multiple courses, with higher number of nodes and more variety in linguistic complexity between them.

In the case of the data capture tools, there have been some limitations that didn’t allow for the inclusion of useful data in the present investigation. For example, the Learning Management System does not capture all the learner’s scores in the assessment activities, which could have been another measure of learner performance in the course. Additionally, the system didn't capture the ID of each activity accessed by a learner, just the activity type. Moreover, certain aspects of learner behavior were also not captured by the LMS, like visiting the overview page for a module or node, or accessing the dashboard to review performance analytics provided to the students. Inclusion of that information in future research will allow investigation on how learning behavior may vary in activities involving different language skills (reading, writing, listening, or speaking), what type of study materials is accessed more depending on the language skill involved, etc. Finally, the learning tactics and strategies implemented by learners may be captured at a higher level of detail.

The sample size for this research study, two student cohorts and 58 students in total, was also restrictive. A larger number of participants, ideally from different institutions instead of only one, would also add value to future research. As an extension to this line of investigation, inclusion of participants from different levels of education (K-12, both elementary and secondary schools) as well as different types of schools such as heritage language schools operating in the afternoons or weekends, could also provide interesting and novel lines of investigation on how language learning behavior may vary. The limited access to the data of the student cohort under investigation because of the ethics approval restrictions was a significant limitation of the research. Hence, in a future iteration, some additional information on the students such as prior knowledge in linguistics and previous language learning experience would allow for investigation of language learning behavior in relation to the learners' individual differences. Another possible line of investigation in future research may involve examination of learning tactics and strategies in relation to various measures of student performance (i.e., the scores achieved in the assessment activities of the course). Finally, an improved clickstream data capture instrument with learner self-report features would provide an additional layer of data about the students' learning behaviors and adopted tactics and strategies.

A future iteration of the research could incorporate fundamental extensions of the research questions, with the goal of further illuminating the learning processes involved when interacting with the online Modern Greek language course. For the estimation of difficulty, my thesis focused entirely on the features and characteristics of the learning content, which served as a foundation for identifying the structural complexity of a node in the course. Establishing this foundation allows further research to bring attention to the learner individual differences. Future research questions addressing this perspective may include an investigation of the impact of learner characteristics such as self-efficacy or prior knowledge (in the language and in language learning in general). Additionally, an interesting investigation opportunity could be an examination of how the LCS may incorporate learner-based metrics, like performance in specific language learning competences (reading, writing, speaking or listening), or in nodes with different Linguistic Complexity Subindices values. Another compelling new direction for the research would be to approach the notion of concept drift, i.e., the change in the learners' study patterns with time, offering a longitudinal perspective in the investigation of online language learning behavior. Finally, future research may provide empirical evidence on the impact of the design modifications suggested in the implications section, such as ordering the presentation of the learning objects and providing nudges or prompts on optimal learning tactics or strategies.

# References

AbuSeileek, A. F. (2009). The effect of using an online-based course on the learning of grammar inductively and deductively. *ReCALL*, *21*(3), 319–336.

Anderson, A., Huttenlocher, D., Kleinberg, J., & Leskovec, J. (2014). Engaging with massive online courses. *Proceedings of the 23rd International Conference on World Wide Web*, 687–698.

Astivia, O. L. O., & Zumbo, B. D. (2019). Heteroskedasticity in Multiple Regression Analysis: What it is, How to Detect it and How to Solve it with Applications in R and SPSS. *Practical Assessment, Research, and Evaluation*, *24*(1), 1.

Bax, S. (2003). CALL—past, present and future. *System*, *31*(1), 13–28.

Bera, P. (2016). How colors in business dashboards affect users' decision making. *Communications of the ACM*, *59*(4), 50–57.

Blake, R. (2013). CALL research and practice: Quo vadis? *International Journal of the Linguistic Association of the Southwest (IJLASSO)*, *32*(1).

Blikstein, P., Worsley, M., Piech, C., Sahami, M., Cooper, S., & Koller, D. (2014). Programming pluralism: Using learning analytics to detect patterns in the learning of computer programming. *Journal of the Learning Sciences*, *23*(4), 561–599.

Brown, H. D. (2000). *Principles of language learning and teaching* (Vol. 4). Longman New York.

Bulté, B., & Housen, A. (2012). Defining and operationalising L2 complexity. *Dimensions of L2 Performance and Proficiency: Complexity, Accuracy and Fluency in SLA*, 23–46.

Chapelle, C. A. (2010). The spread of computer-assisted language learning. *Language Teaching*, *43*(1), 66–74.

Chen, X., Zou, D., Xie, H. R., & Su, F. (2021). *Twenty-five years of computer-assisted language learning: A topic modeling analysis*. http://hdl.handle.net/10125/73454

Chun, D. M. (2013). Contributions of tracking user behavior to SLA research. *CALICO Journal*, 256–262.

Dahl, Ö. (2004). *The growth and maintenance of linguistic complexity* (Vol. 10). John Benjamins Amsterdam.

Debski, R. (2003). Analysis of research in CALL (1980–2000) with a reflection on CALL as an academic discipline. *ReCALL*, *15*(2), 177–188.

Deci, E. L., & Ryan, R. M. (2012). *Self-determination theory.*

DeKeyser, R. (2016). Of moving targets and chameleons: Why the concept of difficulty is so hard to pin down. *Studies in Second Language Acquisition*, *38*(2), 353–363.

Desmarais, L., Laurier, M., & Reni, D. (1998). The analysis of navigation patterns in CALL. *Computer Assisted Language Learning*, *11*(3), 309–315.

Deutscher, G. (2009). Nominalization and the origin of subordination. *Syntactic Complexity: Diachrony, Acquisition, Neuro-Cognition, Evolution*, 199–214.

Deutschmann, M., & Vu, M. T. (2015). *Computer assisted language learning in language education: An overview of theories, methods, and current practices*.

Dewey, J. (1908). What does pragmatism mean by practical? *The Journal of Philosophy, Psychology and Scientific Methods*, *5*(4), 85–99.

Division, C. of E. C. for C. C. E. C. M. L. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge University Press.

Doughty, C. J., & Long, M. H. (2003). *Optimal psycholinguistic environments for distance foreign language learning*.

Duval, E. (2011). Attention please! Learning analytics for visualization and recommendation. *Proceedings of the 1st International Conference on Learning Analytics and Knowledge*, 9–17.

Evseeva, A., & Solozhenko, A. (2015). Use of flipped classroom technology in language learning. *Procedia-Social and Behavioral Sciences*, *206*, 205–209.

Garrett, N. (2009). Computer-assisted language learning trends and issues revisited: Integrating innovation. *The Modern Language Journal*, *93*, 719–740.

Gasevic, D., Jovanovic, J., Pardo, A., & Dawson, S. (2017). Detecting learning strategies with analytics: Links with self-reported measures and academic performance. *Journal of Learning Analytics*, *4*(2), 113–128.

Gelan, A., Fastré, G., Verjans, M., Martin, N., Janssenswillen, G., Creemers, M., Lieben, J., Depaire, B., & Thomas, M. (2018). Affordances and limitations of learning analytics for computer-assisted language learning: A case study of the VITAL project. *Computer Assisted Language Learning*, *31*(3), 294–319.

Gibson, D. C. (2018). Unobtrusive observation of team learning attributes in digital learning. *Frontiers in Psychology*, *9*, 834.

Gillespie, J. (2020). CALL research: Where are we now? *ReCALL*, *32*(2), 127–144.

Gollin, J. (1998). *Deductive vs. Inductive language learning*.

Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers*, *36*(2), 193–202.

Groot, P. J. (2000). *Computer assisted second language vocabulary acquisition*.

Haight, C. E., Herron, C., & Cole, S. P. (2007). The effects of deductive and guided inductive instructional approaches on the learning of grammar in the elementary foreign language college classroom. *Foreign Language Annals*, *40*(2), 288–310.

Haspelmath, M., & Sims, A. (2010a). *Understanding morphology oxford university press*. USA.

Haspelmath, M., & Sims, A. (2010b). *Understanding morphology oxford university press*. USA.

Heift, T., & Chapelle, C. A. (2013). Language learning through technology. In *The Routledge handbook of second language acquisition* (pp. 573–588). Routledge.

Heller, J., Steiner, C., Hockemeyer, C., & Albert, D. (2006). Competence-based knowledge structures for personalised learning. *International Journal on E-Learning*, *5*(1), 75–88.

Hendriks, H., & Watorek, M. (2012). The Role of Conceptual Development in the Acquisition of the Spatial Domain by L1 and L2 Learners of French, English and Polish. *Comparative Perspectives on Language Acquisition. A Tribute to Clive Perdue, Multilingual Matters, Bristol, Buffalo, Toronto*, 401–419.

Housen, A., & Kuiken, F. (2009). Complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics*, *30*(4), 461–473.

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

Hromalik, C. D., & Koszalka, T. A. (2018). Self-regulation of the use of digital resources in an online language learning course improves learning outcomes. *Distance Education*, *39*(4), 528–547.

Hulstijn, J. H., & De Graaff, R. (1994). Under what conditions does explicit knowledge of a second language facilitate the acquisition of implicit knowledge? A research proposal. *AILA Review*, *11*(97–112).

Hwu, F. (2007). Learners' strategies with a grammar application: The influence of language ability and personality preferences. *ReCALL*, *19*(1), 21–38.

Hwu, F. (2013). A review of the use of script-based tracking in CALL research for data sharing: Applications providing meaning aids. *Computer Assisted Language Learning*, *26*(4), 350–370.

Hwu, F., & Tzseng, C. (2013). Challenges and Directions of Script-Based Tracking in Tutorial CALL. *CALICO Journal*, 246–255.

Keskin, S., Şahin, M., & Yurdugül, H. (2019). Online learners' navigational patterns based on data mining in terms of learning achievement. In *Learning technologies for transforming large-scale teaching, learning, and assessment* (pp. 105–121). Springer.

Kovanovic, V., Gašević, D., Dawson, S., Joksimovic, S., & Baker, R. (2015). Does time-on-task estimation matter? Implications on validity of learning analytics findings. *Journal of Learning Analytics*, *2*(3), 81–110.

Lan, Y.-J. (2013). The effect of technology-supported co-sharing on L2 vocabulary strategy development. *Journal of Educational Technology & Society*, *16*(4), 1–16.

Larson, M. B., & Lockee, B. B. (2013). *Streamlined ID: A practical guide to instructional design*. Routledge.

Lee, P., & Lin, H. (2019). The effect of the inductive and deductive data-driven learning (DDL) on vocabulary acquisition and retention. *System*, *81*, 14–25.

Levy, M. (2009). Technologies in use for second language learning. *The Modern Language Journal*, *93*, 769–782.

Levy, M., & Stockwell, G. (2013). *CALL dimensions: Options and issues in computer-assisted language learning*. Routledge.

Li, H., Flanagan, B., Konomi, S., & Ogata, H. (2018). Measuring behaviors and identifying indicators of self-regulation in computer-assisted language learning courses. *Research and Practice in Technology Enhanced Learning*, *13*(1), 1–12.

Lin, C.-H., Zhang, Y., & Zheng, B. (2017). The roles of learning strategies and motivation in online language learning: A structural equation modeling analysis. *Computers & Education*, *113*, 75–85.

Lu, J., Liu, A., Dong, F., Gu, F., Gama, J., & Zhang, G. (2018). Learning under concept drift: A review. *IEEE Transactions on Knowledge and Data Engineering*, *31*(12), 2346–2363.

Ma, W., Adesope, O. O., Nesbit, J. C., & Liu, Q. (2014). Intelligent tutoring systems and learning outcomes: A meta-analysis. *Journal of Educational Psychology*, *106*(4), 901.

Maldonado-Mahauad, J., Pérez-Sanagustín, M., Kizilcec, R. F., Morales, N., & Munoz-Gama, J. (2018). Mining theory-based patterns from Big data: Identifying self-regulated learning strategies in Massive Open Online Courses. *Computers in Human Behavior*, *80*, 179–196.

Mallia, J. G. (2014). Inductive and Deductive Approaches to Teaching English Grammar. *Arab World English Journal*, *5*(2).

Malvern, D. D., & Richards, B. J. (1997). A new measure of lexical diversity. *British Studies in Applied Linguistics*, *12*, 58–71.

Martín-Monje, E., Castrillo, M. D., & Mañana-Rodríguez, J. (2018). Understanding online interaction in language MOOCs through learning analytics. *Computer Assisted Language Learning*, *31*(3), 251–272.

Matcha, W., Gašević, D., Ahmad Uzir, N., Jovanović, J., Pardo, A., Maldonado-Mahauad, J., & Pérez-Sanagustín, M. (2019). Detection of learning strategies: A comparison of process, sequence and network analytic approaches. *Transforming Learning with Meaningful Technologies: 14th European Conference on Technology Enhanced Learning, EC-TEL 2019, Delft, The Netherlands, September 16–19, 2019, Proceedings*, 525–540.

Matcha, W., Gašević, D., Uzir, N. A., Jovanović, J., & Pardo, A. (2019). Analytics of learning strategies: Associations with academic performance and feedback. *Proceedings of the 9th International Conference on Learning Analytics & Knowledge*, 461–470.

McEneaney, J. E. (2016). Simulation-based evaluation of learning sequences for instructional technologies. *Instructional Science*, *44*, 87–106.

Meyers, L. S., Gamst, G., & Guarino, A. J. (2016). *Applied multivariate research: Design and interpretation*. Sage publications.

Nation, P., & Kyongho, H. (1995). Where would general service vocabulary stop and special purposes vocabulary begin? *System*, *23*(1), 35–41.

Nesbit, J. C., & Hunka, S. (1987). A method for sequencing instructional objectives which minimizes memory load. *Instructional Science*, *16*(2), 137–150.

Nikolov, M., & Djigunović, J. M. (2006). Recent research on age, second language acquisition, and early foreign language learning. *Annual Review of Applied Linguistics*, *26*, 234–260.

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

Pandarova, I., Schmidt, T., Hartig, J., Boubekki, A., Jones, R. D., & Brefeld, U. (2019). Predicting the difficulty of exercise items for dynamic difficulty adaptation in adaptive language tutoring. *International Journal of Artificial Intelligence in Education*, *29*(3), 342–367.

Payne, J. S. (2020). Developing L2 productive language skills online and the strategic use of instructional tools. *Foreign Language Annals*, *53*(2), 243–249.

Peirce, C. S. (1998). *The Essential Peirce: Selected Philosophical Writings. 1893-1913. Volume 2* (Vol. 2). Indiana University Press.

Poole, R. (2012). Concordance-based glosses for academic vocabulary acquisition. *Calico Journal*, *29*(4), 679–693.

Révész, A., Kourtali, N.-E., & Mazgutova, D. (2017). Effects of task complexity on L2 writing behaviors and linguistic complexity. *Language Learning*, *67*(1), 208–241.

Rienties, B., Lewis, T., McFarlane, R., Nguyen, Q., & Toetenel, L. (2018). Analytics in online and offline language learning environments: The role of learning design to understand student online engagement. *Computer Assisted Language Learning*, *31*(3), 273–293.

Rotelli, D., & Monreale, A. (2022). Time-on-task estimation by data-driven outlier detection based on learning activities. *LAK22: 12th International Learning Analytics and Knowledge Conference*, 336–346.

Schneider, S., Nebel, S., Beege, M., & Rey, G. D. (2018). The autonomy-enhancing effects of choice on cognitive load, motivation and learning with digital media. *Learning and Instruction*, *58*, 161–172.

Shaffer, C. (1989). A comparison of inductive and deductive approaches to teaching foreign languages. *The Modern Language Journal*, *73*(4), 395–403.

Sharma, S. (1995). *Applied multivariate techniques*. John Wiley & Sons, Inc.

Shirvani Boroujeni, M., & Dillenbourg, P. (2018). Discovery and Temporal Analysis of Latent Study Patterns from MOOC Interaction Sequences. *8th International Learning Analytics and Knowledge Conference (LAK18)*, *CONF*.

Siadaty, M., Gašević, D., & Hatala, M. (2016). Measuring the impact of technological scaffolding interventions on micro-level processes of self-regulated workplace learning. *Computers in Human Behavior*, *59*, 469–482.

Sik, K. (2015). Tradition or modernism in grammar teaching: Deductive vs. inductive approaches. *Procedia-Social and Behavioral Sciences*, *197*, 2141–2144.

Silva, L. H. R., & Roehr-Brackin, K. (2016). Perceived learning difficulty and actual performance: Explicit and implicit knowledge of L2 English grammar points among instructed adult learners. *Studies in Second Language Acquisition*, *38*(2), 317–340.

Slavuj, V., Meštrović, A., & Kovačić, B. (2017). Adaptivity in educational systems for language learning: A review. *Computer Assisted Language Learning*, *30*(1–2), 64–90.

Song, M.-K., Lin, F.-C., Ward, S. E., & Fine, J. P. (2013). Composite variables: When and how. *Nursing Research*, *62*(1), 45.

Sottilare, R. A., Graesser, A., Hu, X., & Holden, H. (2013). *Design recommendations for intelligent tutoring systems: Volume 1-learner modeling* (Vol. 1). US Army Research Laboratory.

Stepp-Greany, J. (2002). *Student perceptions on language learning in a technological environment: Implications for the new millennium*.

Susnjak, T., Ramaswami, G. S., & Mathrani, A. (2022). Learning analytics dashboard: A tool for providing actionable insights to learners. *International Journal of Educational Technology in Higher Education*, *19*(1), 12.

Szmrecsanyi, B. (2016). An informationtheoretic approach to assess linguistic complexity. *Complexity, Isolation, and Variation*, *57*, 71.

Tsai, K.-J. (2019). Corpora and dictionaries as learning aids: Inductive versus deductive approaches to constructing vocabulary knowledge. *Computer Assisted Language Learning*, *32*(8), 805–826.

Van Merriënboer, J. J., & Kirschner, P. A. (2017). *Ten steps to complex learning: A systematic approach to four-component instructional design*. Routledge.

Van Merriënboer, J. J., Kirschner, P. A., & Kester, L. (2003). Taking the load off a learner’s mind: Instructional design for complex learning. *Educational Psychologist*, *38*(1), 5–13.

Veletsianos, G., Kimmons, R., Larsen, R., & Rogers, J. (2021). Temporal flexibility, gender, and online learning completion. *Distance Education*, *42*(1), 22–36.

Wainer, H., & Mislevy, R. J. (1990). Item response theory, item calibration, and proficiency estimation. *Computerized Adaptive Testing: A Primer*, *4*, 65–102.

Warschauer, M. (2000). *CALL for the 21st Century. IATEFL and ESADE Conference, 2 July 2000, Barcelona, Spain*.

Wauters, K., Desmet, P., & Van Den Noortgate, W. (2010). Adaptive item-based learning environments based on the item response theory: Possibilities and challenges. *Journal of Computer Assisted Learning*, *26*(6), 549–562.

Winne, P. H. (2017). Learning analytics for self-regulated learning. *Handbook of Learning Analytics*, *754*, 241–249.

Winne, P., & Hadwin, A. (1998). *Studying as self-regulated learning. Metacognition in educational theory and practice, D. HACKER, J. DUNLOSKY, AND A. GRAESSER, Eds*. Mahwah, NJ: Erlbaum.

Youngs, B. L., Prakash, A., & Nugent, R. (2018). Statistically-driven visualizations of student interactions with a French online course video. *Computer Assisted Language Learning*, *31*(3), 206–225.

Zhou, Y., & Wei, M. (2018). Strategies in technology-enhanced language learning. *Studies in Second Language Learning and Teaching*, *8*(2), 471–495.