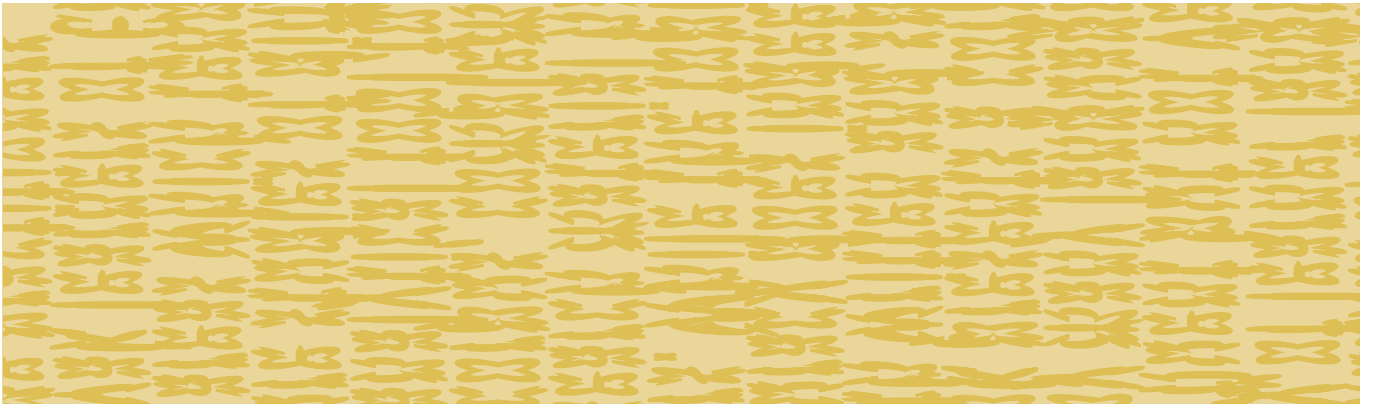
Accuracy Matters
Exploring the Accuracy of Advanced Learner English in Czech Tertiary Education

Simona Kalová

Masaryk University Press

MUNI
PRESS



Accuracy Matters
Exploring the Accuracy of Advanced Learner English in Czech Tertiary Education

Simona Kalová

Masaryk University Press, Brno 2022

KATALOGIZACE V KNIZE – NÁRODNÍ KNIHOVNA ČR

Kalová, Simona, 1965-
Accuracy matters : exploring the accuracy of advanced learner English in Czech tertiary education / Simona Kalová. -- First published. -- Brno : Masaryk University Press, 2022. -- 1 online zdroj. -- (Cizí jazyky a jejich didaktiky : teorie, empirie, praxe ; volume 14)
České resumé
Obsahuje bibliografie, bibliografické odkazy a rejstřík

978-80-280-0335-7 (online ; pdf)

* 811.111 * 378.14 * 81'24 * 81'271.14 * 37.012 * 37.016.026 * (437.3) * (048.8)
– angličtina
– vysokoškolské studium -- Česko
– studium jazyků
– jazyková výuka
– jazykové chyby
– pedagogický výzkum
– předmětová didaktika
– monografie

811.111 - Angličtina [11]

Series: Cizí jazyky a jejich didaktiky: teorie, empirie, praxe
Volume 14

Reviewed by: doc. PhDr. Markéta Malá, Ph.D.
doc. PhDr. Lucie Betáková, M.A., Ph.D.
Mgr. Kristýna Červinková Poesová, Ph.D.

ISBN 978-80-280-0335-7
ISBN 978-80-280-0334-0 (paperback)
https://doi.org/10.5817/CZ.MUNI.M280-0335-2022

Acknowledgements

I would like to express my sincere gratitude to all those people who have guided me in the process of conducting the research described in this book.

First and foremost, my thanks go to doc. Mgr. Martin Adam, Ph.D., for his kind guidance, endless support, and his humour which helped whenever I was losing faith. I would also love to wholeheartedly thank doc. Mgr. Světlana Hanušová, PhD., and prof. PhDr. Věra Janíková, Ph.D., for their enthusiasm and expertise, understanding and encouragement. They will always remain my source of inspiration.

I am indebted to my wonderful colleagues, Chris Rance and David Musk, who have selflessly spent hours with me when analysing students' work and even claimed they were enjoying it. I would like to extend my thanks and appreciation to my colleagues and students from the English department where the research was conducted. Also, I am very grateful to doc. Mgr. Mária Králová, Ph.D. and RNDr. Jiří Jarkovský, Ph.D. for their professional help with statistical analysis.

A very special thanks goes to the reviewers of the manuscript, doc. PhDr. Markéta Malá, Ph.D., doc. PhDr. Lucie Betáková, M.A., Ph.D. and Mgr. Kristýna Červinková Poesová, Ph.D., for their feedback, valuable comments and suggestions.

Last but not least, my most heartfelt gratitude goes to my family, for their patience, tolerance and love with which they supported me, tolerated my grumpy moods and patiently provided advice and loving support whenever necessary.

Table of contents

INTRODUCTION

In the Czech foreign language teaching context, similarly to most former communist countries, the prevailing teaching method used in the early 1990s was the Grammar Translation Method, a form-focused approach with emphasis on correctness. Innovative communicative methodologies with an increased focus on meaning emerging in other countries since the 1980s were not broadly adopted in this country until much later (Hanušová, 2003, p. 17). These days, more communication-oriented teaching methods appear to be widely employed in Czech schools. More opportunities to travel abroad, the availability of foreign language teaching materials, growing numbers of native-speaking teachers, the implementation of digital technologies in education, and overall increased exposure to the target language (TL) might have, at least to a certain extent, also affected learner language, both positively and negatively. It seems that in the 1990s, perhaps as an effect of overwhelmingly adopted focus on form, students were often relatively accurate in their use of grammar and lexis. However, maybe due to lack of opportunities to use the target language in authentic situations, many learners only reached a limited level of fluency; the opposite is often true these days. Grammatical and lexical inaccuracies seem to be a typical feature of lower-level students but also affect, perhaps surprisingly, advanced learner language.

Advanced learners of English as a foreign language tend to be fluent and frequently use complex lexical and grammatical structures, while the accuracy of their performance is frequently problematic (Götz, 2015; 2019; Gráf, 2015; 2017). Some of the typical errors they make are attributed to language transfer, the negative influence of their native language; others are recurrent fossilized errors (Selinker, 1972; Corder, 1981; Ellis & Barkhuizen, 2005; Bestgen et al., 2012; Montrul, 2014). While many other sources of errors can be identified, e.g. false analogies, misconceptions, incomplete rule application, avoidance, omission, or difficulty of target language features (Hendrich, 1988, pp. 367-368;

Ellis & Barkhuizen, 2005, pp. 65-66), it is the native language (L1) that "plays a prominent role in the […] outcomes of second language acquisition" (Montrul, 2014, p. 81).

Aspects of learner language and its development have seen a renewed interest of researchers (Cobb, 2003; MaDonald et al. 2013; Han & Tarone, 2014); studies of second language acquisition focusing primarily at advanced learners and the accuracy of their learner language are, however, often lacking. This book hopes to provide some insights into the analysis of advanced learner English in the Czech tertiary context and explore the possibilities of how the accuracy of learner language can be increased through focus on form.

Accuracy, together with fluency and complexity, creates the three-dimensional model which has been applied in describing learner performance and proficiency in language. It has become increasingly influential in the past few decades and successfully complements the well-established proficiency models. Not only is this model relevant for Second Language Acquisition[1] (SLA) research but it also offers important insights for language practitioners. Both teachers and testers address what seems to be key questions for all: "What makes a second language (L2) learner a proficient language user? And how can L2 proficiency be most adequately (i.e. validly, reliably and feasibly) measured?" (Housen et al., 2012, p. 1). Although, or maybe because, all these terms are used on everyday basis by language practitioners, defining the complexity, accuracy and fluency (CAF) triad is not without challenges.

In both SLA research and foreign language teaching, accuracy is described as correct and appropriate use of language which does not deviate from language norms (Ellis & Barkhuizen, 2005; Housen et al., 2012). Not many areas in second and foreign language teaching have attracted as much attention as the role of errors in the language classroom, and related questions such as error treatment, corrective feedback, focus on form and explicit grammar teaching. Error analysis is not, however, solely an important element of foreign language pedagogy but also a key source of information on the development of learner language, an area explored in second language acquisition research.

1 The term Second Language Acquisition (SLA) has traditionally been used in the literature to describe the acquisition of a second, third and other languages, as well as foreign languages. In this book, the term will therefore be used to describe what is in fact foreign language acquisition – learning and teaching English in the Czech context.

SLA research and second and foreign language[2] teaching have a lot in common, especially their focus on how learners can best be helped to acquire a (second, third, or foreign) language. Despite this, researchers and teachers still do not collaborate as much as they could and there is often a sense of mistrust on both sides (Shimanskaya & Slabakova, 2017, p. 260). Although in the past decades "far more studies have investigated classroom-based language learning […] the question of how research and teaching are related remains a matter of discussion and disagreement" (Lightbown, 2017, p. 105). Even though teachers often feel that researchers fail to produce any findings truly relevant for and applicable in classroom practice or do little more than confirm well-known facts or common-sense knowledge, SLA research has undoubtedly a marked influence on teaching and its positive impact should be recognized.

The benefits of research informing teaching practice can be viewed from two different perspectives. First, research can confirm the existing beliefs teachers hold and give them "more confidence in their pedagogical choices when they see these are reinforced and explained by research evidence" (Lightbown, 2017, p. 105). Second, it can inspire teachers and encourage them to adopt novel approaches: "substantial research evidence that goes against the tide of popular opinion […] may prompt teachers to examine their beliefs and try something different in their classrooms, […] help teachers avoid inflexible pedagogical habits and contribute to their professional growth" (Lightbown, 2017, p. 105). In addition to other aims, this book is also an attempt to bridge the gap between research and classroom practice, by identifying some of the problem areas of learner language which require more focused attention from learners and teachers alike.

The main purpose of the research in the field of Instructed Second Language Acquisition[3] (ISLA) presented in this book is to explore the accuracy of advanced learner language. It intends to determine

2 In the Czech context, English is taught as a *foreign language*, unlike in the countries of the *outer* or *norm-developing circle* (Kachru, 1985; 2009) in which English is often official language, language of instruction or otherwise, and is referred to as a *second language* (Tarone & Swierzbin, 2009). In the literature, the two terms are often used interchangeably. This book, analysing the situation in the Czech Republic, a country of *the expanding* or *norm-dependent circle* (Kachru, 1985; 2009), deals with foreign language teaching and this is the term to be used here.

3 *Instructed Second Language Acquisition* is "a subdomain of second language acquisition (SLA) that emerged in the early 1980s" and explores second and foreign language instruction, together with benefits and drawbacks of L2 acquisition in the language classroom (Sok et al., 2018).

the extent to which accuracy can be influenced by targeted educational intervention aimed at minimising errors and raising learners' awareness of error-prone areas. In the research, a quantitative design was adopted; with one-group pre-test-post-test design used in the pilot study, and quasi-experimental pre-test-post-test control and experimental group design used in the main study. The participants in all groups were undergraduate university students of English philology. The pilot and experimental groups received educational intervention in the form of a 13-week blended learning course[4] focused on the most problematic areas affecting the accuracy of learner language; the control group did not receive any intervention. All participants in the three groups were tested for accuracy and asked to express their certainty in answering. After the pre-test, the pilot and experimental groups attended the course as part of the intervention, while the control group did not. Three sets of data collected from all participants were analysed: Grammaticality Judgement Tests[5] (GJT), Certainty-Based Marking[6] (CBM) and samples of their spoken and written production. To collect these samples, students in the experimental and control groups, in addition to taking GJT and CBM, were also asked to write a short essay and participate in oral interviews which were video recorded. After the intervention, all groups were tested again, and the efficacy of the intervention was evaluated by comparing the scores in pre-test and post-test for all groups. The data collected from spoken and written production were used to compile a corpus of learner data which was analysed to provide insights into advanced learner language.

The book consists of two main parts, theoretical and empirical. The former includes chapter 1, in which theoretical dimensions of the research are laid out, and chapter 2, with an overview of the relevant

4 *Blended learning* refers to the combination of contact classes and online study activities enhanced with the use of a variety of forms of digital technology; learners can engage in these activities both in the classroom and individually via the internet, interactive whiteboards, or some other form of digital technology. Blended learning successfully complements face-to-face learning and teaching, making use of online communication, both synchronous and asynchronous. (Sharma, 2007)

5 *Grammaticality Judgement Test or Task* is "a task in which speakers of a language are presented with linguistic stimuli (typically sentences) and asked to judge whether they are correct in the language. Such tasks are widely used in linguistic theory to formulate and refine claims about a speaker-hearer's internal grammar or competence." (Richards & Schmidt, 2010, p. 254)

6 *Certainty-Based Marking* is a Moodle based testing tool which requires that respondents, in addition to answering test questions, also express how certain they are about the correctness of their answers. This is believed to eliminate guessing in answering, boost confidence of respondents and enhance reliability of test results (Gardner-Medwin & Curtin, 2007).

literature and current research related to the topic. The empirical part is composed of three chapters. In chapter 3, educational intervention in the form of a blended learning course is detailed. Research methodology, aims, questions, design, participants, tools, data collection and analysis used in this research are outlined in chapter 4, while in chapter 5 research results and analyses are presented. In chapter 6, the answers to research questions are addressed and the findings are compared with and contrasted to similar Czech and international studies. In the final section, the findings of the research are summarised and conclusions of the research are drawn. Pedagogical implications for foreign language teaching are also discussed, together with the limitations of the study and suggestions for further research.

List of abbreviations

CA	contrastive analysis
CAE	Certificate in Advanced English,
CAF	complexity, accuracy and fluency
CBM	Certainty-Based Marking
CC	communicative competence
CEA	computer-aided error analysis
CEFR	Common European Framework of Reference for Languages
CIA	Contrastive Interlanguage Analysis
EA	error analysis
EFL	English as a foreign language
EIL	English as an International Language
EK	explicit knowledge
ELF	English as a Lingua Franca
ENL	English as a native language
ESL	English as a second language
ELT	English Language Teaching
EU	European Union
FFI	form-focused instruction
FLLOC	French Learner Language Oral Corpora
FLT	Foreign Language Teaching
FRIDA	French Interlanguage Database
F2F	face-to-face
GJT	Grammaticality Judgement Test/Task
ICLE	International Corpus of Learner English
IELTS	International English Language System
IK	implicit knowledge
IJLCR	International Journal of Learner Corpus Research
IL	Interlanguage
ISLA	Instructed Second Language Acquisition

LCR	Learner Corpus Research
LINDSEI	Louvain International Database of Spoken English Interlanguage
LC	learner corpus/corpora
LL	learner language
L1	first language
L2	second language
ME	metalinguistic explanation
MFI	meaning-focused instruction
NA	not applicable
NL	native language
NS	native speaker
NNS	non-native speaker
RQ	research question
SLA	Second Language Acquisition
SLT	Second Language Teaching
TL	target language
VOICE	Vienna-Oxford Corpus of English

1 THEORETICAL FRAMEWORK

In the first chapter, an outline of the underlying theoretical concepts and theories influencing second language acquisition and second and foreign language teaching and learning addressed in the research will be presented and discussed. Also, key terms used throughout the book will be defined.

1.1 Learner language

The main aim of this book is to explore advanced *learner language* and how its development, especially in terms of accuracy, can be enhanced. This key theoretical concept and its exact meaning as discussed in the book is therefore the first to be clarified.

As Ellis and Barkhuizen suggest, "learner language is the oral and written language produced by learners" (Ellis & Barkhuizen, 2005, p. 5). Analysing learner language provides both teachers and researchers in SLA with invaluable information about how the target language is acquired, serving as "the primary data for the study of L2 acquisition" (ibid). It should also be taken into account that "learner language is not a monolithic phenomenon but rather highly variable" (ibid). Learner language reflects learners' current level of linguistic knowledge of which errors and inappropriate usage are inevitable and typical features. Examining learner language and its characteristics, its development through instruction, as well as individual differences among learners establish the framework of instructed second language acquisition research to which the research project detailed in this book belongs.

Learner language is related to *interlanguage*, a term coined by Selinker in his eponymous seminal paper (1972), based on earlier concepts, especially Corder's *idiosyncratic dialects* and *transitional*

competence (1967). In what has become known as *the Interlanguage Hypothesis*, Selinker explains the learning process from a psycholinguistic perspective. In his view, interlanguage is characterized as "a separate linguistic system based on the observable output which results from a learner's attempted production of a TL norm" (1972, p. 214). Apart from identifying the existence of three independent autonomous linguistic systems: native language (NL), target language (TL), and interlanguage (IL), he also describes phenomena characterising interlanguage, especially overgeneralization, language transfer and transfer of training, fossilization and communication strategies. According to Selinker, foreign language speakers on their way from L1 to L2 "create interlanguage when attempting to express meanings in a second language" (2014, p. 223); so rather than one universal interlanguage, there are individual interlanguages characterized as "non-native languages which are created and spoken whenever there is language contact" (ibid).

In the current literature, the two terms – learner language and interlanguage − seem to be often used interchangeably. In the thriving area of corpus-based research, terms like *learner corpus, learner data, learner complexity*, and *learner fluency*, all derived from *learner language* seem to prevail; other researchers analysing learners' production in the target language, however, prefer to use *interlanguage development, spoken* and *written interlanguage*. This inconsistency is also apparent in the names of language corpora, regardless of the target language; with e.g. FLLOC − French Learner Language Oral Corpora[7], and ICLE – International Corpus of Learner English[8], on the one hand, and FRIDA – French Interlanguage Database[9], and LINDSEI – Louvain International Database of Spoken English Interlanguage[10] on the other. Selinker comments on this inconsistent usage of the two terms, saying that "there are puzzling questions about different types of interlanguage, with too many equating 'learner language' with interlanguage" (2014, p. 229). In order to avoid ambiguity, the terms *learner language* and *learner English* will be used in this book to indicate learners'

7 http://www.flloc.soton.ac.uk

8 http://www.uclouvain.be/en-cecl-icle.html

9 http://www.uclouvain.be/en-cecl-frida.html

10 http://www.uclouvain.be/en-cecl-lindsei.html

production in the target language, while *interlanguage* will be used to refer to the underlying linguistic system or linguistic competence.

Rather than presenting a theory in its own right, Selinker in his Interlanguage Hypothesis raised a number of questions regarding processes occurring in L2 acquisition and addressed topics that "have continued to resurface for the last 40 years in research on second-language acquisition" (Han & Tarone, 2014, p. 9), and still resonate in current empirical studies. These are in particular questions about how native language influences the acquisition of the target language, why some areas of learner language tend to fossilize, and what data should be elicited from learners to provide reliable information on their interlanguage. In the following section, theoretical principles guiding the choice of data to be analysed when exploring learner language will be addressed.

1.2 Implicit and explicit knowledge in language learning and acquisition

It has been widely acknowledged that "there are two kinds of language knowledge at work in the mind of the adult L2 learner", *explicit* and *implicit* (Han & Tarone, 2014, p. 14). An important distinction must be drawn between how these two kinds of knowledge are reflected in (language) learning. Slightly simplifying the difference, it could be asserted that while *explicit learning* is conscious, accessible through controlled processing, and involves the use of working memory, *implicit learning* is unconscious, unintentional, available through automatic processing and cannot be reported about by learners (Ellis et al., 2009, pp. 14-17). The debate about the existence of the two systems, and especially the ways in which they interact, is not without controversy in SLA research. "Just about all theories of L2 acquisition acknowledge the distinction between implicit and explicit knowledge" (Ellis & Shintani, 2014, p. 14). This concept of two aspects of knowledge and learning was first established in the area of cognitive psychology (Ellis et al., 2009), and has been reflected both in research into SLA, and language pedagogy: "Running throughout the history of SLA has been a concern for the role that explicit L2 knowledge plays in learning" (Ellis & Shintani, 2014, p. 14).

The issue of what kind of learner data should be examined to inform us about learners' underlying linguistic knowledge, one of the questions posed by Selinker, was revisited in research more than forty years later (Han & Tarone, 2014, p. 14). In his original paper, Selinker (1972) claimed that only data gained from learners' free spoken production when focused on meaning are relevant, and dismissed those data elicited in drills, Grammaticality Judgement Tests and any other production focused on accuracy and form rather than meaning. The question Han explores is whether learners' free production "when focused on meaning is quite different from that produced when […] focused on formal accuracy" (Han & Tarone, 2014, p. 14). In order to be able to address this question, we must first focus on what processes are related to language learning and acquisition.

As Rod Ellis et al. (2009) reported in their overview of research into implicit and explicit knowledge, a number of studies have been conducted to assess how effective each of the two is in second language learning. Such research is, however, not without problems; the main challenge is how exactly explicit and implicit knowledge should be operationalised and measured. This may be one of the reasons why the studies have often generated conflicting outcomes. While no convincing results have been demonstrated about implicit learning, it appears, according to many studies (Ellis, 1993; Rosa & O'Neill, 1999; Gass & Mackey, 2002) that "explicit learning is more effective than implicit", especially when not very complex grammar structures are concerned (Ellis et al., 2009, pp. 10-11). Some other studies (for example Doughty, 1991; Shook, 1994), however, reported "no difference between implicit and explicit learning" (Ellis et al., 2009, p. 10). One of the reasons for these opposing outcomes might have been that the studies were only short-term and adopted tests, e.g. Grammaticality Judgement Tests (GJT), favouring explicit learning (ibid). The main differences between the concepts of implicit and explicit knowledge, summarised by Ellis and Shintani (2014, p. 13), are shown in Table 1 below.

The table demonstrates that while implicit knowledge is tacit and intuitive, explicit knowledge is conscious and aware. Similarly to our knowledge of L1, unconscious and proceduralised implicit knowledge is fully internalised and "available for automatic use in spontaneous production" (Ellis & Barkhuizen, 2005, p. 5). Explicit knowledge, on the other hand, is declarative and metalingual; learners can comment on their

use of language with or without using metalanguage, applying "lexical knowledge of technical and non-technical linguistic terminology" (Ellis & Barkhuizen, 2005, p. 5). Implicit knowledge is formulaic and rule-based, drawing on the unconscious knowledge of "stored and ready-made chunks […] realizable lexically in an indefinite number of sentences" (Ellis & Barkhuizen, 2005, p. 5). On the contrary, explicit knowledge is often imprecise and inaccurate, and improves "as proficiency increases" (Ellis et al., 2009, p. 15). In relation to age, while there seem to be "age constraints on the ability of learners to fully learn an L2 implicitly" (Ellis et al., 2009, p. 14), explicit knowledge appears to be learnable without any age limit. Some researchers claim, however, that contrary to common belief, gaining implicit knowledge of the language is possible even past the 'critical period' (Ellis & Shintani, 2014, p. 12). The fact which is vital for research is that implicit knowledge is not open to direct inspection and the only thing that can be examined is learners' verbal behaviour.

Table 1
Key characteristics of implicit and explicit knowledge (Ellis & Shintani, 2014, p. 13, shortened)

Characteristics	Implicit knowledge (IK)	Explicit knowledge (EK)
Awareness	learner has no conscious awareness of linguistic norms but does intuitively know what is correct	learner is consciously aware of linguistic norms
Type of knowledge	IK is 'procedural'; available for automatic processing	EK is 'declarative'; consists of 'facts' about language that are only available through controlled processing
Systematicity	IK is variable but systematic	EK is often anomalous and inconsistent as learners may have only a partial understanding of a linguistic feature
Use of L2 knowledge	IK is only evident when learners use it in communication	EK is used to monitor L2 production; used when learners lack the requisite implicit knowledge
Self-report	IK cannot be directly reported	EK can be reported; reporting requires access to metalanguage
Learnability	there may be age limits on learners' ability to acquire IK ('critical period')	EK is learnable at any age

A question which is fundamental in both SLA research and language pedagogy is how the two aspects of learner knowledge are related and how they interact. While some propose that the systems underlying explicit and implicit knowledge operate independently and even "reside in neuro-anatomically distinct systems" (Ellis et al., 2009, p. 16), others argue that there is just "a single knowledge source" (Ellis et al., 2009, p. 10). These conflicting views on understanding processes of learning have driven SLA research in the last three decades (Ellis & Shintani, 2014, pp. 11-13). The former approach, held especially by Krashen (1989), has been among the most influential and will be outlined in more detail.

In his Monitor Theory, Krashen distinguishes between the process of conscious *learning*, often as part of formal instruction in the classroom, and *acquisition* as an unconscious process, similar to how a mother tongue is acquired in childhood, occurring without any conscious effort from the users of language. The Monitor, as understood by Krashen, has a role in checking learners' output in L2 by consciously applying the learned rules and affecting what is being produced either in speech or in writing. There are, of course, individual differences among learners in what ways and to what extent they apply the Monitor. A number of mutually interdependent factors are at play here[11]: the time available to learners when they plan their performance; the linguistic knowledge they possess; how much their performance is focused on form; their language aptitude, i.e. predispositions to learn a language; and individual personality factors, such as personality types, anxiety levels, motivation, and the overall attitude to the culture of the target language (Krashen, 2002, pp. 12-39).

While very influential, Krashen's hypotheses have also inspired debate among researchers and have been subjected to criticism by many (Gregg, 1984; McLaughlin, 1987; Horner, 1987; Schmidt, 1990; Zafar, 2011). One of the criticised aspects of these hypotheses was applying Chomskyan principles of first language acquisition to the processes of SLA in adult speakers of L2, an approach perceived as overly simplistic, ignoring other factors, e.g. the influence of critical period on second language acquisition. Rather fuzzy definitions of the key concepts, e.g. learning and acquisition, explicit and implicit knowledge,

11 The following list is far from exhaustive, as this is not the main focus of the present book.

comprehensible input, or affective filter made it almost impossible to test the hypotheses empirically (Zafar, 2011, pp. 140-145). The fact that there was not enough evidence from research to support them has been regarded as a major flaw of Krashen's theories (Zafar, 2011, p. 141). Also, the hypotheses failed to take into account many important factors influencing SLA, e.g. both positive and negative influence of the mother tongue on L2 acquisition. It was equally unable to explain how some adult L2 learners manage to achieve native-like competence in using most features in the target language, while failing to acquire others (Zafar, 2011, p. 144). Alternative hypotheses aiming at explaining language learning and acquisition have appeared and some of them took a completely opposing view of the processes, e.g. Schmidt's Noticing Hypothesis (1990). According to this hypothesis, no language features are acquired without being noticed first, in other words without conscious targeted attention paid to them (Schmidt, 1990); a claim which is in direct opposition to those expressed by Krashen. Despite many controversies, however, Krashen's hypotheses have been of immense importance as they started an important discussion among researchers regarding the processes of SLA.

In what has become known as the non-interface position, Krashen hypothesises that the two systems, learning and acquisition, operate independently in adult learners' minds and "are interrelated in a definite way: subconscious acquisition appears to be far more important" (Krashen, 2002, p. 1). Other researchers (for example DeKeyser, 1995; Ellis, 2005), believe that the two systems do interact, supporting the strong and weak interface positions, with varying degrees of overlap between explicit and implicit knowledge and learning (Ellis & Shintani, 2014, pp. 11-13).

Considering the distinction from the learning and teaching perspectives, most of the above mentioned views agree on the fact that explicit and implicit "learning processes are correlated to some degree at least" and "interact at the level of performance" (Ellis et al., 2009, p. 17). Similarly, it has become a commonly accepted view that "a learner's *implicit knowledge* (*competence*) is not open to direct inspection […], thus, by large, researchers are forced to infer competence from some kind of *performance*" (Ellis & Barkhuizen, 2005, p. 6). This is related to an important question raised by researchers: what kind of learner data should be collected for analysis.

Unlike Selinker's original standpoint that the only data that should be analysed "are utterances produced by second language learners when they are trying to communicate meaning in the target language in unrehearsed situations" (Tarone, 2013, p. 4) and all data gathered in the classroom practice and/or from learners' reports about their learning should be ignored (ibid); other researchers disagree. Both Selinker's contemporary Corder (1981) and more recently Ellis and Barkhuizen (2005) claim that in order to get a complex picture of learner language, we need a variety of data. These could be elicited from learners as samples of both their, ideally free, spoken and written production in the target language, together with their intuitions about language which can be measured, e.g. by Grammaticality Judgement Tests. Both datasets were therefore collected for this research study. A detailed account of what research tools were adopted in this study and how the data were collected and analysed are provided in chapter 4.

1.3 Competence, performance and proficiency

Drawing on earlier theories and empirical research, Chomsky (1965) was the first to distinguish between *competence* and *performance*. He defines them as two opposing terms characterising the unconscious knowledge of language – competence, and how language is actually used in everyday communication – performance (Brown et al., 1996, p. 2). Chomsky understood competence from a cognitive perspective, focusing almost entirely on linguistic competence.

It is the primary aim of SLA research to provide a "description and explanation of L2 learners' *competence* and how this develops over time" (Ellis & Barkhuizen, 2005, p. 5), as well as understand "the relationship between competence and *performance*" (ibid). A number of competence models have evolved over the years. Despite different approaches to what aspects of competence should be reflected in them, all these models always involve "underlying systems of linguistic knowledge" (Ellis & Barkhuizen, 2005, p. 5). As the possibilities to directly explore competence are still rather limited, although functional brain imaging studies in SLA research have been on the rise (Chee et al., 1999; Ullman, 2004), exploring learners' performance in the target language is still the core of SLA research.

Learners' production in L2 is an area of interest shared by SLA researchers and language teachers. There is a difference, however, in "different conceptualizations of the products of L2 acquisition" (Ellis & Barkhuizen, 2005, p. 362). Richards and Schmidt (2010) define competence not just as "the implicit system of rules that constitutes a person's knowledge of a language" but they also refer to "a person's ability to create and understand sentences, including sentences they have never heard before […], and the ability to recognize ambiguous and deviant sentences" (p. 103). Performance is perceived as "a person's actual use of the language" and is often investigated as an indirect indication of the learner's competence (p. 428).

While competence and performance are terms favoured in SLA research, especially when conducted from the psycholinguistic perspective, language practitioners, both teachers and experts developing testing and teaching materials tend to speak about *language proficiency* rather than performance (Ellis & Barkhuizen, 2005, p. 362). This concept can be defined as "the degree of skill with which a person can use a language, such as how well a person can read, write, speak or understand language" (Richards & Schmidt, 2010, p. 321). The skill as encompassed in the definition is measurable through tests.

Taylor (1988) points to the often interchangeable use of the two terms, and distinguishes between "what a speaker knows and what he does" (1988, p. 166), where learner's knowledge as a static concept is competence, and proficiency as a dynamic concept is "the ability to use competence" (ibid). In other words, competence is knowing the language, and proficiency is being able to use it.

In this book, written within the framework of instructed second language acquisition, all three terms – competence, performance and proficiency − will be used. The terms competence and performance will be applied when referring to learner language analysis from the linguistic viewpoint, while proficiency will be used when the teaching perspective prevails.

1.3.1 Communicative competence

It has become widely accepted that the main goal of language learning and teaching is the development of *communicative competence* in the target language (Larsen-Freeman & Anderson, 2011, p. 115; Richards

& Rodgers, 2014, pp. 83-87). This does not merely mean being able to use the language correctly from the linguistic viewpoint, but also using it appropriately and adequately in accordance with the situation and in interaction with other, both native and non-native, users of the language. In other words, apart from the linguistic aspect of communication, its social, pragmatic, cultural and intercultural aspects, as well as their interaction, have to be accounted for. Over the years, a number of models attempting to provide a concise description of communicative competence have been developed. Some of the most influential models are briefly introduced and discussed below.

When defining competence and performance, Chomsky (1965) relied on an abstract ideal user of language, possessing a perfect linguistic knowledge without any constraints from unfavourable conditions. This approach had soon sparked criticism from other theoreticians of language. In reaction to Chomsky, Hymes (1972) acknowledged the fact that language should not be studied in isolation but rather as a means of communication within society. *Communicative competence*, the term he coined in 1972, has been in use ever since (Richards & Rodgers, 2014, p. 87). Hymes broadened Chomsky's understanding of purely linguistic competence accounting for the importance of context and the need of appropriate use of language in a variety of social contexts.

In foreign language teaching, the works of two theoreticians, Halliday and Widdowson, were of utmost importance. Halliday (1977), whose theory of language was frequently reflected in communicative language teaching, looked at communicative competence from a functional perspective, specifying basic functions of language. Widdowson (1978) attempted to bridge the gap between researchers of language and language practitioners by trying to provide teachers of language with clear guidelines on how the then new communicative theory of language should be incorporated in everyday teaching practice and reflected in the teaching materials produced.

In the 1980s, Canale and Swain (1980) and Canale (1983) elaborated on Hymes's conception and developed what is now regarded as the first model of communicative competence. It consists of four key competences: it adds discourse and strategic competences to the previously described grammatical (linguistic) and sociolinguistic competences. *Grammatical competence*, including the linguistic knowledge of grammar structures and lexis, as well as the knowledge of rules governing word order,

pronunciation and orthography, refers to language as code. *Sociolinguistic competence* stresses the importance of social factors in communication, especially the awareness of using the language appropriately depending on the social context of communication. *Discourse competence* is the ability to produce and understand language through coherent and cohesive utterances. It indicates that the speaker is able to understand how discourse or spoken and written texts, are organized beyond the level of sentences. *Strategic competence* has to do with how effective communication can be best achieved. It is compensatory in nature, i.e. it is only activated when other competences cannot be applied (Skehan, 1998) and "involves the knowledge of how to use verbal and non-verbal communication strategies to handle breakdowns in communication" (Usó-Juan & Martínez-Flor, 2006, p. 11).

While very influential, Canale and Swain's model was not without flaws, especially in that it failed to address how individual competences interact; this was addressed by Savignon (1983) whose model attempted to illustrate how the interaction between the four competences causes improvement of the whole communicative competence, when only one competence improves (Usó-Juan & Martínez-Flor, 2006, p. 11).

Building upon Canale and Swain (1980), Bachman (1990) and Bachman and Palmer (1996) proposed a more thorough model. Bachman uses the term *communicative language ability* and defines it as "consisting of both knowledge, or competence, and the capacity for implementing, or executing the competence in appropriate, contextualized communicative language use" (Bachman, 1990, p. 84). Apart from a detailed description of the existing competences, he subdivided language competence into *organizational*: including *grammatical* and *textual competences*, and *pragmatic competence,* consisting of *illocutionary* and *sociolinguistic competences* (Bachman, 1990, p. 87). In Bachman's model, *strategic competence* was viewed as being "central to all communication" (Skehan, 1998, p. 161). Communication is perceived as a dynamic process in which learner's active approach and use of communication strategies and the role of context are recognised (Bachman, 1990, p. 98).

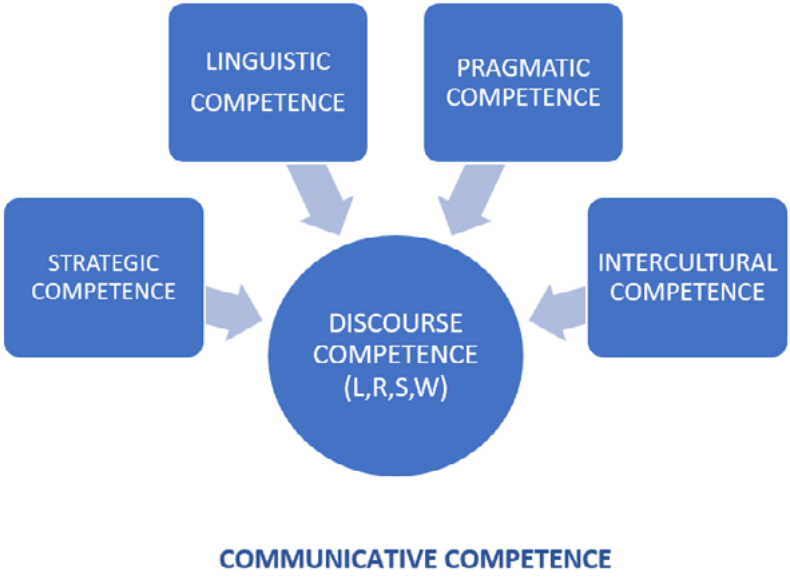
The models of communicative competence described above, as well as many others, were revisited by Usó-Juan and Martínez-Flor (2006, p. 14). In order to cater for cultural aspects fostering communication, they added *intercultural communicative competence* to the existing

models (2006, p. 12). Having critically evaluated all existing models, they proposed a new model, based on the four skills. This model seems to successfully reflect everyday realities of the language classroom. It regards "discourse as the key competence with the rest of the competencies (i.e. linguistic, pragmatic, intercultural and strategic) shaping it" (Usó-Juan & Martínez-Flor, 2008, p. 168). Incorporated in this model are the four skills, reading, writing, listening and speaking, centrally located in discourse competence, making it clear that "all components cannot be developed in isolation […] an increase in one component interacts with the other components to produce an increase in the whole construct of communicative competence" (Usó-Juan & Martínez-Flor, 2006, p. 16). This integrative model, illustrated in Figure 1 below, seems to be the most appropriate for this research project, as it is well suited for both SLA and classroom driven research.

While communicative competence should be perceived as an interaction of all its components which are considered as equally important, it is beyond the scope of this research to explore them all. The current study will focus on *linguistic competence*; other aspects, pragmatic, strategic or intercultural, will not be analysed. Linguistic competence will be narrowed down to aspects concerning grammar and lexis; phonological aspects will not be included in the analysis. This decision is supported by the fact that not only spoken but also written language is analysed, and in order to enable comparisons, the same criteria should be applied in the analysis.

The main focus of the research presented in this book is exploring learner language, with particular focus on its accuracy. Examining how learners perform tasks in the target language and analysing their production, both spoken and written, in terms of complexity, accuracy and fluency (CAF) is believed to provide a complex picture of learner language (Ellis & Barkhuizen, 2005, p. 134; Housen et al., 2012, p. 1; Ellis & Shintani, 2014, p. 148). Adopting this approach to learner language analysis has become widely accepted in research and appears to be a legitimate choice for this project. Complexity, accuracy and fluency as aspects of L2 performance are detailed in the next subchapter.

Figure 1
Communicative competence model as proposed by Usó-Juan & Martínez-Flor, adapted (2006, p. 161)



1.4 The CAF model

Describing learner performance and proficiency in language in terms of complexity, accuracy and fluency (CAF) has become increasingly influential in the past three decades. Understanding the three concepts commonly used by teachers, remains, however, rather fuzzy. Despite this, using the CAF model for learner language description enables researchers to reflect the multifaceted nature of L2 proficiency. It has been adopted as "a notable complement to other established proficiency models such as the traditional four-skills model and sociolinguistic and cognitive models of L2 proficiency" (Housen et al., 2012, p. 1). By analysing accuracy, fluency and complexity, "a broader and more balanced picture of learner language" is provided (Ellis & Barkhuizen, 2005, p. 139).

As both researchers and teachers agree, "L2 proficiency is not a unitary construct but, rather […] multicomponential in nature" (Housen et al., 2012, p. 1). This construct was originally viewed

as two-dimensional, taking into account *accuracy* and *fluency* of learner language (Hartmann & Stork, 1976; Brumfit & Johnson, 1979; Fillmore et al., 1979). In the 1990s, a third dimension – *complexity* – was added by Skehan (1996), creating a 3-dimensional model well-suited for descriptions of learner language. When trying to describe accuracy and fluency, Brumfit (1984) was aware of the pitfalls of creating clear and concise definitions. At the same time, he considered understanding the polarity of the two as vital for a better understanding of the processes involved in second language acquisition (p. 52).

There is relatively little disagreement among researchers when defining *accuracy*. In the narrow sense, it is traditionally perceived as "the ability to produce target-like and error-free language" (Housen et al., 2012, p. 2). Accuracy or correctness, or rather the lack of it, is defined as a deviation from a certain standard form of the target language. Straightforward as it may sound, it is still a very complex notion, particularly when exact definitions of *error* and *standard* should be formulated. It is in particular the question of standard form or norm in ELT that is very complex, especially when considering the changing roles of the English language in the world. These notions will be addressed in more detail in the following text, in subchapter 1.5.

Defining the two remaining concepts – fluency and complexity – is not without difficulties. There have been different approaches to defining *fluency*; e.g. Fillmore (1979) stresses the speed with which learners are able to produce an utterance, as well as coherence and semantically rich expression, appropriateness to context, and creativity (Fillmore et al., 1979). Definitions of fluency tend to be very broad, and often refer to the overall ability of learners to use the language naturally and effectively in both written and spoken communication, to use it with ease, be close to native-like performance but not necessarily without errors (Richards & Schmidt, 2010, p. 223). These days, a more focused approach to fluency has been adopted in SLA research, defining fluency as "mainly a phonological phenomenon" in which three subcategories are distinguished: speed fluency or rate of speech, breakdown fluency – pauses, and repair fluency, self-corrections and false starts (Housen et al., 2012, pp. 4-5).

Complexity refers to the use of sophisticated lexis and advanced grammar structures. This requires a certain level of risk-taking from the learner in using the "language that is at the upper limit of their interlanguage systems, and thus not fully automated" (Ellis & Barkhuizen,

2005, p. 139). It is often perceived as both *linguistic complexity*, i.e. the use of L2 forms and meanings, and *cognitive complexity*, "determined in part by the learners' individual backgrounds (e.g. their aptitude, motivation, stage of L2 development, L1 background)" (Housen et al., 2012, p. 4).

Another challenge is to determine how the three aspects of learners' proficiency in L2, CAF, are interrelated and to what extent they interact. Some researchers, (e.g. Meisel et al., 1981; VanPatten, 1990) believe that it is difficult for learners to be equally attentive to all three simultaneously, and it can be assumed that they consciously or unconsciously decide "to prioritize one aspect of the L2 over another" (Ellis & Barkhuizen, 2005, p. 140). This might be influenced to a certain extent by the nature of the task, whether it is focused more on meaning or form, "meaning is reflected in fluency, while form is manifested in either accuracy (if control is prioritized) or complexity" (Ellis & Barkhuizen, 2005, p. 142). Especially researchers taking the psycholinguistic perspective, (e.g. Skehan, 1996; Foster & Skehan, 1996; Bygate, 1999), an area addressed in more detail above in part 1.2, believe that limited attention span and working memory capacity cause fluency to "[compete] for attentional resources with accuracy, while accuracy in turn competes with complexity" (Housen et al., 2012, p. 7). Others, (for example Robinson, 2001; 2003), however, disagree, claiming that the development of complexity, accuracy and fluency is linear and "all three components may in principle jointly increase or decrease in L2 performance" (Housen et al., 2012, p. 7). Individual differences among learners also play a role: while some of them might prefer taking risks and use complex structures they might not have fully mastered yet, others prefer to strictly adhere to familiar grammar and lexis in order to avoid errors.

This section has attempted to describe how analysing learner language in terms of complexity, accuracy and fluency has become increasingly popular in SLA research, with results that are also relevant for language learning and teaching. Such research, however, is not without challenges. In order to eliminate inconsistencies in studies adopting the CAF triad, to ensure their comparability and facilitate their replication, careful attention must be paid to how these three components are operationalised; it is equally important to consider their interdependence as identified in SLA research (Housen et al., 2012, pp. 3-10). As this book focuses on exploring the accuracy of learner

language, different ways of measuring this aspect of learner performance will now be described. Also, an explanation of how accuracy will be operationalised for this research will be provided.

1.4.1 Measuring and operationalising accuracy

The CAF model can be adopted for analysing both written and spoken language. While in measuring accuracy and complexity, identical measures can be applied for either spoken or written production, different measures need to be applied when assessing fluency. A variety of methods have been applied in research, "ranging from holistic and subjective ratings to objective quantitative measures of L2 production" (Housen et al., 2012, p. 8).

Over the years, different approaches to measuring grammatical and lexical accuracy have been employed, using both specific and general measures of accuracy. Some of the studies in accuracy focused on concrete aspects of the target language, especially those which are believed to be problematic for learners in general and also for speakers of a particular L1 (Wierszycka, 2013; Hamed, 2014; Sun, 2014; Long & Hatcho, 2018). Such studies investigated for example the use of articles, plural forms, noun-verb agreement, verb tenses and erroneous use of lexis, especially if caused by the influence of learners' native languages. These studies will be presented in more detail in the review of the literature, chapter 2.

Another approach adopted in the analysis of learners' production in L2 is applying general measures of accuracy, taking into account numbers of self-corrections, the percentage of error-free clauses and errors per 100 words (Ellis & Barkhuizen, 2005, p. 150). While early L2 research was mostly in favour of specific measures, there has been a growing tendency to adopt "general measures, either because they provide a more comprehensive picture of performance in each of the CAF areas or because they seem to be more sensitive in discriminating between broad proficiency levels or at detecting treatment effects between groups" (Housen et al., 2012, p. 8).

It is for the reasons mentioned above that a global measure of accuracy has been adopted in this research project in which the effect of educational intervention is assessed. While using a general measure that counts percentage of error-free clauses poses another challenge in how a clause should be defined, using the number of errors

per 100 words seems to solve the problem (Ellis & Barkhuizen, 2005, p. 151). Despite the relative straightforwardness of such measure of accuracy, there were still quite a few decisions to be taken before the analysis could be completed. These will be addressed in chapter 4, where a detailed account of how the research was conducted is provided. In the next section, two terms closely related to learner language analysis, language norm and, most importantly, error, will be addressed.

1.5 Norm in foreign language teaching

Error in language use is typically defined as "a certain undesirable deviation from language norm […] reliably distinguishing a native speaker from a non-native speaker"[12] (Hrdlička, 2012, p. 89). It is, therefore, essential to make it clear what a norm is, how it should be specified and what kind of norm should be applied in foreign language teaching.

Defining language norms is a very complex endeavour which requires many factors, not purely linguistic, but also historical, social and even political, to be considered. It is even more complicated when the language for which the norm should be defined is English which is, unlike other languages and due to a variety of reasons, in an unprecedented situation. It is developed both by its native speakers (NS) and non-native speakers (NNS), used in communication between NS and NNS, but also, increasingly, as a lingua franca to enable understanding between non-native speakers of English with different native languages (Kachru, 1986; Crystal, 2003; Jenkins, 2007; Seidlhofer, 2009). As a consequence, when looking for the most appropriate standard to be adopted in English language teaching, there is a growing tendency to abandon native-speaker norms and look for new models of the target language.

Some of the experts, for example (Bamgbose, 1998; Crystal, 2003; Seidlhofer, 2004; Jenkins, 2007; Seidlhofer & Berns, 2009; Seidlhofer & Widdowson, 2017; Medgyes, 2019), have argued that using native-speaker models as the only acceptable norm in ELT is no longer desirable and other options should be considered: "Traditionally, native speakers of English have been regarded as providing the authoritative standard and the best teachers. Now, they may be seen as presenting an obstacle to

12 All Czech quotations have been translated by the author of the book.

the free development of global English" (Graddol, 1997, p. 114). There have been growing tendencies to look for alternative models to be adopted in foreign language teaching. The supporters of English as a lingua franca (ELF)[13], for example, call for the creation of a new paradigm in which ELF might exist alongside the other now respected models, as one of the appropriate options in the future (Jenkins, 2007, pp. 19-28). While this attitude has been gaining support, it is hardly possible to adopt ELF as a new model for language teaching since as yet it has not been fully described and codified. It is first necessary to provide "a description of salient features of English as a lingua franca (ELF), alongside English as a native language (ENL)" (Seidlhofer, 2004, p. 209). In this respect, compiling large corpora of non-native English might provide the data necessary for such a description. This has already started with large corpora, such as "English as an International Language (EIL), English as a Lingua Franca (ELF)" (Granger et al., 2002, p. 29) and especially the Vienna-Oxford Corpus of English (VOICE)[14] being compiled. Before the corpus of English as a lingua franca has been completed, however, the problem remains unresolved and we are faced with what Seidlhofer calls "a conceptual gap" (2004).

Despite all the controversies related to applying native-speaker norms in foreign language teaching, a situation aptly described as being "torn between the norms" (Bamgbose, 1998), it is still mostly regarded today as the preferred model in foreign language teaching. It seems that both teachers and learners also prefer native-speaker standards to be applied in ELT. This has been supported by the results of an international survey study carried out by Timmis (2002) in which he explored teachers' and learners' attitudes to native-speaker norms as opposed to international English. Almost 600 responses from students of English as a foreign language, and native and non-native teachers of English from 45 countries were analysed. Overall, the results indicated that 68% of students preferred to conform to native-speaker models in grammar and 67% in pronunciation. For teachers, the responses showed a bigger tendency to move away from native-speaker norms but still the majority was in favour of this model in grammar, with 54% preferring to adhere to native-speaker norm. In pronunciation, however, 39% of teachers

13 ELF is now increasingly used and has replaced the older terms, e.g. *English as an International Language – EIL, Global Language* (Crystal, 2003) or *Global English* (Jenkins, 2007, pp. 3-4).

14 https://www.univie.ac.at/voice/

chose a NS norm for their students, 27% preferred a non-native speaker model, and 34% showed no preference. It seems, however, that they were choosing this option as being "the more realistic rather than the more desirable outcome" (Timmis, 2002, p. 243). A similar survey study was carried out by the author of this book among Czech and Slovak students in 2017. The results demonstrated that the participants were even more decidedly in favour of native-speaker (NS) norm to be applied in ELT: 71% were in favour of English native-speaker standard in grammar and 74% in pronunciation (Kalová, 2017). Comparable results in the Czech educational context have been reported by authors examining learners' beliefs and attitudes towards native and non-native accents confirming that native-speaker model remains the one most learners aim to achieve (Quinn Novotná, 2012; Jakšič & Šturm, 2017; Brabcová & Skarnitzel, 2018), as reported by Lancová and Červinková (Červinková Poesová & Lancová, 2021). One of the most significant findings emerging from Lancová and Červinková's research study conducted among pre-service English teachers revealed that despite general growing tendencies to propose English as a lingua franca as a model in pronunciation, "almost all participants expressed an explicit desire to acquire native(-like) accent" (2021, p. 105).

As regards European language policy in relation to norms in ELT, it does not seem to be explicitly specified. No clear guidelines are provided in the official documents about what norms should be applied in foreign language teaching, apart from the requirement of "exposure to words and fixed expressions in authentic spoken and written texts" as specified in the Common European Framework of Reference for Languages (2001, p. 149) in order to develop required linguistic competences. It is only when discussing errors, when the notion of norms is mentioned: "When the learner makes errors, his performance truly accords with his competence, which has developed characteristics different from those of L2 norms" (2001, p. 155). It can therefore be assumed that native-speaker norms are implied as a norm of reference for teaching foreign languages within the EU.

The situation seems to be very similar in the Czech Republic: in the official curricular document – *Rámcový vzdělávací program (Framework Education Programme)* (2007) – no explicit requirement of language norm is specified but it seems that, similarly to the CEFR specifications, native-speaker (NS) norm is implied. British Standard English as an appropriate model for ELT is mentioned in the official

curricular document for elementary schools published by the Czech Ministry of Education (1996, p. 17, 56). The preferred norm in the majority of textbooks used in Czech schools is also British English, although other NS varieties are also occasionally included, especially American English.

This section has attempted to provide a brief outline of guidelines for preferred language norm applied in ELT in the Czech Republic. It can be concluded that despite a number of controversies and somewhat unclear language policy in this respect, native-speaker norms seem to be the prevailing norm of reference, in particular the standard British English model. As this is also the preferred model of most participants in the research, it seems to be a legitimate choice to be applied in research presented in this book[15].

1.6 Analysing errors in learner language

Analysing learner language for errors lies at the core of SLA research and is equally important in foreign language learning and teaching. *Error* is a fundamental theoretical concept in this book, as the main focus of this research is exploring the accuracy of learner language and describing its characteristics. Errors are important both for linguists, as part of their study of learner language, and for language practitioners, because of their importance in second/foreign language teaching. For a better understanding of the concept of error in both areas, a precise definition is necessary. In the history of SLA, errors have been approached from different perspectives; the most influential approaches relevant for this research will be briefly described and discussed below. Precise error classification is also vital for conducting corpus-based analysis of learner language, an important part of the research project presented in this book.

Errors in learner language have always been a primary concern of foreign language teaching. While focus on form and errors was not favoured in some approaches and methods in language teaching, especially in the strong version of Communicative Language Teaching, and to some extent rejected in others, e.g. Krashen and Terrell's Natural Approach,

15 The issue of the norm as related to Contrastive Interlanguage Analysis (CIA) has also been addressed by Granger (2015, pp. 15-16). For a more detailed account of questions of norm in ELT in the Czech Republic, see also Kalová (2017; 2018) and Červinková Poesová & Lancová (2021).

in today's post-method era it has been recognised as important, (see for example Kumaravadivelu, 2006; Richards & Rodgers, 2014). A number of classroom-based research studies, whose results were summarized in three thorough meta analyses (Norris & Ortega, 2000; Spada, 2011; Sok et al., 2018), have been conducted over the past three decades in the effectiveness of meaning-focused and form-focused instruction. These studies have provided evidence that "learners benefit from both meaning-focused and language-focused instructional elements. Teaching approaches that exclude – or virtually exclude – either element deprive learners of opportunities to reach their potential for language development" (Lightbown, 2017, p. 113). As a result, a *counterbalanced approach* (Lyster, 2007) in which error-prone areas of L2 receive attention from both teachers and learners and thus enhance learning, has now been largely adopted.

A number of perspectives on errors have been adopted both in SLA and ELT research. These have gradually shifted from perceiving error as "a sinful act that must be prevented from occurring" to acknowledging its positive role "as an indicator of the mental processes that take place during the learning and acquisition of the target language" (Bitchener & Ferris, 2012, p. 6). Before the varied theoretical perspectives on the role of errors will be described, the definition of error and error classification will first be detailed.

1.6.1 Error definition

One of the first to draw researchers' and teachers' attention to the significance of learners' errors was Corder, who, in his eponymous paper (1967), distinguished between two major approaches to error in learner language: error as an inadequate use of the target language on the one hand, and error as an inevitable feature of learning on the other (1967, pp. 162-163). In his seminal work, written from the nativist perspective, he compares processes occurring when children acquire their L1 to those occurring when a foreign or second language is learned. Corder draws a distinction between *mistakes* and *errors*: while mistakes are defined as non-systematic errors of performance, also frequently committed by native speakers of the language, errors are described as "systematic errors of the learner from which we are able to

reconstruct his knowledge of the language to date, i.e. his *transitional competence*" (1967, p. 167). The users of the language are able to correct their mistakes either immediately as they realize they have committed them, or when alerted. In doing so, they adopt what Krashen (1989) has described as the Monitor, the conscious control of their production based on the learned rules about the language; an area addressed in more detail in section 1.2. These mistakes are often caused by the current mental state of the speaker, e.g. tiredness, stress, memory problems, strong emotions, and other factors; and affect not just learners but also native speakers of the language. Corder calls such incidences "slips of the tongue (or pen)" (1981, p. 10). Errors, on the other hand, are those which learners are unable to correct, as they are unaware of making them and do not possess the knowledge to be able to do so. It is the errors that are significant in teaching and learning: for teachers of foreign languages because they get information about their students' progress (or the lack of it) in language; for learners to test their hypotheses about the target language; and for researchers in second language acquisition as evidence of how language is learned and acquired (Corder, 1967, p. 167; Edge, 1989, pp. 10-11).

Despite the fact that errors are an inevitable feature of learner language, providing a clear and concise definition of error is far from easy (Kulič, 1971, p. 91; Ellis & Barkhuizen, 2005, p. 56). Typically, errors in learner language are perceived as violating the target language norm, which is rather problematic considering the ongoing debate about what variety of English should serve as a norm in ELT; a question addressed in more detail above, in part 1.5. The controversy related to defining error in English language teaching lies in the fact that there is no unanimity in understanding what norm should be used as reference. Richards and Schmidt (2010) avoid this controversy by defining error committed by a foreign language learner in their spoken or written production as follows: "[error is] the use of a linguistic item (e.g. a word, a grammatical item, a speech act, etc.) in a way which a fluent or native speaker of the language regards as showing faulty or incomplete learning" (2010, p. 201). This definition implies that it is not necessarily just a native speaker who could be an appropriate norm-creating model but acknowledges that anyone with a fluent knowledge of language could serve as such model. What exactly is meant by 'a fluent speaker', however, is not specified.

It is also significant that when analysing errors, usually just those in learner production are explored in research and errors of comprehension are often ignored. To a certain extent, this is understandable as "comprehension errors are difficult to detect as it is often impossible to locate the precise linguistic source of an error" (Ellis & Barkhuizen, 2005, p. 51). As a result, it is almost entirely learners' free spoken and written production that is analysed for errors (Selinker, 1972; Ellis & Barkhuizen, 2005; Han, 2014).

Another decision to be taken when attempting to define error is "whether *grammaticality* or *acceptability* should serve as the criterion" (Ellis & Barkhuizen, 2005, p. 56). Looking at errors from the grammaticality viewpoint provides a relatively objective method; related to the distinction between *overt* and *covert* errors. While overt errors are identifiable on the sentence level, a covert error is not apparent at first sight and can only be identified when "a larger stretch of the discourse is considered" (ibid). A broad context also needs to be considered when looking at errors when the acceptability criterion is applied. This is often believed to be far too subjective, however, to be suitable for research (ibid.).

In order to fully understand the notion of error and provide a definition to be adopted in this book, it is first necessary to mention the different taxonomies according to which errors can be classified. These are described in the section that follows.

1.6.2 Error classification

In the previous part, the distinction between errors and mistakes as errors of competence and performance, and the difference between overt and covert errors have been described. Ellis and Barkhuizen use the error versus mistake distinction as well. Unlike Corder, however, they believe that errors and mistakes are equally valid in research "both practically and theoretically" (2005, p. 62), and not just errors but mistakes too should be explored. They claim that mistakes occur partly because an erroneous feature or language item has not been fully acquired and learners are therefore unable to use them correctly in their free production (2005, pp. 62-64); they might, however, be able to recognise errors, e.g. in Grammaticality Judgement Tests.

There are many different criteria which can be adopted in error classification. In the most frequently applied approach, learner language forms are compared with those in the target language (Ellis & Barkhuizen, 2005; James, 2013; Börjars & Burridge, 2019). Based on such comparisons, errors in learner language need to be identified and described. For this, a system of categories – error taxonomy – is required and the frequency of errors in the categories can then be explored; the two approaches are not mutually exclusive and tend to be combined. Two error taxonomies are usually applied: linguistic, based on well-defined grammatical and lexical categories, and surface structure taxonomy (Ellis & Barkhuizen, 2005, pp. 60-61). The former "indicates on what level of language the error is located: in phonology, graphology, grammar, lexis, text or discourse" (James, 2013, p. 105), and which grammatical system it affects (ibid). In the latter, errors of *omission, addition, misinformation, misordering* and *blends* are identified (Ellis & Barkhuizen, 2005, p. 61). Corder (1981), however, warns against superficial linguistic error analysis, claiming that while such classification is important, it should be regarded as "only a starting point for systematic analysis" (1981, p. 36). Surface structure taxonomy is also "of less obvious practical use as grammar teaching is organized in terms of traditional descriptive categories" (Ellis & Barkhuizen, 2005, p. 62). Error taxonomy based on linguistic description appears to be more appropriate in analysing learner language as it enables a more concise and therefore more objective error description. Such taxonomy is widely applied in corpus-based analyses (Ellis & Barkhuizen, 2005) and will therefore also be used in the research presented in this book.

In the Czech educational context, three major criteria are usually applied in error taxonomy: 1) the seriousness of error; 2) the linguistic aspect; and 3) the causes of errors (Choděra, 2013). Regarding error gravity, three degrees of intensity are defined, from grave to minor errors. The criterion of seriousness of error is similar to Corder's distinction between mistake, lapse and error. Apart from Corder's term *lapse*, indicating a mistake caused by internal factors, e.g. fatigue, stress, or nervousness (1981, p. 10), Edge (1989, pp. 10-11) uses the term *attempts* to indicate what learners are trying to express without knowing the exact rules. It should also be taken into consideration whether or not errors impede communication; the more they do, the graver the errors are (Richards & Schmidt, 2010, p. 202). Such errors are

labelled as *interpretive* when "misunderstanding of a speaker's intention or meaning" (2010, p. 201) occurs, and *pragmatic* when "the wrong communicative effect" is produced (ibid). Similarly, Brown (2007), in relation to writing refers to *local* and *global* errors; the former are errors of grammar and discourse that do not impede communication, the latter are those that need to be addressed immediately as they might cause misunderstanding (p. 426).

When the linguistic criterion is adopted, errors are classified according to what aspect of language is affected by error: *lexical, grammatical, orthographic*, and *phonetic* (Choděra, 2013, pp. 163-164). Regarding the causes of errors as a psychological aspect, Choděra (2013), similarly to Corder (1967), classifies errors caused by generalizations and false analogies, as well as errors caused by internal factors such as stress-induced errors, or errors caused by tiredness or inattentiveness (pp. 163-164).

Classifying errors according to their causes might be rather challenging, as it is not always clear why errors occur: "while error analysis has the advantage of describing *what* learners actually do […] it does not always give us clear insights into *why* they do it" (Lightbown & Spada, 2013, p. 45). Despite the challenge, identifying causes of errors in learner language is the major focus of SLA research and it is relevant for teachers and learners as well. Two major causes of errors can be identified, *interlingual* and *intralingual*. The former are related especially to the influence of learners' L1 but also other languages, which could either facilitate learning, then it is usually referred to as *transfer*, or make it more difficult, then it is called *interference*, see for example (Ellis & Barkhuizen, 2005, pp. 64-65; Saville-Troike, 2006, pp. 35-36; Janíková, 2013, p. 47). Intralingual errors, on the other hand, occur universally in all learners, regardless of their L1 and are similar to developmental errors in children learning English as their mother tongue (Ellis & Barkhuizen, 2005; James, 2013).

Transfer is the key concept of the Interlanguage Hypothesis as formulated by Selinker (1972), who identified five major areas affected by transfer: "1) language transfer, 2) transfer-of-training, 3) strategies of second-language learning, 4) strategies of second-language communication, 5) overgeneralization of TL linguistic material" (pp. 215-221). Errors of transfer are not just those directly related to the influence of the mother tongue, but also teacher or training induced

errors, which are caused by how a particular language feature is presented (Richards & Schmidt, 2010, p. 279). This could be caused by oversimplified rules presented either by teachers or in teaching materials. In the case of teachers, problems might arise with both NS teachers who might fail to explain the rules as studying grammar is often not part of their education, and NNSs whose knowledge might in some cases be deficient, and who, especially if they share the same NL with the learners might not be aware of some of the typical problem areas.

Transfer is not, however, limited to the relationship and influence of L1 and L2 in language learning and acquisition; knowledge of other language or languages has an effect too. Such influence is referred to as *cross-linguistic influence* (Lightbown & Spada, 2013, p. 59) and it is the key concept in multilingualism (Janíková, 2013, pp. 47-57). Similarly to transfer affecting L1 and L2, cross-linguistic influence is a major factor in language learning and acquisition, and its influence is multidirectional: L1 influences L2, L3 and other learned languages but it is also the influence of L2 that affects L3 acquisition, and L3 that affects L2 and L1. Within the concept of multilingualism, errors caused by the knowledge of other languages are examined, but it is especially the positive role of transfer as a proactive and facilitating factor of acquisition that is the focus of research in multilingualism (Janíková, 2013, pp. 52-57). As multilingualism is not the main focus of this book, the influence of L3, L4 and other languages on L2 acquisition will not be further explored.

In his Interlanguage Hypothesis, Selinker (1972) also addressed errors caused by *fossilization*. Fossilized errors are caused by the cessation of development of learner's interlanguage, when "incorrect linguistic features become a permanent part of the way a person speaks or writes" (Richards & Schmidt, 2010, p. 230). While Selinker believed that it was the whole system of learner's interlanguage that fossilizes, fossilization is now perceived as operating selectively and affecting only some parts of learners' language, depending on a variety of learner related factors, both internal and external (Han, 2013; 2014).

Intralingual errors are often similar to errors made by native speakers of the target language when they learn the language as their mother tongue; such errors are called developmental errors. Some theoreticians suggest that the order in which certain language features are acquired is almost identical for first and second language learners.

Krashen's Natural Order Hypothesis, first formulated in the 1970s (Saville-Troike, 2006, p. 45; Lightbown & Spada, 2013, p. 106) is based on this assumption. This claim has also been supported by research, especially the morpheme studies conducted in the 1960s and 1970s (Saville-Troike, 2006, pp. 43-44; Lightbown & Spada, 2013, pp. 7-9). Other intralingual errors are caused by erroneous application of the rules or drawing false analogies about them; these are usually classified as *overgeneralizations* (or false analogies), *under-generalizations* (or incomplete rule application), *misanalysis of rules*, applying *simplified rules*, *communication-based errors, errors of avoidance and overproduction* (Richards & Schmidt, 2010, p. 201; Lightbown & Spada, 2013, pp. 65-66). Errors caused by overgeneralization appear when grammar rules are incorrectly applied, typically in forming the plural forms in nouns when regular and irregular forms are confused, and similarly in forming past tense forms. This process does not occur in the SLA context only but also when children acquire English as their first language.

In addition to the above-mentioned causes of errors, objectively difficult features of language also tend to be typical error-prone areas. These are, for example, incorrectly used quantifiers (number vs amount, (a) few, (a) little), third conditionals, errors in punctuation and very formal lexis, relatively frequently misused even by native speakers of English. Besides, grammatical patterns which do not exist in the learners' first language are also often affected by errors, e.g. the use of articles by speakers of L1 which does not have articles in its system (Díez-Bedmar & Papp, 2008; Sun, 2014; Long & Hatcho, 2018). It must be borne in mind, however, that defining difficulty in language is not without challenges and remains rather fuzzy (Ellis, 2006). Despite this, it cannot be denied that language learners tend to either deliberately avoid such features in their production, entirely omit them, or use them in the wrong way.

Although this survey of error classification is inevitably far from exhaustive, it provides sufficient theoretical basis for the research specified further on. In this book, exploring the accuracy of advanced learner language and its typical features, errors, will be analysed from a linguistic perspective. Error is defined as an erroneous usage of the target language structures and lexis; pragmatic errors and errors of pronunciation will not be included in the analysis. Native-speaker norms will be used as a norm of reference for identifying errors; the reasons guiding this

choice are specified above, in section 1.5. The exact description of how the data were collected and which steps were taken in the error analysis are described in sections 4.5 and 4.7 respectively. In the next subchapter, the most influential approaches to learner language analysis and details of corpus-based analysis will be presented.

1.6.3 Perspectives on errors in SLA theories

Four major theories can be identified in the history of language acquisition: behaviourist, innatist, cognitivist and sociocultural or interactionist (Lightbown & Spada, 2013, p. 103; Šebesta et al., 2016). In these theories, different approaches to understanding the role of errors in SLA have been adopted; out of these two have played a prominent role, contrastive analysis and error analysis.

In the early approaches of the 1960s, based on behaviourism and structuralism, contrastive analysis comparing L1 and L2 was largely adopted. This was later gradually rejected as inadequate in favour of Chomsky's innatist perspective which focused on error analysis. The original concepts of contrastive analysis and error analysis have been revisited and adopted in research again, especially since computer-based learner language analysis appeared in the late 1980s and generated new possibilities of investigating learner language. These different approaches to researching errors will now be briefly outlined.

Contrastive analysis

Contrastive analysis (CA) (Lado, 1957), based largely on the behaviourist theory of learning and structural approach to the theory of language, tried to determine and predict problem areas in the target language largely based on comparisons of L1 and L2. From the behaviourist perspective, errors were viewed negatively and the main task of language learning and teaching was to prevent them entirely; there was zero tolerance for errors in learning (Bitchener & Ferris, 2012, p. 4; Lightbown & Spada, 2013, p. 104). Language acquisition was perceived as a result of habit formation and "errors were therefore predicted to be the result of the persistence of existing mother tongue habits in the new language [and] errors were ascribed to interference" (Corder, 1981, p. 1). According to CA, the most important influence in the acquisition of the target language was

the first language. This could be both positive and facilitate acquisition – transfer, and negative, when errors are committed – interference; as was mentioned above in section 1.6.2 (Saville-Troike, 2006, p. 35; Bitchener & Ferris, 2012, p. 4; Lightbown & Spada, 2013, p. 42).

Comparative studies analysing errors in learner language revealed, however, that this assumption cannot always be verified. Neither the positive influence of L1 facilitating L2 acquisition, nor all the predicted errors of transfer always appeared (Bitchener & Ferris, 2012, pp. 5-6; Lightbown & Spada, 2013, p. 104). CA was also unable to explain those errors which were apparently not caused by language transfer (Ellis & Barkhuizen, 2005, pp. 52-53) but occurred in the interlanguage of learners with different first languages. These errors, labelled as *developmental errors* (Saville-Troike, 2006, p. 39; Lennon, 2008, p. 53; Lightbown & Spada, 2013, p. 44), were not unlike errors committed by children when they learn the language as their mother tongue.

Apparently, in this strong version predicting learners' errors, contrastive analysis failed to produce the expected explanation of the processes occurring in acquisition of the target language (Saville-Troike, 2006, pp. 34-35; Bitchener & Ferris, 2012, pp. 4-5). Also, within CA, the influence of L1 interference was overestimated and other factors influencing L2 acquisition were not taken into consideration, especially internal factors now considered important, such as the age of learners, their motivation, or knowledge of other languages (Šebesta et al., 2016, pp. 34-35). It was mostly for the reasons mentioned above that this approach to SLA, popular mainly in the 1950s and 1960s, was largely abandoned in the following years and replaced by error analysis (EA) in the late 1960s and early 1970s (Dagneaux et al., 1998, pp. 164-165).

Error analysis

In reaction to the drawbacks of CA and behaviourism, error analysis based on the innatist theory of language appeared in the 1970s. It was a marked improvement that error analysis, unlike contrastive analysis, paid more attention "to actual learner errors in L2, rather than idealized linguistic structures attributed to native speakers of L1 and L2 (as in CA)" (Saville-Troike, 2006, p. 37). Also, rather than considering learner language affected with errors as "incorrect version of the target language" as it

was typical of CA (Lightbown & Spada, 2013, p. 41), error analysis perceived errors as "windows onto learners' interlanguage" (Dagneaux et al., 1998, p. 164), providing information about learner language. Errors were also regarded as an important sign of L2 development and evidence of how learners test their hypotheses about L2 (ibid). Regarding the sources of errors, EA revealed that "the majority of L2 errors do not come from the learner's L1 or the L2 and that they must, therefore, be learner-internal" (Bitchener & Ferris, 2012, p. 5). This was a major shift in our understanding of the processes occurring in learners' minds during acquisition.

EA was, however, later disregarded by researchers and language practitioners, due to a number of factors, both methodological and conceptual. Five major drawbacks of EA were identified:

1) EA is based on heterogeneous learner data;
2) EA categories are fuzzy;
3) EA cannot cater for phenomena such as avoidance;
4) EA is restricted to what the learner cannot do;
5) EA gives a static picture of L2 learning.

(Dagneaux et al., 1998, p. 164)

The first two limitations relate to the methodology of EA research and stress the importance of collecting reliable data providing relevant information. It is especially important to precisely define error categories; with ill-defined, often overlapping and highly subjective categories, interpretation, replication and comparison of studies are rendered almost impossible. What is a major problem with EA is that, especially by attending to erroneous use of language only and not taking correct use into consideration, and ignoring the dynamic nature of learner language, it does not provide a realistic picture of learner language (Dagneaux et al., 1998, p. 164).

It was especially for the reasons listed above that EA was largely abandoned: "once a very popular enterprise, error analysis (EA) is now out of favor with most SLA/FLT circles" (Granger, 2003, p. 466). This attitude was also partly caused by the shift of focus in language teaching in the 1980s, from focus on form to focus on meaning and communication, in which error analysis and corrective feedback were no longer regarded as important and were even perceived as harmful (Lightbown, 2017, p. 112). This does not mean, however, that the interest of researchers

and teachers in the concept of error completely disappeared; it was clear, however, that EA had to be revised and a new approach adopted. This occurred in the late 1980s and early 1990s when new digital technologies enabled the exploration of large amounts of learner data and the use of once abandoned methods of examining errors was revived.

Contrastive interlanguage analysis and computer-aided error analysis

The study of errors has always played an important role in SLA research as well as in language learning and teaching: for researchers, it serves as a source of information about how acquisition and learning occur; for teachers, it shows what their students have already mastered and what requires more attention in the classroom; and for learners, it identifies areas they need to focus on in their studies (Corder, 1967; 1981; Ellis & Barkhuizen, 2005, p. 51). Despite the many limitations of contrastive analysis and error analysis, it would be short-sighted to dismiss them entirely; their positive role in learner langauge analysis should be recognised. Both these traditional approaches have been revisited and continue to provide powerful tools in SLA research. New possibilities of CA and EA opened with the onset of digital technology in SLA research in the late 1980s.

During the past 40 years, Selinker's (1972) original concepts of interlanguage and its fossilization have reappeared in research (Han, 2004; 2006; Han & Cook, Wei, 2009; Montrul, 2014). It has been acknowledged that "native language influence is the major shaping force in fossilizable speech behavior and, when combined with other factors, solidifies fossilization" (Han, 2013, p. 137). Comparing the two language systems, L1 and L2, and the related questions of language transfer and fossilization, as two major themes reflected in the Interlanguage Hypothesis (addressed in part 1.1 of this book), are central in contrastive analysis and have continuously attracted attention of researchers. Rather than independently, they should be explored as inter-connected and inter-dependent; and since they are best manifested in learners' free production, such data should be collected and analysed. Numerous empirical studies have provided "a wide range of evidence in the last four decades for the major role of transfer in SLA" (Han, 2013, p. 137). In these studies, corpus-based analysis of learner language has often been adopted,

(for example Wierszycka, 2013; Götz, 2015; Gráf, 2015; 2017; Long & Hatcho, 2018). This type of analysis enabled large-scale comparisons between the production of native speakers and that of learners with different L1 backgrounds, revisiting some of the concepts of contrastive analysis, and gave rise to "a new research paradigm of *contrastive interlanguage analysis*" (Barlow, 2005, p. 342), a term coined by Granger in 1998 (ibid). This newly adopted method of linguistic research has gained in popularity among researchers and "has spawned a large and highly diversified body of research" (Granger, 2015, p. 9).

Compiling large corpora of both native and non-native language provides researchers with a very rich source of data in amounts hardly possible before. This has sparked a renewed interest in error analysis. State-of-the-art software tools enable relatively quick and easy large-scale analysis of both written and spoken language. Most importantly, this *computer-aided error analysis* (CEA) enables researchers to build on the principles of traditional error analysis but overcomes most of its limitations. Learner language can be analysed as a whole, when both erroneous and good usage of the target language can be explored. Corpus-based analysis also makes it possible to analyse errors in context rather than in isolation (Dagneaux et al., 1998; Barlow, 2005; Granger, 2009; Granger et al., 2015). Since the early 1990s, there has been a growing body of studies on errors as well as their causes in SLA research, enabled by using data from large electronic corpora:

> Learner corpora have a lot to contribute to SLA research. They lead researchers to a better understanding of how foreign languages are learned and can help them to answer questions at the heart of SLA research, such as the as yet unresolved issue of the exact *role of transfer* in second language acquisition and the notion of avoidance. (Granger, 2009, pp. 268-269).

Computer-aided error analysis offers novel possibilities in learner language analysis. Unlike the previous, now largely discredited, traditional approach to error analysis, CEA enables analysing large quantities of texts, both spoken and written which can be error tagged, and annotated for error-types. One major advantage of this type of analysis is that "errors are not isolated from the texts in which they originated, as was the case in traditional EA, but rather are studied in context alongside cases of correct use and over- and under-use" (Granger, 2009, p. 268). Another important advantage is the possibility of adding metadata.

These can be inserted in the corpus together with the annotated texts and enable a variety of approaches to the analysis, based on e.g. age, gender, proficiency of learners, L1 backgrounds, types of texts, etc.

Based on the reasons specified in this section, it can therefore be concluded that computer-aided error analysis is a legitimate tool to be adopted when analysing learner language for errors, which is the main focus of the research presented in this book.

1.7 Summary

The aim of this chapter was to clarify major theoretical concepts and how they were operationalized in the research addressed in this book. Notions of learner language, roles of implicit and explicit knowledge in foreign language learning and second language acquisition were described, differences between competence, performance and proficiency were delineated. Definitions of key concepts − error and norm in foreign language learning and teaching − and some controversies related to them were discussed. The CAF model was adopted in the research, with focus on accuracy which was defined as error-free language, without deviations from standard native speaker norms, and operationalised as a number of errors per 100 words. Different approaches to error analysis were detailed, with special attention paid to corpus-based error analysis which was adopted in the research described in this book. All details of how the analysis was conducted in this research are specified in chapter 4 focusing on the methodology applied in this research. Relevant contemporary studies conducting corpus-based analysis will be discussed in chapter 2, in which literature review is provided.

2 REVIEW OF THE LITERATURE

In this chapter, a critical analysis of current studies investigating the accuracy of learner language will be presented. These studies often explore the influence of the mother tongue on the acquisition of the target language, L1-induced errors, as well as characteristics of learner language described in terms of accuracy, fluency and complexity. A novel approach frequently adopted in them is computer-based analysis of learner language and learner corpora.

There are two major areas where learner corpora can be used in foreign language teaching research: 1) analysing different aspects of learner language, both spoken and written, and 2) methods of developing learner proficiency in the target language by using corpora in the second and foreign language classroom (Smirnova, 2017, p. 302). The main focus of this book is the former, and for this reason, articles on corpus-driven teaching and learning have not been included in this review.

For the review of the literature, both Czech and international scholarly journals were searched for relevant articles. The search was carried out according to the following pre-set criteria: the article was published from 2013 to 2019[16] in international educational journals, and in Czech linguistic and educational journals. Due to the fact that terminology in the researched area often tends to be used inconsistently, manual search of relevant studies was used instead of database search. First, reputed academic journals with a high impact factor were reviewed and articles chosen according to the titles and key words. This choice was followed by the analysis of abstracts and the most relevant studies were then analysed in more detail. The following key words were used

16 Originally, five years, from 2015 to 2019, were reviewed for relevant articles but as quite a few relevant studies had been published in the years 2013-2015, articles published in this period were also included in the review.

to fine-tune the search: corpus-based analysis of learner language; computer-aided error analysis of learner language; L1-induced errors; native language (L1) influence on the acquisition of the target language; positive/negative transfer; the accuracy of learner language. Based on these criteria, fourteen international scholarly journals, six Czech educational and four linguistic journals were reviewed (for the complete list of reviewed journals, see Appendix 1).

2.1 International journals

Using learner corpora in SLA research and classroom-driven FLT/SLT research has become a thriving area, a fact clearly manifested by many research studies adopting this methodology. A multidisciplinary scholarly journal specialising in these topics, the *International Journal of Learner Corpus Research* (IJLCR), has been published since 2015. As the whole journal aims at corpus-based research of learner language all articles meet the search requirements described above. Therefore, for this particular journal, an overall outline of the most frequently addressed topics was conducted first, and special attention was paid to the topics relevant for this study: corpus-based analyses of learner language taking into account the influence of the native language on the acquisition of the target language.

The total of 54 articles have been published in the ten issues of the journal since 2015. According to their main focus, these texts can been divided into five areas[17]: general texts on learner corpora (LC) and learner language (LL); corpus development methodology; aspects of learner language; role of tasks, learning context and/or proficiency in learner language; and L1 influence on the acquisition of the target language. The most frequently addressed area (explored in 21 articles) was different aspects of learner language, both general, e.g. academic phraseology, and more specific, e.g. innovative verb-to-noun-conversion. Eight articles provided theoretical background to developing learner corpora and guidelines and rules to be observed when conducting corpus-based research. Seven general and introductory texts explored how corpora are applied in learner language research. Six articles examined how tasks, learning context and proficiency levels as research variables influence

17 Book reviews and interviews have been excluded from this outline.

second language acquisition. An outline of the topics and related articles published in the IJLCR till now is provided in Appendix 2, Table A1[18]. Twelve articles related to the area discussed in this book, all dealing with native language influence on learner language, will be analysed below. First, however, an important theoretical study on Contrastive Interlanguage Analysis (CIA), highly relevant for this research, will be addressed (Granger, 2015).

Granger's (2015) paper summarises the main characteristics of CIA as a method of corpus-based research, addresses some of its frequent criticisms and advocates the suitability of this powerful tool in exploring typical traits of learner language. According to Granger, there are two main advantages of CIA: first, the possibility of exploring language features in context rather than in isolation, and second, the fact that two or more language systems, typically learner language and the target language, or learner languages used by speakers from different L1 backgrounds, can be compared and contrasted (2015, pp. 7-8). Unlike traditional SLA research, focused mostly on lower proficiency levels and spoken language, CIA enables analysing advanced learner language, both spoken and written (2015, pp. 10-12). Two major limitations of CIA are addressed in the paper: the question of the norm in learner language analysis, an area addressed thoroughly in part 1.5 of this book, and *comparative fallacy*. Comparative fallacy, a term coined by Bley-Vroman (1983), refers to the fact that constant comparison of learner language to the target language might cause that the former is perceived as deficient and such comparisons should therefore be avoided. Granger, however, argues against this claim. While she admits that it is also possible to explore learner language without contrasting it with the TL, learner language is not just the focus of theoretical SLA research but is important for language practitioners as well. Teachers' perspectives of learner language research must, therefore, also be taken into account: "from a pedagogical point of view, the benefit of L1-L2 comparisons is […] obvious, as they provide language teaching professionals with precious information on what learners do right or wrong or partly wrong in a particular skill or task" (2015, p. 14). Presenting a new model of CIA, Granger promotes a new approach to be adopted in these comparisons, one that acknowledges "the notion of 'varieties': reference language

18 All tables included in the Appendices, are indicated as Table A1, A2, and A3. The tables included in the main text are indicated as Table 1, Table 2, etc.

varieties on the one hand, and interlanguage varieties on the other". She addresses the question of what norm should be used as reference, claiming that "there are a number of different reference points against which learner data can be set" (2015, p. 17), not necessarily based entirely on NS norms as the only option. In the model, in addition to "the traditional inner circle varieties […] outer circle varieties as well as corpora of competent L2 user data" (p. 17) are applied. In her view, this approach enables "researchers to take variability in learner language even more into account" and to fine-tune learner corpus research studies investigating large number of variables, not just learners' mother tongue, but also their proficiency levels in the TL, their motivation, etc. (p. 18). Granger also clarifies the terms *overuse* and *underuse* frequently appearing in interlanguage analyses as bearing no negative connotations and being purely descriptive in nature (2015, pp. 18-19). The model, newly proposed by Granger, appears to be an appropriate research tool to be adopted in learner language analysis.

As studies on determining how L1 can influence TL acquisition are most relevant for the topics addressed in this book, they will now be analysed in more detail. Twelve studies published in the IJLCR focused on the relationship between the native language (four Asian and ten European languages[19]) and the acquisition of the target language (ten studies explored English as the TL, two studies French and Dutch, and Spanish and German were focused on in one study each [20]). Five studies analysed written language, six spoken, and in one study the focus was not specified. Very few studies provided detailed information about the participants or research methods but most gave details of the corpus used in the analysis. Five articles explored grammatical choices made by learners, three explained lexical choices, three discussed fluency and focused on pronunciation. All studies reported some kind of native language influence: this influence was manifested on the linguistic level, in both lexis (Kyle et al., 2015), grammar (Crosthwaite, 2016; Schneider & Gilquin, 2016; Brunner et al., 2016; Stormbom, 2018; Deshors, 2018; Hendrikx et al., 2019; Gilquin, 2019), and pronunciation (Belz et al., 2017; Gósy et al., 2017; Lecumberri et al., 2017). While varying degrees of both positive (Hendrikx et al., 2019) and negative (Crosthwaite, 2016)

19 Four Asian languages: Chinese Mandarin, Hindi, Korean, and Thai; ten European languages: Bulgarian, Dutch, Finnish, French, German, Hungarian, Italian, Russian, Spanish and Swedish

20 Five target languages: English, Dutch, French, German and Spanish

transfer were reported in these studies, a finding relevant for this book as a whole, none of them provided an account of research directly related to this research study. For this reason, the articles were not analysed any further. Details of studies on the influence of L1 and learner language published in the International Journal of Learner Corpus Research are detailed in Appendix 2, Tables A2-A4.

In addition to the IJLCR, 13 different international journals[21] were reviewed adopting the same search criteria. Eight studies were relevant for this research and their detailed account is provided in Appendix 2, Tables A5 and A6. Most of these studies were small-scale projects ranging from 16 to 68 participants, with one bigger study in which 126 students from five universities took part. The target language was mostly English as a foreign language, with one study where the TL was French as a second language. Levels of proficiency in the TL were mostly advanced (five studies), intermediate (one study), or mixed (two studies). Six studies focused on written and two on spoken learner language; the focus was not specified in one study. The L1 backgrounds of the participants were seven European languages: English, French, German, Latvian, Norwegian, Polish, and Spanish, and four non-European languages: Arabic, Japanese, Mandarin Chinese, and Turkish.

All the studies with one exception (Uçar & Yükselir, 2017), a linguistic meta-analysis of 20 articles written by NNS scholars and published in linguistic journals, were conducted in tertiary education institutions. One half of the reviewed texts focused on error analysis, and the other half on specific language features, e.g. phrasal verbs, conjunctions. In five studies, corpus-based analysis of the collected data was adopted.

Six of these studies (MacDonald et al., 2013; Hamed, 2014; Sun, 2014; Karazoun, 2016; Shimanskaya & Slabakova, 2017; Long & Hatcho, 2018) have conclusively shown that L1 is a decisive factor in L2 acquisition and negative language transfer appears to be an important cause of errors and a strong influence on error types in the target language; "L1 is more of a factor than many EFL teachers realize, and thus, this information should be highlighted to students as to how their L1 might be impacting their L2" (Long & Hatcho, 2018, p. 119). A number of pedagogical implications are mentioned in the findings,

21 For the complete list of reviewed journals see Appendix 1.

requiring teachers in the foreign language classroom to pay more attention to the influence of L1, as "drawing parallels and highlighting differences between the L1 and L2 grammatical meanings" can have beneficial effects on learning in the context where most learners share the same L1 (Shimanskaya & Slabakova, 2017, p. 274).

In the reviewed studies, different approaches to error definition were adopted. MacDonald et al. (2013) define error as "a form or structure in the learner's production that is identifiable as being deviant […] in comparison to a native speaker of the target language attempting to say the same in an identical linguistic and communicative context" (2013, p. 39). The Louvain error tagging system (Dagneaux et al., 1998) was implemented in the learner corpora they created. In the study, two learner corpora of written English were created, based on the analysis of documents written collaboratively by 126 university students from five different L1 backgrounds (German, Norwegian, Spanish, Latvian, and French) in synchronous (online conferences) and asynchronous (emails) communications in English as the target language. The participants' proficiency levels ranged from intermediate to advanced, with the minimum required level of B1 according to the CEFR. These corpora were then analysed for errors with the following results: as anticipated, more errors were identified in the synchronous communication, in which learners tend to focus more on fluency and meaning than on accuracy. Different types of errors appeared in the two modes of communication, while in the synchronous mode, more errors in form and grammar were reported, errors in lexis and style were more frequent in the asynchronous mode. One of the possible explanations offered by the authors of the study is that this form of communication "encourages students to spend more time planning their messages, allowing them to take 'language risks' […] and to exploit a wider variety of lexical choices. This, in turn, may lead to more errors being produced in this category" (MacDonald et al., 2013, p. 49). Another important finding was that frequency of error types varied within each different L1 background of the participants, which indicates that L1 clearly influences specific error types. A question for further research has been posed regarding the extent to which L1-specific errors are due to negative interference from the mother tongue (MacDonald et al., 2013).

In his study of ungrammatical patterns in the written production of 30 undergraduate Chinese students of English philology, Sun (2014) argues that errors that impede understanding, especially erroneous use

of verb tenses, should receive more attention from language teachers than errors which do not prevent effective communication. Such errors should be tolerated even though learners' awareness of them should also be raised (p. 176). Neither a definition of error, nor the norm used as reference in the study are explicitly defined; it seems, however, that NS norm is applied in the analysis without the variety of English being specifically mentioned (p. 177). Perhaps unsurprisingly, the most frequent grammatical error type resulting from the analysis of free writing was the misuse of determiners. This was followed by erroneous L1-induced 'Chinese-English pattern' (Chinglish)[22], tense errors, and misuse of prepositions, lack of subject-verb agreement and misuse of adverbials (pp. 177-181). Sun regards the influence of the learners' native language as the most important factor causing these error types; this claim is strengthened by the fact that unlike English, Chinese Mandarin does not have articles and as a result, their correct use is highly problematic for Chinese learners of English, which had been proven by large-scale corpus studies as well, for example (Díez-Bedmar & Papp, 2008).

A study published by Shimanskaya and Slabakova (2017) assesses the effect of targeted instruction in teaching pronouns to Anglophone learners of French. Their study, similarly to this book, attempts to combine research with language pedagogy by addressing a topic relevant for both researchers in SLA and teachers (pp. 259-261). The results of this study indicate "that the process of learning an L2 starts with an attempt to impose native language categories on the new language" (p. 274), a fact that should be reflected when presenting new features in L2 instruction. It has also been demonstrated in this study that learners can benefit from comparisons of the linguistic systems of their L1 and the TL which enable users "to predict […] transfer patterns and specific errors" (ibid); this is a finding which should be taken into consideration by language teachers. Pointing out the similarities and differences between L1 and L2 can help to make the teaching more effective, especially when learners share the same native language (p. 272). Also, "being able to predict acquisitional difficulties allows FL practitioners to tailor metalinguistic explanations to the needs of the students" (p. 273), making the instruction better targeted and more effective. Perhaps the

22 "Wei and Fei (2003) define Chinese English (Chinglish) as an interlanguage, usually manifested as Chinese--style syntax with English words, Chinese phonological elements in pronunciation or grammatical variations that attempt to follow Standard English rules but miss the mark" (Sun, 2014, p. 179).

most important finding of the study is proving that "meaningful focus on form exercises that take into account transfer effects might speed up acquisition and allow learners to be more efficient in noticing" (p. 274) the differences between L1 and L2 and learning from them. For obvious reasons (Shimanskaya and Slabakova's paper was published a year later than the research described in this book was conducted), the study was not a direct inspiration for this research. Its outcome, however, has been of utmost importance as it supports the design of the current research study and justifies the focus on transfer patterns typical of learners with the same L1 background.

The topic explored in the last reviewed international paper was the grammatical accuracy of Japanese EFL learners (Long & Hatcho, 2018). In the study, the most frequent error types and their causes were identified in the spontaneous spoken production of 61 university students at B1 level. A learner corpus of spoken English was devised based on the transcripts of conversations which were analysed with the following results: the most frequent error type was the use of articles, followed by erroneous verb tenses, prepositions, errors due to omission, errors in the use of modifiers and in subject-verb agreement. Regarding the errors, interlingual errors are reported as the most frequent, accounting for 51% of errors, 35% were intralingual errors and 12.5% errors of unspecified origin. It is not, however, specified how interlingual and intralingual errors were distinguished in the analysis, nor are the details of the analysis provided. These results indicate that L1 seems to be a major factor in the grammatical accuracy of learner language. According to the authors, more attention to form should be paid in the language classroom, focusing in particular on the features which tend to be problematic for Japanese learners. What is stressed as important is the actual use of correct grammar forms in meaningful communicative context rather than "just learning about the correct usage of these forms" (Long & Hatcho, 2018, p. 119).

The evidence reviewed here seems to suggest a significant role of native language on the acquisition of the target language. The studies presented above, however, remain rather narrow in their focus on some aspects of learner language only, without dealing with it as a whole or providing its thorough analysis. Also, most of the studies focus either on written or spoken language only, and in some the focus is not specified. Overall, the studies provide analysis of learner language but do

not attempt to improve it or eliminate any of the anticipated problems. In view of all this, it is hoped that the research presented in this book exploring possibilities of how the accuracy of learner language could be improved, together with analysing both spoken and written language, might generate fresh insights into learner language and L1-induced errors.

2.2 Czech journals

As corpus-based analysis of language is one of the tools adopted both by applied linguists in the field of SLA, and by researchers in education, both educational and linguistic journals published in the Czech Republic were reviewed (for the complete list of reviewed journals, see Appendix 1). The search was based on the criteria mentioned above, in part 2.1.

The analysis of abstracts in the pedagogical journals did not reveal many relevant texts published from 2013 to 2019 related to the topics dealt with in the current research. Only two texts focused on the role of transfer in ELT: Göbel and Vieluf's article on how positive language transfer can be implemented in teaching English in the German secondary school context (Göbel & Vieluf, 2018), and Konečný's text on native language interference in teaching Russian as a foreign language. Its main focus was on developing sociocultural competence and exploring a lack of cultural knowledge as a possible source of errors and misunderstandings (Konečný, 2014). As this aspect of communicative competence is not the aim of the present book, it was not reviewed any further.

One of the few relevant studies applying corpus-based analysis of learner language was published by Tůma (2013). In his study, he reports about how the communicative competence of learners was developed in a blended learning course in the EFL context. Unlike most of the international studies reviewed in part 2.1, both error and norm are clearly delineated in this study. Error is defined as "a deviation from the language norm" (Tůma, 2013, p. 96) and standard British English is applied as the primary norm of reference, with American English forms also regarded as acceptable. In his research, Tůma focused on the development of linguistic competence in writing and the accuracy in the use of verb forms. The sample, 18 undergraduate students at A2 proficiency level according to the CEFR, participated in three online asynchronous discussion fora. A learner corpus was compiled from two

texts written by the students in two online discussions and annotated for verb-tense related errors in pre-defined domains: the analysis was limited to morphological, orthographic, lexical and syntactic errors in verb clauses. Tůma reports that while CC improved overall, the comparison of the two texts revealed an increase in error rate in all the analysed domains. In his view, this can be explained by the nature of the tasks in which different topics required more frequent use of advanced verb tenses in the second task. Unlike in the first task, in the second task intention was expressed more frequently which, according to Tůma (2013, p. 105), might have led to an increased error rate affecting verb tenses. This fact rendered comparisons between the two tasks difficult and the results related to accuracy remained therefore inconclusive. One of the most important findings of the study is that morphology at A2 level remains unstable and the opportunities for eliciting data from free production at this level of knowledge are limited (Tůma, 2013, p. 107). This study thus focuses on verb-tense related error analysis of Czech learners of English at A2 level, exploring one aspect of learner language only. One of the findings relevant for this research (for details of the study see Appendix 2, Table A7) is the fact that corpus-based analysis appears to be an effective method to be applied in learner language error analysis. Not enough details about how the corpus was compiled and analysed are, however, provided. What is especially relevant for the current study is the need to pay careful attention to the selection of the topic when soliciting free production data from learners as it may affect the error rate considerably. This finding played an important role in optimising the design of the spoken and written tasks adopted in the present research study; this process is detailed in part 4.3.

Due to the fact that the review of the literature published in Czech educational journals did not produce any other relevant articles, linguistic journals were also reviewed. Seven studies examining language corpora were published in *Slovo a slovesnost* (2016-2019). Most of them explored general aspects of creating corpora, use of corpus-based analysis in literary studies, and focused on the Czech language analysis. For these reasons, they are not further analysed in this review. The journal *Naše řeč* published a monothematic issue in 2014 on the use of corpora in linguistics and translation, and one text on corpus-based analysis of poetry in 2018; all these texts explored the Czech language and are therefore not relevant for this study.

A relatively large body of texts applying corpus-based analysis appeared in *Acta Universitatis Carolinae Philologica* in 2013, when eight articles were published presenting a variety of linguistic analyses of three Romance languages, Italian, Portuguese and Spanish, in parallel and diachronic corpora; one text was also published in 2014 on the use of corpora in linguistics and lexicography, one study from 2015 focused on linguistic analysis of German, and one diachronic corpus study of Spanish was published in 2014. A study which is highly relevant for this research was published in this journal by Gráf in 2017 and will be described in more detail in the following passage. The last Czech linguistic journal which was reviewed for this paper was *Časopis pro moderní filologii*. An impressive array of 17 articles, mostly on corpus-based analysis of contemporary Czech, with one text on historical Spanish and one on Czech-Polish parallel corpus was published between 2013 and 2019. Apart from Gráf's article (2017) which will be detailed below, none of the above mentioned corpus-based studies focused on learner language analysis. This is why they are not reflected in the outline (see Appendix 2, Table A7) in which the two relevant articles and one dissertation are presented.

The two studies most relevant for this research were conducted by Gráf in 2015 and 2017 (see Appendix 2, Table A7). His corpus-based error analysis of the accuracy of advanced spoken learner language focused on verbal categories (2017), developing further his previously published dissertation on accuracy and fluency of advanced learner English (2015). The latter was also, to a certain extent, inspiration for the current research. Having collected data from three spoken tasks produced by 50 students of English philology, Gráf compiled a Czech sub-corpus of the LINDSEI spoken learner corpus[23] and explored it in terms of accuracy and fluency. Analysing the dataset for errors, he first identified the most error-prone areas of Czech learners of English. These were grammar errors, especially erroneous omission of either the definite or indefinite articles, followed by the wrong use of verb tenses, especially the present perfect. This error domain was followed by lexical errors, especially in the use of prepositions. The analysis of learner fluency revealed that both native and non-native fluency is affected by task design, but there is no statistically relevant correlation between

23 LINDSEI – Louvain International Database of Spoken English Interlanguage, LINDSEI_CZ – the Czech subcorpus of LINDSEI

accuracy and fluency (Gráf, 2015, pp. 101-153). In the conclusion to his dissertation, Gráf argues that when comparing the learner language accuracy of German and Czech LINDSEI sub-corpora, "the differences between the two nations appear in the frequency of particular error types, which indicates the possible effect of language transfer" (Gráf, 2015, p. 155). In his later study (Gráf, 2017) based on the same spoken learner corpus, he provided a more thorough analysis of verb-related errors. Inaccuracies in the use of the present perfect tense and noun verb agreement were identified as the key areas of difficulty for advanced Czech learners of English. Gráf proposes that preventative measures should be taken in the advanced language classroom as it is often the teaching of grammar focusing on decontextualized practice and using potentially confusing rules that triggers many of these errors. He also identifies the negative influence of language transfer as a potential cause of these recurrent errors. One of the most important pedagogical implications of Gráf's research studies is that the error-prone areas identified by his research should be systematically addressed by teachers and advanced learners, and novel ways of approaching their elimination should be considered. Gráf's findings helped to provide the basis for this research study, especially in helping to select relevant language features which require attention when teaching advanced Czech learners of English, which was reflected in designing targeted educational intervention and devising the testing materials. From the methodological viewpoint, the Louvain error tagging system used by Gráf (2015; 2017) was also suggested to be used in the current study.

2.3 Summary

The review of the literature presented above has provided ample evidence that a learners' mother tongue is a decisive force in the acquisition of the target language. Its negative influence has been identified as the major cause of errors in (advanced) learner language and it has been demonstrated that typical error prone areas differ depending on the L1 background of the learners. In the conclusions to the studies, the need of focused attention on these areas from teachers and learners was stressed (MacDonald et al., 2013; Sun, 2014; Karazoun, 2016; Belz et al., 2017; Lecumberri et al., 2017; Shimanskaya & Slabakova, 2017; Long & Hatcho, 2018). Some of the articles reviewed here have indicated that learners

benefit from focused comparisons between L1 and L2 (Shimanskaya & Slabakova, 2017). Regarding the methodology of research, corpus-based analysis seems to be a suitable option and the Louvain error tagging system a useful tool for learner language analysis (Götz, 2015; 2019; Gráf, 2015; 2017). In most studies, standard British English was adopted as the norm of reference; and it therefore appears to be a legitimate choice for this research as well. All these findings were reflected in the design of this research project in which conducting focused educational intervention aimed at increasing the accuracy of advanced learner language and raising learners' awareness of the most problematic areas was proposed. While all the studies reviewed concentrated on either spoken or written language, in this book a variety of data was analysed. In addition to the data elicited from free spoken production of advanced learners of English, their written language, as well as the Grammaticality Judgement Tests were analysed in order to provide a complex picture of their language. The research focus, design and findings will be detailed in the following empirical part of the book, chapters 4-6. First, educational intervention whose effectiveness was explored in this research study will be delineated in the following chapter.

3 EDUCATIONAL INTERVENTION

In this chapter, educational intervention the impact of which is examined in this book is presented. The intervention was designed in order to improve learners' accuracy of the target language and raise their awareness of its potential problem areas. The intervention took the form of a one-term elective course for undergraduate students of English philology held in spring and autumn terms of 2016. Within the course, special attention was paid to form-focused instruction: grammar input presented in meaningful context, explicit grammar practice, focus on pre-selected linguistic forms, metalinguistic explanations, error correction and corrective feedback.

When examining the relevant literature, we are often faced with conflicting views on focus on form in language education; while there are ardent advocates of such instruction who claim it has positive effects on both implicit and explicit language knowledge (Schulz, 1996; Ferris,1999; Shintani & Ellis, 2013), others express vehement opposition (Truscott, 1996; Gray, 2004; Krashen, 2004). As both sides in this somewhat controversial issue have been supported with research evidence, it seems that a balanced approach is advisable. What is important for teachers is which aspect of learner performance they aim to develop, fluency or accuracy: "there is now ample evidence to show that meaning-focused instruction (MFI) is highly effective in enabling learners to develop fluency and confidence in using an L2 [however] MFI does not guarantee high levels of linguistic accuracy. Thus, to be effective, instruction must also direct attention onto form" (Ellis & Shintani, 2014, pp. 22-23). As the main purpose of the course presented here was developing the accuracy of learner language, the focus on form seems to be a legitimate choice.

The details of the course aims, format, description and content, as well as relevant results of empirical studies guiding the choices made when designing the course are provided below.

3.1 Course aims

The main objective of the course was to develop the accuracy of advanced learner language as one of the three aspects of proficiency in language, and raise students' awareness of those areas of lexis and grammar that tend to be typically problematic for Czech and Slovak learners of English (these areas are specified below, in part 3.3).

Upon completion of the course, students were expected to be able to identify erroneous language use and provide correct alternatives. As a result, the focus on accurate and correct forms was expected to be demonstrated in an increased accuracy of their own production in the target language, both spoken and written.

Another area of concern was developing learners' independence in identifying erroneous language use outside the classroom environment by providing samples of such use and commenting on it. This was believed to help to develop learners' autonomy and increase focus on accuracy in the TL. Teamwork among the students was promoted through assignments prepared in groups and presented in class. Students received written feedback on these tasks from the teachers on the course, and learned to receive and provide peer-feedback on the content and form of these assignments. The decision to include corrective feedback was driven by recent empirical research which has reported positive influence of feedback on accuracy of learners' free production in the TL, both spoken and written (Bitchener & Ferris, 2012; Ellis & Shintani, 2014). A thorough account of the course content is provided below, in part 3.4

3.2 Course format

The form of the intervention was a one-term blended learning course – a combination of contact classes and intensive online support. The reason why a blended learning course was chosen was based on research findings published in the literature, both international and Czech.

As reported by Blake (2011), who summarised findings from multiple experimental studies, the efficacy of online learning especially in tertiary education is decisively positive. Blake's claims are among others based on a meta-analysis evaluating the effects of online learning conducted by the U.S. Department of Education (Means et al., 2010) which indicated the positive influence of web-based instruction as compared with traditional face-to-face (F2F) instruction. According to this report (Means et al., 2010) in which 50 relevant studies contrasting online and F2F instruction published from 1996 through 2008 were analysed, students of online courses outperformed those in traditional contact classes. This impact was even more profound in blended learning courses: "in recent […] studies contrasting blends of online and face-to-face instruction with conventional face-to-face classes, blended instruction has been more effective" (Means et al., 2010, p. xviii). Another important finding is related to developing the autonomy of learners through online learning, as "online learning can be enhanced by giving learners control of their interactions with media and prompting learner reflection" (Means et al., 2010, p. xvi).

Similar results were reported in the Czech context by Hubáčková (2013, pp. 167-169), who confirmed the efficacy of hybrid learning in comparison to F2F instruction and pure e-learning both in the outcomes of learning and in the positive attitudes of learners to it. Her findings are highly relevant for this research as she investigated the impact of blended learning on language instruction in the context of Czech tertiary education.

In view of these findings, choosing a blended learning course seemed to be a logical option for the intervention described in this research. The details of the course design and development are provided below.

3.3 Course description and learning resources

The course was designed as an elective one-term Moodle-based[24] language course for undergraduate students of English philology. In the pilot run of the course, conducted from February 2016 to June 2016,

24 Moodle is a flexible, freely available multilingual learning platform enhancing learner-centred approach to teaching. Its development was guided by social constructionist pedagogy. Available from https://docs.moodle.org/

students in the second and third years of their studies could enrol. For the main study, running from October 2016 to January 2017, the course was only open to students in the third term of studies. More details about the participants both in the pilot and the main run of the course, the criteria for enrolment and other details important for the research are provided below, in the research methodology sections 4.4.1 and 4.4.2.

For the reasons described above, the course was organised as a combination of face-to-face sessions held bi-weekly and individual online learning. Altogether, seven contact classes, one introductory, five regular, and one final session were held. These were complemented by five online lessons. The topics covered on the course were divided into five modules; each module consisted of one F2F and one contact class and covered the same topics. This way, students were able to prepare for in-class activities by first working online, as well as revise potential problem areas according to their needs, after these were addressed in the classroom. This format enabled them to proceed at an appropriate pace and engage in more detail in studying the topics especially relevant for them.

All study materials on the course were accessible to the students at all times. They were offered in a variety of formats in order to address different learning styles. The core course material, a revised version of Sparling's reference book (1991), was available both in the original form with Czech explanations and examples in English, and in a revised, updated and complemented version from 2015[25] with new entries, corpora-based examples and English explanations. In both these materials, the entries were organized alphabetically in order to enable easy searching. In the modules, however, a different approach was taken. All entries were grouped according to topics, rather than alphabetically, which enabled more logical sequencing. The core material was also transformed into a mind map to cater for more visually oriented learners.

In order to motivate students and better target the instruction, a list of topics to be covered in each module was announced prior to class and students were asked to select up to five topics and areas they regarded as important and challenging. Based on the poll the students voted in,

25 The updated version of the original *English or Czenglish* (Sparling, 1991) was prepared by Christopher Rance, Irina Matusevich and Simona Kalová, the Czech text was translated into English by Dušan Kolcún. The revised text was proofread by Don Sparling. This revised version has been approved for the use on the course by Don Sparling.

the areas gaining most votes were then addressed in class in a thorough way, and more materials for online self-study were provided.

Apart from the core resource material described above, one more compulsory self-study reference book for advanced learners (Hewings, 2009) was used on the course to ensure that relevant areas were given enough attention and provided students with further practice. Eight optional learning resources were recommended to learners to complement the two required reference books. For the complete list of learning resources used on the course see Appendix 4.

3.4 Course content

On the course, students received targeted input in those language features which are perceived as problematic by Czech learners of English. Apart from this, they worked with samples of authentic language and learned to identify errors and provide correct forms. Also, they were encouraged to recognize occurrences of erroneous use of the TL outside the classroom, especially in publicly available materials translated from Czech into English. They shared these occurrences in an online forum on the course, together with comments on the errors and suggestions for correct usage. The main focus was on grammatical and lexical features of language, while some pragmatic features were also addressed, e.g. the language of apologies and using appropriate register. Areas that received most attention were the use of articles and determiners, verb tenses, word order, subject verb agreement, complementation, false friends, collocations, etc. Their choice was based on both the study of the literature on learner language (Sparling, 1991; Swan & Smith, 2001; Swan, 2005) and the findings of research carried out in the Czech Republic (Gráf, 2015; 2017). In his research, Gráf identified errors in the use of articles and tenses as the most challenging areas for advanced Czech learners of English (for details see chapter 2). Other problematic language features were also complemented by the two teachers on the course, a native British speaker and a Czech teacher, both with more than 20 years of experience with teaching Czech and Slovak advanced students of English. Contributions from students were invited and the error-prone areas they chose as relevant were also addressed.

Throughout the course, students were engaged in an analysis and practice of advanced grammar and lexical points, presented by the teachers

on the course in context and in a variety of formats. Students also actively contributed to the course content by selecting areas to be addressed in class, and by preparing and delivering presentations of selected problem areas, which were presented in class for their peers, accompanied by practice. They were free to choose both the content and form of the task, and worked in groups to prepare it. Written feedback on each assignment was provided both by the teachers in the course and the other students.

All in-class activities, as well as all presentations delivered in the face-to-face sessions, were available online for students to revise from and use for further practice. In the online part of the course, students engaged in exercises in which they had to decide which sentences were grammatical and which ungrammatical. The access to these exercises was unlimited and students could take them repeatedly. The exercises were in a test format, with a final score provided to inform students on which structures they had mastered, and which required further attention. On completion of the test, students could see all their answers corrected. Detailed comments were provided on each of their answers, with explanations of the problematic language features and examples of correct usage. Such metalinguistic explanation (ME) comments have been reported to develop explicit language knowledge and improve "learners' understanding of the target structure […] as measured by a Grammaticality Judgement Test" (Shintani & Ellis, 2013, p. 290), and "ME also led to increased accuracy in a new piece of writing" (Shintani & Ellis, 2013, p. 300). In view of these findings, both Grammaticality Judgement Test and test in free written production were used in the research, as specified in chapter 4.

3.5 Summary

In this chapter, it has been described how educational intervention forming a part of the research project aimed at improving the accuracy of advanced learners of English as a foreign language was designed and executed. Some current research findings that were taken into consideration when devising the course have been briefly outlined to provide the rationale behind the choices made in the process. The chapter that follows will move on to a detailed account of the research methodology utilized in this book.

4 METHODOLOGY OF RESEARCH

Research methodology and the rationale behind its choice are detailed in chapter 4. First, the aims of the investigation are presented and research questions formulated. Then the research design, theoretical concepts influencing the choice of methods, tasks and procedures, the process of data collection and their treatment are outlined.

4.1 Research aims

The main goal of this research project is to examine the possibilities of boosting the accuracy of learner language through targeted educational intervention, which was detailed in the previous chapter. The efficacy of the intervention was explored by measuring learners' ability to identify and correct ungrammatical features of language, their certainty in answering, as well as the accuracy of their own production in the target language. These were tested prior to the intervention and after it had finished, and the pre-test and post-test scores were analysed and compared.

This research study also seeks to examine aspects of advanced learner language in terms of accuracy. To identify the main problematic areas for advanced Czech and Slovak learners of English as a foreign language, samples of spoken and written language were collected, error tagged, corrected, and a learner corpus was created. The data from the corpus were analysed as a whole and then separately for spoken and written language in order to identify the most error-prone areas in the whole corpus, and differences, if any, between the corpus of spoken and written learner language.

4.2 Research questions

The main research question this research project aims to explore is as follows:

What is the overall impact of educational intervention focused on eliminating typical errors of advanced Czech and Slovak students of English on the accuracy of their learner language?

In relation to the main focus of the study, specific research questions were formulated. The first question was designed to assess the impact of the targeted intervention on the accuracy of learner language and students' ability to identify errors and provide correct versions of erroneous use. The second question explored how the intervention affects the participants' certainty in answering. The third question seeks to determine relationships, if any, between the entry test scores, accuracy and certainty in answering. Question four is based on a corpus-based analysis of participants' spoken and written language and tries to identify whether and to what extent the intervention affects error rates in learners' free production. Questions five and six elicit information on the most problematic areas of advanced learner language analysed as a whole, and identify differences, if any, between the two modalities, written and spoken. The six research questions were formulated as follows:

RQ1) Does intervention focused on accuracy affect students' ability to identify and correct errors?

This area was addressed in the analysis of Grammaticality Judgement Test by comparing and contrasting pre-test and post-test results.

RQ2) Does intervention focused on accuracy affect students' certainty in identifying and correcting errors?

Certainty-Based Marking provided an answer to this question. In the analysis, pre-test and post-test results were analysed.

RQ3) Are there any identifiable correlations between students' entry test scores, their accuracy and certainty in answering?

Comparing the scores from entry tests, the Grammaticality Judgement Tests and Certainty-Based Marking, and analysing their relationship offered explanations to the question posed here.

RQ4) Does intervention focused on accuracy affect the rate of errors in students' spoken and written production?
This answer was elicited from the corpus-based analysis of spoken and written students' production in the experimental group, which received an intervention, and in the control group, which did not receive any intervention. These results were then compared and contrasted.

RQ5) What areas of the advanced learner language of Czech and Slovak students of English as a foreign language are most frequently affected by errors?
To identify these areas and determine the frequency of different types of errors, a corpus-based analysis of samples of spoken and written learner language was carried out.

RQ6) Do these areas differ in spoken and in written learner language? If they do, what are the main differences?
The frequency of errors was first assessed in the whole learner corpus, then separately for spoken and written students' production. Error rates and error types in the two domains, spoken and written, were examined.

4.3 Research design

Quasi-experimental methodology in the natural setting was employed in this research project. The research was conducted in two stages: first, a pilot study was carried out from February 2016 to June 2016, followed by the main study from October 2016 to January 2017.

In the pilot study, a one group pre-test-post-test design was used, and a pre-test-post-test non-equivalent group design was adopted in the main study, with two groups, experimental and control. The pilot and experimental groups received treatment in the form of educational intervention, while the control group did not receive any special treatment. This intervention has been detailed in chapter 3 above. All groups were tested twice, and their results were analysed and compared. While such design is not without flaws, it is a compromise frequently used in educational research where true experiments in controlled conditions or random allocations to groups are impossible (Cohen et al., 2011, pp. 322-323).

Three sets of data were collected in the pilot run of the study: Grammaticality Judgement Tests (GJT), Certainty-Based Marking (CBM) and questionnaires with open-ended answers. In the main study, these data were complemented with samples of learners' spoken and written production. For the outline of the research design in the pilot and main study, see Tables 2 and 3 below.

Table 2
Research design in the pilot study

PILOT GROUP
pre-test (n = 29) *(GJT, CBM, questionnaire)* *intervention* *(13-week blended learning course)* *post-test (n = 26)* *(GJT, CBM, questionnaire)* *data analysis + changes suggested for the main study*

Table 3
Research design in the main study

MAIN STUDY	
EXPERIMENTAL GROUP	*CONTROL GROUP*
pre-test (n = 32) *(GJT, CBM, questionnaire, samples of spoken and written language)* *intervention* *(13-week blended learning course)*	*pre-test (n = 16)* *(GJT, CBM, questionnaire, samples of spoken and written language)*
post-test (n = 30) *(GJT, CBM, questionnaire, samples of spoken and written language)* *data analysis*	*post-test (n = 14)* *(GJT, CBM, questionnaire, samples of spoken and written language)* *data analysis*

4.4 Research participants

The research project used a convenience sample of 70 undergraduate university students of English language and literature: 26 in the pilot, 30 in the experimental and 14 in the control group. They were all learners of EFL and their proficiency levels in English were comparable. This

was ensured by the fact that they had all passed the same standardized oral and written language exam at C1 level according to the CEFR in the second term of the bachelor's programme. The average score obtained by students in the exam was 77.62% in the pilot group (with the median value 78.71%), 82.25% (median 81.67%) in the experimental group and 82.36% in the control group (median 83.04%). Taking the median values in the entrance test into consideration when comparing the groups, some differences can be identified: the median score in the pilot group was by 2.96% lower than in the experimental group and by 4.33% lower than that in the control group. When comparing the experimental and control groups, however, the difference was a mere 1.37%. This difference was not significantly high, and it is therefore not likely that the initial proficiency levels of the participants might have significantly affected the results of this research study.

Trying to collect data in the population of the same proficiency level was of utmost importance; it has been shown by research that comparable proficiency levels are one of the key factors influencing the type and number of errors in production. As Bestgen, Granger, and Thewissen report in the conclusion of their corpus-based study of learners' errors, "differences in proficiency levels […] influence error frequency and, consequently, the subsequent discriminant analysis results" (2012, p. 146).

A detailed description of the research design and participants in the two stages of the study, pilot and main, will follow below, in sections 4.4.1. and 4.4.2.

4.4.1 Pilot study

The pilot study was performed in order to verify the overall efficacy of educational intervention and to test the reliability of the main research tools – the Grammaticality Judgement Test and Certainty-Based Marking. It was held from February 2016 to June 2016, using the one-group pre-test-post-test design. 29 students participated in the pre-test, scheduled in February 2016, then received educational intervention in the form of a 13-week course, described above in chapter 3. On completion of the course, 26 students took a post-test. The results of those students who only took one test were excluded from the overall analysis.

The first language of most participants in the pilot study was Czech – 85%, and 15% were native speakers of Slovak. Regarding the gender of participants, most of them were women 77%, and 23% were men; such gender ratio is typical in language focused degrees.

The participants of this trial run were undergraduate students majoring in English, studying the same degree programme but in different years – some of them were in the fourth term, others in the sixth or eighth terms of the bachelor's programme. This could have influenced the students' proficiency levels and consequently also their results on the test. For this reason, one of the suggested changes for the main study was to limit access to the course to third-term students only in order to eliminate different length of the study as one important influencing factor.

Another proposed change regarded Certainty-Based Marking, a tool largely unknown by students. When first administered as part of pre-testing, it seemed to be a distracting factor for some of the students and might have had a negative impact on their answers. For this reason, it was suggested that a more detailed explanation and clearer instructions before administering this part of the test were used when administering the main study.

In the pilot study, one-group pre-test-post-test design was used, and research tools were tested. After the results of the pilot study had been evaluated, several changes were suggested for the main study. The most important change proposed for the main part of the research was collecting more learner data. Apart from GJT and CBM testing, it was recommended that samples of spoken and written language should also be elicited from the participants, in order to provide a more complex picture of their learner language. The rationale guiding this decision will be specified in section 4.6. The design adopted in the main study is detailed below.

4.4.2 Main study

Resulting from the analysis of the pilot study, a pre-test-post-test non-equivalent group design was adopted in the main part of the research, with two groups, experimental and control. The original idea of selecting a randomized sample from the whole population of students in their second year of studies, approximately 120 students, and allocating them

randomly into two groups, experimental and control, was impossible due to the rules imposed by the institution in which the research was conducted. As a result, non-probability convenience sampling had to be adopted. Students chose to either enrol on a one-term blended learning course which represented targeted educational intervention, or decided to participate in the testing only, as members of the control group. Educational intervention was only implemented in the experimental group. Both groups were tested before the intervention in the experimental group and then again, 13 weeks later after the intervention had finished, and a thorough analysis of their results was performed.

48 students were recruited for the main study and participated in pre-testing held at the beginning of October 2016, with 32 students in the experimental and 16 in the control group. Students from the experimental group were then subjected to the intervention − they attended a blended learning course aimed at increasing accuracy, which was held during the winter term 2016. Students from the control group did not attend any such course. Apart from the intervention, students from both groups attended other courses as prescribed in the syllabus for the 3rd term. At the end of the term, in January 2017, all students from both groups took the post-test. Four students, two and two in each of the groups, did not participate in post-testing, so the total of 14 students in the control group and 30 in the experimental group participated in both pre-test and post-test. The results of the students who only took one series of testing were excluded from the statistical evaluation of the data but were included in the overall analysis of learner language.

In the experimental group, 70% of the participants spoke Czech as their first language, and 30% were native speakers of Slovak. Most of the respondents in this group were women – 73%, and 27% were men. In the control group, the situation was similar: 71% of respondents were native speakers of Czech, 29% of respondents were Slovak native speakers. There were slightly more male respondents in this group: 36%, and 64% female.

In order to ensure the comparability of the two groups and make it possible to assess the impact of the intervention, primary inclusion criteria for either of the groups were specified, as described below. All participants:

- were undergraduate students of English philology;
- were in the third term of the bachelor's programme;
- had the same level of language proficiency (for details, see section 4.4).

4.5 Research data collection

In order to gain reliable and varied information about learners' knowledge of language, it is advisable to use different methods of data collection. These can be roughly divided into three main areas: "(1) non-linguistic performance data, (2) samples of learner language, and (3) reports from learners about their own learning" (Ellis & Barkhuizen, 2005, p. 15).

In the research presented in this book, both naturalistic data, samples of learner spoken and written language, and learners' reports about their learning were collected. A learner corpus was compiled from analysed essays and recordings of learner language, and learners' intuitions about language were explored using a Grammaticality Judgement Test. As a result, within the research project, four sets of quantitatively analysed data have been gathered: GJT and CBM on the one hand, and students' essays and recorded interviews providing samples of their spoken and written production on the other. In addition to these datasets, students also answered questionnaires with open-ended questions in order to provide contextual information.

A combination of these different types of data meets the requirement of *triangulation* in educational research and thus provides a multifaceted picture of the phenomena under scrutiny: "triangulation methods in the social sciences attempt to map out, or explain more fully, the richness and complexity of human behaviour by studying it from more than one standpoint, […] by making use of both quantitative and qualitative data" (Cohen et al., 2011, p. 195). Most of the data collected in this research were analysed quantitatively and were complemented by information from close and open-ended questionnaires.

In the pilot stage of the research, only GJT and CBM were administered. In the main study, these two tests were complemented with data from interviews and short essays on everyday topics appropriate for this level as samples of spoken and written learner language. Invaluable additional information for the qualitative part of the research was elicited

from answers given by the participants in questionnaires. These focused especially on the participants' motivation to take part in the research, on their perception of how the accuracy of their language had changed in the course of the research, and, in the experimental and pilot groups, also on their beliefs about the effectiveness of the intervention.

Samples of learner language were collected from participants in the experimental and control groups, not from those in the pilot study. Two types of tests were administered: a test in speaking and a test in writing. Both tests were taken twice by each student in the control and experimental groups, before the intervention as a pre-test, and after the intervention as a post-test. Pre-testing, in which 32 students in the experimental and 16 in the control group participated, was conducted in October 2016. After the pre-test, participants from the experimental group received educational intervention, while those in the control group did not. The intervention was a 13-week blended learning course which combined bi-weekly held 90-minute contact lessons with intensive online support for individual study and further materials for practice. The course aims, design, format and other relevant facts are detailed in chapter 3 above.

Three months after the pre-test, in January 2017, after the intervention had finished, a series of post-testing was conducted. The format of the tests was the same as in the pre-test. 30 students in the experimental and 14 in the control group participated in the post-test.

A thorough description of the research tools together with the underlying theoretical principles guiding their choice for this study will be provided in the following section.

4.6 Research tools and underlying theoretical concepts

This subchapter is concerned with the research tools used in this project. While Grammaticality Judgement Tests and Certainty-Based Marking provide relevant information on specific features of learner language which is hard to elicit otherwise, it is equally important to analyse samples of free spoken and written learner production. This is in accordance with Ellis and Barkhuizen's claim that "the primary data for investigating L2 acquisition should be samples of learner language" (2005, p. 21) because "production is seen as providing the clearest evidence of what a learner has acquired" (ibid).

Samples of learner language collected in the study were used to compile a learner corpus. Throughout the research project, this corpus has become a very important research tool, enabling a thorough analysis of learner language. The process of creating this corpus as well as the other research tools are detailed below.

4.6.1 Grammaticality Judgement Test

Grammaticality Judgement Tests, in which learners are asked to distinguish between grammatically correct and incorrect sentences, have been widely used in second language acquisiton research (Ellis & Barkhuizen, 2005, p. 17), often, however, with conflicting views on their reliability: "research to date on the reliability issue of L2 Grammaticality Judgement Tests have yielded mixed results" (Han, 2006, p. 61). Recent research into the reliability of GJT methodology has indicated, however, that if certain conditions are met, the tests can yield both reliable and valid data. These conditions can be summarised as follows:

- test-takers should not merely decide on the grammaticality of the sentences, they also need to provide correct answers (Ellis & Barkhuizen, 2005, pp. 18-19; Han, 2006, p. 64);
- a time limit should be imposed on the tests (Ellis & Barkhuizen, 2005, p. 18);
- Grammaticality Judgement Tests should be used with learners whose proficiency in the TL is high, in which case GJT provide more consistent data than with learners of low proficiency (Ellis & Barkhuizen, 2005, p. 20).

All the above-mentioned principles have been taken into consideration in the current study: participants were asked to first decide whether the sentences were correct or incorrect, and then they had to provide correct versions of those they believed were erroneous. A time limit of 35 minutes was imposed on the test. All test-takers were proficient users of English, ranging from C1-C2 levels according to the CEFR as they had all passed a proficiency exam of that level (for details, see part 4.4 above).

It can be assumed that Grammaticality Judgement Testing is an appropriate research tool for this study. This claim can be further strengthened by the fact that unlike samples of free production, GJT can effectively test knowledge of strictly defined areas of language, especially its advanced features, which learners often tend to avoid and the knowledge of which is therefore rather difficult to elicit (Han, 2006, p. 62). A Grammaticality Judgement Test, in combination with the analysis of samples of learners' free production, promises to provide enough information about advanced learner language.

The test was based on the study of the relevant literature, both Czech (Sparling, 1991) and international (Swan, 2005), focused on typical errors of advanced learners of English. It also drew on the results of published research into the language of advanced learners of English conducted in the Czech Republic (Gráf, 2015; 2017) as detailed in chapter 2.2 above. In his corpus-based study, Gráf worked with students of English philology; it can therefore be assumed that the problems he identified can be expected in the language of a similar population of students with the same L1. The test was designed as a combination of 30 correct and incorrect sentences. Test-takers were first asked to express their intuitions about their grammaticality, deciding whether a sentence was correct or incorrect, and in the latter case correct the error and provide a correct version. Ten test items dealt with errors in the use of articles, five with countability issues, five with erroneous use of lexis, especially typical 'false friends' and collocations, wrong prepositions, verb tenses and word order. The full version of the test is provided in Appendix 5.

In order to assess the reliability of the test, demonstrating its internal consistency as a research tool, Cronbach's alpha was used (Cohen et al., 2011, p. 201, 640). Internal reliability was analysed in each group, pilot, experimental and control, and each time, pre-test and post-test. The analysis was computed using SPSS 25.0.0.1 (IBM Corporation, 2019)[26]. The results, as shown in Table 4 below, indicate that the internal consistency of the Grammaticality Judgement Test was ranging from an alpha coefficient of 0.840 to 0.919 which is very high. The test appears to be a reliable research tool, suitable for use in the research project.

26 The analysis of test reliability was conducted by RNDr. Jiří Jarkovský, Ph.D., Masaryk University, Faculty of Medicine, Institute of Biostatistics and analyses.

Table 4
Internal reliability of GJT in all groups and pre-test and post-test

Cronbach's alpha	Pilot (n = 29)	Experimental (n = 32)	Control (n = 16)
pre-test	0.840	0.909	0.886
Cronbach's alpha	Pilot (n = 26)	Experimental (n = 30)	Control (n = 14)
post-test	0.919	0.886	0.900

4.6.2 Certainty-Based Marking

Another research tool adopted to complement Grammaticality Judgement Testing is Certainty-Based Marking (CBM). It has been introduced in testing in order to get information about how sure respondents are about the correctness of their answers and to eliminate unnecessary risk-taking in answering. It requires test-takers to not only answer the question as accurately as possible, but also to express how sure they are about their answers being correct. A 3-level scale is adopted, ranging from 1 for the lowest level of certainty, 2 for middle, and 3 for the highest level. Correctly answered questions with high level of certainty are rewarded, e.g. a correctly answered question for which a highest level of certainty is chosen gains 3 points, and incorrectly answered questions with high level of certainty mean losing points, e.g. an incorrectly answered question with the highest level of certainty means a loss of 6 points. This marking scheme is illustrated in Table 5 below.

The use of CBM in testing seems to be beneficial for a number of reasons, as reported by Gardner-Medwin & Curtin: "CBM differentiates between different students who give the same answers in a test […] they stimulate deeper learning by the fact that students need to prepare thoroughly" (2007, pp. 3-4). Another reason in favour of CBM marking is its practicality: it is easily implemented with existing tests and is one of the tools available in Moodle-based courses.

Table 5
Marking scheme for CBM (Gardner-Medwin & Curtin, 2007, p. 1)

Degree of certainty	C = 1 (low)	C = 2 (middle)	C = 3 (high)	No reply
Mark if correct	1	2	3	0
Penalty if wrong	0	−2	−6	0

A major advantage of this relatively infrequently used method is minimising the level of guessing in answering during tests. Despite initial difficulties during the first series of testing, it has proven out to be an effective tool in this research, especially as raising the awareness of typical problem areas in language and therefore increased certainty in answering was one of the foci of the present study.

4.6.3 Corpus-based analysis of samples of spoken and written language

In order to provide a complex picture of learner language, apart from GJT and CBM marking, samples of free learner production are necessary. Such samples can be either language used naturally in real-life situations, which is both difficult and impractical to obtain, or samples elicited by researchers under controlled conditions. The level of control depends on whether the research is focused on eliciting general language or examples of strictly defined structures: "in the case of a general sample, the elicitation instrument is designed to provide a context for learners to speak or write in the L2 in a purposeful manner" (Ellis & Barkhuizen, 2005, p. 30). Data collection for this part of the study is detailed in part 4.5. The essays as samples of written language and transcribed oral interviews were used to create a corpus of learner language which was analysed for errors.

Unlike native language corpora, learner corpora require a different approach to how the data are treated (Granger et al., 2002, p. 18). Before computer-aided error analysis of learner language can be carried out, two important elements are required: errors must be annotated and a learner corpus based on the annotated data and metadata compiled. Approaches to error tagging differ, depending on what criteria are chosen to classify errors: "One major decision to make is whether to tag errors in terms of their nature (grammatical, lexical, etc.) or their source (interlingual, intralingual, etc.). The former is arguably preferable in that it involves less subjective interpretation and is therefore likely to be applied with greater consistency and reliability by different analysts" (Granger et al., 2002, p. 19). In order to maintain the objectivity of the analysis and support the consistency of tagging, which is vital if a team of assessors work together, errors in this analysis were tagged by their nature.

Prior to the analysis, two important decisions had to be taken; how accuracy would be measured, and what norm would be applied in error correction. From a number of approaches to measuring the accuracy of learner language, a general measure of accuracy rather than the use of specific measures was adopted in this study. Accuracy was expressed by the number of errors per 100 words, as it is believed to be a reliable tool, generating required information about learner language (Ellis & Barkhuizen, 2005, pp. 150-151). More details about the operationalising and measuring accuracy in this research project are provided above, in section 1.4.1.

Despite the controversy related to the traditional native-speaker norms, as described in part 1.5, they will be used as a reference norm in the present study. The reasons for this choice are as follows: as the participants of the research are students of English philology who are likely to work as English language professionals in the future, they naturally aspire to get as close to the NS norms as possible. This assumption was validated by the results of a survey among the participants of the study prior to the analysis of the collected data, in which 73% of respondents expressed their preference for native-speaker norms (Kalová, 2017).

4.6.4 Questionnaires providing contextual information

One of the caveats of (quasi)experimental design in educational research is the impossibility of excluding factors which might influence the observed change but are beyond the researcher's control. Being unable to exclude these extraneous variables might affect the validity of such research (Cohen et al., 2011, pp. 322-323). With this in mind, two questionnaires were devised in order to elicit relevant contextual information about factors which might potentially affect the results of the research. These questionnaires were distributed to the experimental and control groups in the main study, not in the pilot group.

The first questionnaire, administered online immediately after the pre-test, was identical for both groups and consisted of one open-ended question. It was designed to find out why students in the experimental group had decided to join the course and why students in the control group had decided to participate in testing.

The second questionnaire, administered online to all participants in both groups immediately after they had finished the post-test, consisted of 18 questions, 14 open-ended and 4 multiple-choice questions. (For the full versions of both questionnaires, see Appendix 8). Students were asked to compare the pre-test and post-test in terms of content and difficulty and provide comments at both times on all parts of the test, Grammaticality Judgement Test, essay and the oral interview. They were also asked to describe what they had done in the three-month period between pre- and post-tests to improve their English language skills. The data elicited from the analysis of the two questionnaires provided important contextual information and complemented the data elicited from the analysis of the GJT, CBM, essays and oral interviews.

4.7 Data analysis

In this subchapter, the analysis of all data collected in this study is presented. In order to assess the influence of the intervention, the scores gained in the pre-test were compared to those in the post-test and the differences for each group of participants were analysed. Based on the study of the relevant literature (Ellis & Barkhuizen, 2005, p. 57), the following steps were taken in the analysis:

1) collection of a sample of learner language;
2) identification of errors;
3) description of errors;
4) explanation of errors;
5) evaluation of errors.

4.7.1 Analysing Grammaticality Judgement Test and Certainty-Based Marking

Grammaticality Judgement Testing and certainty-based marking scores are features of a computer-administered test in Moodle, as detailed above, in sections 4.6.1 and 4.6.2. The answers were automatically graded by Moodle but were then checked again by the researcher and some of the answers were regraded, e.g. in case of obvious typos or erroneous punctuation which were not tested in the quiz. Also, because

CBM marking was used, Moodle automatically generated a combination of the two marks – accuracy and certainty – as a final mark in the test. For the purpose of this study, it was necessary to analyse the two scores separately in order to be able to determine what changes occurred in the span of one term, i.e. between the pre- and post-tests, in accuracy and certainty, and to assess what the relationship between the two variables is.

Binary distinction, correct – incorrect, was used in the analysis of accuracy of the Grammaticality Judgement Test and the scores were expressed in percentages, with 100% for absolute accuracy. The values of Certainty-Based Marking were in an interval from −6 to +3 points for each answer (for details see Table 5 above). The scores of those participants who only took one of the tests were not included in the analysis.

4.7.2 Analysing samples of spoken and written production

For a more thorough exploration of learner language, samples of learners' spoken and written production in the TL were collected. To elicit these samples, the participants in the experimental and control groups were asked to write a short essay and take part in an oral interview. These oral and written tests were part of the pre-test as well as the post-test, as detailed above in Table 3. The elicited data were used to compile a learner corpus.

Topics for the pre-test and post-test written part were chosen from a list of freely available topics prepared by experts from International Language Testing System (IELTS)[27]. The main reason for this choice was that these tests are standardized and offer topics of comparable difficulty, appropriate for the level of proficiency of the participants, C1 according to the CEFR. Both written tasks were argumentative essays, with a minimum required length of 250 words. The submitted texts varied in length, from 137 to 573 words, with the average length of 294 words. The time limit imposed on the written task was 30 minutes. A detailed account of the written test is provided in Appendix 7.

The oral part of the task was modelled on a standardized oral test at C1 proficiency level according to the CEFR and the questions used were adapted from the materials published by Cambridge Assessment

27 https://www.ielts.org/

English[28]. The interview was taken by pairs of students, in case of odd numbers three students were interviewed at the same time. The tasks included both individual turns and an interactive task. The interlocutor, the Czech teacher on the course and the researcher, asked questions and provided visual prompts but did not interfere any further in the answers, so that the respondents were not limited in their production and so that the whole process was as close to a real-life situation as possible. All materials used in the interview are provided in Appendix 6. The interviews were video recorded in a studio, transcribed and analysed for errors. The average length of a transcribed interview was 719 words, ranging from 393 to 1,369 words. The analytical process will be addressed in greater detail in part 5.5 below.

Altogether, 92 essays as samples of written language, and 84 transcripts of interviews as samples of spoken language were analysed. All data were strictly pseudonymised[29] – each participant was assigned a unique code which was a combination of letters. The first letter indicated the group, E for the experimental and C for the control group. The following letters in the code indicated the respondent. All samples were analysed for errors and used to compile a learner corpus of spoken and written learner language. The corpus-based analysis of spoken and written learner data elicited from the recordings and essays will now be described in more detail.

Error Tagging

The Louvain error-tagging system was adopted in the analysis. This decision was taken early on in the project, as Louvain tagging seemed to meet the requirements of the analysis very well. Not only is this system well described in the literature (Dagneaux et al., 1998; Bestgen et al., 2012, pp. 155-153) but it is also flexible and versatile, and new categories can easily be added. The system has also been used in other studies on learner language, (for example Götz, 2015; 2019; Gráf, 2015; 2017), so it enables comparing the results.

28 https://www.cambridgeenglish.org/exams-and-tests/advanced/

29 Due to the nature of the present research, strict anonymisation of the data was not possible during the analytical process. The identities of the participants were only known to the researcher throughout the analysis. All participants signed informed consent by which they agreed to have their data analysed by the researcher. In the analysis, all data were pseudonymised according to a key. Once the analysis is complete, this key will be destroyed and all data will thus become strictly anonymised. This is important in order to avoid any possible negative repercussions for the participants.

Table 6
The Louvain error tagging system, adapted and complemented (Bestgen et al., 2012, p. 130)

Error domain	definition	description	tag	example
F	Formal errors	Spelling or morphological errors	FM	*unpossible
G	Grammatical errors	Errors that break the general rules of English grammar, e.g. in the use of articles	GAMD	He is *teacher.
L	Lexical errors	Errors involving the semantic properties of words and phrases, e.g. conceptual, collocational or connotative	LS	*a university absolvent
X	Lexico-grammatical errors	Errors that violate the lexico-grammatical properties of words, e.g. erroneous dependent prepositions, noun countability, complementation patterns	XVPR	they stare *to their phones
Q	Punctuation errors	Errors that target punctuation, e.g. missing or redundant punctuation markers	QR	he knew*, that she was wrong
W	Word redundant/ missing/ order errors	Unnecessary, missing or misordered words	WO	principles of how *does everything work
S	Style errors	Sentence fragments and incomprehensible or too long sentences	SU	* we are not prepared to get into the real environment
Z	Infelicities	Inappropriate register, problems with stylistics	ZIR	*tons of books

The system is hierarchical – the tag consists of a chain of letters indicating the type of error rather precisely; the first letter – error domain – provides general information about the area affected by the error, and the following letters contain more specific information about the nature of the error. In the original tagging system there were 8 main domains of errors:

F − formal, G − grammatical, L – lexical, X – lexico-grammatical, Q – punctuation, W − word redundant, missing or errors in word order, S – style, and Z – infelicities, to mark errors in inappropriate register, problems with political correctness and stylistics (Bestgen et al., 2012, p. 130; Gráf, 2015, p. 76). The following letters in the tag provide more detailed information about the error, e.g. the tag GAMD indicates an error in grammar (G), erroneous use of the article (A), which is missing (M) and it is the definite article that should have been used (D). The complete set of the main error domains used in the study together with their definitions, descriptions and examples is provided in Table 6 above. For the complete list of all 73 error types used in this learner corpus, see Appendix 9.

The exact error tagging procedure also followed the steps recommended by the Louvain Centre for English Corpus Linguistics (Dagneaux et al., 1998, p. 165). All collected data were first manually marked for errors and problematic areas. The assessors were two native speakers, both of British origin and both with a long experience in teaching English as a foreign language in the Czech Republic. Dagneaux et al. suggest that in order to conduct a comprehensive and in-depth analysis of learner language, a team of "ideally two researchers – native and non-native – should work in close collaboration" (1998, p. 165) because as they claim, "a bilingual team heightens the quality of error correction" (1998, p. 165). They also recommend that the non-native expert has a very good command of the target language and shares the same L1 background as the learners whose language is analysed. Based on these guidelines, two native speakers and a non-native speaker, a C2 proficient user of English, whose L1 is Czech, formed a team. They met once a week for a period of one year, from June 2018 to June 2019, in order to discuss all the errors and their corrected versions. These meetings were aimed primarily at eliminating inconsistencies in error correction and ensuring inter-rater reliability. For the same reason, a list of rules which all assessors strictly adhered to was created based both on the assessors' experience and guidelines from the literature (Štindlová & Čurdová, 2015, pp. 196-199). The rules for error annotation were formulated as follows. In the analysis, it is important:

- to correct errors clearly and unambiguously;
- to take context into consideration;

- to apply the minimum intervention rule; the aim of the analysis is to identify and correct erroneous forms and lexis in learner language, not to provide stylistically perfect formulations close to the target language as if they were produced by a native speaker;
- to prevent bias in annotation; obvious typos should not be regarded as errors, e.g. **theoretcal* − theoretical, **imrpovement* – improvement. Those errors which manifest a lack of knowledge, e.g. spelling, **aproach* – approach, or morphology, **unpossible* – impossible, count as errors of the appropriate category;
- to consider overlapping errors carefully; the most likely correction and error tag should be used, i.e. the one all three assessors agree on;
- to consult the English Web 2015 corpus in Sketch Engine[30] and / or Google Books Ngram Viewer[31] to verify the frequency of the correct option in case of a disagreement among the assessors.

Strictly adhering to the above outlined guidelines helps to make the corpus-based analysis of learner language as objective as possible and produce reliable results.

Compiling Learner Corpus

In order to be able to conduct a detailed analysis, annotation of the learner data needs to be carried out. Traditionally used techniques of annotation need to be combined with new ones reflecting the special nature of learner data (Granger et al., 2002, p. 18). Learner corpora annotation usually consists of two stages, error-tagging, i.e. assigning the error codes, and emendation, a process in which correct versions of the erroneous part of learner language are added (Štindlová, 2011, p. 5). Once all the data were tagged for errors and corrections added, the annotated text was used to create a learner corpus. For this purpose, Sketch Engine was used as a corpus-building tool. It enables users to accurately search and filter queries in language corpora. Its functions are based on mathematical and statistical computations (Baroni et al., 2006; Kilgarriff et al., 2015). It also makes it possible to create one's

30 https://app.sketchengine.eu/

31 https://books.google.com/ngrams

own corpus from either data downloaded from the internet, or from any texts that are to be analysed. The main reasons, both methodological and practical, for the choice of this research tool in the present study are summarised below.
Sketch Engine:

- enables using one's own data to create a learner corpus;
- enables using tags to refine searches;
- makes it possible to retrieve both linguistic and metalinguistic information;
- is suitable for carrying out quantitative analyses;
- provides both online and personal technical support;
- provides free access to university students, teachers and researchers.

To enable a thorough analysis of the collected data, in addition to the samples of learner language, metadata were also inserted in the compiled corpus. These provide additional information vital for a detailed analysis and also enable comparisons of the two groups of respondents participating in the study. They inform the researcher not only about the documents compiled in the corpus: the type of document – spoken or written, type of test – pre-test or post-test, types of errors, types of corrections; but also about the respondents: the group each respondent belonged to – experimental or control, the respondents' L1, their gender, and unique codes that enable the identification of each individual respondent. This way, the analysis can be conducted from a variety of different perspectives.

A learner corpus containing 106,013 words was compiled based on the data elicited from the students in the experimental and control groups. Altogether, 176 documents were inserted in the corpus, providing samples of both written – 92 essays, and spoken learner language – 84 transcripts. The compiled learner corpus is relatively small, but it must be taken into consideration that the process of transcribing the recorded interviews, correcting errors, error-tagging and creating a corpus from one's own data is very complex and demanding in terms of time, especially for a sole researcher. As Granger et al. claim, however, "there is also great value in collecting smaller in-house

corpora" (2002, p. 27) compiled by teachers from their students' work. Such learner corpora can then be used for a variety of purposes in the language classroom, from creating tailor-made study materials to error analysis carried out by teachers and/or students. It is therefore believed that despite its relatively small size, the learner corpus compiled in the present study will generate some valuable insights into learner language of Czech and Slovak students of English.

4.8 Statistical analysis

The primary purpose of the statistical analysis was to describe the samples themselves (pilot group, experimental group and control group) via descriptive statistics, and to show targeted educational intervention effects on learners' outcomes. In addition, in order to generalise findings to a hypothetical population of all comparable students, statistical tests were performed. The design of the research required tests developed for two different settings. The first group of tests examined the intervention effect on the identical group of students. For this setting (pre-intervention and post-intervention outcomes) either the paired t-test or Wilcoxon signed-rank test were applied[32].

The t-test is a parametric statistical test frequently adopted in a pre-test-post-test experimental design in order to determine whether the differences between the means of two groups are statistically significant. It uses "parametric data drawn from random samples with a normal distribution" (Cohen et al., 2011, p. 642), and has two different variants (ibid). If the two groups are not related, the t-test for independent samples is applied; while the paired t-test, also referred to as the t-test for related samples, is adopted when "the same sample group is measured on two occasions (e.g. the pre-test and the post-test) […] or the same variable is measured at two points in time" (Cohen et al., 2011, p. 644). In this research, the dependent variable under scrutiny was the accuracy of learner language and its changes at two different times were examined individually for the pilot, experimental and control groups. The two tests, pre-test and post-test, were conducted approximately three months

32 STATISTICA software package was adopted in the statistical analysis. The analysis was conducted in collaboration with doc. Mgr. Maria Králová, Ph.D., Masaryk University, Faculty of Economics and Administration Department of Applied Mathematics and Computer Science.

apart. Such a delay is believed to be long enough to produce reliable results even when the same test is used (Šamalová, 2018, p. 108).

The Wilcoxon signed-rank test is adopted as a non-parametric alternative to the paired t-test, used for a statistical analysis of ordinal data applied when the differences between pairs of data are not normally distributed. The test is able to detect even slight differences between the two related samples and is applied for repeated measurements (Chráska, 2016, p. 85; Cohen et al., 2011, pp. 655-657).

When comparing two independent samples (e.g. experimental group and control group), another group of tests is to be used, and in the analyses, two-sample t-tests were conducted. Also, correlation tests to evaluate the significance of correlation coefficients between pairs of examined variables were performed.

4.9 Summary

In chapter 4, an account of how the research was conducted is given. A detailed description of the research aims, methods, questions, design, and participants is provided. The aim of the research was to explore advanced learner English in terms of accuracy and identify its main error-prone areas. Quasi-experimental research design was used to evaluate the impact of educational intervention aimed at improving accuracy of learner language. The process of collecting samples of authentic spoken and written language from the participants, 70 undergraduate students of English, and their analysis were detailed. In the analysis, a variety of research tools was utilized: Grammaticality Judgement Test, Certainty-Based Marking and corpus-based analysis of learner language. Descriptive statistical methods were applied in the analysis. The elicited information was complemented with questionnaires, providing relevant contextual information. In the chapter that follows, the results of the pre- and post-test analyses of GJT and CBM, as well as corpus-based analysis of advanced learner language will be presented.

5 RESEARCH RESULTS AND ANALYSES

In this chapter, all the results generated by the research are presented and described in detail. First, the impact of educational intervention on the accuracy of advanced learner language is examined by comparing the results in the pre-tests and post-tests conducted in all groups of participants – pilot, experimental and control. In the analysis, the scores reached in the Grammaticality Judgement and Certainty-Based Marking tests are compared, and a statistical analysis conducted separately for each group is presented. An account of the results in the pilot group is provided first, followed by a description and analysis of the results in the experimental and control groups. The comparison of GJT and CBM scores in the three groups is outlined and statistically analysed. The results in speaking and writing in the experimental and control groups are compared and a statistical analysis is performed. Contextual information from the questionnaires distributed among students in the experimental and control groups is provided in part 5.4.5.

In the final part of this chapter, the results of the corpus-based analysis of the samples of written and spoken learner language elicited from the students in the experimental and control groups are detailed. The major aim of this analysis is to identify the main problem areas in the advanced learner language of Czech and Slovak speakers of English, the main error-prone language domains and the most frequent error types. Differences, if any, between spoken and written learner language are detected. A summary of all results emerging form the analyses is provided in the last section, 5.6.

5.1 Results in the pilot group

The first part of the analysis consisted in comparing the results in GJT and CBM pre- and post-test scores for 26 participants in the pilot group.

In addition to descriptive statistics of the results, inferential statistical analysis was conducted and all the relevant tests were performed at the $p = 0.05$ level of significance.

5.1.1 Grammaticality Judgement Test results in the pilot group

First, the difference between pre-test and post-test results in the Grammaticality Judgement Test was calculated. Due to the relatively symmetrical distribution of the difference data (see the histogram in Figure 2 below), a normal distribution of the data is plausible, and thus the t-test for dependent samples was adopted in the analysis. The scores in the Grammaticality Judgement Test were expressed in per cent, with 100% for absolute accuracy on the test with all answers correct. After the intervention, the value of GJT in the pilot group increased by 23.21, from the average score of 33.19 in the pre-test to 56.41 in the post-test, which indicates a positive impact of the intervention (for details, see Table 7 below).

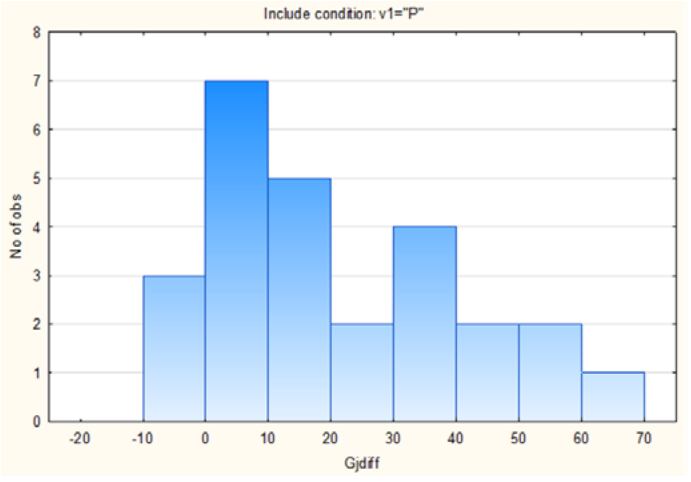
If we tried to generalise this average result from the pilot group sample to the whole population of Czech and Slovak students of English philology, the difference between pre-test and post-test would be statistically significant at the 0.05 level of significance, with $p = 0.000002$. The p-value was calculated for a two-tailed paired t-test. If we were to test that as a result of the intervention the scores in GJT increased, this increase would be statistically significant at $p = 0.000001$ level of significance (one-tailed paired test).

Table 7
Comparison of GJT pre-test and post-test results for the pilot group

	Marked differences are significant at $p < 0.05000$ Include condition: V1: "P"									
Variable	Mean	Standard deviation	N	Difference	Standard deviation difference	t	Diff.	p	Confidence −95%	Confidence +95%
GJ2	**56.41**	**20.61**								
GJ1	**33.20**	**10.31**	**26**	**23.22**	**19.51**	**6.07**	**25**	**0.000002**	**15.34**	**31.09**

Note. GJ1 refers to the Grammaticality Judgement Test taken as a pre-test; GJ2 refers to the Grammaticality Judgement Test taken as a post-test

Figure 2
Histogram of GJT scores difference in pre-test and post-test for the pilot group



5.1.2 Certainty-Based Marking results in the pilot group

The values of Certainty-Based Marking ranged from −6 to +3 points for each answer (for details see section 4.6.2 above). The scores of those participants who only took one of the tests were not included in the analysis. Again, the difference between pre-test and post-test results in Certainty-Based Marking was calculated first.

When analysing CBM scores in pre- and post-testing in the pilot sample, the Wilcoxon signed-rank test for two related samples was applied. This was due to the fact that the difference in CBM scores in pre- and post-testing was far from a normal distribution, as shown in the histogram in Figure 3 below, and the sample was relatively small (n = 26). As a result, a paired t-test could not be applied in the analysis. In the signed-rank Wilcoxon test which was applied in the analysis, median pre- and post-test scores were used instead of mean scores. The median score in the pre-test was −0.25, and in the post-test, conducted after the intervention, it reached 0.78, with the median difference of 1.03. Based on the Wilcoxon test, this difference is statistically significant with p = 0.000097.

Apparently, Certainty-Based Marking in the pilot group before and after the intervention improved in a statistically significant way, as can be seen from the increased median after the intervention as opposed to the median value before the intervention. If we wanted to hypothesise that certainty in answering had increased, which occurred in the sample, this increase at the whole relevant population would be statistically significant at the p = 0.0000485 level of significance.

Figure 3

Histogram of CBM scores difference in pre-test and post-test for the pilot group

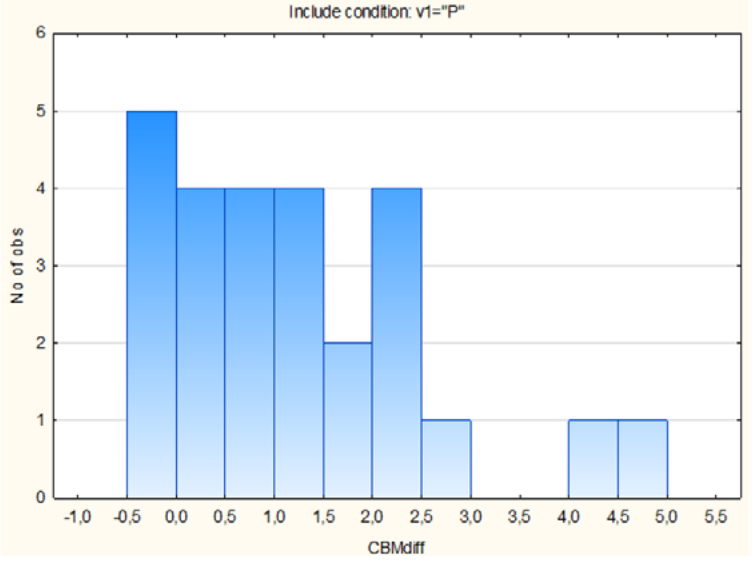


Table 8
Comparison of CBM test and post-test results for the pilot group

	Include condition: V1: "P"								
Variable	Valid N	Mean	Median	Minimum	Maximum	Lower Quartile	Upper Quartile	Quartile Range	Standard Deviation
CBM2	26	0.9092	0.78	−1.60	4.00	0.03	1-63	1.60	1.173507
CBM1	26	−0.3419	−0.25	−1.57	0.50	−0.70	0.07	0.77	0.54784
CBMdiff	26	1.2511	1.03	−0.43	4.57	0.26	2.03	1.77	1.309874

Note. CBM1 refers to the Certainty-Based Marking test taken as a pre-test; CBM2 refers to the Certainty-Based Marking test taken as a post-test, CBM diff refers to the calculated difference between the two scores, CBM1and CBM2

5.1.3 Correlations between selected variables in the pilot group

The results of the correlation analysis of three variables, the proficiency level of knowledge of the participants at the entry, Grammaticality Judgement Test scores differences and Certainty-Based Marking scores differences are presented in Table 9, with statistically significant result in bold print.

As can be seen from the table below, a very low correlation of r = 0.07 between the entry score and GJT difference indicates that there is almost no association between the initial level of knowledge as manifested in the entry score and GJT scores. Interestingly, a positive sample correlation of 0.25 was found between CBM difference and entry score. It can be seen that in the pilot sample, CBM increase positively correlated with entry score; in other words, the higher the scores in the entry test, the bigger the increase in CBM. However, this correlation was not found to be statistically significant. The correlation between the entry score and CBM difference is not significant, with r = 0.25.

There is a statistically significant positive correlation of r = 0.66 between GJT difference and CBM difference scores. Hypothetically, if we wanted to generalise this result to the whole population, this correlation at the p = 0.05 level of significance would be statistically significant with p = 0. This means that the higher the improvement in GJT scores, the higher the improvement in CBM scores.

Table 9

Correlations between entry score, GJT difference and CBM difference in the pilot group

	Marked correlations are significant at p < 0.05000 N=26 (Case wise deletion of missing data) Include condition: V1: "P"				
Variable	Means	Standard Deviation	Entry score	GJT difference	CBM difference
Entry score	77.6192	9.8418	1.0000	0.0707	0.2503
GJT difference	23.2154	19.4955	0.0707	1.0000	**0.6579**
CBM difference	1.2512	1.3099	0.2503	**0.6579**	1.0000

Table 10
Correlations between entry score, GJT difference and CBM difference in the pilot group and the relevant p-values

Variable	Marked correlations are significant at p < 0.05000 N=26 (Case wise deletion of missing data) Include condition: v1: "P"		
	Entry score	GJT difference	CBM difference
Entry score	1.0000	0.0707	0.2503
	p = ---	p = 0.731	p = 0.218
GJT difference	0.0707	1.0000	**0.6579**
	p = 0.731	p = ---	**p = 0.000**
CBM difference	0.2503	**0.6579**	1.0000
	p = 0.218	**p = 0.000**	p = ---

Unlike the experimental and control groups, no data related to the written and spoken production were elicited from the students in the pilot group.

5.2 Results in the experimental group

The data collected in the experimental group were analysed using the same methodology as in the pilot group. First, the scores in GJT and CBM pre- and post-testing for 30 participants in the experimental group were compared. This analysis was followed by a presentation of possible associations between the variables. The accuracy of spoken and written production of the students in the experimental group is the last to be described in this subchapter.

5.2.1 Grammaticality Judgement Test results in the experimental group

After the intervention, the mean GJT score in the experimental group improved: while the mean GJT score in the pre-test was 40.78, the mean score achieved in this group in the post-test was 46.89, with an increase of 6.107. If we wanted to generalise this result to the whole population of comparable students, the improvement (post-test minus pre-test is a positive number) would be statistically significant with p = 0.00378/2 = 0.00189. The p-value was calculated using the paired t-test for dependent samples. For details, see Figure 4 and Table 11 below.

Figure 4

Histogram of GJT scores difference in pre-test and post-test for the experimental group

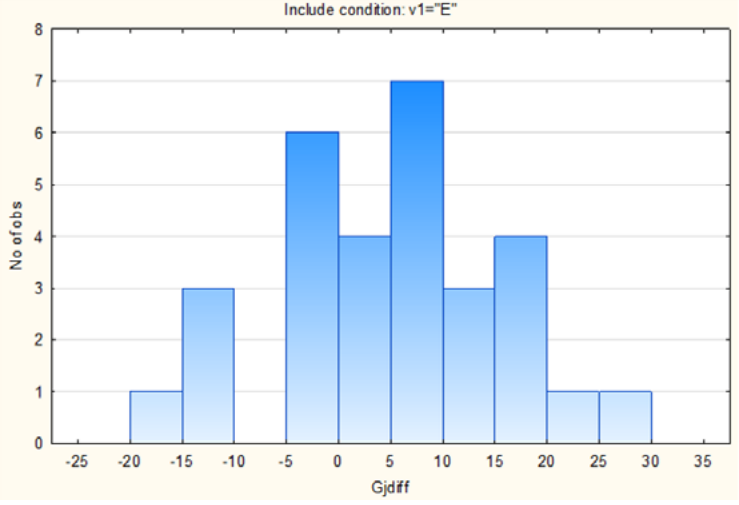


Table 11
Comparison of the GJT pre-test and post-test results for the experimental group

	Marked differences are significant at $p < 0.05000$ Include condition: V1: "E"									
Variable	Mean	Standard deviation	N	Difference	Standard deviation difference	t	Diff.	p	Confidence −95%	Confidence +95%
GJ2	**46.8900**	**13.8721**								
GJ1	**40.7833**	**15.7791**	**30**	**6.1067**	**10.6227**	**3.1497**	**29**	**0.0038**	**2.1401**	**10.0732**

Note. GJ1 refers to the Grammaticality Judgement Test taken as a pre-test; GJ2 refers to the Grammaticality Judgement Test taken as a post-test

5.2.2 Certainty-Based Marking in the experimental group

Following the intervention, CBM scores increased in the experimental group, from a negative value of −0.169 to a positive result of 0.113. The average increase was 0.283. With normal distribution of the difference data, see Figure 5, the paired t-test was used in the analysis. If we wanted to show that certainty in answering increased, this conclusion when generalised to the whole comparable population of students would be statistically significant at the level of significance $p = 0.017663/2 = 0.00883$, see Table 12. The p-value was calculated using the paired t-test.

5.2.3 Correlations between selected variables in the experimental group

In this part, possible associations between selected variables in the experimental group, as shown in Table 13 below, will be discussed.

There is no statistically significant correlation between the scores from the entry test and any of the variables presented in the table. However, sample statistics themselves can help to filter ideas for future research and thus they are relevant for exploration. There is a negative sample correlation between GJT difference variable and entry test scores, with $r = -0.2368$. This is a rather surprising outcome of the analysis, indicating that the higher the score in the entry test, the smaller the improvement in GJT. This could mean that the intervention has a less positive impact on the students with a better knowledge on entry, and a more pronounced effect on those with worse entry test scores who seem to benefit from the intervention more. But it could also indicate that the level of knowledge these students manifested was already high, so the improvement was not that pronounced.

The negative correlation between CBM difference and entry score is $r = -0.0553$, indicating almost no relationship between the initial knowledge and the difference in the levels of certainty. There is a positive correlation between GJT difference and CBM difference, described with the positive correlation coefficient $r = 0.4377$. If inferring from sample to the population, this correlation, at the level of significance of 0.05 would be statistically significant with $p = 0.032$. Such a relationship is of moderate strength and indicates that the higher the GJT scores, the higher the CBM scores.

All results of the correlation analysis conducted for the experimental group are summarised in Table 13 below, with significant correlations in bold print.

Figure 5

Histogram of CBM scores difference in pre-test and post-test for the experimental group

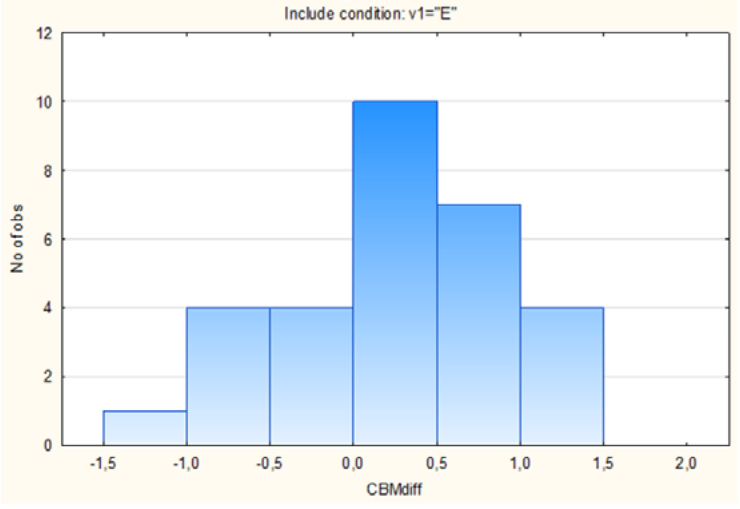


Table 12
Comparison of CMB pre-test and post-test results for the experimental group

	Marked differences are significant at p < 0.05000 Include condition: V1: "E"									
Variable	Mean	Standard deviation	N	Difference	Standard deviation difference	t	Diff.	p	Confidence −95%	Confidence +95%
CBM2	**0.1133**	**0.8235**								
CBM1	**−0.1697**	**0.7429**	**30**	**0.2830**	**0.6161**	**2.5159**	**29**	**0.0177**	**0.0529**	**0.5131**

Note. CBM1 refers to the Certainty-Based Marking test taken as a pre-test; CBM2 refers to the Certainty-Based Marking test taken as a post-test

Table 13
Correlations between entry score, GJT difference and CBM difference in the experimental group

	Marked correlations are significant at p < 0.05000 N = 24 (Case wise deletion of missing data) Include condition: V1: "E"					
Variable	Entry score	Motivation	Writing difference	Speaking difference	GJ difference	CBM difference
Entry score	1.0000 p = ---	−0.2005 p = 0.348	0.1936 p = 0.365	0.1632 p = 0.446	0.2368 p = 0.265	0.0553 p = 0.797
GJT difference	0.2368 p = 0.265	−0.1966 p = 0.357	0.1075 p = 0.617	0.0664 p = 0.758	1.0000 p = ---	**0.4377** **p = 0.032**
CBM difference	0.0553 p = 0.797	0.2646 p = 0.212	0.2326 p = 0.274	0.1904 p = 0.373	**0.4377** **p = 0.032**	1.0000 p = ---

5.2.4 The accuracy of spoken language in the experimental group

Unlike the pilot group, which was only tested in GJT and CBM, participants from the experimental and control groups were also tested in the accuracy of their free spoken and written production. Samples of spoken and written language were elicited from learners in the two groups and were used to devise a learner corpus. A thorough analysis of the corpus data will be provided in section 5.5 below. At this point, the results of the two groups, experimental and control, in pre-test and post-test in speaking and writing will be compared and contrasted, and a statistical evaluation of the yielded data will be provided. In the analysis, a general measure of accuracy, the number of errors per 100 words, was used. When comparing the results of pre- and post-testing, the lower average number of errors in the post-test was desirable, as it indicated improvement.

For the experimental group, the average number of errors per 100 words in speaking was 3.93 in the pre-test and 3.18 in the post-test. This indicates a lower average number of errors in the post-test by 0.74 errors per 100 words. This difference between the pre-test number of errors and post-test number of errors is statistically significant with $p = 0.0108$. The p-value was calculated using the paired t-test for dependent samples. This result was calculated for 26 students because, unlike the tests in GJT, CBM and writing, not all students in the experimental group took both the pre- and post-test in speaking.

If we wanted to demonstrate that the accuracy in speaking increased, i.e. that the number of errors decreased, this conclusion when generalised to the whole comparable population of students would be statistically significant at the level of significance $p = 0.0054$. For details, see Figure 6 and Table 14 below.

Figure 6

Histogram of the accuracy of spoken language difference in pre-test and post-test for the experimental group

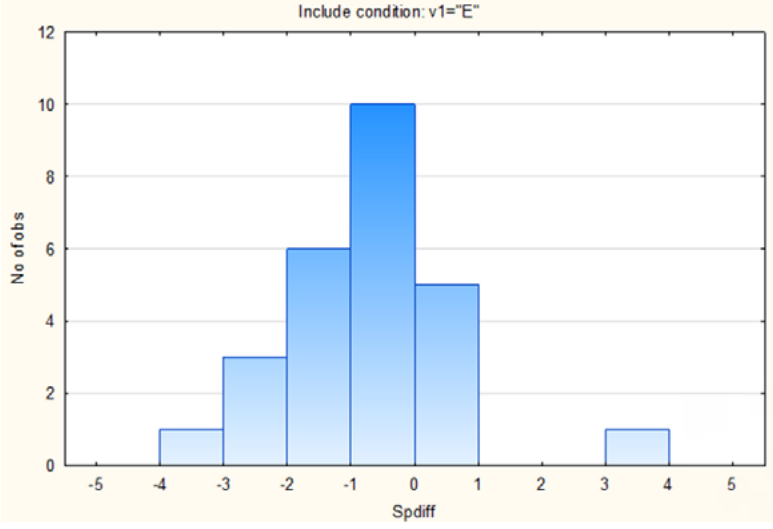


Table 14
Comparison of the accuracy of spoken language pre-test and post-test results for the experimental group

	Marked differences are significant at $p < 0.05000$ Include condition: V1: "E"									
Variable	Mean	Standard deviation	N	Difference	Standard deviation difference	t	Diff.	p	Confidence −95%	Confidence +95%
SP2	**3.1889**	**1.1997**								
SP1	**3.9292**	**1.3818**	**26**	**−0.7404**	**1.3714**	**−2.7527**	**25**	**0.0109**	**−1.2943**	**−0.1865**

Note. kSP1 indicates the result in the speaking pre-test; SP2 indicates the result in the speaking post-test

5.2.5 The accuracy of written language in the experimental group

The most surprising aspect emerging from the analysis is the comparison of the accuracy of written production in pre- and post-testing in the experimental group. This was calculated for 30 students who participated in both series of testing. From 5.69, the average number of errors per 100 words in the pre-test, the score increased in the post-test, to an average of 8.14 errors per 100 words. That is an average increase in the number of errors after the intervention by 2.45, a result which was certainly not anticipated.

A paired t-test for dependent samples was adopted in the analysis, with a statistically significant p value, p = 0.000406 showing that there is a difference between pre- and post-test scores. However, from the researcher's perspective, the improvement, not the general difference itself, is interesting. Thus, if the results were generalised to the whole population, an improvement, i.e. a lower average number of errors, would not be indicated, with p = 1 – 0.000406/2 = 0.999797.

The findings reported in this section, which are illustrated below, in Figure 7 and Table 15, are both surprising and unexpected. In chapter 6, these findings will be addressed in more detail.

Figure 7
Histogram of the accuracy of written language differences in pre-test and post-test for the experimental group

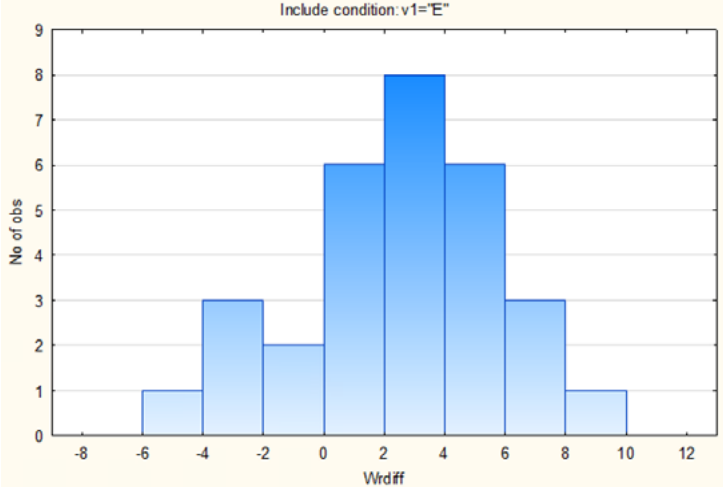


Table 15
Comparison of the accuracy of written language pre-test and post-test results for the experimental group

	Marked differences are significant at $p < 0.05000$ Include condition: V1: "E"									
Variable	Mean	Standard deviation	N	Difference	Standard deviation difference	t	Diff.	p	Confidence −95%	Confidence +95%
WR2	**8.1367**	**2.8034**								
WR1	**5.6897**	**2.1240**	**30**	**2.4470**	**3.3555**	**3.9942**	**29**	**0.0004**	**1.1940**	**3.6999**

Note. WR1 indicates the result in the writing pre-test; WR2 indicates the result in the writing post-test

5.3 Results in the control group

Descriptive statistical analysis was carried out for the control group, with 14 participants completing both pre- and post-testing. In the section below, GJT results will be outlined first, then CBM marking, followed by the correlations between the variables in the control group. An account of the changes in the accuracy of spoken and written production will be described in sections 5.3.4 and 5.3.5 below.

5.3.1 Grammaticality Judgement Test results in the control group

GJT scores in the control group improved, rising from the average GJT score in the pre-test, 35.96, to 45.72, that is by 9.76. This difference between GJT pre- and post-testing is statistically significant at the level of significance of 0.05, with p = 0.00244. If we wanted to generalise GJT improvement from these results to the whole population, it would be statistically significant with p = 0.000122. The p-value was calculated using the paired t-test for dependent samples. For details, see Figure 8 and Table 16 below.

Figure 8

Histogram of GJT scores differences in pre-test and post-test for the control group complemented with normality test p-value which does not reject normality of CBM differences

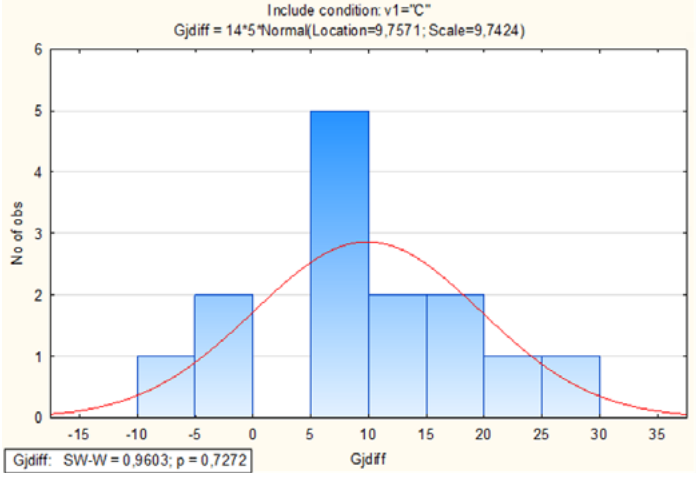


Table 16
Comparison of GJT pre-test and post-test results for the control group

	Marked differences are significant at p < 0.05000 Include condition: V1: "C"									
Variable	Mean	Standard deviation	N	Difference	Standard deviation difference	t	Diff.	p	Confidence −95%	Confidence +95%
GJ2	**45.7214**	**16.9190**								
GJ1	**35.9643**	**14.8531**	**14**	**9.7571**	**9.7424**	**3.7473**	**13**	**0.0024**	**4.1321**	**15.3822**

Note. GJ1 refers to the Grammaticality Judgement Test taken as a pre-test; GJ2 refers to the Grammaticality Judgement Test taken as a post-test

5.3.2 Certainty-Based Marking results in the control group

In the control group, CBM scores after the intervention were higher by 0.547, rising from the negative value of −0.625 to −0.078. The CBM difference was statistically significant at the level of significance of 0.05, with p = 0.0272. If we wanted to generalise CBM improvement from these results to the whole population, it would be statistically significant with p = 0.0136. The p-value was calculated using a paired t-test for dependent samples. For details of the data, see Figure 9 and Table 17 below.

Figure 9

Histogram of CBM pre-test and post-test differences for the control group complemented with normality test p-value which does not reject normality of CBM differences

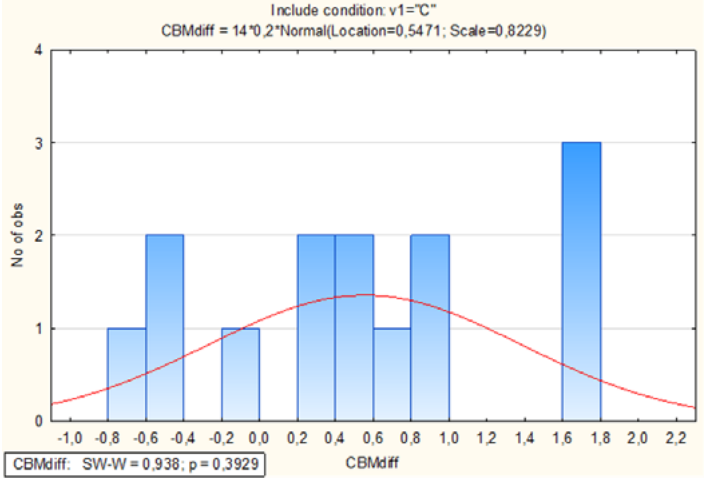


Table 17
Comparison of CBM pre-test and post-test results for the control group

	Marked differences are significant at $p < 0.05000$ Include condition: V1: "C"									
Variable	Mean	Standard deviation	N	Difference	Standard deviation difference	t	Diff.	p	Confidence −95%	Confidence +95%
CBM2	**−0.0786**	**0.8155**								
CBM1	**−0.6257**	**0.8063**	**14**	**0.5471**	**0.8229**	**2.4877**	**13**	**0.0272**	**0.0720**	**1.0223**

Note. CBM1 refers to the certainty-based marking test taken as a pre-test; CBM2 refers to the Certainty-Based Marking test taken as a post-test

5.3.3 Correlations between selected variables in the control group

Correlation analysis was conducted to determine if any associations exist between the variables in question, as shown in Table 18 below, with significant correlations marked in bold print.

The analysis of the control group results revealed that entry test scores positively correlate with GJT difference and CBM difference and showed that also other pairs of variables in the sample reach high correlations, though not statistically significant.

GJT difference and entry test score correlate positively, with $r = 0.621$. This result can be interpreted as follows: the higher the score on the entry test, the bigger the difference in the GJT pre-test and post-test, and thus a more noticeable improvement in GJT. This result is statistically significant with $p = 0.023$.

There is also a positive correlation between CBM difference and entry score, with $r = 0.6127$. This means that with an increased value of entry test score, CBM difference between pre-test and post-testing increased, and therefore a more pronounced improvement occurred. This result is statistically significant with $p = 0.026$.

The most striking result is the association between GJT difference and CBM difference which is described with the correlation coefficient $r = 0.8064$. When generalising, the result would be at the level of significance of 0.05, with $p = 0.001$. This association is strong and means that the more significant the improvement in GJT, the more substantial the improvement of CBM.

Table 18
Correlations between entry score, writing difference, speaking difference, GJT difference and CBM difference in the control group

Marked correlations are significant at p < 0.05000
N=13 (Case wise deletion of missing data)
Include condition: V1: "C"

Variable	Entry score	Motivation	Writing diff.	Speaking diff.	GJ diff.	CBM diff.
Entry score	1.0000 p = ---	−0.1246 p=0.685	0.1745 p=0.569	0.3122 p = 0.299	**0.6211** **p = 0.023**	**0.6127** **p = 0.026**
Motivation	−0.1246 p = 0.685	1.0000 p = ---	−0.1618 p = 0. 597	−0.4467 p = 0.126	−0.1761 p = 0.565	−0.0567 p = 0.854
Writing difference	0.1745 p = 0.569	−0.1618 p = 0.597	1.0000 p = ---	0.4999 p = 0.082	0.1127 p = 0.714	−0.1211 p = 0.693
Speaking difference	0.3122 p = 0.299	−0.4467 p = 0.126	0.4999 p = 0.082	1.0000 p = ---	0.0613 p = 0.842	−0.1247 p = 0.685
GJ difference	**0.6211** **p = 0.023**	−0.1761 p = 0.565	0.1127 p = 0.714	0.0613 p = 0.842	1.0000 p = ---	**0.8064** **p = 0.001**
CBM difference	**0.6127** **p = 0.026**	−0.0567 p = 0.854	−0.1211 p = 0.693	0.1247 p = 0.685	**−0.8064** **p = 0.001**	1.0000 p = ---

The results of the correlation analysis conducted for the control group are summarised below, in Table 18. In this section, only some associations, those between GJT, CBM and entry test, are presented, the other variables, namely accuracy in speaking and writing, will be addressed later.

5.3.4 The accuracy of spoken language in the control group

In speaking, both groups have improved: the improvement is, however, less pronounced for the control group, with an average number of errors of 3.37 in the pre-test and 3.29 in the post-test, with a small decrease of 0.08 errors. This result, however, is not statistically significant, with p = 0.8056. The analysis did not show any significant change in speaking in the control group. These results are illustrated in Figure 10 and Table 19 below.

Figure 10

Histogram of the accuracy of spoken language pre-test and post-test results for the control group complemented with normality test p-value which does not reject normality of speaking differences

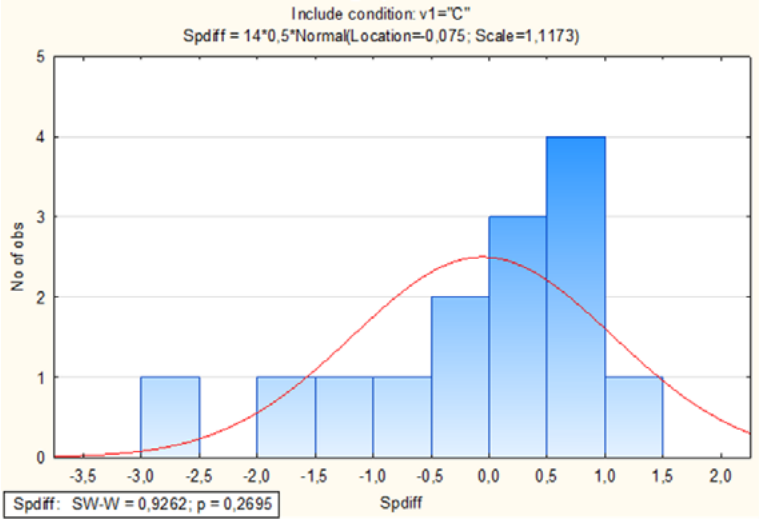


Table 19
Comparison of the accuracy of spoken language pre-test and post-test results for the control group

	Marked differences are significant at $p < 0.05000$ Include condition: V1: "C"									
Variable	Mean	Standard deviation	N	Difference	Standard deviation difference	t	Diff.	p	Confidence −95%	Confidence +95%
SP2	3.2986	0.8817								
SP1	3.3736	1.4134	14	−0.075	1.1173	−0.2511	13	0.8256	−0.7201	0.5701

Note. SP1 indicates the result in the speaking pre-test; SP2 indicates the result in the speaking post-test

5.3.5 The accuracy of written language in the control group

Unlike in the experimental group, a small increase of accuracy manifested as a lower average number of errors per 100 words was reported for the control group. The average number of errors was 5.93 in the pre-test and 5.90 in the post-test, with a very small improvement of 0.03. The paired t-test for dependent samples adopted in the statistical analysis did not show any statistically significant difference in writing for the control group, with p = 0.92. For details, see Figure 11 and Table 20 below.

Figure 11
Histogram of the accuracy of written language pre-test and post-test results for the control group complemented with normality test p-value which does not reject normality of writing accuracy differences

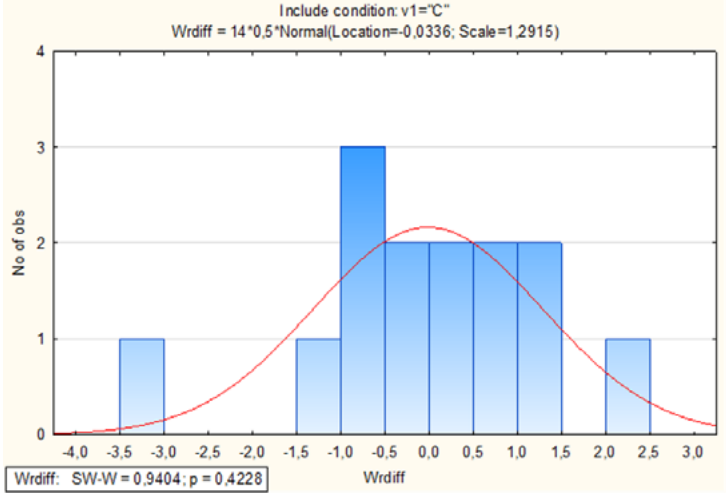


Table 20
Comparison of the accuracy of written language pre-test and post-test results for the control group

	Marked differences are significant at $p < 0.05000$ Include condition: V1: "C"									
Variable	Mean	Standard deviation	N	Difference	Standard deviation difference	t	Diff.	p	Confidence −95%	Confidence +95%
WR2	5.9014	1.7853								
WR1	5.9350	1.6411	14	−0.0336	1.2915	−0.0973	13	0.9240	−0.7793	0.7121

Note. WR1 indicates the result in the writing pre-test; WR2 indicates the result in the writing post-test

5.4 Comparisons between the pilot, experimental and control group

In this section of chapter 5, comparisons between the three groups will be summarised, and similarities and differences outlined. For a detailed account of the final scores in the pilot group, see section 5.1 above. The summary of the scores in GJT tests, CBM marking, written and spoken production in the experimental and control groups, resulting from the statistical analysis, is detailed below.

In the analysis of the results in writing, it was not possible to use the two-sample t-test because the assumption of equality of variances, which is an assumption of the two-sample t-test, was not met. In the statistical analysis of the other variables – GJT, CBM and speaking – two-sample t-tests for independent samples were used. The difference between the experimental and control groups was not, however, statistically significant. The results for the two-tailed t-test failed to show any significant difference between the experimental and control groups. (The p-value for writing is not relevant, as the assumption of the equality of variances is not met. This is manifested by the right p-value = 0.00077.)

5.4.1 Grammaticality Judgement Test results comparison in the pilot, experimental and control groups

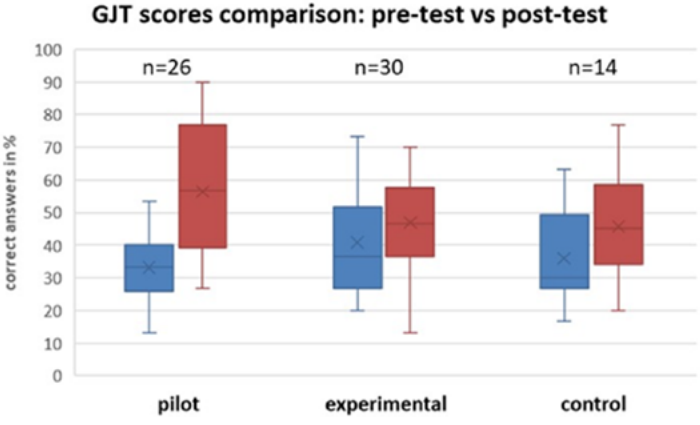
The results obtained from the analysis of GJT scores and the differences between the pre-test and post-tests for each of the groups are illustrated in Figure 12 below. From the graph, it can be seen that all groups improved, irrespective of the intervention. This result, however, might be to a certain extent biased by the low number of participants in the control group (n = 14), as compared to the other two groups, pilot (n = 26), and experimental (n = 30), which both received the intervention.

What stands out is the fact that the pilot group was the most accurate in their answering in the post-test, with a mean score in the post-test of 56.41, as opposed to 33.19 in the pre-test, indicating an increase by 23.21. This difference was statistically significant with p = 0.000002. If we test the improvement itself, then the p-value equals to 0.000001, indicating a positive impact of the intervention.

When comparing the experimental and control group results in the GJT, it can be seen that both groups improved in the post-test. A more pronounced improvement occurred in the control group, with an average increase of 9.76, while in the experimental group the increase was 6.1. The improvement in the GJT in both groups was statistically significant. The difference in these increases between the experimental and control groups was not, however, statistically significant, with $p = 0.282$.

The results of the analysis indicate a statistically significant improvement in Grammaticality Judgement Testing in all groups. For details, see Table 21.

Figure 12
Comparison of GJT scores in the pilot, experimental and control groups



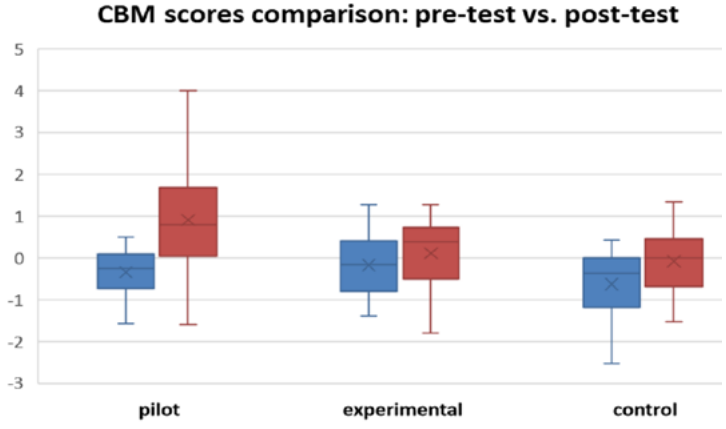
5.4.2 Certainty-Based Marking results comparison in the pilot, experimental and control groups

It is apparent from Figure 13 below that when comparing certainty-based marking results, ranging from the negative of −6 to +3 for each answer, between pre-and post-tests, all groups improved. The most significant change, similarly to GJT test, occurred in the pilot group, with the median difference between pre-and post-test of +1.03. This difference was statistically significant with $p = 0.000097$. As the assumption of the paired t-test was not met, the Wilcoxon signed-rank test based on median values rather than means was performed.

The results obtained from the control and experimental groups indicate that the control group improved by 0.547. This result is statistically significant with p = 0.0136. The increase in the Certainty-Based Marking in the experimental group was lower, 0.283, and statistically significant at the level of significance p = 0.00883. This difference in improvement between the two groups is not statistically significant, with p = 0.24.

Overall, the analysis has shown that all groups have improved in Certainty-Based Marking and this outcome is statistically significant. For details, see Table 17 above.

Figure 13
Comparison of CBM scores in the pilot, experimental and control groups.



5.4.3 Comparison of the accuracy of spoken language in the experimental and control groups

When comparing the average scores in the pre-test and post-test in the accuracy of spoken production, it is obvious from the graph in Figure 14 and Table 21 below, that both groups, experimental and control, have improved. In the experimental group, the difference between pre- and post-test scores was −0.74, i.e. the decrease of the average number of errors per 100 words of 0.74. This result was statistically significant at p = 0.0108. For the control group, there was also a decrease in the average number of errors; this was, however, not so pronounced, with a mean value of −0.08. More importantly, this result was not statistically

significant, with p = 0.8056. This indicates a larger and statistically significant improvement in speaking in the experimental group.

5.4.4 Comparison of the accuracy of written language in the experimental and control groups

What stands out from the comparison of the accuracy in written production between the experimental and control groups, as illustrated in the graph below, Figure 15, is the fact that while the control group has slightly improved, the experimental has not. In the experimental group, the average number of errors in writing has increased from 5.69 to 8.14, i.e. by 2.44. This result is statistically significant at p = 0.000406. In the control group, the average number of errors slightly decreased, from 5.93 to 5.90, by 0.03. As it was impossible to carry out a two-sample t-test, due to the reasons explained above, it is obvious that the control group achieved a better result in writing. This result, with p = 0.92, was not, however, statistically significant.

The results presented in this part are summarised in Table 21 below. What stands out in the table is the fact that the overall impact of the intervention was positive, with a positive effect in the increase of accuracy in the GJT and CBM in the pilot and experimental groups. The experimental group also improved in speaking in a statistically significant way but did not improve in writing; this result was also statistically significant. The control group improved in the GJT and CBM. The small improvement in speaking and writing in the control group was not statistically significant. In the table, the plus and minus signs indicate the positive or negative impact of the intervention, respectively, while the numbers in brackets refer to the increase of the variable in question. In the third column, the correlations between the GJT and CBM scores are set out; the plus sign indicates a positive correlation. Statistical tests adopted in the analysis are specified in the brackets. Together with all the other findings, these results and their possible explanations will be addressed more thoroughly in chapter 6.

Figure 14

Comparing accuracy of spoken language for the experimental and control groups, in pre-test and post-testing

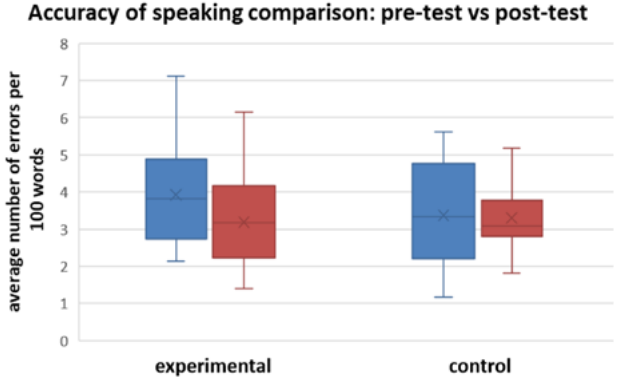


Figure 15

Comparing accuracy of written language for the experimental and control groups, in pre-test and post-test

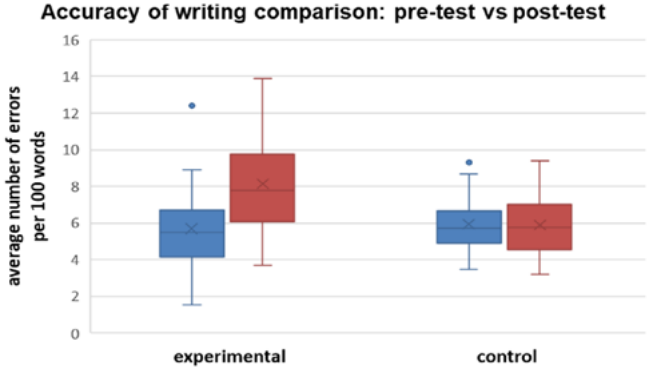


Table 21
Group comparison: summary of the selected statistical analysis results in the three groups

GROUP	GJT	CBM	Correlations GJT vs CBM	SPEAKING	WRITING
PILOT	+	+	+	NA	NA
n = 26	(23.21)	(1.03)	r = 0.65		
	p = 000002	p = 000097	p = 0		
	(paired t-test, two-tailed)	(Wilcoxon signed-rank test, two-tailed)	(correlation test, two-tailed)		
EXPERIMENTAL	+	+	+	+	–
n = 30	(6.107)	(0.283)	r = 0.4377	(fewer errors 0.74)	(more errors 2.45)
	p = 0.00378	p = 0.01767	p = 0.032	p = 0.0108	p = 0.000406
	(paired t-test, two-tailed)	(paired t-test, two-tailed)	(correlation test, two-tailed)	(paired t-test, two-tailed)	(paired t-test, two-tailed)
CONTROL	+	+	+	+	+
n = 14	(9.76)	(0.547)	r = 0.8064	(fewer errors 0.08)	(fewer errors 0.03)
	p = 0.00244	p = 0.0272	p = 0.001	p = 0.8056	p = 0.92
	(paired t-test, two-tailed)	(paired t-test, two-tailed)	(correlation test, two-tailed)	(paired t-test, two-tailed)	(paired t-test, two-tailed)

Note: + indicates a positive impact of the intervention or positive correlation, – indicates a negative impact of the intervention or negative correlation; numbers in brackets indicate the average improvement; *n* indicates the number of participants in the groups; *p* stands for the level of statistical significance; *r* is the coefficient of correlation; *NA* means not applicable

In the next section, an analysis of the questionnaires distributed to students after the post-test will be outlined. The open-ended answers might provide some valuable insights into the learners' views on testing, their motivation to participate in the research and could help to better understand the results of the research.

5.4.5 Analysis of questionnaires providing contextual information

Two questionnaires were distributed among the participants from the control (C) and experimental (E) groups with the aim of eliciting relevant contextual information; the first immediately after the pre-test and the second after the post-test. For full details of both questionnaires and the topics for written tasks and oral interviews in pre-and post-tests, see Appendices 8, 7 and 6 respectively.

The first questionnaire asked respondents in both groups to complete one open-ended question: *Why did you decide to take this course?* For the experimental group, *the course* consisted in both participating in the two series of testing and in attending the one-term blended learning course which represented educational intervention described in chapter 3. For the control group, the question was related to participating in the two series of testing only.

In the experimental group, almost all participants, 29 out of 30, provided answers in the questionnaire. Some students mentioned more than one reason for taking the course. Three recurrent areas emerged from the analysis of their free answers[33].

What I expect from attending the course is:

- *being able to identify my own as well as others' typical (L1-induced) errors* (13 students)
- *improving my English language skills* (11 students)
- *being able to avoid (L1-induced) errors in the future* (10 students)

The response rate in the control group was 100%, with 14 students who answered this open-ended question. They gave a variety of reasons why they had decided to participate in the research; some of them only provided one reason, others up to four different ones. Their responses can be summarised as follows:

I decided to participate in the series of testing because I would like to:

- *know how much my English language skills improve in the span of one term* (6 students)
- *find out what my problematic areas in English are* (6 students)

33 Students' answers have not been quoted verbatim but have been paraphrased by the author of the book in order to report similar patterns in them.

- *gain this interesting experience* (5 students)
- *get the one credit awarded for the participation* (5 students)
- *help the department in research* (4 students)

The second questionnaire, distributed after the post-test, required respondents to compare pre- and post-tests and comment on all parts of testing. The response rate was high, almost 100% in both groups; 14 students from the control group and 29 from the experimental provided their answers. In the first question, 76% of the participants from the C group and 72% of the participants from the E group believed that GJT was the same as in the pre-test, 24% in the E and 12% in the C group could not remember, and 3% in the E and 12% in the C thought it was different. Most of the respondents claimed recalling the pre-test. In the open-ended comment on GJT (questions 2 and 3), however, they mostly stated that despite remembering the form of the test and some of the questions, they could not remember the answers and they found the post-test equally challenging as the pre-test. As one of the respondents (code EK) said: *I believe that the questions in the test were the same or at least some of them were. However, I did not really remember the answers, so it was like taking a new test today.* They also stated that when taking the post-test, they felt slightly more confident than when taking the test for the first time, mostly because they were already familiar with the format of the test and knew what to expect.

When comparing the difficulty of the written task (question 12), 69% of students from the experimental and 59% from the control group regarded the tasks equally difficult, 7% in the E group and 18% in the C group thought the post-test was more difficult, and 24% in both groups remained undecided. Generally, they claimed that the topic in the post-test was more relevant and easier for them to relate to than the one in the pre-test. Also, some of them (seven in the E group) mentioned that in the pre-test they were facing technical difficulties which might have caused distress. A surprisingly high number of participants, 52% from the E and 79% in the C group, claimed they could not recall the topic in the pre-test at all. Some of them complained about the short time limit for both written tests (4 in the E and one in the C group).

Regarding the difficulty of the oral interview, most students were either undecided (55% in the E and 18% in the C group) or thought the level of difficulty in both tests was the same (45% in the E and 71% in the

C group). Only 12% of the students in the C group thought the oral post-test was more difficult than the pre-test, especially because they found the topics more challenging. When comparing the two interviews in the pre- and post-tests, about one third of the students from both groups reported having been slightly more nervous in the first one, mostly because they did not know what to expect. Surprisingly few participants, just one in each group, reported the presence of the camera as a distracting factor, and one student complained about the unpleasant brightness of the lights when being video recorded. Although some students stated that they had liked the post-test interview better (five in the E and two in the C group), others felt it was almost the same (seven in the E and one in the C group). Ten students from both groups expressed their dislike of any oral interviews, while six appreciated a relaxed atmosphere of the two interviews.

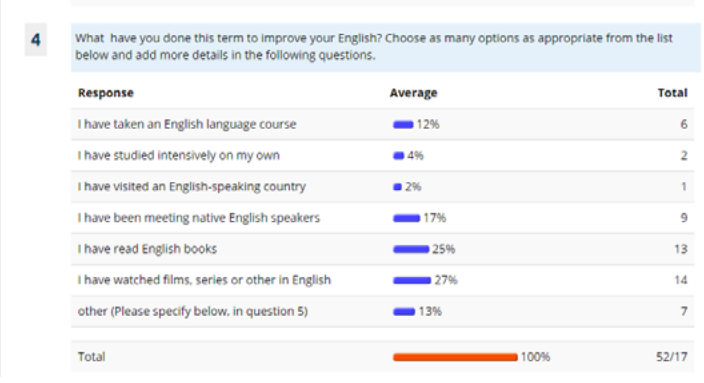
Question 4 required respondents to give information on how they tried to improve their English language proficiency in the period of one term between the pre- and post-tests. They were first asked to choose any of the activities listed in the multiple-choice question and then to provide further comments in open-ended answers. One point was awarded for each of the selected responses in this question, with the maximum possible gain of seven points. This number was then regarded as an indicator of the level of their motivation. It can be seen from the data in Figure 16 below that the two groups were mostly comparable in this area. The only slight difference was in the amount of individual studying where only 4% in the E group reported doing this, while 11% of respondents in the C group reported studying on their own.

Taken together, the results of the analysis of two questionnaires provided in this subchapter indicate that the two groups under scrutiny were very similar in all respects. Apart from understandably different motivation to participate in the research, their evaluation of all parts of the tests was comparable. What is very important is the fact that their activities during the term in which the research was conducted did not manifest any significant differences. Considering the fact that all participants were limited by a set of criteria on entering the study (as specified above in 4.4.2), it can therefore be assumed that the control and experimental groups were comparable and any differences between pre- and post-test scores can be attributed to the influence of educational intervention.

In the section that follows, a corpus-based analysis of learner language based on the samples of written and spoken production elicited from the participants in the experimental and control groups will be presented.

Figure 16
Comparing answers to question 4 about learners' motivation for experimental and control group

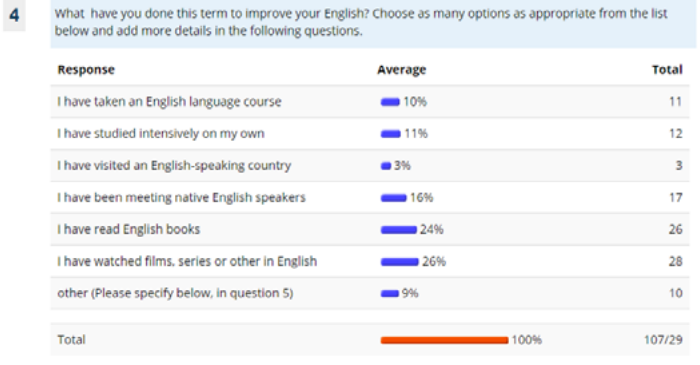
Experimental group



4 What have you done this term to improve your English? Choose as many options as appropriate from the list below and add more details in the following questions.

Response	Average	Total
I have taken an English language course	12%	6
I have studied intensively on my own	4%	2
I have visited an English-speaking country	2%	1
I have been meeting native English speakers	17%	9
I have read English books	25%	13
I have watched films, series or other in English	27%	14
other (Please specify below, in question 5)	13%	7
Total	100%	52/17

Control group



4 What have you done this term to improve your English? Choose as many options as appropriate from the list below and add more details in the following questions.

Response	Average	Total
I have taken an English language course	10%	11
I have studied intensively on my own	11%	12
I have visited an English-speaking country	3%	3
I have been meeting native English speakers	16%	17
I have read English books	24%	26
I have watched films, series or other in English	26%	28
other (Please specify below, in question 5)	9%	10
Total	100%	107/29

5.5 Corpus-based analysis of learner language

The analysis of the data from the learner corpus, compiled from 92 essays as samples of written language, and 84 transcripts of interviews as

samples of spoken language, proceeded in the following steps. First, all data collected in the corpus were analysed for error types and their frequencies to establish the characteristics of learner language. Then, spoken and written language were analysed separately, in order to find out whether the error types and frequencies differ. This was followed by comparing the results in control and experimental groups and establishing differences between pre-test and post-test results in both groups.

The overall analysis of the corpus revealed that out of the eight error domains, four were affected by errors considerably more than the remaining four. In Table 22 below, the exact numbers of errors in each domain are provided, both with the total of errors in the whole corpus, and in the breakdown according to their occurrence in spoken and written language.

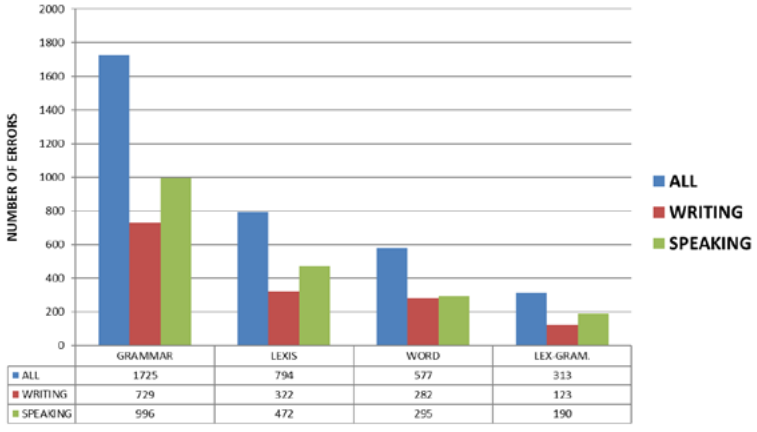
As can be seen from the table, the most error-prone domain was grammar, with the total of 1,725 errors, followed by errors affecting lexis – 794, errors in the word domain, with words either missing or redundantly used, or erroneous word order – 577, and lexico-grammatical errors, with the total of 313 errors. The domain least affected by errors was style – 48 errors overall, followed by infelicities – 69, punctuation – 132, and form – 208.

Table 22
Error frequencies in the eight error domains

	error frequency (count)		
Error domain	**all data**	**writing**	**speaking**
form	208	125	83
grammar	1,725	729	996
lexis	794	322	472
lexico-grammatical	313	123	190
word	577	282	295
style	48	35	13
punctuation	132	132	NA
infelicities	69	66	3

The graph in Figure 17 below presents the four most frequently affected error domains and the breakdown of the data collected from samples of spoken and written learner language. The most numerous error types resulting from the analysis will be now discussed in more detail.

Figure 17
General corpus-based analysis – error frequency in the four domains most affected by errors



	GRAMMAR	LEXIS	WORD	LEX-GRAM.
ALL	1725	794	577	313
WRITING	729	322	282	123
SPEAKING	996	472	295	190

5.5.1 Frequency of error types in the whole corpus

The most frequent error type in the whole corpus is lexical error affecting a single word coded as LS according to the Louvain error tagging system (for the details of the Louvain error tagging system see part 4.7 and Table 6 above). These errors are typically wrong collocations, false friends and word-for-word translations from Czech: e.g. **stipend*[34] – scholarship, people with a **title* – people with a degree, an **absolvent* – a graduate, **accommodate* to the changes – adapt to the changes, interactive **table* – interactive whiteboard. Lexical errors affecting the whole phrase or its part (LP) were also the third most frequent type of error, e.g. **break in nerves* – have a breakdown, **undergo the maturita* – take the school-leaving exam, **according to me* – in my opinion.

Perhaps unsurprisingly, the second most frequent type of error is the use of articles, the 2nd, 4th and 5th most frequent error type overall. The articles were either used redundantly (GARD – grammar article redundant), especially when the definite article was erroneously used with abstract words, e.g. **the society* needs people – society needs people, **the formal qualifications* might not be needed – formal qualifications might not be needed, *aspects of *the human nature* – aspects of human

34 Erroneous usage is indicated with an asterisk and italics.

nature. The second most often occurring error in the use of articles was a missing indefinite article (GAMI), e.g. **being waiter* – being a waiter, **from poor family background* – from a poor family background, **since young age* – since a young age.

The next error type was W, indicating errors in word order (WO), the redundant use of a single word (WRS) or a single missing word (WMS). Errors in word order were most often indirect questions: **principles of how does everything work* – principles of how everything works, followed by misplaced adverbs **it probably doesn't yet exist* – it doesn't probably exist yet, and inverted sentences **not only we are becoming* – not only are we becoming.

Errors in the plural forms of nouns (GNN) were the 9th most frequent type of error. These were mostly errors in irregular plural forms, e.g. **childrens* – children, abstract nouns used in the plural: presence of **technologies* – presence of technology. Whenever the assessors were doubtful about the corrections in the regular and irregular plural noun forms, English Web 2015[35] and Google Books Ngram viewer[36] were consulted and the frequency of both plural forms viewed and compared. This was for example the case with **social medias* – social media. The search in English Web 2015 showed 5,314,257 hits for 'media' and 8,692 for 'medias', which was confirmed in Google Books Ngram viewer. While it seems that the use of 'medias' is on the rise, 'media' is still used far more frequently by native speakers of English and was therefore regarded as the only correct option.

The 10th most frequent position in the error frequency count was occupied by the tag XVPR, used for verbs with an erroneous, missing, or redundant dependent preposition, as well as for wrong particles in phrasal verbs, e.g. they **stare to* their phones – they stare at their phones, they are mostly **paid* – they are mostly paid for, to **balance between* those two skills – to balance those two skills, I **dress up* – I get dressed.

Another ten error types ordered by the frequency of occurrence were as follows: errors in tenses – GVT, errors in prepositions – LSP, redundant use of the indefinite article – GARI, FS – spelling errors, GSVA – grammar errors in subject verb agreement, WRM – multiple

35 English Web 2015, or enTenTen15, is an English corpus made up of texts from the Internet, compiled using technology specialized in collecting only linguistically valuable content. It is available from the Sketch Engine web page https://www.sketchengine.eu/ententen-english-corpus/

36 Google Books Ngram viewer is an online search engine providing a comparison of frequencies of any expressions over time.

missing words, ZIR – infelicities, i.e. errors of style, inappropriate register and politically incorrect expressions, XVCO – lexico-grammatical errors in verb complement, GAWD – erroneously used definite article instead of indefinite, GVMOD – grammar errors affecting modal verbs. The most frequently occurring error types in the whole learner corpus are detailed below, in Figure 18.

5.5.2 Comparing errors in spoken versus written learner language

When comparing error frequencies in spoken and written learner language, it can be observed that the differences between the two are not significant. The first five most frequently occurring error types in both sets of data, written and spoken, are identical, with one exception: WRS – a single redundant word which occupies the 4th position in writing but 9th in speaking. All the other error types in the first five places are the same in both datasets, occupying, however, different positions. While the most frequent error type in speaking was lexical errors affecting a single word (LS), this type was the second most frequently occurring in writing. Grammar errors in the use of articles, both missing and redundant, occupied three of the first five positions in spoken and two in written language. Missing indefinite articles (GAMI) occupied the 2nd position in speaking, and 7th in writing. Missing definite articles (GAMD) were the 5th most frequent error type in both sets of data. A redundant use of the definite article (GARD) was the 3rd most frequent in speaking and the 1st in writing.

Error types occupying the 6th to 10th positions differ for speaking and writing. The only exception is GNN error type, i.e. using the wrong plural form in nouns. This was comparably frequent both in speaking and writing, 10th and 9th respectively. Apart from the already mentioned redundant use of single words (WRS) in the 9th position, word order issues (WO) were the 6th most frequently affected area in speaking, followed by lexical errors in prepositions (LSP), erroneous use of prepositions with verbs (XVPR), and wrong plural forms in nouns (GNN). In writing, spelling errors (FS) were the 6th most frequent error type, followed by missing indefinite articles (GAMI), single missing words (WMS), and errors in inappropriate register (ZIR). For details of the comparison of error frequencies in spoken and written learner language, see Figure 19 below.

Figure 18
Corpus-based analysis – frequency list of error types in the learner corpus

1	LS	373	11	GVT	113
2	GARD	351	12	LSP	108
3	LP	282	13	GARI	88
4	GAMI	249	14	FS	82
5	GAMD	199	15	GSVA	73
6	WO	169	16	WRM	72
7	WRS	162	17	ZIR	65
8	WMS	130	18	XVCO	60
9	GNN	125	19	GAWD	59
10	XVPR	119	20	GVMOD	50

Figure 19
Corpus-based analysis – comparing error frequencies in spoken and written language

spoken language

	Err.type	Frequency	Relative [%]
1	LS	220	3,045.7
2	GAMI	171	3,546.2
3	GARD	165	2,427.4
4	LP	147	2,691.8
5	GAMD	118	3,062
6	WO	113	3,452.7
7	LSP	89	4,255.4
8	XVPR	83	3,601.6
9	WRS	69	2,199.4
10	GNN	62	2,561.2

written language

	Err.type	Frequency	Relative [%]
1	GARD	186	3,098.6
2	LS	153	2,398.5
3	LP	135	2,799.3
4	WRS	93	3,356.8
5	GAMD	81	2,380.1
6	FS	81	5,776.1
7	GAMI	78	1,831.7
8	WMS	74	3,328.5
9	GNN	63	2,947.1
10	ZIR	62	5,577.5

When comparing the results in writing and speaking, it is also important to mention that there are two error types which were only present in written and not in spoken language. These are errors of form affecting spelling (FS) and errors in punctuation (QC, QM, QR). For obvious

reasons, these errors can only appear in writing and not in speaking. Spelling errors accounted for 4.5% of the total number of errors in writing and errors in punctuation for 7.27%. This had, however, no significant influence on the ratio of errors between the pre- and post-test scores in either of the groups.

5.6 Summary

This chapter has attempted to provide a detailed description of the results obtained from the research. First, statistical analysis examining the efficacy of educational intervention in the pilot, experimental and control groups was provided.

In all groups, the ability of respondents to identify and correct errors was tested in the Grammaticality Judgement Tests, and their certainty in answering was measured through Certainty-Based Marking. In the pilot group, the average increase in the GJT was 23.21; this result was statistically significant at the 0.05 level of significance. Similarly, the median difference between CBM pre- and post-testing increased by 1.03 and this difference was statistically significant with $p = 0.00097$. There was also a statistically significant positive correlation between GJT and CBM increases, indicating that with increased levels of accuracy certainty in answering increased as well, which is one of the most important findings of the research.

In the experimental and control groups, in addition to GJT and CBM, the accuracy in learners' free written and spoken production in the target language was analysed. Both groups improved in GJT and CBM scores; this improvement was statistically significant in both groups at the level of significance of 0.05. The difference in the increases between the two groups was not statistically significant, with $p = 0.282$ for the GJT and $p = 0.24$ for CBM. As regards speaking, improvement in accuracy was observed in both groups. While this improvement was statistically significant in the experimental group, with $p = 0.0108$, it was not the case in the control group, with $p = 0.8056$. The most unexpected result was that in writing; the experimental group worsened by an average of 2.45 errors per 100 words, with $p = 0.000406$, while the control group slightly improved, by 0.03 errors per 100 words; for details, see subchapters 5.2.5 and 5.3.5. This improvement in the

control group was not, however, statistically significant; more details are provided in part 5.4.4.

In summary, the results suggest that a positive and statistically significant change in the accuracy and certainty of answering occurred in the pilot group, which suggests the positive impact of the intervention. Together, the results in the experimental and control groups indicate that an improvement can be observed in both groups in GJT, CBM and speaking. In writing, more errors were detected in the experimental than in the control group. No statistically significant differences were found, however, between the two groups.

Another important part of the analysis addressed in this chapter was a corpus-based analysis of learner language. Learners' spoken and written production was analysed for errors and the major error-prone areas of advanced Czech learner language were identified. These were especially grammar errors, followed by errors affecting lexis and word order. In the domain of grammar, the most frequently identified error type was the (mis)use of both the definite and indefinite articles, an error frequently attributed to the negative influence of learners' mother tongue.

The next chapter moves on to provide a more elaborate commentary on the findings in which the results of the research as addressed in the research questions will be described. Also, the findings from the current study will be compared to similar studies conducted in the Czech Republic and abroad.

6 DISCUSSION

The results and analyses presented in the previous chapter will now be interpreted and answers to the research questions formulated in 4.2 will be addressed. First, the overall impact of intervention as formulated in the main research question will be discussed, followed by a more detailed analysis provided in the answers to the six sub-questions. These analyse the changes caused by intervention in GJT and CBM, identify how these results correlate with entry test scores, and compare changes in accuracy of learners' free spoken and written production. Questions 5 and 6 provide the outcomes of the corpus-based analysis of learner language. These findings are compared and contrasted with those resulting from similar corpus-based studies.

6.1 Answering the research questions and interpreting the findings

The main research question sought to determine the efficacy of the intervention and was formulated as follows:

What is the overall impact of educational intervention focused on eliminating typical errors of advanced Czech and Slovak students of English on the accuracy of their learner language?

The impact of educational intervention was explored in two groups, pilot and experimental. The scores from the experimental group were contrasted with the control group, which was comparable with the experimental (for details of how the participants were selected see subchapter 4.4). Descriptive statistical analysis of all results from the three groups was conducted (for details, see chapter 5).

Taken together, there was a marked statistically significant accuracy improvement in the pilot group, which suggests a positive impact of the educational intervention. This improvement also occurred in the experimental group, where the GJT, CBM and speaking scores

improved in a statistically significant way. In writing, a positive impact of the intervention was not achieved. In the control group, which did not receive any intervention, statistically significant improvement was observed in the GJT and CBM scores. In writing and speaking, there was a slight improvement, which was not, however, statistically significant. Overall, these results indicate a statistically significant improvement in the GJT, CBM and speaking in the experimental group, which indicates a positive impact of the intervention in these three areas. In the control group, a statistically significant improvement occurred in two areas only, GJT and CBM, and was not indicated in speaking or writing.

RQ1) Does intervention focused on accuracy affect students' ability to identify and correct errors?

The scores from the GJT reached by the respondents from all three groups in the pre- and post-testing were compared, in order to assess the effect of the intervention. The most obvious finding to emerge from the analysis in the pilot group is a statistically significant improvement in identifying and correcting grammar errors. It was also reported that the improvement in accuracy positively correlated with the level of certainty in answering. This analysis confirmed a positive impact of the intervention in the GJT in the pilot group. In the experimental and control groups, the results indicate a statistically significant improvement in the GJT in both groups. Similarly to the pilot group, a positive correlation was also identified between GJT and CBM in both groups. It can thus be concluded that the positive impact of intervention as indicated in the pilot group has also been documented in the experimental group. Surprisingly, this improvement also occurred in the control group where it cannot be ascribed to the intervention.

RQ2) Does intervention focused on accuracy affect students' certainty in identifying and correcting errors?

Similarly to the resulting GJT scores, in Certainty-Based Marking the most pronounced positive effect of the intervention was observed in the pilot group; this result was also statistically significant. In the experimental and control groups an increase in CBM which was statistically significant also occurred. The difference between the experimental and control groups is not statistically significant. This finding can be interpreted as follows: while the educational intervention

positively impacted CBM in the pilot group where students felt more certain when identifying and correcting their mistakes, in the main study, the increase in certainty was not detected.

RQ3) Are there any identifiable connections between students' entry test scores, their accuracy and certainty in answering?

Correlation analysis of three variables, entry test scores, GJT and CBM, revealed some interesting relationships between them. The most important finding to emerge from the analysis is that GJT and CBM positively correlate in all three groups. This positive correlation is strong in the control group ($r = 0.81$), as well as in the pilot group ($r = 0.66$) and moderate in the experimental group ($r = 0.43$). This means that correctness of answers increases with certainty in answering.

The relationship between the entry test score and increase in the accuracy of answering was statistically significant in the control group only, with $r = 0.62$. In the pilot group, this relationship was very low, with $r = 0.07$. The most surprising result was in the experimental group, with a negative correlation of -0.24. This indicates that while in the control group, the students with a higher entry test scores improved the most, in the experimental group, the improvement was lower with higher entry scores. It appears that the impact of the intervention in the experimental group was more significant in students with lower knowledge on entry and students with more accurate knowledge on entry benefitted from the intervention less. In the pilot group, there was almost no association between the entry score and improved scores in GJT, with $r = 0.07$. The result in identifying the association between the entry test score and GJT scores suggests that there is no identifiable relationship between entry test scores and the accuracy of answering caused by intervention.

Similarly, when examining the relationship between entry test scores and CBM scores, correlation analysis produced mixed results. While in the control group there was a strong positive correlation between entry test scores and certainty in answering, with $r = 0.61$, the connection was very low in the pilot group, with $r = 0.25$, and negative in the experimental group, with $r = -0.05$. Again, these results indicate that there might be no clear relationship between the entry test score and certainty in answering.

In summary, the most tangible outcome resulting from correlation analysis is that with increased accuracy in identifying and correcting erroneous sentences, the certainty of the respondents' increases as well.

RQ4) Does intervention focused on accuracy affect the rate of errors in students' spoken and written production?

The fourth question sought to determine the effect of targeted intervention on the accuracy of learners' free production in the target language. The average accuracy was expressed by the number of errors per 100 words. When comparing the results, the bigger the difference in the number of errors between the pre- and post-tests, the bigger the improvement. The two resulting scores for both groups were compared separately for speaking and writing.

In speaking, both groups improved; this improvement was, however, more pronounced and statistically significant in the experimental group. In the control group, the improvement was lower and not. statistically significant.

The analysis of writing in the experimental group revealed that the average number of errors between the pre- and post-testing increased, which means a worse result in the post-test. In the control group, a slight improvement was detected in writing; this improvement was not, however, statistically significant, with $p = 0.92$.

To sum up, the effect of intervention in speaking was positive and the improvement was statistically significant. Contrary to expectations, the effect was adverse in writing. This finding must be interpreted with caution. There are several factors at play, both internal and external, when considering the effects of intervention and its analysis; these factors will be addressed below.

One of the aspects which might influence students' results is the nature of the task, especially its modality, speaking and writing, and time available for planning. Research on the influence of task type and planning on the complexity, accuracy and fluency of learners' free production has yielded mixed results (Housen et al., 2012, pp. 111-113). In writing, complexity and fluency have been reported to benefit from pre-task planning, while the results for accuracy have been mixed (Ellis & Yuan, 2004). The outcome of Weissberg's research (2000) indicates that increased accuracy in writing as opposed to speaking depends on learners' preferences for written or spoken language, and is also influenced by individual differences among learners. When exploring

the influence of the modality of the task on CAF, Granfeldt (2007), on the other hand, reported that while lexical complexity was higher in writing than in speaking and that there were no significant differences between written and spoken language in syntactic complexity, accuracy in writing was lower than in speaking. Contrary to expectations, both grammatical accuracy and complexity of grammar structures in writing were affected, with a larger number of errors (Granfeldt, 2007). This finding is in accordance with the result of this research study in which the average number of errors in writing in both the experimental and the control groups is higher than in speaking (see 5.4.3 and 5.4.4 for more details).

When comparing the pre-and post-test results it should also be considered to what extent they can be influenced by the relative difficulty of the topic. This is not, however, likely to be the case in the written post-test, as more than a half of the participants, 59% in the experimental and 56% in the control group, stated in the questionnaires distributed after the written post-test that the topic was easier for them to write about and relate to than the one in the pre-test. Also, it has been reported in the literature that topic does not show any considerable effect on the development of accuracy, fluency or complexity (Vercellotti, 2017), so it does not seem that the choice of topics in the written task could have had any effect on the results if this research.

Another possible explanation of the unexpected result in writing is the method adopted in the analysis, in which only accuracy was explored. As this is just one of the three aspects of learners' performance, analysing the other two, complexity and fluency, might complement the current findings. In other words, in accuracy studies the focus is on *what students cannot do* rather than on *what they can do*. Looking at the complexity of learner language might reveal that while the accuracy of written language in the experimental group decreased, the complexity of their language might have increased. Such interdependence between accuracy and complexity has been supported by some researchers (Foster & Skehan, 1996; Skehan, 1996; 1998; Bygate, 1999), while others claim that all three aspects of L2 performance develop simultaneously and evenly (Robinson, 2001; 2003; Vercellotti, 2017). The main focus of this research was on the accuracy of learner language; analysis of its complexity is, therefore, beyond the scope of this book and would require further work to establish whether there is any relationship between the accuracy and complexity of advanced learner language.

It can also be hypothesised that students in the experimental group were more aware of the typical problem areas and this led them to contemplate more about the language they used, sometimes with the undesirable effect of 'overthinking'. This claim can be illustrated by some of the comments made by the participants of the study from the experimental group in the questionnaires distributed after the post-test. One of the students from the experimental group (coded as EK) described the reason why s/he committed the errors as follows: *I wouldn't probably make these mistakes when I'm speaking, maybe I would make them while writing, because when I'm thinking too much about what is wrong, I almost always make a mistake*, and another (coded as EAF) said: *I tend to focus on things that are not the problem, but the more I look at them, the more irrational mistakes I create.* This may be, to a certain extent, related to another important factor and that is the duration of intervention. It can be hypothesized that raising awareness of error-prone areas in L2 might result in temporary insecurity and therefore lower accuracy of learner language, but the long-term effect of intervention might still be positive. This would require further investigation in a long-term project.

Another plausible cause of the decrease in accuracy of written language in the experimental group could be attributed to learner differences. While some learners benefit from form-focused instruction and it helps them to eliminate errors in their production, it might have an adverse effect on others. While research has provided ample evidence of the benefits of form-focused instruction in general (Norris & Ortega, 2000; Spada, 2011; Ellis & Shintani, 2014), there might be differences among learners. Further studies analysing the impact of different learning styles on FFI would need to be undertaken.

RQ5) What areas of the advanced learner language of Czech and Slovak students of English as a foreign language are most frequently affected by errors?

Based on the data elicited from learners' spoken and written production, a corpus of advanced learner language was compiled. In order to identify the main error-prone areas and determine the frequency of errors, all data were error-tagged and corpus-based analysis was carried out. The results emerging from the analysis indicate that most of the errors occur in the grammar domain, especially in the use of articles, both missing and redundant. The second most frequently affected domain is lexis,

followed by errors in word order, redundant use of single words and missing words (for more details see section 5.5).

Despite the fact that the most error-prone domain was grammar, the most numerous error type in the corpus was lexical errors affecting single words (coded as LS according to the Louvain error tagging system) and the third most frequent error type was lexical errors affecting phrases (LP). These were often errors resulting from literal translations from L1, or false friends. In the grammar domain, errors in the use of articles were by far the most frequently occurring error type in the whole corpus. As the use of articles is a frequently observed flaw of Czech learner English, this result is not unexpected.

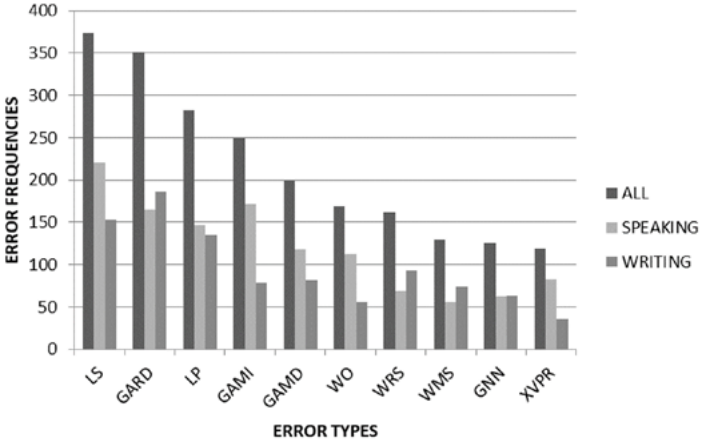
The high incidence of these error types suggests that language transfer might be a considerable influence on the accuracy of L2, even at advanced levels. These results support evidence from previous studies, both international and Czech, as reported above in chapter 2. Lexical errors affecting single words and phrases, and errors in the use of articles were identified as the most problematic areas of advanced Czech and Slovak learners of English as a foreign language in this research study.

RQ6) Do these areas differ in spoken and in written learner language? If they do, what are the main differences?

Following the analysis of learner language as a whole, error frequencies in spoken and written language were explored separately, in order to determine the differences, if any, between the two. This analysis revealed only minor differences between spoken and written learner language. While the occurrence of error types in the domains was identical for both datasets, with the grammar domain being the most frequent, followed by the lexical, word and lexico-grammatical domains, a more thorough analysis of error types revealed some minor differences, as specified above in section 5.5.2. A difference was observed in the redundant use of single words (WRS) which was the 9th most frequent error type in speaking but the 4th in writing. This difference could have been caused by adopting slightly different criteria when correcting errors in speaking, as opposed to writing. In speech, redundancy is, to a certain extent, a natural language feature and might not have been perceived as wrong, which is not the case in writing. An error type which only appeared among the first ten most numerous types in writing, but not in speaking, was inappropriate use of register (ZIR). This result is probably not

unexpected, as in written language stricter rules are commonly applied to using appropriate register than there are in speaking. The differences between error types and their frequencies in the whole corpus and in the written and spoken language are illustrated in Figure 20 below.

Figure 20
Corpus-based analysis – comparing error types and error frequencies in the whole corpus, in spoken and written language



Note: error type codes used in the figure: LS indicates a lexical error in a single word, GARD – a redundant use of indefinite article, LP – a lexical error in a phrase, GAMI – a missing indefinite article, GAMD – a missing definite article, WO – erroneous word order, WRS – a single redundant word, WMS – a single missing word, GNN – erroneous noun plural form, XVPR – a verb used with a wrong preposition

6.2 Comparing and contrasting the results of the corpus-based analysis to similar studies

The results of the current corpus-based analysis were compared with the findings reported in two similar studies, both using the Louvain error-tagging system: Gráf's (2015), which was conducted in the Czech Republic, and Götz's, which was carried out in Germany (2015). Both research studies were corpus-based error analyses of advanced spoken English. For this reason, it is only possible to compare and contrast data from spoken language. Similarly to the research presented here, the participants of Gráf's study were students of English philology; it

can therefore be assumed that their proficiency levels in English were comparable with those analysed in this study. The participants of Götz's research were German university students of English in their third or fourth years; it can be expected that their proficiency levels in English will also be comparable. In some respects, the current research supports evidence from these two studies.

Unlike the current study, Gráf analysed two aspects of advanced learner language, accuracy and fluency. The comparison of the results between his study and this one is, therefore, limited to the accuracy of spoken language. In accord with Gráf's results, the same two most error prone domains have been identified in this research, grammar and lexis. The frequency of errors in the grammar domain was comparable in both studies, with 55% in Gráf's research (labelled as CZ 1) and 49% in the research presented here (CZ 2). Errors affecting lexis were the second most frequent error domain in both studies, with 33% in CZ 1, and significantly lower, with 23% in CZ 2. Lexico-grammatical errors in CZ 2 were almost twice as numerous as in CZ 1; they occurred in 9% of all identified errors, as opposed to 5% in CZ 1. Taken together, lexical and lexico-grammatical errors occurred more frequently in CZ 1 with 38% as opposed to 32% in CZ 2. Interestingly, the third most frequent error domain in CZ 2 with 15% of all errors, unlike Gráf's study, was errors in the word domain, indicating erroneous word order, words missing or redundant. Error frequencies in the other domains were comparable in both studies (for details, see Table 23).

In her study of advanced learner English, Götz (2015) (whose study is indicated below as GE) reported the grammar domain as being the most affected by errors, with 47% of all errors, a result which is very close to the current study with 49% of errors affecting grammar. In accord with CZ 1 and CZ 2, the second most numerous error domain was lexical, with 36% of all errors. Unlike CZ 1 but in agreement with the present results, the third error domain was affecting word order and errors affecting words, both single and word phrases. This was 9% in GE, and 15% in CZ 2.

The most significant finding emerging from this comparison is that grammar is the domain most frequently affected by errors in all three studies, followed by errors in lexis. The third most frequent error domain in GE and CZ 2 is errors affecting words, in CZ 1 is lexico-grammatical errors (for the exact error frequencies, see Table 23 below).

Overall, however, the results obtained from this comparison indicate that grammar is the most error-prone area of advanced learner language in all three studies.

Table 23
Comparing error frequencies in six error domains reported in three studies, adapted and complemented (Gráf, 2015, p. 126)

error domain	GE	Götz	CZ 1	Gráf	CZ 2	Kalová
	count	%	count	%	count	%
form[37]	24	2%	3	0%	83	4%
grammar	**627**	**47%**	**710**	**55%**	**996**	**49%**
lexico-grammatical	67	5%	71	5%	190	9%
lexical	480	36%	434	33%	472	23%
word	114	9%	51	4%	295	15%
infelicities	23	2%	30	2%	3	0%
total	1335		1299		2039	

A more detailed analysis of the most frequent error types in the grammar domain has revealed some interesting findings. The most frequent error type in both Czech advanced learner corpora was errors in the use of articles, while in the German corpus it was the use of tenses. In CZ 1, the second most frequent error type was in the use of tenses, while in CZ 2 it was errors affecting word order and over or underuse of single words. The 4th and 5th most frequently occurring error types in all three corpora were errors in lexical phrases, especially 'false friends', and erroneous use of prepositions (for details, see Table 24).

The findings summarised above indicate that types of errors in the grammar domain might be attributable to the influence of the L1 of the speakers. Czech speakers of English, whose mother tongue does not use articles, display significantly more errors in their correct use than German speakers, whose language uses articles similarly to the target language. This finding has been reported previously for Czech, Chinese, and other languages with missing or different use of articles (Koppel et al., 2005; Wong & Dras, 2009). Further research should be undertaken to reveal some more L1-induced differences in error types among speakers of different mother tongues.

37 This error domain is referred to as morphological errors in Gráf's study.

Table 24
Comparing error type frequencies in the grammar domain reported in three studies

	GE	CZ 1	CZ 2
	GVT	GA	GA
	LS	GVT	LS
error type	GA	LS	W
	LP	LP	LP
	LSP	LSP	LSP

Note: error type codes used in the figure: GVT indicates a grammar error in tense, GA indicates an erroneous use of the article, LS indicates a lexical error in a single word, LP indicates a lexical error affecting a phrase, LSP indicates a lexical error affecting a preposition, W indicates erroneous word order, words missing or used redundantly.

6.3 Summary

In this chapter, answers to the research questions and possible explanations for the findings have been provided. The increase in GJT and CBM scores in the pilot group was the most pronounced and statistically significant, indicating positive impact of educational intervention. The main study has been unable to demonstrate that the improvement in GJT and CBM was due to the intervention as both groups, experimental and control, improved in a statistically significant way. In speaking, which was only analysed in the experimental and control groups, it was only the experimental group in which statistically significant improvement occurred, while there was no statistically important change in the control group. One unanticipated finding was that in writing, it was the control group that improved; this slight positive change was not, however, statistically important.

As it is common in educational research, these findings are not easy to interpret. There are, however, quite a few interesting outcomes. The intervention seemed to have a pronounced effect on the accuracy of spoken production but not on written production. It appears that increasing accuracy of writing would require an intervention longer than a one-term course. On completion of the course, the respondents may

still be processing and internalizing what they had learned on the course. Also, focused attention on accuracy and raising awareness of typical errors may lead to increased levels of certainty when identifying errors (GJT and CBM) but to temporary uncertainty in learners' production. This claim can be supported by data elicited from the questionnaires distributed in the experimental and control groups after the post-test.

chapter 6 also attempted to identify main error-prone areas of advanced learner English and decide whether or not any differences in error domains and error types are identifiable in advanced written and spoken language. Comparison of the findings with those of other studies (Götz, 2015; 2019; Gráf, 2015; 2017) confirms that grammatical errors, followed by errors in vocabulary are the most error-prone domains. In accord with a comparable Czech study (Gráf, 2015) it is errors in the use of articles that result as the most frequent error type. Unlike Gráf's study (2015) in which errors in the use of tenses were identified as the second most frequent error type, the research presented in this book indicates that it is errors in lexis affecting single words.

Answers to some of the issues emerging from the above-described comparison relate directly to foreign language learning and teaching. In the final chapter, summarising the research project described in this book, some of its pedagogical implications will be explored, its limitations will be outlined, and some suggestions for further research will be suggested.

CONCLUSION

The main goal of this research was to determine how the accuracy of advanced learner language can be affected by targeted educational intervention. The accuracy of Czech and Slovak advanced English as a foreign language was explored from different perspectives using a variety of research tools, as specified in subchapter 4.6. First, students' ability to identify and correct errors was measured, then the accuracy of their free production in the target language was examined through a corpus-based analysis. An annotated corpus was devised as a research tool in order to identify the most error-prone areas of advanced learner language.

The efficacy of the intervention was first explored in the pilot group, whose accuracy and certainty in answering were tested in the Grammaticality Judgement Test and Certainty-Based Marking. In the main study, the experimental and control groups were also tested for accuracy and certainty, and, in addition, their free spoken and written language was analysed. The pilot and experimental groups both received targeted educational intervention in the form of a one-term blended learning course aimed at improving their accuracy in English as the target language; the details of the intervention are specified in chapter 3. All groups were pre- and post-tested and the results of testing were then analysed using statistical methods, compared and contrasted.

The analysis revealed that the most pronounced beneficial effect of the intervention was on learners' ability to identify and correct errors and to enhance their certainty in answering. These two aspects have improved in a statistically significant way in both the pilot and the experimental groups. It has also emerged from the analysis that the accuracy of spoken language in the experimental group has improved; this increase in accuracy was statistically significant. A result which was unexpected was the adverse effect on learners' written production. Contrary to the expected improvement, more errors occurred in the post-test than in the pre-test. In the control group, statistically significant

improvement occurred in the accuracy and certainty of answering; the results in speaking and writing in this group were not, however, statistically significant. Details of the analyses conducted for all groups are provided in chapter 5.

The results obtained from the comparison between experimental and control groups manifest that the experimental group has improved in three of the four aspects under scrutiny. Compared to the control group, only in the experimental group has the accuracy of spoken language increased in a statistically significant way. This indicates that the overall impact of targeted intervention aimed at increasing the accuracy of learner language was positive.

An important part of this investigation was collecting samples of spoken and written learner language and their analysis conducted in order to identify the main problematic areas for Czech and Slovak learners of English as a foreign language. The samples were also assessed individually to determine whether there were any differences in the frequency of error types between spoken and written language. Two major error-prone areas of advanced learner language have been identified: grammar errors, especially in the use of articles, either redundant or missing, and lexical errors, affecting both single words and phrases or parts of phrases. Errors affecting word order and lexico-grammatical errors were also relatively frequent, while errors in style, form and punctuation were rare. Similarly to the results reported in other studies, as detailed in chapter 2, many of these errors can be attributed to the negative influence of learners' first language. One of the contributions of this research is the fact that it analysed both spoken and written language. No major differences have been identified between the two, the impact of intervention, however, differed.

Confirming some of the previous findings, described in 6.2, this book offers some important insights into the nature of advanced learner language, especially by identifying the areas which seem to be problematic for advanced Czech and Slovak learners of English. One of the major contributions of this research has been compiling a corpus of advanced learner English which was annotated for errors and complemented with metadata. This corpus can be used for further research of different aspects of learner language. Identifying areas most frequently affected by errors in speakers of English whose L1 is Czech or Slovak is not important only for further research. Targeted treatment

of such errors can be successfully implemented in the language classroom, whether in devising tailor-made remedial materials, or raising learners' awareness of typical errors. The findings from this research provide a solid empirical foundation on which a textbook of advanced English for Czech and Slovak learners addressing their typical error-prone areas could be based. Despite some limitations, as detailed below, the present research provides evidence of the benefits of form-focused instruction in the advanced language classroom. These findings suggest that focus on form and errors in foreign language teaching should not be rejected as even learners at advanced proficiency levels can benefit from them.

Pedagogical implications

One of the outcomes of the research presented in this book was also providing teachers of English with relevant data about advanced learner language of Czech and Slovak speakers of English which would be easily applicable in the foreign language classroom. An important pedagogical implication of this study is a possible use of the data collected from samples of spoken and written production of students to prevent typical, often L1-induced errors. This can be done both in *syllabus design* and *devising tailor-made remedial study materials* for advanced Czech and Slovak learners of English. Identified problem areas could also be reflected in how grammar and lexis are taught and practised. Currently available teaching materials are far too often designed for a 'typical global' learner rather than 'tailor-made', designed to meet the needs of students with well-defined problematic areas of language that require focused attention. It has been reported in research (Seidlhofer, 2002; Han & Cook, Wei, 2009; Han & Tarone, 2014; Götz, 2015; 2019) that the errors (advanced) learners make are often closely related to their mother tongues and should, therefore, be addressed accordingly in the English language classroom (Seidlhofer, 2002; Thewissen, 2015). As mentioned above (in section 6.2), Czech learners of English manifest typical error patterns different from speakers of other languages and this should be reflected both in the classroom and in developing teaching resources. For years, one of the few error-focused reference materials specifically designed for Czech learners was Don Sparling's *English or Czenglish?* first published in 1989, when no corpus-based analysis

of learner language was available. It is one of the concrete outcomes of the research detailed in this book that an updated and thoroughly revised edition of the *English or Czenglish* reference book (Sparling et al., 2021) addressing error-prone areas typical of Czech and Slovak speakers of English as identified in this research study into advanced learner language was published in 2021.

Another possible area that is worth considering further is *the use of learner corpora in the foreign language classroom.* This has been the subject of intense debate, with many conflicting views expressed, and regarded as "a highly controversial issue", as reported by Granger (Granger et al., 2002, p. 26). This controversy lies especially in the fact that it is still firmly believed by some researchers (Truscott, 1996; Choděra, 2000; Gray, 2004) and teachers alike that exposing learners to errors can have harmful effects on their acquisition of the target language, and that they do not benefit from corrective feedback. Others, however, increasingly accentuate the need for focused attention paid to error treatment and to raising awareness of fossilized errors in particular (Schulz, 1996; Ferris, 1999; Bitchener & Ferris, 2012; Shintani & Ellis, 2013; Ellis & Shintani, 2014). This seems to be important especially with advanced learners, as it is beneficial for their learning to help them to "notice the gap between their own and target language forms" (Granger et al., 2002, p. 26). Advanced learners are believed to be able to work with errors without perceiving them as stigmatising and use them as an opportunity to learn and cultivate their target language.

Regarding the use of learner corpora in the advanced language classroom, Seidlhofer suggests a novel approach in which samples of learner language are analysed by the same learners who produced them, becoming "both participants in and analysts of their own language use" (Seidlhofer, 2002, p. 213). Using 'learning-driven data' in teaching makes the learning experience more personalized and meaningful, raising not just learners' awareness of their own problematic areas of language but also their motivation to learn (ibid). A focused course designed to address error-prone areas of L2 typical of a particular group of learners sharing the same L1, together with specific materials addressing the problem areas and engaging learners, might therefore yield favourable results.

Limitations of the study

The generalizability of the findings presented in this book is naturally subject to certain limitations. Unlike controlled experiments conducted in laboratory settings, using experimental design in educational research is not without flaws. First, random allocations of participants to groups in this research project were impossible, due to constraints imposed by the rules of the institution in which it was conducted. Second, the groups under scrutiny were of different sizes, which caused some difficulty in the comparison. This was, at least partly, overcome by setting strict criteria for the selection of participants and making the conditions of testing in the experimental and control groups as similar as possible. It can be expected that by imposing these criteria, the influence of factors other than the intervention was limited to a minimum, which enabled the comparison of the results and assessment of the efficacy of intervention. At the same time, however, it might have also negatively impacted the numbers of students applying for participation in the control group, which resulted in a much smaller size of the control group. Therefore, it is important to bear in mind that these findings cannot be extrapolated to all Czech and Slovak advanced learners of English. The findings produced by descriptive statistics therefore only relate to the sample presented here.

Further, the study was limited by exploring one aspect of performance – accuracy – only, without taking the other two – fluency and complexity – into account. The process of compiling and analysing the research data, however, was so demanding that equal analysis of all three dimensions, at least for a sole researcher, would have been unrealistic.

Notwithstanding these limitations, the study adds to our knowledge of learner language by identifying its most problematic areas for advanced Czech and Slovak learners, as well as providing some practical implications for foreign language teaching. It also raises questions which could be addressed in future research studies, some of which will be outlined in the following section.

Suggestions for further research

This book has attempted to answer questions related to accuracy of learner language. It has also posed many questions which have not been addressed here but might be a fruitful area to explore further.

This study has examined how accuracy, as one of the three aspects of learner language, can be impacted by educational intervention. Further research should be carried out to establish how the other two aspects, complexity and fluency, are affected. Such analysis could provide a more complex picture of advanced learner language, and changes resulting from the intervention. It might also determine the relationship between the three aspects of learner language, how they interact and whether learners' focus on accuracy results in decreased complexity or fluency, or whether they all develop simultaneously. Such research would help both SLA researchers and FLT teachers to understand how advanced learner language evolves.

What could also be assessed is how other factors influencing learner language, and changes in it caused by intervention, affect the results of testing. One of the important external factors worth examining is the length of the intervention; intervention longer than the one presented in this research might yield a more favourable effect on the development of complexity, accuracy and fluency of advanced learner language. Also, exploring the characteristics of the task used in testing both in terms of its cognitive difficulty, its modality, writing and speech, and of planning the task might shed some more light on the changes occurring in advanced learner language. Internal factors, especially learners' motivation and differences in their learning styles could generate some insights into how effective form-focused instruction is in different learners.

Most importantly, the annotated learner corpus which was compiled in this research could provide enough data for further analysis. In addition to analysing accuracy, complexity and fluency of learner language, it might also be fruitful to examine other aspects of learner language, e.g. its pragmatic features. Gender-related differences and L1 influence could also produce interesting findings. Such research would be a logical sequence of the current study as the most demanding task of compiling the learner corpus has already been completed and complemented with the metadata which would allow for a variety of aspects of advanced learner language to be explored.

References

Bachman, L. (1990). *Fundamental considerations in language testing* (1 ed.). Oxford University Press.

Bachman, L., & Palmer, A. (1996). *Language testing in practice: designing and developing useful language tests* (1 ed.). Oxford University Press.

Bamgbose, A. (1998). Torn between the norms: innovations in world Englishes. *World Englishes, 17*(1), 1-14. https://doi.org/10.1111/1467-971X.00078

Barlow, M. (2005). Computer-based analyses of learner language. In R. Ellis & G. Barkhuizen, *Analysing learner language* (5th, pp. 335-357). Oxford University Press.

Baroni, M., Kilgarriff, A., & Rychlý, P. (2006). WebBootCaT: Instant domain-specific corpora to support human translators. In *Proceedings of EAMT: 11th Annual Conference of the European Association for Machine Translation* (1st, pp. 247-252). NA.

Belz, M., Sauer, S., Lüdeling, A., & Mooshammer, C. (2017). Fluently disfluent?. *International Journal of Learner Corpus Research, 3*(2), 118-148. https://doi.org/10.1075/ijlcr.3.2.02bel

Bestgen, Y., Granger, S., & Thewissen, J. (2012). Error Patterns and Automatic L1 Identification. In S. Jarvis & S. Crossley (eds.), S. Jarvis, S. Crossley, *Approaching language transfer through text classification: explorations in the detection-based approach* (1 ed., pp. 127-153). Multilingual Matters.

Bitchener, J., & Ferris, D. (2012). *Written Corrective Feedback in Second Language Acquisition and Writing* (1 ed.). Routledge. https://scholar.google.com/

Blake, R. (2011). Current Trends in Online Language Learning. *Annual Review of Applied Linguistics: Topics in Second Language Pedagogy, 31*(), 19-35. https://doi.org/10.1017/S026719051100002X

Bley-Vroman, R. (1983). The Comparative Fallacy in Interlanguage Studies: The Case of Systematicity 1. *Language Learning, 33*(1), 1-17. https://doi.org/10.1111/j.1467-1770.1983.tb00983.x

Börjars, K., & Burridge, K. (2019). *Introducing English grammar* (Third edition). Routledge.

Brabcová, K., & Skarnitzel, R. (2018). Foreign or native-like?: The attitudes of Czech EFL learners towards accents of English and their use

as pronunciation models. *Studie z aplikované lingvistiky – Studies in Applied Linguistics, 9*(1), 38-50.

Brown, G., Malmkjær, K., & Williams, J. (eds.). (1996). *Performance and competence in second language acquisition* (2004 ed.). Cambridge University Press.

Brown, H. (2007). *Teaching by principles: an interactive approach to language pedagogy* (3rd ed). Pearson Education.

Brumfit, C. (1984). *Communicative methodology in language teaching: the roles of fluency and accuracy* (1st). Cambridge University Press.

Brumfit, C., & Johnson, K. (1979). *The communicative approach to language teaching* (1st). Oxford University Press.

Brunner, M., Diemer, S., & Schmidt, S. (2016). "It's always different when you look something from the inside". *International Journal of Learner Corpus Research, 2*(2), 323-350. https://doi.org/10.1075/ijlcr.2.2.09bru

Bygate, M. (1999). Quality of language and purpose of task: patterns of learners' language on two oral communication tasks. *Language Teaching Research, 3*(3), 185-214. https://doi.org/10.1177/136216889900300302

Canale, M. (1983). From communicative competence to communicative language pedagogy. In J. Richards & R. Schmidt, *Language and communication. Longman*. (1st, p. 288). Routledge.

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics, 1*(1), 1-47. https://doi.org/10.1093

Chee, M., Wan, E., & Thiel, T. (1999). Mandarin and English Single Word Processing Studied with Functional Magnetic Resonance Imaging. *Journal of Neuroscience, 19*(8), 3050-3056. https://doi.org/https://doi.org/10.1523/JNEUROSCI.19-08-03050.1999

Choděra, R. (2000). *Výuka cizích jazyků na prahu nového století* (Vyd. 1). Ostravská univerzita – Pedagogická fakulta.

Choděra, R. (2013). *Didaktika cizích jazyků: úvod do vědního oboru* (Vyd. 2). Academia.

Chomsky, N. (1965). *Aspects of the Theory of Syntax* (1st). MIT Press.

Chráska, M. (2016). *Metody pedagogického výzkumu: základy kvantitativního výzkumu* (2., aktualizované vydání). Grada.

Cobb, T. (2003). Analyzing late interlanguage with learner corpora: Quebec replications of three European studies. *Canadian Modern Language Review-Revue Canadienne Des Langues Vivantes, 2003 Mar, 59*(3), 393-423. https://doi.org/0008-4506

Cohen, L., Manion, L., Morrison, K., & Bell, R. (2011). *Research methods in education* (7th ed). Routledge.

Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR). (2001) (1st). CUP. https://www.coe.int/en/web/common-european-framework-reference-languages/home (2001) (1oth). Cambridge University Press. http://assets.cambridge.org/97805218/03137/frontmatter/9780521803137_frontmatter.pdf

Corder, S. (1967). The Significance of Learner's Errors. *IRAL – International Review of Applied Linguistics in Language Teaching, 5*(1-4), 161-170. https://doi.org/10.1515/iral.1967.5.1-4.161

Corder, S. (1981). *Error Analysis and Interlanguage* (1st). Oxford University Press.

Crosthwaite, P. (2016). L2 English article use by L1 speakers of article-less languages. *International Journal of Learner Corpus Research, 2*(1), 68-100. https://doi.org/10.1075/ijlcr.2.1.03cro

Crystal, D. (2003). *English as a Global Language* (1st). Oxford University Press.

Červinková Poesová, K., & Lancová, K. (2021). Cause for optimism. Non-native pre-service teachers' attitudes towards and beliefs about accent. In A. Kirkova-Naskova, A. Henderson & J. Fouz-González, *English Pronunciation Instruction: Research-based insights* (1st, pp. 91-117). John Benjamins Publishing Company.

Dagneaux, E., Denness, S., & Granger, S. (1998). Computer-aided error analysis. *System, 26*, 163-174.

DeKeyser, R. (1995). Learning Second Language Grammar Rules. *Studies in Second Language Acquisition, vol. 17*(issue 3), 379-410. https://doi.org/10.1017/S027226310001425X

Deshors, S. (2018). Does the passé composé influence L2 learners' use of English past tenses?. *International Journal of Learner Corpus Research, 4*(1), 23-53. https://doi.org/10.1075/ijlcr.17007.des

Díez-Bedmar, M., & Papp, S. (2008). The use of the English article system by Chinese and Spanish learners. *Linking up contrastive and learner corpus research, 66*(1), 147-175. https://doi.org/10.1163/9789401206204_007

Doughty, C. (1991). Second Language Instruction Does Make a Difference. *Studies in Second Language Acquisition, vol. 13*(issue 4), 431-469. https://doi.org/10.1017/S0272263100010287

Edge, J. (1989). *Mistakes and correction* (1st). Longman.

Ellis, N. (1993). Rules and instances in foreign language learning: Interactions of explicit and implicit knowledge. *European Journal of Cognitive Psychology, 5*(3), 289-318. https://doi.org/10.1080/09541449308520120

Ellis, N. (2005). At the Interface: dynamic interactions of explicit and implicit knowledge. *Studies in Second Language Acquisition, vol. 27*(issue 02), 305-352. https://doi.org/10.1017/S027226310505014X

Ellis, R. (2006). Current Issues in the Teaching of Grammar: An SLA Perspective. *TESOL Quarterly, 40*(1), 83-107. https://doi.org/10.2307/40264512

Ellis, R., & Barkhuizen, G. (2005). *Analysing learner language* (1st pub). Oxford University Press.

Ellis, R., Loewen, S., Elder, C., Erlam, R., Philp, J., & Reinders, H. (2009). *Implicit and explicit knowledge in second language learning, testing and teaching* (1st). Multilingual Matters.

Ellis, R., & Shintani, N. (2014). *Exploring language pedagogy through second language acquisition research* (First published). Routledge/Taylor & Francis Group.

Ellis, R., & Yuan, F. (2004). The effect of planning on fluency, complexity, and accuracy in second language writing. *Studies in Second Language Acquisition, 26*(1), 59-84. https://doi.org/10.1017/S0272263104026130

Ferris, D. (1999). The case for grammar correction in L2 writing classes: A response to truscott (1996). *Journal of Second Language Writing, 8*(1), 1-11. https://doi.org/10.1016/S1060-3743(99)80110-6

Fillmore, C., Kempler, D., & Wang, W. (eds.). (1979). *Individual Differences in Language Ability and Language Behavior* (1st). Academic Press.

Foster, P., & Skehan, P. (1996). The Influence of Planning and Task Type on Second Language Performance. *Studies in Second Language Acquisition, 18*(3), 299-323. https://doi.org/10.1017/S0272263100015047

Gardner-Medwin, T., & Curtin, N. (2007). Certainty-Based Marking (CBM) for reflective learning and proper knowledge assessment. *Available from: https://www.researchgate.net/publication/228648846_Certainty-Based_Marking_CBM_for_reflective_learning_and_proper_knowledge_assessment [accessed Sep 01 2019].*

Gass, S., & Mackey, A. (2002). Frequency Effects and Second Language Acquisition. *Studies in Second Language Acquisition, 24*(2), 249-260. https://doi.org/10.1017/S0272263102002097

Gilquin, G. (2019). Light verb constructions in spoken L2 *English. International Journal of Learner Corpus Research, 5*(2), 181-206. https://doi.org/10.1075/ijlcr.18003.gil

Göbel, K., & Vieluf, S. (2018). Specific Effects of Language Transfer Promoting Teaching and Insights Into the Implementation in EFL-Teaching. *ORBIS SCHOLAE, 11*(3), 103-122. https://doi.org/10.14712/23363177.2018.279

Gósy, M., Gyarmathy, D., & Beke, A. (2017). Phonetic analysis of filled pauses based on a Hungarian-English learner corpus. *International Journal of Learner Corpus Research, 3*(2), 149-174. https://doi.org/10.1075/ijlcr.3.2.03gos

Götz, S. (2015). Tense and aspect errors in spoken learner English: Implications for language testing and assessment. In M. Callies & S. Götz (eds.), M. Callies, S. Götz, *Learner Corpora in Language Testing and Assessment: Studies in Corpus Linguistics 70* (1 ed., pp. 191-216).

John Benjamins Publishing Company. https://www.researchgate.net/publication/297685507_Tense_and_Aspect_Errors_in_Spoken_Learner_English

Götz, S. (2019). Filled pauses across proficiency levels, L1s and learning context variables. *International Journal of Learner Corpus Research, 5*(2), 159-180. https://doi.org/10.1075/ijlcr.17018.got

Graddol, D. (1997). *The Future of English?: A guide to forecasting the popularity of the English Language in the 21st century* (2nd). The British Council. https://doanbangoc.files.wordpress.com/2012/07/the-future-of-english.pdf

Gráf, T. (2015). *Accuracy and fluency in the speech of the advanced learner of English* [Dizertační práce]. Univerzita Karlova v Praze Filozofická fakulta Ústav anglického jazyka a didaktiky Filologie.

Gráf, T. (2017). Verb Errors in Advanced Spoken English. *Acta Universitatis Carolinae Philologica, 2017*(1), 131-149. https://www.ceeol.com/search/article-detail?id=550602

Granfeldt, J. (2007). *Speaking and Writing in L2 French: Exploring Effects on Fluency, Complexity and Accuracy* (1st). Contactforum.

Granger, S. (2003). Error-tagged learner corpora and CALL: A promising synergy. *CALICO Journal, 20*(3), 445-480.

Granger, S. (2009). Learner corpora. *In Corpus linguistics: an international handbook* (1 ed., pp. 259-274). Walter de Gruyter.

Granger, S. (2015). Contrastive interlanguage analysis: A reappraisal. *International Journal of Learner Corpus Research, 1*(1), 7-24. https://doi.org/10.1075/ijlcr.1.1.01gra

Granger, S., Gilquin, G., & Meunier, F. (eds.). (2015). *The Cambridge Handbook of Learner Corpus Research* (1st). Cambridge University Press.

Granger, S., Hung, J., & Petch-Tyson, S. (eds.). (2002). *Computer learner corpora, second language acquisition, and foreign language teaching* (1st). John Benjamins.

Gray, R. (2004). Grammar Correction in ESL/EFL Writing Classes May Not Be Effective. *The Internet TESL Journal*, (11), NA. http://iteslj.org/Technique/Gray-WritingCorrection.html

Gregg, K. (1984). Krashen's Monitor and Occam's Razar. *Applied Linguistics, 5*(2), 79-100. https://doi.org/10.1093/applin/5.2.79

Halliday, M. (1977). Learning how to mean: explorations in the development of language. (Explorations in Language Study Series, edited by Peter Doughty and Geoffrey Thornton). London. *Language in Society, 6*(1), 114-118. https://doi.org/10.1017/S0047404500004942

Hamed, M. (2014). Conjunctions in Argumentative Writing of Libyan Tertiary Students. *English Language Teaching, 7*(3), 108-120. https://doi.org/10.5539/elt.v7n3p108

Hanušová, S. (2003). Pregraduální příprava učitelů anglického jazyka na fakultách vysokých škol. *Pedagogická orientace, 13*(1), 17-33.

Han, Z. (2004). Fossilization: five central issues. *International Journal of Applied Linguistics, 14*(2), 212-242. https://doi.org/10.1111/j.1473-4192.2004.00060.x

Han, Z. (2006). Fossilization: Can Grammaticality Judgement Be a Reliable Source of Evidence?. In Z. Han & T. Odlin (eds.), Z. Han, T. Odlin, *Studies of Fossilization in Second Language Acquisition* (1 ed., pp. 56-82). Multilingual Matters.

Han, Z. (2013). Forty years later: Updating the Fossilization Hypothesis. *Language Teaching, 46*(2), 133-171. https://doi.org/10.1017/S0261444812000511

Han, Z. (2014). From Julie to Wes to Alberto. In Z. Han & E. Tarone, *Interlanguage: Forty years later* (1st, p. Revisiting the construct of fossilization). John Benjamins Publishing Company.

Han, Z., & Cook, V. (eds.), Wei, L. (2009). Interlanguage and fossilization: Towards an analytic model. *Contemporary applied linguistics*, 137-162.

Han, Z., & Tarone, E. (eds.). (2014). *Interlanguage: Forty years later* (1st). John Benjamins Publishing Company.

Hartmann, R., & Stork, F. (1976). *Dictionary of Language and Linguistics* (1st). John Wiley & Sons.

Hendrich, J. (1988). *Didaktika cizích jazyků* (1. vyd). Státní pedagogické nakladatelství.

Hendrikx, I., Van Goethem, K., & Wulff, S. (2019). Intensifying constructions in French-speaking L2 learners of English and Dutch. *International Journal of Learner Corpus Research, 5*(1), 63-103. https://doi.org/10.1075/ijlcr.18002.hen

Hewings, M. (2009). *Grammar for CAE and proficiency: with answers: self-study grammar reference and practice* (1 ed.). Cambridge University Press.

Horner, D. (1987). Acquisition, learning and the monitor: A critical look at Krashen. *System, 15*(3), 339-349. https://doi.org/10.1016/0346-251X(87)90008-X

Housen, A., Kuiken, F., & Vedder, I. (2012). *Dimensions of L2 performance and proficiency: complexity, accuracy and fluency in SLA* (first). John Benjamins Pub. Co.

Hrdlička, M. (2012). Jazyková chyba a práce s ní v jazykovém vyučování. In *Čeština – cílový jazyk a korpusy* (1 ed., pp. 89-109). Technická univerzita v Liberci.

Hubáčková, Š. (2013). *Využití moderních komunikačních technologií ve výuce cizího jazyka* [Dissertation]. UK FF.

Hymes, D. (1972). On Communicative Competence. In *Sociolinguistics: selected readings* (pp. 269-293). Penguin.

Jakšič, J., & Šturm, P. (2017). Accents of English at Czech Schools: Students' Attitudes and Recognition Skills. *Research in Language, 15*(4), 353–369.

James, C. (2013). *Errors in Language Learning and Use: Exploring Error Analysis* (5th). Routledge.

Janíková, V. (2013). *Mnohojazyčnost a didaktika cizích jazyků* (c2013). Masarykova univerzita.

Jenkins, J. (2007). *English as a Lingua Franca: Attitude and Identity* (5 ed.). Oxford University Press.

Kachru, B. (1985). Standards, Codification and Sociolinguistic Realism: The English Language in the outer circle. In R. Quirk, H. Widdowson & Y. Cantu, *English in the world: teaching and learning the language and literatures* (1 ed., p. NA). Cambridge University Press for the British Council.

Kachru, B. (1986). *The Alchemy of English: The Spread, Functions, and Models of Non-native Englishes* (NA). University of Illinois Press.

Kachru, B., Kachru, Y., & Nelson, C. (2009). *The Handbook of World Englishes* (1st pub). Wiley-Blackwell.

Kalová, S. (2017). Problematika jazykové normy ve výuce angličtiny. In *Acta Facultatis Paedogogicae Universitatis Comenianae Bratislavensis: Acta Doctorandorum: Filologické štúdie 3* (1 ed., pp. 108-126). KIRSCH-Verlag.

Kalová, S. (2018). Úloha chyb, korektivní zpětné vazby a normy v cizojazyčné výuce s důrazem na výuku angličtiny. *Pedagogická orientace, 28*(1), 72-110. https://doi.org/10.5817/PedOr2018-1-72

Karazoun, G. (2016). A Linguistic Analysis on Errors Committed by Jordanian EFL Undergraduate Students: A Case of News Headlines in Jordanian Newspapers. *English Language Teaching, 9*(8), 170-189. https://doi.org/10.5539/elt.v9n8p170

Kilgarriff, A., Marcowitz, F., Smith, S., & Thomas, J. (2015). Corpora and Language Learning with the Sketch Engine and SKELL. *Revue Française de Linguistique Appliquée, 20*(1), 61-80. https://www.researchgate.net/publication/282751109_Corpora_and_Language_Learning_with_the_Sketch_Engine_and_SKELL/references

Konečný, J. (2014). К ВОПРОСУ ИНТЕРФЕРЕНЦИИ И ПОЛОЖИТЕЛЬНОГО ПЕРЕНОСА В ПРОЦЕССЕ ОБУЧЕНИЯ РУССКОМУ ЯЗЫКУ В ЧЕШСКОЙ ШКОЛЕ. *Lingua viva: Odborný časopis pro teorii a praxi vyučování cizím jazykům a češtině jako cizímu jazyku*. X (18), 32-42. http://www.pf.jcu.cz/research/lingua_viva/LV18_2014.pdf

Koppel, M., Schler, J., & Zigdon, K. (2005). *Determining an Author's Native Language by Mining a Text for Errors.* Association for Computig Machinery.

Krashen, S. (1989). *Language acquisition and language education* (1st). Prentice Hall International.

Krashen, S. (2002). *Second Language Acquisition and Second Language Learning* (First internet edition). University of Southern California.

Krashen, S. (2004). Applying the Comprehension Hypothesis: Some Suggestions. *The International Journal of Foreign Language Teaching, 1*(1), 21-29. http://ijflt.com/images/ijflt/IJFLTWinter041.pdf

Kulič, V. (1971). *Chyba a učení: funkce chybného výkonu v učení a jeho řízení* (1st). Státní pedagogické nakladatelství.

Kumaravadivelu, B. (2006). *Understanding language teaching: from method to post-method* (1st). Mahwah, N.J.: Lawrence Erlbaum Associates.

Kyle, K., Crossley, S., & Kim, Y. (2015). Native language identification and writing proficiency. *International Journal of Learner Corpus Research,* 1(2), 187-209. https://doi.org/10.1075/ijlcr.1.2.01kyl

Lado, R. (1957). *Linguistics across Cultures: Applied linguistics for language teachers* (1 ed.). University of Michigan Press: Ann Arbor.

Larsen-Freeman, D., & Anderson, M. (2011). *Techniques & principles in language teaching* (3rd ed). Oxford University.

Lecumberri, M., Cooke, M., Wester, M., Cooke, M., & Wester, M. (2017). A bi-directional task-based corpus of learners' conversational speech. *International Journal of Learner Corpus Research, 3*(2), 175-195. https://doi.org/10.1075/ijlcr.3.2.04gar

Lennon, P. (2008). Contrastive Analysis, Error Analysis, Interlanguage. In S. Gramley & V. Gramley, *Bielefeld introduction to applied linguistics: a course book* (Bielefelder Schriften zu Linguistik und Literaturwissenschaft, pp. 51-60). Aisthesis. http://wwwhomes.uni-bielefeld.de/sgramley/CA-ErrorAnalysis-Interlang-Lennon.pdf

Lightbown, P. (2017). SLA Research and Foreign-Language Teaching. In E. Hinkel, *Handbook of Research in Second Language Teaching and Learning: Volume III* (1 ed., pp. 103-116). Routledge, Taylor & Francis Group. https://doi.org/NA

Lightbown, P., & Spada, N. (2013). *How languages are learned* (Fourth edition). Oxford University Press.

Long, R., & Hatcho, Y. (2018). The First Language's Impact on L2: Investigating Intralingual and Interlingual Errors. *English Language Teaching, 11*(11), 115-121. https://doi.org/10.5539/elt.v11n11p115

Lyster, R. (2007). *Learning and Teaching Languages Through Content: A counterbalanced approach* (1st). John Benjamins.

MacDonald, P., García-Carbonell, A., & Carot-Sierra, J. (2013). Computer Learner Corpora: Analysing Interlanguage Errors in Synchronous and Asynchronous Communication. *Language Learning & Technology: A refereed journal for second and foreign language scholars and educators, 17*(2), 36–56. https://doi.org/10251/65848

McLaughlin, B. (1987). *Theories of second language learning* (1st). Edward Arnold.

Means, B., Toyama, Y., Murphy, R., Bakia, M., & Jones, K. (2010). Evaluation of Evidence-Based Practices in Online Learning: A Meta-Analysis and Review of Online Learning Studies. U.S. Department of Education, Office of Planning, Evaluation, and Policy Development. https://repository.alt.ac.uk/629/

Medgyes, P. (2019). *The Non-Native Teacher* (1st updated). Swan.

Meisel, J., Clahsen, H., & Pienemann, M. (1981). On determining developmental stages in natural second language acquisition. *Studies in Second Language Acquisition, vol. 3*(issue 2), 109-135. https://doi.org/10.1017/S0272263100004137

Montrul, S. (ed.). (2014). *In Interlanguage: forty years later* (1 ed., pp. 75-104). John Benjamins Publishing Company.

Norris, J., & Ortega, L. (2000). Effectiveness of L2 Instruction: A Research Synthesis and Quantitative Meta-analysis. *Language Learning, 50*(3), 417-528.

Quinn Novotná, V. (2012). *World Englishes and English as a lingua franca: A reflection of global paradigmatic changes in the Czech Republic* [Unpublished doctoral dissertation]. Charles University.

Richards, J., & Rodgers, T. (2014). A*pproaches and methods in language teaching* (Third edition). Cambridge University Press.

Richards, J., & Schmidt, R. (2010). *Longman dictionary of language teaching and applied linguistics* (4th ed). Longman.

Robinson, P. (2001). Task complexity, task difficulty, and task production: exploring interactions in a componential framework. *Applied Linguistics, 22*(1), 27-57. https://doi.org/10.1093/applin/22.1.27

Robinson, P. (2003). The cognitive hypothesis, task design, and adult task-based language learning. *Second Language Studies, 21*(2), 45-107. http://hdl.handle.net/10125/40656

Rosa, E., & O'Neill, M. (1999). Explicitiness, Intake, and the Issue of Awareness: Another Piece to the Puzzle. *Studies in Second Language Acquisition, 21*(4), 511-556. https://doi.org/10.1017/S0272263199004015

Savignon, S. (1983). *Communicative Competence: Theory and Classroom Practice. Texts and Contexts in Second Language Learning.* (1st). Addison-Wesley Publishing Company.

Saville-Troike, M. (2006). *Introducing Second Language Acquisition* (1st). Cambridge University Press.

Seidlhofer, B. (2002). Pedagogy and local learner corpora: working with learning-driven data. In *Computer learner corpora, second language acquisition, and foreign language teaching* (1 ed., pp. 213-234). John Benjamins Pub.

Seidlhofer, B. (2004). 10. Research Perspectives on Teaching English as a Lingua Franca. *Annual Review of Applied Linguistics, 24*(1), 209-239. https://doi.org/10.1017/S0267190504000145

Seidlhofer, B. (2009). Common ground and different realities: world Englishes and English as a lingua franca. *World Englishes, 28*(2), 236-245. https://doi.org/10.1111/j.1467-971X.2009.01592.x

Seidlhofer, B., & Berns, M. (2009). Perspectives on English as a lingua franca: introduction. *World Englishes, 28*(2), 190-191. https://doi.org/10.1111/j.1467-971X.2009.01586.x

Seidlhofer, B., & Widdowson, H. (2017). Thoughts on independent English. *World Englishes, 36*(3), 360-362. https://doi.org/10.1111/weng.12274

Selinker, L. (1972). Interlanguage. *IRAL: International Review of Applied Linguistics in Language Teaching, 10*(3), 209-241.

Selinker, L. (2014). Interlanguage 40 years on: Three themes from here. In E. Tarone, *Interlanguage: Forty years later* (1st, pp. 231-240). John Benjamins Publishing Company.

Sharma, P. (2007). *Blended learning: using technology in and beyond the language classroom* (5th). Macmillan.

Shimanskaya, E., & Slabakova, R. (2017). L1–L2 differences in the L2 classroom: Anticipating Anglophone learners' difficulties with French pronoun interpretation. *Language Teaching Research, 23*(2), 259-277. https://doi.org/10.1177/1362168817739650

Shintani, N., & Ellis, R. (2013). The comparative effect of direct written corrective feedback and metalinguistic explanation on learners' explicit and implicit knowledge of the English indefinite article. *Journal of Second Language Writing, 22*(3), 286-306. https://doi.org/10.1016/j.jslw.2013.03.011

Shook, D. (1994). FL/L2 Reading, Grammatical Information, and the Input-to-Intake Phenomenon. *Applied Langauge Learning, 5*(2), 57-93. https://eric.ed.gov/?id=EJ508428

Schmidt, R. (1990). The Role of Consciousness in Second Language Learning. *Applied Linguistics, 11*(2), 129-158. https://doi.org/10.1093/applin/11.2.129

Schneider, G., & Gilquin, G. (2016). Detecting innovations in a parsed corpus of learner English. *International Journal of Learner Corpus Research, 2*(2), 177-204. https://doi.org/10.1075/ijlcr.2.2.03sch

Schulz, R. (1996). Focus on Form in the Foreign Language Classroom: Students' and Teachers' Views on Error Correction and the Role of Grammar. *Foreign Language Annals, 29*(3), 343-364. https://doi.org/10.1111/j.1944-9720.1996.tb01247.x

Skehan, P. (1996). A Framework for the implementation of task-based instruction. *Applied linguistics, 17*(1), 38-62. https://doi.org/10.1093/applin/17.1.38

Skehan, P. (1998). *A cognitive approach to language learning* (1st). Oxford University Press.

Smirnova, E. (2017). Using Corpora in EFL Classrooms: The Case Study of IELTS Preparation. *RELC Journal, 48*(3), 302-310. https://doi.org/10.1177/0033688216684280

Sok, S., Kang, E., & Han, Z. (2018). Thirty-five years of ISLA on form-focused instruction: A methodological synthesis. *Language Teaching Research, 23*(4), 403-427. https://doi.org/10.1177/1362168818776673

Spada, N. (2011). Beyond form-focused instruction: Reflections on past, present and future research. *Language Teaching, 44*(2), 225-236. https://doi.org/10.1017/S0261444810000224

Sparling, D. (1991). *English or Czenglish?: jak se vyhnout čechismům v angličtině* (2. vyd). Státní pedagogické nakladatelství.

Sparling, D., Kalová, S., & Rance, C. (2021). *English or Czenglish: Avoiding Czechisms in English* (1st revised and updated edition). Masarykova univerzita.

Stormbom, C. (2018). Epicene pronouns in intermediate to advanced EFL writing. *International Journal of Learner Corpus Research, 4*(1), 1-22. https://doi.org/10.1075/ijlcr.16016.sto

Sun, X. (2014). Ungrammatical Patterns in Chinese EFL Learners' Free Writing. *English Language Teaching, 7*(3), 176-183. https://doi.org/10.5539/elt.v7n3p176

Swan, M. (2005). *Practical English usage* (3rd ed., New international student's ed). Oxford University Press.

Swan, M., & Smith, B. (2001). *Learner English: a teacher's guide to interference and other problems* (2nd ed). Cambridge University Press.

Šamalová, M. (2018). *Using Pedagogical Translation in English Language Teaching at Grammar Schools: Current State and the Possibilities of Using the Contemporary Concept of Pedagogical Translation* [Dissertation, PedF MU]. https://is.muni.cz/auth/th/kc2f0/?lang=en

Šebesta, K., Lehečková, E., Pierścieniak, P., & Šormová, K. (2016). *Aplikovaná lingvistika: Příručka pro studenty Bc. studia ČJL* (1 ed.). Univerzita Karlova. www.karolinum.cz ebooks@karolinum.cz

Štindlová, B. (2011). *Evaluace chybové anotace v žákovském korpusu češtiny* [Dissertation, Ústav českého jazyka a teorie komunikace, Filozofická fakulta, Univerzita Karlova]. https://is.cuni.cz/webapps/zzp/detail/25046

Štindlová, B., & Čurdová, V. (2015). MERLIN: Multilingvální platforma pro evropské referenční úrovně. *Časopis pro moderní filologii, 97*(2), S. 190–200.

Tarone, E. (2013). Interlanguage. In C. Chapelle (ed.), C. Chapelle, *The encyclopedia of applied linguistics* (1st, pp. 1-7). Wiley-Blackwell.

Tarone, E., & Swierzbin, B. (2009). *Exploring Learner Language* (5th). Oxford University Press.

Taylor, D. (1988). The Meaning and Use of the Term 'Competence' in Linguistics and Applied Linguistics. *Applied Linguistics, 9*(2), 148-168. https://doi.org/https://doi.org/10.1093/applin/9.2.148

Thewissen, J. (2015). *Accuracy across proficiency levels: A learner Corpus Approach* (1st). UCL Presses Universitaires.

Timmis, I. (2002). Native-speaker norms and International English: a classroom view. *ELT Journal, 56*(3), 240-249. https://doi.org/10.1093/elt/56.3.240

Truscott, J. (1996). The Case Against Grammar Correction in L2 Writing Classes. *Language Learning, 46*(2), 327-369. https://doi.org/10.1111/j.1467-1770.1996.tb01238.x

Tůma, F. (2013). Korpusová analýza jazykové přesnosti v písemném projevu studentů angličtiny jako cizího jazyka. *Pedagogická orientace, 23*(1), 92-111. https://doi.org/10.5817/PedOr2013-1-92

Uçar, S., & Yükselir, C. (2017). A Corpus-Based Study on the Use of the Logical Connector 'Thus' in the Academic Writing of Turkish EFL Learners. *English Language Teaching, 10*(2), 64-72. https://doi.org/10.5539/elt.v10n2p64

Ullman, M. (2004). Contributions of memory circuits to language:: the declarative/procedural model. *Cognition, 92*(1-2), 231-270. https://doi.org/https://doi.org/10.1016/j.cognition.2003.10.008

Usó-Juan, E., & Martínez-Flor, A. (2006). *Current trends in the development and teaching of the four language skills: Approaches to language learning and teaching: Towards acquiring communicative competence through the four skills* (1st). M. de Gruyter.

Usó-Juan, E., & Martínez-Flor, A. (2008). Teaching Intercultural Communicative Competence through the Four Skills. *Revista Alicantina de Estudios Ingleses, 21*(2), 157-170. https://rua.ua.es/dspace/bitstream/10045/10400/1/RAEI_21_09.pdf

VanPatten, B. (1990). Attending to Form and Content in the Input. *Studies in Second Language Acquisition, vol. 12*(issue 3), 287-301. https://doi.org/10.1017/S0272263100009177

Vercellotti, M. (2017). The Development of Complexity, Accuracy, and Fluency in Second Language Performance: A Longitudinal Study. *Applied Linguistics, 38*(1), 90-111. https://doi.org/10.1093/applin/amv002

Weissberg, B. (2000). Developmental relationships in the acquisition of English syntax: writing vs. speech. *Learning and Instruction, 10*(1), 37-53. https://doi.org/10.1016/S0959-4752(99)00017-1

Widdowson, H. (1978). *Teaching Language as Communication* (5th). Oxford University Press.

Wierszycka, J. (2013). Phrasal Verbs in learner English: a semantic approach.: A study based on a POS tagged spoken corpus of learner English. *Research in Corpus Linguistics, 2013*(1), 81-93. http://www.academypublisher.com/ricl

Wong, S., & Dras, M. (2009). Contrastive Analysis and Native Language Identification. *The Association for Computational Linguistics*, 53-61.

Zafar, M. (2011). Monitoring the 'Monitor': A Critique of Krashen's Five Hypotheses. *Dhaka University Journal of Linguistics, 2*(4), 139-146. https://doi.org/10.3329/dujl.v2i4.6903

RESUMÉ

Hlavním cílem výzkumu představeného v této knize bylo ověřit, jak může cílená didaktická intervence ovlivnit přesnost mezijazyka pokročilých mluvčích angličtiny. Ke zkoumání přesnosti různých aspektů pokročilé angličtiny českých a slovenských studentů byla použita řada výzkumných nástrojů, tak jak je podrobně popsáno v podkapitole 4.6. Nejprve byla testována schopnost studentů identifikovat a opravit chyby, poté byla měřena přesnost jejich volné produkce v cílovém jazyce, a to pomocí korpusové analýzy. Byl sestaven anotovaný korpus žákovského jazyka, sloužící jako výzkumný nástroj k identifikaci těch oblastí pokročilého žákovského jazyka, které jsou chybami nejvíce zasaženy.

Účinek intervence byl nejprve ověřen u respondentů z pilotní skupiny, jejichž přesnost a jistota odpovídání byly testovány pomocí testu posuzování gramatické správnosti *Grammaticality Judgement Test (GJT)*, a dále dotazováním na míru jistoty při odpovídání *Certainty-Based Marking (CBM)*. V hlavní studii byly stejně testovány dvě skupiny, experimentální a kontrolní, a navíc byly u těchto skupin analyzovány volná mluvená a psaná produkce v cílovém jazyce. Pilotní a experimentální skupina se zúčastnila cílené didaktické intervence, která měla podobu jednosemestrálního kurzu kombinujícího přímou a online výuku. Tento kurz byl zaměřen na zlepšení přesnosti angličtiny jako cílového jazyka a podrobnosti o jeho průběhu jsou uvedeny v kapitole 3. Všechny skupiny absolvovaly pre-test a post-test, a výsledky tohoto testování byly porovnány a podrobeny statistické analýze.

Analýza odhalila, že nejpříznivější dopad měla intervence na zlepšení schopnosti studentů identifikovat a opravit chyby a na jistotu jejich odpovídání. Tyto dva aspekty se zlepšily statisticky významně jak u pilotní, tak u experimentální skupiny. Z analýzy rovněž vyplynulo, že se zlepšila přesnost mluveného projevu u experimentální skupiny, a toto zvýšení přesnosti bylo statisticky významné. Nečekaným výsledkem byl nepříznivý dopad intervence na přesnost psaní, v němž u experimentální

skupiny došlo oproti očekávání v post-testu ke snížení přesnosti a zvýšení chybovosti. V kontrolní skupině se statisticky významně zlepšila přesnost a jistota odpovídání, zatímco výsledky analýzy mluveného a psaného projevu nebyly statisticky průkazné. Podrobnosti analýzy všech sledovaných aspektů ve třech skupinách respondentů jsou uvedeny v 5. kapitole a diskuse a možná interpretace výsledků výzkumu jsou popsány v kapitole 6.

Výsledky srovnání mezi experimentální a kontrolní skupinou ukazují, že experimentální skupina se zlepšila ve třech ze čtyř sledovaných aspektů. Ve srovnání se skupinou kontrolní došlo pouze u experimentální skupiny ke statisticky významnému zlepšení přesnosti mluveného projevu. To ukazuje na celkově pozitivní dopad intervence na přesnost žákovského jazyka v experimentální skupině.

Důležitou součástí výzkumu bylo vytvoření databáze pokročilého mluveného a psaného žákovského jazyka v podobě žákovského korpusu. Tento korpus byl využit k analýze, jejímž cílem bylo identifikovat ty oblasti angličtiny, které u pokročilých českých a slovenských mluvčích vykazují nejvyšší chybovost. Psaný a mluvený žákovský jazyk byl rovněž analyzován samostatně, aby bylo možné stanovit, zda mezi psanou a mluvenou podobou existují rozdíly v typech a frekvenci výskytu chyb. Z analýzy vyplývá, že nejčastěji chybami postižené oblasti pokročilé žákovské angličtiny jsou gramatika, zejména chyby v užití členů, ať už jejich absence nebo nadužívání, a chyby lexikální, a to jak na úrovni jednotlivých slov, tak vět nebo jejich částí. Relativně časté byly rovněž chyby slovosledu a chyby lexikálně-gramatické, zatímco chyby stylistické, morfologické a chyby interpunkce se vyskytovaly zřídka. V souladu s výsledky, které vyplývají z dříve provedených a podobně zaměřených studií, blíže popsaných v kapitole 2, mnohé z těchto chyb lze přičítat negativnímu vlivu mateřského jazyka mluvčích. Přínosem tohoto výzkumu je mimo jiné i to, že zkoumal nejen mluvený, ale i psaný jazyk. Z hlediska chybovosti nebyly odhaleny žádné významné rozdíly mezi psaním a mluvením, dopad intervence na ně však byl na odlišný.

Vedle potvrzení některých dříve publikovaných výsledků, jak bylo popsáno v části 6.2, přináší tento výzkum nová důležitá zjištění o povaze pokročilého žákovského jazyka, zvláště identifikaci jazykových rysů, jejichž osvojení je pro české a slovenské studenty angličtiny problematické. Za jeden z největších přínosů tohoto výzkumu lze považovat vytvoření korpusu pokročilé žákovské angličtiny, v němž byla provedena chybová

anotace, a který byl doplněn o metadata, upřesňující údaje o respondentech. Tento korpus lze tedy do budoucna využít k dalšímu výzkumu různých aspektů pokročilého žákovského jazyka. Identifikace oblastí, v nichž se u českých a slovenských mluvčích angličtiny vyskytuje nejvíce chyb, má význam nejen pro další výzkum. Cílená práce s chybami se může stát nedílnou součástí výuky, a to zejména ve formě výukových materiálů, navržených s cílem typické chyby eliminovat a zvýšit povědomí o nich. Zjištění prezentovaná v tomto výzkumu tak mohou poskytnout solidní empirický základ pro navržení učebnice pro pokročilé české a slovenské studenty angličtiny, která se bude zaměřovat na problematické oblasti pro ně typické. Výsledky zde popsané analýzy pokročilého žákovského jazyka již byly úspěšně využity při tvorbě nového, zcela přepracovaného vydání oblíbené jazykové příručky English or Czenglish, která byla publikovaná v roce 2021 (Sparling et al.).

LIST OF FIGURES

LIST OF TABLES

APPENDICES

Appendix 1 The list of international and Czech scholarly journals reviewed in chapter 2

International journals

1) *Advances in Language and Literary Studies*
2) *English Language Teaching (ELT)*
3) *International Education Studies*
4) *International Journal of Corpus Linguistics (IJCL)*
5) *International Journal of Learner Corpus Research (IJLCR)*
6) *International Review of Applied Linguistics in Language Teaching (IRAL)*
7) *Journal of Education and Training Studies (JETS)*
8) *Language Learning and Technology: A Journal for Second and Foreign Language Educators (LLT)*
9) *Language Teaching Research (LTR)*
10) *Language Testing (LT)*
11) *Research in Corpus Linguistics (RiCL)*
12) *Research in the Teaching of English (RELC)*
13) *Second Language Research (SLR)*
14) *Teaching English with Technology (TEwT)*

Czech journals

1) *Acta Universitatis Carolinae Philologica*
2) *Časopis pro moderní filologii*
3) *e-Pedagogium*
4) *Lingua Viva*
5) *Naše řeč*
6) *Orbis Scholae*
7) *Pedagogická orientace*
8) *Pedagogika*
9) *Slovo a slovesnost*
10) *Studia pedagogica*

Appendix 2 Outlines of relevant research areas as published in the literature

Table A1
The outline of research areas addressed in the International Journal of Learner Corpus Research

Volume, year, issue	general texts on LC and LL	LC development and research	aspects of LL	tasks, context, proficiency	L1 influence and LL
1, 2015, 1	**Callies**[38] **Granger Gries**	**Tracy-Ventura**	**Maden-Weinberger Alexopoulou**		
1, 2015, 2		Ivaska	Pezik Gnevsheva		Kyle
2, 2016, 1		Vyatkina	Deshors Garner		Crosthwaite
2, 2016, 2	Deshors		Koch Callies Edwards Horch Rosen	Van Rooy	Schneider Brunner
3, 2017, 1		Paquot	Van Vuuren Rankin		Belz
3, 2017, 2	Trouvain	Detey	Niebuhr	Gut	Gósy Lecumberri
4, 2018, 1		Kwon	Schanding Kreyer		Stormbom Deshors
4, 2018, 2	Fuchs		Meriläinen Rautionaho Zhao Tracy-Ventura	Fuchs	
5, 2019, 1		Larsson		Lester	Hendrikx
5, 2019, 2	Brezina	Gablasova	Castello	Römer Pérez-Paredes	Götz Gilquin
TOTAL	7	8	21	6	12

Note. LC learner corpora, LL learner language

38 Only the first author of each text is mentioned.

Table A2

Details of studies on the influence of L1 and learner language published in the International Journal of Learner Corpus Research

year	2015	2016	2016	2016
author	**Kyle, K.**	**Crosthwaite, P.**	**Schneider, G.**	**Brunner, M.-L.**
research focus	the impact of lexical and phrasal choices by members of five L1 groups on native language identification (NLI)	L2 English article use by L1 speakers of article-less languages	verb + preposition structures (including phrasal verbs) and adjective + preposition structures	linguistic creativity in informal Skype conversations in ELF context
research participants	[39]			university students
research data (type, collection and analysis)	essays at two proficiency levels	575 written essays	measuring overuse by collocation measures, similarities and differences between EFL and ESL, quantitative and qualitative analyses	200–300 transcribed Skype conversations
L1 background	five L1 groups	Mandarin Chinese, Korean,Thai		German,Bulgarian, Italian, Spanish, Swedish, Finnish, French, Dutch
L2	English	English	English	English
corpus type		ICNALE	ICLE, BNC (reference)	CASE
language focus	written	written	written	spoken
proficiency		4 proficiency levels, not specified which		
research findings	lower proficiency writers make lexical and phrasal choices that are more similar to other lower proficiency writers that share an L1	overproduction of indefinite/ definite articles problematic for all three groups regardless of L1 background and L2 proficiency	more verb/adjective + preposition combinations than previous studies and significant roles of analogy and transfer	functionally accepted innovations, distinguishing instances of L1 influence, which seems to have a positive influence on the communicative setting, e.g. illustrated by code-switching

39 Not all types of information are provided in the abstracts for all studies. If that is the case, the window is left blank.

Table A3
Details of studies on the influence of L1 and learner language published in the International Journal of Learner Corpus Research

year	**2017**	**2017**	**2017**	**2018**
author	**Belz, M.**	**Gósy, M.**	**Lecumberri, M.L.G.**	**Stormbom, C.**
focus	frequency, form, and place of silent and filled pauses as well as self-repairs	forms and position of all filled pauses, and the durations and the formants of vocalic filled pauses	cross-language comparability, first language influences and non-native speech using the corpus is illustrated by means of acoustic, segmental, suprasegmental, and conversational phenomena	the choice of epicene pronouns, i.e. pronouns that refer to both sexes, in EFL contexts, the use of the epicene pronouns he, he or she, and they
research participants		30 young learners		
research data (type, collection, analysis)		corpus-based analysis	picture-based task	
L1 background	English	Hungarian	English, Spanish	
L2	German	English	English, Spanish	
corpus type		HunEng-D learner corpus	corpus of task-based conversational speech	two L1 and L2 corpora of student writing
language focus	spoken	spoken	spoken	written
proficiency		various		
research findings	learners adhere to the pattern of their native language English, for some aspects of disfluencies learners can adapt to a native-like pattern, while others are imported from the L1	forms of filled pauses were similar in both languages, irrespective of level of language proficiency, significantly longer vocalic filled pauses in basic and intermediate learners in their L2 relative to more advanced peers, similar articulatory configurations for all vocalic filled pauses	different types of interactions between the first language, the second language, and non-nativeness revealed	L2 English speakers use he significantly more than the L1 speakers, whereas the L1 speakers use they more. Variation found in the L2 subcorpora seems to be partly related to L1 influence: Those using he the most are speakers of gendered L1s with a traditional practice of masculine generics.

Table A4

Details of studies on the influence of L1 and learner language published in the International Journal of Learner Corpus Research

year	**2018**	**2019**	**2019**	**2019**
author	**Deshors, S.**	**Hendrikx, I.**	**Götz, S.**	**Gilquin, G.**
research focus	the uses of the present perfect (PP) and simple past (SP) by French learners of English and how those uses differ from those in native English and those of the passé composé (PC) in native French which, semantically, overlaps with PP and SP	the cross-linguistic influence and the (longitudinal) impact of Content and Language Integrated Learning (CLIL) on the acquisition of intensifying constructions	filled pauses across proficiency levels	light verb constructions (LVCs), e.g. take a walk, make a choice in spoken L2
research participants			1,244	1,244
research data (type, collection and analysis)	3,000 contextualized occurrences of PP, SP and PC, and includes cluster and collostructional analyses		corpus-based analysis	corpus-based analysis
proficiency			B1-C2	B1-C1
L1 background	French, English	English, French, Dutch	Italian, Spanish, Russian, Chinese, Hindi	
L2	English, French	English, Dutch	English	English
corpus type		corpora of French, Dutch, and English productions by L1 speakers, and L2 English and L2 Dutch produced by French-speaking learners in CLIL and FLT	the Trinity Lancaster Corpus Sample (the TLC)	the Trinity Lancaster Corpus Sample (the TLC)
language focus		written	spoken	spoken

year	2018	2019	2019	2019
author	**Deshors, S.**	**Hendrikx, I.**	**Götz, S.**	**Gilquin, G.**
research findings	advanced learners have integrated the uses of past tenses and that the influence of the PC is relatively weak, at an upper-intermediate to advanced proficiency level, learners have integrated the fine-grained contextual information characteristic of the use of English past tenses	cross-linguistic similarities between English and French result in fewer mistakes in the English learners' use of intensifiers, more input through CLIL correlates positively with a more target-like use of intensifiers	the number of filled pauses significantly different across proficiency levels, other learning context variables, such as the country of origin, the age of acquisition or the examiners' experience have a much stronger effect on the learners'/ speakers' use of filled pauses	frequency of LVCs increases with higher proficiency levels, ESL speakers use LVCs more than NS

Table A5
An outline of relevant international studies

year	**2013**	**2013**	**2014**	**2014**
author	**Wiersycka, Joanna**	**MacDonald, García-Carbonell, Carot-Sierra**	**Hamed, Muftah**	**Sun, X.**
research focus	underuse of phrasal verbs (PVs) by Polish learners of English	- computer-aided EA of IL errors in synchronous and asynchronous communication - comparing amount and type of errors in 1)synchronous and asynchronous communication, and 2) depending on the L1 backgrounds	use of conjunctions in argumentative writing of Libyan tertiary students	ungrammatical patterns in Chinese EFL learners' free writing
research participants	50 university students, majoring in English	126 students at 5 European tertiary education institutions	16 fourth-year undergraduates majoring in English	30 undergraduates majoring in English
research data (type collection and analysis	comparison of PVs usage between NSs and NNSs: contrastive interlanguage analysis	- computer-mediated communication: forum discussions and IDEELS simulations: -synchronous exchanges (on-line conferences held under real-time conditions) - asynchronous exchanges (e-mail type communication)	- questionnaire: participants' demographic data - two argumentative essays, 200-250 words, written by each student (n = 16) on two different topics	free writing of about 250-word on free topics
research methodology	partially automatic extraction and a subsequent manual filtering of PVs from a POS-tagged NNS corpus and its reference NS corpus	- online teleconferences and email communication error-coded - the programme used: the Louvain error editor + the error-coding guide	- coding system based on Halliday+Hassan's taxonomy (1976) of conjunctions - appropriacy of conj.s discussed	- NS evaluation for errors - typical error patterns are identified:using the error analysis model of Keshavarz (1994)
proficiency	advanced	intermediate (minimum B1) – advanced	advanced	advanced

year	2013	2013	2014	2014
author	**Wiersycka, Joanna**	**MacDonald, García-Carbonell, Carot-Sierra**	**Hamed, Muftah**	**Sun, X.**
L1 background	Polish	German, Norwegian, Spanish, Latvian, French	Arabic	Mandarin Chinese
L2	English (FL)	English (FL)	English (FL)	English (FL)
corpus type	the PLINDSEI the LOCNEC	the MiLC	none	none
corpus size	95,906 words in the learner corpus 118,554 in the NS corpus	42,059 words – synchronous 42,625 words asynchronous	NA	NA
language focus	spoken	written	written	written
research findings	- native use of PVs: linear - Polish learners underuse PVs, especially those idiomatically opaque PVs - this is not related to proficiency levels but might be related to how PVs are taught	- overall, more errors in the synchronous communication, learners focus more on fluency+meaning than on accuracy - in the synchronous: more errors in form +grammar - in the asynchronous: more errors in lexis and style - frequency of error types varied within each different L1: L1 clearly influences specific error types - question for further research: to what extent are L1-specific errors due to negative interference from the native language?	most participants: inappropriate use of conjunctions in writing, confusing their semantic functions -causes: attributable to L1 (Arabic) negative transfer – more research needed	- misuse of determiners is the most frequent grammatical error - followed by wrong L1-induced'Chinese-English pattern' (Chinglish)[40], tense errors, and misuse of prepositions, lack of subject-verb agreement and misuse of adverbials - one important cause of these errors is the inappropriate Language transfer
relevance	medium	high	medium	high

40 "Wei and Fei (2003) define Chinese English (Chinglish) as an interlanguage, usually manifested as Chinese-style syntax with English words, Chinese phonological elements in pronunciation or grammatical variations that attempt to follow Standard English rules but miss the mark" (Sun, 2014, p. 179).

Table A6
An outline of relevant international studies

year	2016	2017	2018	2019
author	**Karazoun, Al**	**Uçar, Serpil**	**Long, Hatcho**	**Shimanskaya, Slabakova**
research focus	a linguistic analysis of errors made by Jordanian EFL learners in translating newspaper headlines	a corpus-based-study on the use of the logical connector "thus" in the academic writing of Turkish EFL learners	grammatical accuracy of Japanese EFL learners - most frequent error types - causes of these errors	L1-L2 differences in the L2 classroom: anticipating Anglophone learners' difficulties with French pronoun interpretation
research participants	60 undergraduate students majoring in English	20 Turkish speakers of EFL publishing articles in linguistics	61 university students, aged 18-21	- 68 learners of French as SL - 43 native controls – all university students or university graduates
research data (type, collection and analysis)	a test composed of (30) English news headlines and (30) Arabic ones covering various areas of news occurring in a large corpus of Jordanian newspapers	corpus-based analysis of written data collected from 2 NS corpora + 1 NNS corpus	61 transcripts of spontaneous speech	- picture selection task - related test in ambiguous pronouns
research methodology	two major categories of errors lexical-semantic and syntactic-morphological	- raw frequencies of 'thus', frequencies per million words, frequencies per text and log-likelihood ratio were measured – varieties compared across the three corpora	error analysis	- statistical analysis: a repeated-measures ANOVA - post hoc tests for each answer using the Bonferroni correction
proficiency	advanced	advanced	B1 (IELTS 3-4)	beginners (n = 38) intermediate (n = 27) advanced (n = 22)

year	2016	2017	2018	2019
author	**Karazoun, Al**	**Uçar, Serpil**	**Long, Hatcho**	**Shimanskaya, Slabakova**
L1 background	Arabic	Turkish	Japanese	English
L2	English (FL)	English (FL)	English (FL)	French (SL)
corpus type	a sample of (30) English and (30) Arabic news headlines randomly selected from a large learner corpus: - corpus of Jordanian newspapers[41]	experts' corpora: - the COCA - NES[42] – 20 scientific articles of native speakers as used as control corpora[43] - NNES – 20 scientific articles of Turkish advanced EFL learners	the Japanese University Student Corpus (JUSC)[44]	none
corpus size	not specified	COCA 91,066,191 NES 273.560 NNES 257.848	51,061 words	none
language focus	written (translation)	written	spoken	NA

41 The corpus is not further specified in the reviewed research paper.

42 The articles were gathered from these journals: Journal of Pragmatics, Lingua, and Cognition.

43 The articles were gathered from these journals: Journal of Pragmatics, Lingua, System and Journal of Second Language Writing, and written by Turkish NNS, published between 2005-2015

44 The corpus can be found at www.genderfluency.com.

year	2016	2017	2018	2019
author	**Karazoun, Al**	**Uçar, Serpil**	**Long, Hatcho**	**Shimanskaya, Slabakova**
research findings	- most frequent error types: grammatical, discourse and lexical - inadequate knowledge of the English headlines rules - causes: the most syntactic-morphological error: the use of typical Arabic construction in English and errors in articles are likely to be due to inter-lingual interference	- Turkish EFL learners underuse the connector 'thus' in their academic prose compared to native speakers - significant differences found between NS and learner corpora - connectors troublesome not only for Turkish advanced learners but also for other learners from different linguistic backgrounds	- the primary errors were articles, verb tense, prepositions, omission, modifier errors, and subject-verb agreement -these results indicate that L1 is a major factor in grammatical accuracy	- the study demonstrated that the process of learning an L2 starts with an attempt to impose native language categories on the new language - careful linguistic analysis allows to predict transfer patterns and specific errors - efficient grammar teaching to learners who share the same L1 should take advantage of these findings by drawing parallels and highlighting differences between the L1 and L2 grammatical meanings
relevance	medium	medium	very high	very high

Table A7
An outline of relevant Czech studies

year	**2013**	**2015**	**2017**
author	**Tůma, František**	**Gráf, Tomáš**	**Gráf, Tomáš**
research focus	communicative competence in writing, focus on language accuracy	accuracy and fluency of advanced learner language	inaccuracies in the use of verbal categories
research participants	18 university students, not majoring in English	50 university students majoring in English	50 university students majoring in English
research data (type, collection and analysis)	3 online discussion tasks, written in asynchronous fora	3 tasks: a monologue on a preselected topic, a dialogue and a picture-based narrative	3 tasks: a monologue on a preselected topic, a dialogue and a picture-based narrative
research methodology	learner corpus error analysis of selected aspects, verbs	Computer-aided Error Analysis within the larger framework of Learner Corpus Research	Computer-aided Error Analysis within the larger framework of Learner Corpus Research
proficiency	A2	C1-C2	C1-C2
L1 background	Czech	Czech	Czech
L2	English (FL)	English (FL)	English (FL)
corpus type	a learner corpus compiled from the texts written by the students in two online discussions	spoken learner corpus LINDSEI_CZ	spoken learner corpus LINDSEI_CZ
corpus size	13,622 words	125,000 tokens	125,000 tokens
language focus	written	spoken	spoken

year	2013	2015	2017
author	**Tůma, František**	**Gráf, Tomáš**	**Gráf, Tomáš**
research findings	- overall level of communicative competence increased - analysis of accuracy of conclusive- learners' accuracy was linked to the communicated content	- task design has an effect on spoken production - particular error types indicate the possible effect of Language transfer - basic errors frequent in advanced learner language - most frequent error-types: use of articles and tenses - pedagogical implications: "the use of articles and tenses must be targeted at all proficiency levels including the most advanced" (p. 162) - supplementary teaching materials should be devised addressing these problematic areas - raising awareness, practice and explicit instruction is effective	- the key area of difficulty is the use of tenses and tense agreements, especially the use of the present perfect - other error-prone aspects are also described - a number of triggers of errors identified: - deficiencies in the teaching of grammar, mainly too much focus on decontextualized practice, use of potentially confusing rules, - causes of errors: the effect of negative Language transfer or the existence of various triggers or false associations - pedagogical implications: work with advanced learners should systematically target these error-prone areas not leaving the L1 aside, and also a reconsideration of how these problematic areas are approached in the teaching process
relevance	medium	very high	very high

Appendix 3 The list of corpora used in the reviewed research papers

1) COCA – the Corpus of Contemporary American English
2) JUSC – the Japanese University Student Corpus
3) LINDSEI – the Louvain International Database of Spoken English Interlanguage
4) LINDSEI_CZ – the Czech subcorpus of LINDSEI, spoken learner corpus
5) LOCNEC – the Louvain Corpus of Native English
6) MiLC Corpus – a multilingual learner corpus of texts written by language learners from different language backgrounds
7) MCSAW – the Malaysian Corpus of Students' Argumentative Writing
8) PLINDSEI – the Polish component of LINDSEI

Appendix 4 The complete list of required and optional resources used on the course described in chapter 3

Required materials:

1) Hewings, M. (2009). *Grammar for CAE and proficiency with answers: self-study grammar reference and practice.* 1st ed. Cambridge: Cambridge University Press. ISBN 9780521713757.
2) Sparling, T. D. (1991). *English or Czenglish? Jak se vyhnout čechismům v angličtině.* 2*nd* ed. Praha: SPN, 1992. 250 pp. ISBN 80-04-25969-3.

Recommended materials:

1) French, A. (2002). *CAE testbuilder: with answer key.* Oxford: Macmillan. 192 pp. ISBN 1-4050-1400-8.
2) Hewings, M. (1999). (Editor). *Advanced grammar in use: a self-study reference and practice book for advanced learners of English.* 1st ed. Cambridge: Cambridge University Press, 340 pp. ISBN 0-521-49868-6.
3) McCarthy, M.; O'Dell, F.(2006). *English Vocabulary in Use, Advanced.* Cambridge: Cambridge University Press. 315 pp. ISBN 0-521-67746-7.
4) Moore, J. (2005). *Common Mistakes at Proficiency…and How to Avoid Them.* 1st ed. Cambridge: Cambridge University Press ISBN-10: 0521606837
5) Murphy, R. (1995). (Editor). *English Grammar in Use: a self-study reference and practice book for intermediate students.* 2nd ed. Cambridge: Cambridge University Press. 350 pp. ISBN 0-521-43680-X.

6) Powell, D. (2006). (Editor). *Common Mistakes at CAE: and How to Avoid Them.* 3rd ed. Cambridge: Cambridge University Press, 64 pp. ISBN 978-0-521-60377-5.

7) Side, R.; Wellman, G. (2006). *Grammar and Vocabulary for Cambridge Advanced and Proficiency: fully updated for the revised CPE.* 6th ed. Harlow: Longman. 288 pp. ISBN 0-582-51821-0.

8) Swan, M. (2005). *Practical English usage.* 3rd ed., New international student's ed). Oxford: Oxford University Press.

Appendix 5 Grammaticality Judgement Test – full version

A (1-10 Articles, 11-15 Countability)

Correct the errors in the following sentences if necessary, rewriting them in the box below. If you think the sentence is correct, write ok.

1. Each spring students send their designs for homes *in the space* for judging by NASA engineers.
   a. Each spring students send their designs for homes *in space* for judging by NASA engineers.
   b. English Web 2012[45]

2. Being on duty for seventy hours does give you a taste of what *the life* as a doctor is like.
   a. Being on duty for seventy hours does give you a taste of what *life* as a doctor is like.
   b. Grammar, p. 133

3. She delivered a talk on *the life of* two women who gave up their family wealth to embark upon *a life of* voluntary poverty.
   a. OK
   b. English Web 2012

4. Under any circumstances can *the violence* ever be justified?
   a. Under any circumstances can *violence* ever be justified?
   b. Grammar, p. 128

5. I have been particularly interested *in nature of* computer literacy at the university level.
   a. I have been particularly interested *in the nature of* computer literacy at the university level.
   b. English Web 2012

6. Modern man has lost all touch with *the nature* and with the sources of supply of his basic needs.
   a. Modern man has lost all touch with *nature* and with the sources of supply of his basic needs.
   b. English Web 2012

45 The examples described as "English Web 2012" have been taken from Sketch Engine: https://app.sketchengine.eu/#dashboard?corpname=preloaded%2Fententen12_1

7. The main requirement for writing a convincing essay, aside from a command of *English language*, is to be who you really are.
   a. The main requirement for writing a convincing essay, aside from a command of the *English language*, is to be who you really are.
   b. English Web 2012

8. The Masini Hotel Forli is a modern and prestigious property set in *the city centre*, near all the attractions.
   a. OK
   b. English Web 2012

9. He gained his doctorate with a thesis on the seagull.
   a. OK
   b. Grammar

10. The artists chosen for the most sought-after Valentine's banquet are musicians Dana Jones *on violin* and Kevin Crane *on piano*.
    a. OK
    b. English Web 2012

11. The scope and type of work vary but will involve active participation in *an ongoing research*.
    a. The scope and type of work vary but will involve active participation *in ongoing research*.
    b. English Web 2012

12. The subtle energies of healing are not meant as a substitute for seeking medical *advices* from a qualified medical practitioner.
    a. The subtle energies of healing are not meant as a substitute for seeking medical *advice* from a qualified medical practitioner.
    b. English Web 2012

13. It is amazing to see how fast *the good news travel*.
    a. It is amazing to see how fast the good news travels.
    b. English Web 2012

14. Conditions in the prisons remain atrocious, and human rights *abuses* a serious problem.
    a. OK
    b. English Web 2012

15. One evening, driving through *a heavy traffic* I made up a story to entertain my kids.
    a. OK
    b. English Web 2012

B (Recurrent errors)

Complete the second sentence so that it has a similar meaning to the first sentence. You must use between three and six words, including the word in capitals, without changing it.

16. If your goal is to get targeted visitors, this is how you should do it.

a. If your goal is to get targeted visitors, this …*is the way to do*…… it. (WAY)
b. English Web 2012

17. If the cake tastes good, who cares *about its appearance*?
    a. If the cake tastes good, who cares ……*what it looks*……. like? (LOOKS)
    b. English Web 2012

18. It was a new experience for me – I had never cooked for so many people before.
    a. It was a new experience for me – I ….. *was not used to cooking for*……. so many people. (USED)
    b. Swan, p. 605 – adapted

19. They need to decide how they're going to travel to Rome. (DECISION)
    a. They need to ……*make/take a decision about how to*….travel to Rome.
    b. Advanced Trainer, p. 144 – adapted

20. This pub used to be a lot more popular before they opened the new one just round the corner. (LESS)
    a. Since the new pub round the corner was opened, this one …..*has been a lot less*…. popular.

C (False friends)

Correct the errors in the following sentences if necessary, rewriting them in the box below. If you think the sentence is correct, write ok.

21. At least once every five years *must all investigators complete* training in ethics.
    a. At least once every five years *all investigators must complete* training in ethics.
    b. based on an erroneous example from Pokorná's thesis: *Every 2 years is renewed one third of Senate*
    c. sentence from English Web 2012 – adapted

22. Let me quote the last paragraph of the paper, which *according to me*, is the most disturbing of them all.
    a. Let me quote the last paragraph of the paper, which *in my opinion*, is the most disturbing of them all.
    b. English Web 2012

23. Should I tell him? she thought. *He has a birthday*, after all, and it would be a great present telling him I'm in love with him.
    a. Should I tell him? she thought. It is his birthday, after all, and it would be a great present telling him I'm in love with him.
    b. English Web 2012

24. If used correctly, my contacts last longer than usually before getting all itchy and dry.
    a. If used correctly, my contacts last longer than usual before getting all itchy and dry.
    b. English Web 2012

25. They have years of experience, providing students with the most actual and practical knowledge of the field.
    a. They have years of experience, providing students with the most up-to-date and practical knowledge of the field.
    b. English Web 2012

D (Prepositions)

Correct the errors in the following sentences if necessary, rewriting them in the box below. If you think the sentence is correct, write ok.

26. Twenty stores will also host sessions where experts will be on hand to better *explain* householders how to use the new forms of lighting.
    a. Twenty stores will also host sessions where experts will be on hand to better *explain to* householders how to use the new forms of lighting.
    b. English Web 2012

27. The room has a large panoramic arched window with *a view on* the first terrace.
    a. The room has a large panoramic arched window with *a view of* the first terrace.
    b. English Web 2012

28. Over 95% of *visitors of* London are familiar with the exploits of the great detective.
    a. Over 95% of *visitors to* London are familiar with the exploits of the great detective.
    a. English Web 2012

29. Our attorney is available to sit down and *discuss about* the details of your case.
    a. Our attorney is available to sit down and *discuss* the details of your case.
    b. English Web 2012

30. I arrived at the checkout completely out of breath.
    a. OK
    b. Grammar, p. 232

The following resource materials were used in devising the test:

1) Corpus: English Web 2012

2) Side, R.; Wellman, G.: *Grammar and Vocabulary for CAE and CPE*, Pearson Educational Limited, Longman 2001

3) Sparling, D. (1989). *English or Czenglish? Jak se vyhnout čechismům v angličtině?*

4 Swan, M. (2001). *Learner English: a teacher's guide to interference and other problems.* Cambridge: CUP

Appendix 6 Questions used in pre-test and post-test in speaking

The test in speaking was modelled on Cambridge advanced English (CAE) oral exam, and consisted of three parts: 1) introductory individual turn, 2) collaborative task and 3) long individual turn. In the pre-test, conducted in October 2016, the following topics were used:

1) daily routine/ English studies 2) technology /how people learn 3) which areas money can affect most. In the post-test, conducted and video-recorded three months later, in January 20107, the following topics were discussed: 1) likes and dislikes about the CR/learning something new 2) city versus country / how people communicate 3) how to cope with stress (for details, see below).

PART 1: INTRODUCTIONS (2+2+2+2 minutes)

First of all, I'd like to know a little about you.

A Could you please tell me something about your daily routine?

(Are you an early or a late riser? Do you work best in the morning or in the evening? What's the busiest time of the day for you? And finally, what would you like to have more time for?)

B Could you please tell me something about your English studies?

(How long have you been studying English? What do you enjoy most about learning English? Have you been to any English-speaking countries? Can you speak any other languages?)

C Could you please tell me something about the place where you live?

(What is the best and worst thing about your town? Would you like to spend your whole life in this place? What are the advantages of living abroad? If you could live in another country, where would you choose?)

D Could you please tell me who has had the greatest influence on your life so far?

(Has this adapted over the years? Are you easily influenced by other people? What makes a good friend? Which teacher will you always remember?)

Appendix 7 Questions used in pre-test and post-test in writing

The test in writing was a short, 250-word argumentative essay. The topics were chosen from the International English Language Testing System database (IELTS). This is an authoritative test measuring language proficiency, recognized worldwide by more than 700 universities and tertiary education institutions, taken annually by 2 million applicants. The tests are freely available at http://www.ielts.org/teachers.aspx.

Essay topic used in pre-test

> Write a short essay (at least 250 words) about the following topic:
> In today's job market it is far more important to have practical skills than theoretical knowledge. In the future, job applicants may not need any formal qualifications.
> To what extent do you agree or disagree?
> You have 30 minutes to complete the task.

Essay topic used in post-test

> Write a short essay (at least 250 words) about the following topic:
> With the development of technology children are now living in a world that is completely different to what it was 50 years ago.
> What problems does this cause for society and the family?
> You have 30 minutes to complete the task.

Appendix 8 Questionnaires providing contextual information

Pre-test questionnaire

> My motivation to take this course. Why did you decide to take this course? Please, specify.

Post-test questionnaire

1) Was the test you took today the same as the one you took in October?
   a) Yes, it was the same.
   b) No, it was different.
   c) I can't remember.
2) Please, comment on the test you took today.
3) Please, comment on the test you took in October.
4) What have you done this term to improve your English? Choose as many options as appropriate from the list below and add more details in the following questions.
   a) I have taken an English language course.
   b) I have studied intensively on my own.
   c) I have visited an English-speaking country.
   d) I have been meeting native English speakers
   e) I have read English books

f) I have watched films, series, or other in English
g) Other (please, specify below, in question 5).

5) What else have you done this term to improve your English?

6) If you have taken an English language course, please, specify the type of course, level, how frequently classes were held and add any other relevant comments.

7) If you have studied intensively on your own, please specify how you have studied, e.g. what textbooks you have been using, how frequently you have studied, what you have focused on, etc.

8) If you have visited an English-speaking country, please specify the length of your stay, the country you visited, etc.

9) If you have been meeting native English speakers, please give any relevant details.

10) If you have read English books, please give any relevant details (type of books, number of books, etc.).

11) If you have watched films, series or other in English, please give any relevant details.

12) Was the essay you wrote today more difficult than the one in October?
a) Yes, it was the same.
a) No, it was different.
a) I can't remember.

13) Please, comment on the essay you wrote today.

14) Please, comment on the essay you wrote in October.

15) Was the oral interview you took today more difficult than the one in October?
a) Yes, it was the same.
b) No, it was different.
c) I can't remember.

16) Please, comment on the oral interview you took today.

17) Please, comment on the oral interview you took in October.

18) Please, add any more comments if you like.

I appreciate that you've taken the time and energy to share your thoughts. I promise that I deal with the information you were kind enough to provide with care and respect.

THANK YOU!

Simona Kalová

Appendix 9 The Louvain error tagging system (adapted from Bestgen et al., 2012, p. 130 and complemented for this study)

ERROR DOMAIN	CODE	CATEGORY	DESCRIPTION
F	**FORM**		spelling and morphological errors
	FM	morphology	prefixes, suffixes
	FMADJ	adjectives	Also used if a wrong part of speech is used with the same root as the required word: e.g.
	FMADV	adverbs	
	FMD – new	determiners	
	FMN	nouns	Not the errors in number, e.g. missing plural ending, Not for irregular plural errors, e.g. – *womans* – woman, then it is GNN. // error tagging: <err type="FMN">genuinety</err><corr type="FMN">genuineness</corr> //**pickpocketers* – pickpockets //*creation – creativity //Also used when parts of speech are confused, e.g. *educational system – education system (CN1S).
	FMV	verbs	
	FS	spelling	double letters, also capital letters (errors in punctuation, e.g. apostrophes, inverted commas, are Q, not F)
	(FT)	typos	if they are clearly typos, ignore and correct in the text
G	**GRAMMAR**	errors affecting	
	GADJ	adjectives	
	GADJCS		comparative/superlative (also the use of comparative where not appropriate)

ERROR			
DOMAIN	**CODE**	**CATEGORY**	**DESCRIPTION**
F	**FORM**		spelling and morphological errors
	GADV – not used as such, but with additional letters indicating e.g. confusion, comparative and/or superlative errors	adverbs	missing
	GADVC	new	confusion
	GA – not used as such, but with additional letters indicating e.g. articles missing, wrong, or redundant	articles	
	GAMD		missing definite
	GAMI		missing indefinite
	GAR – not used as such, but with additional letters indicating in/ definite articles		redundant
	GARD		redundant definite
	GARI		redundant indefinite
	GAWD		wrong definite instead of indefinite
	GAWI		wrong indefinite instead of definite
	GD	determiners	FOLLOWED BY A NOUN (He gave me this diamond ring. – Macmillan Dictionary)
	GDD		demonstrative (missing, redundant, wrong)
	GDP		possessive (missing, redundant, wrong)
	GDI		indefinite (missing, redundant, wrong) other pronouns (other, another)

ERROR DOMAIN	CODE	CATEGORY	DESCRIPTION
F	**FORM**		spelling and morphological errors
	GDQ	new	expressing quantity and intensity e.g. much, many, (a) lot(s) of, plenty, (a) great deal (of) enough, too, all, both, several, couple, SUCH, SO BUT not when caused by errors in countability, then it is XNUC
	GN	nouns	
	GNN		number, also errors in irregular plural forms; e.g. – womans – woman
	GP	pronouns	WITHOUT A FOLLOWING NOUN (This is the photograph you asked for. Macmillan Dictionary)
	GPP		personal (missing, redundant, wrong)
	GPI		indefinite (missing, redundant, wrong) + other pronouns (other, another)
	GPD		demonstrative
	GPF		reflexive + reciprocal (missing, redundant, wrong)
	GPR		relative and interrogative (missing, redundant, wrong)
	GPU		unclear or erroneous reference
	GV	verbs	
	GVAUX		auxiliary – wrong or missing or redundant
	GVMOD	new	modal – wrong or missing or redundant, also in 'if clauses'
	GVNF		non-finite/finite
	GVT		grammar verb tense
	GVTS	new	simple form instead of progressive
	GVTPR	new	progressive form instead of simple
	GVV		voice also with participles
	GSVA	new	subject verb agreement
L	**LEXICAL**		prepositions phrases /collocations
	LCC		conjunction confusion
	LCM		conjunction missing

ERROR DOMAIN	CODE	CATEGORY	DESCRIPTION
F	**FORM**		spelling and morphological errors
	LS		single words (false friends and others)
	LSP		single word which is a preposition (not dependent but e.g. in front of a noun), erroneous/missing/ redundant use of the preposition.
	LP		lexical phrase (e.g. collocations, false friends, Czenglish)
X	**LEXICO-GRAMMATICAL**		
	XADJPR		adjectives used with wrong or missing dependent preposition
	XADJCO	new	adjectives used with wrong complementation
	XADVPR	new	adverbs used with wrong/ missing dependent preposition
	XADVCO	new	adverbs used with wrong complementation
	XNCO		nouns used with wrong complementation
	XNPR		nouns used with wrong/missing dependent preposition
	XNUC		errors on countable / uncountable nouns also errors caused by this, affecting e.g. quantifiers and determiners, ALSO verbs
	XVPR		verbs used with wrong/ missing/ redundant dependent preposition // also wrong particle in phrasal verbs
	XVCO		wrong verb complementation
W	**WORD**		words missing or redundant / erroneous word order
	WRS		word redundant single; including fillers in spoken language
	WRSC		word redundant single in Czech/Slovak
	WRM		word redundant multiple; including fillers in spoken language
	WMS		word missing single

ERROR			
DOMAIN	**CODE**	**CATEGORY**	**DESCRIPTION**
F	**FORM**		spelling and morphological errors
	WMM		word missing multiple
	WO		word order
S	**STYLE**		
	SU		sentence unclear
	SI		sentence incomplete
	SR	new	sentence redundant
	STL	new	sentence too long
Q	**PUNCTUATION**		
	QC		confusion e.g. mixing up Czech and English inverted commas 2 sentences need to be joined into 1 using incorrect punctuation
	QM		missing
	QR		redundant
Z	**INFELICITIES**		
	ZIR		inappropriate register
	ZPC		problems with political correctness
	ZS		stylistics

INDEX

Accuracy Matters
Exploring the Accuracy of Advanced Learner English in Czech Tertiary Education

PhDr. Simona Kalová, Ph.D.

Series: Cizí jazyky a jejich didaktiky: teorie, empirie, praxe
Volume 14

Published by Masaryk University Press, Žerotínovo nám. 617/9, 601 77 Brno, CZ
Cover design by Bc. Alena Poláčková
Typesetting by Bc. Klára Šimková
First electronic edition, 2022

ISBN 978-80-280-0335-7

Advanced learners of English tend to be fluent but not very accurate is surely an observation made by many language teachers (not exclusively) at universities. This book presents a corpus-based analysis of advanced learner English, both written and spoken, in an attempt to provide evidence that accuracy does indeed need attention from both learners and teachers. While the topic has been frequently discussed, authentic research material has been missing and thus the annotated learner corpus, complemented with metadata which was devised as an analytical tool for this research, provides opportunities for future studies of different aspects of learner language. Identifying areas which tend to be problematic for advanced Czech and Slovak learners of English offers some valuable insights for language practitioners. The book also has some international implications, looking into the influence of L1 on the acquisition of L2 and the influence of a targeted educational intervention on the accuracy of learner language.

Cizí jazyky a jejich didaktiky: teorie, empirie, praxe

MUNI PRESS

MUNI PED